arXiv:1402.2840v3 [cs.LO] 30 Jun 2014

Robust Synchronization in Markov Decision Processes*

Laurent Doyen[1], Thierry Massart[2], and Mahsa Shirmohammadi[1,2]

[1] LSV, ENS Cachan & CNRS, France
[2] Université Libre de Bruxelles, Belgium

Abstract. We consider synchronizing properties of Markov decision processes (MDP), viewed as generators of sequences of probability distributions over states. A probability distribution is p-synchronizing if the probability mass is at least p in some state, and a sequence of probability distributions is weakly p-synchronizing, or strongly p-synchronizing if respectively infinitely many, or all but finitely many distributions in the sequence are p-synchronizing.

For each synchronizing mode, an MDP can be (i) *sure* winning if there is a strategy that produces a 1-synchronizing sequence; (ii) *almost-sure* winning if there is a strategy that produces a sequence that is, for all $\varepsilon > 0$, a $(1\text{-}\varepsilon)$-synchronizing sequence; (iii) *limit-sure* winning if for all $\varepsilon > 0$, there is a strategy that produces a $(1\text{-}\varepsilon)$-synchronizing sequence.

For each synchronizing and winning mode, we consider the problem of deciding whether an MDP is winning, and we establish matching upper and lower complexity bounds of the problems, as well as the optimal memory requirement for winning strategies: (a) for all winning modes, we show that the problems are PSPACE-complete for weak synchronization, and PTIME-complete for strong synchronization; (b) we show that for weak synchronization, exponential memory is sufficient and may be necessary for sure winning, and infinite memory is necessary for almost-sure winning; for strong synchronization, linear-size memory is sufficient and may be necessary in all modes; (c) we show a robustness result that the almost-sure and limit-sure winning modes coincide for both weak and strong synchronization.

1 Introduction

Markov Decision Processes (MDPs) are studied in theoretical computer science in many problems related to system design and verification [24,15,10]. MDPs are a model of reactive systems with both stochastic and nondeterministic behavior, used in the control problem for reactive systems: the nondeterminism represents the possible choices of the controller, and the stochasticity represents the uncertainties about the system response. The controller synthesis problem is

* This work was partially supported by the Belgian Fonds National de la Recherche Scientifique (FNRS), and by the PICS project *Quaverif* funded by the French Centre National de la Recherche Scientifique (CNRS).

to compute a control strategy that ensures correct behaviors of the system with probability 1. Traditional well-studied specifications describe correct behaviors as infinite sequences of states, such as reachability, Büchi, and co-Büchi, which require the system to visit a target state once, infinitely often, and ultimately always, respectively [3,2].

In contrast, we consider symbolic specifications of the behaviors of MDPs as sequences of probability distributions $X_i : Q \to [0, 1]$ over the finite state space Q of the system, where $X_i(q)$ is the probability that the MDP is in state $q \in Q$ after i steps. The symbolic specification of stochastic systems is relevant in applications such as system biology and robot planning [18,6,14], and recently it has been used in several works on design and verification of reactive systems [21,9,1]. While the verification of MDPs may yield undecidability, both with traditional specifications [5,17], and symbolic specifications [21,13], decidability results are obtained for *eventually synchronizing* conditions under general control strategies that depend on the full history of the system execution [14]. Intuitively, a sequence of probability distributions is eventually synchronizing if the probability mass tends to accumulate in a given set of target states along the sequence. This is an analogue, for sequences of probability distributions, of the reachability condition.

In this paper, we consider an analogue of the Büchi and coBüchi conditions for sequences of distributions [11]: the probability mass should get synchronized infinitely often, or ultimately at every step. More precisely, for $0 \le p \le 1$ let a probability distribution $X : Q \to [0, 1]$ be p-synchronized if it assigns probability at least p to some state. A sequence $\bar{X} = X_0X_1 \dots$ of probability distributions is (a) *eventually p-synchronizing* if X_i is p-synchronized for some i; (b) *weakly p-synchronizing* if X_i is p-synchronized for infinitely many i's; (c) *strongly p-synchronizing* if X_i is p-synchronized for all but finitely many i's. It is easy to see that strongly p-synchronizing implies weakly p-synchronizing, which implies eventually p-synchronizing. The qualitative synchronizing properties, corresponding to the case where either $p = 1$, or p tends to 1, are analogous to the traditional reachability, Büchi, and coBüchi conditions.

We consider the following qualitative (winning) modes, summarized in Table 1: (i) *sure* winning, if there is a strategy that generates a {eventually, weakly, strongly} 1-synchronizing sequence; (ii) *almost-sure* winning, if there is a strategy that generates a sequence that is, for all $\varepsilon > 0$, {eventually, weakly, strongly} $(1-\varepsilon)$-synchronizing; (iii) *limit-sure* winning, if for all $\varepsilon > 0$, there is a strategy that generates a {eventually, weakly, strongly} $(1 - \varepsilon)$-synchronizing sequence.

For eventually synchronizing deciding if a given MDP is winning is PSPACE-complete, and the three winning modes form a strict hierarchy [14]. In particular, there are limit-sure winning MDPs that are not almost-sure winning. An important and difficult result in this paper is that the new synchronizing modes are more robust: for weak and strong synchronization, we show that the almost-sure and limit-sure modes coincide. Moreover we establish the complexity of deciding if a given MDP is winning by providing tight (matching) upper and lower bounds:

	Eventually	Weakly	Strongly
Sure	$\exists\alpha \quad \exists n \;\; \mathcal{M}_n^\alpha(T) = 1$	$\exists\alpha \;\; \forall N\, \exists n \geq N \; \mathcal{M}_n^\alpha(T) = 1$	$\exists\alpha \;\; \exists N\, \forall n \geq N \; \mathcal{M}_n^\alpha(T) = 1$
Almost-sure	$\exists\alpha \;\; \sup_n \mathcal{M}_n^\alpha(T) = 1$	$\exists\alpha \;\; \limsup_{n\to\infty} \mathcal{M}_n^\alpha(T) = 1$	$\exists\alpha \;\; \liminf_{n\to\infty} \mathcal{M}_n^\alpha(T) = 1$
Limit-sure	$\sup_\alpha \sup_n \mathcal{M}_n^\alpha(T) = 1$	$\sup_\alpha \limsup_{n\to\infty} \mathcal{M}_n^\alpha(T) = 1$	$\sup_\alpha \liminf_{n\to\infty} \mathcal{M}_n^\alpha(T) = 1$

Table 1. Winning modes and synchronizing objectives (where $\mathcal{M}_n^\alpha(T)$ denotes the probability that under strategy α, after n steps the MDP $\mathcal{M}$ is in a state of T).

for each winning mode we show that the problems are PSPACE-complete for weak synchronization, and PTIME-complete for strong synchronization.

Thus the weakly and strongly synchronizing properties provide conservative approximations of eventually synchronizing, they are robust (limit-sure and almost-sure coincide), and they are of the same (or even lower) complexity as compared to eventually synchronizing.

We also provide optimal memory bounds for winning strategies: exponential memory is sufficient and may be necessary for sure winning in weak synchronization, infinite memory is necessary for almost-sure winning in weak synchronization, and linear memory is sufficient for strong synchronization in all winning modes. We present a variant of strong synchronization for which memoryless strategies are sufficient.

Related Works and Applications. Synchronization problems were first considered for deterministic finite automata (DFA) where a *synchronizing word* is a finite sequence of control actions that can be executed from any state of an automaton and leads to the same state (see [25] for a survey of results and applications). While the existence of a synchronizing word can be decided in polynomial time for DFA, extensive research efforts are devoted to establishing a tight bound on the length of the shortest synchronizing word, which is conjectured to be $(n-1)^2$ for automata with n states [8]. Various extensions of the notion of synchronizing word have been proposed for non-deterministic and probabilistic automata [7,19,20,12], leading to results of PSPACE-completeness [22], or even undecidability [20,13].

For probabilistic systems, a natural extension of words is the notion of strategy that reacts and chooses actions according to the sequence of states visited along the system execution. In this context, an input word corresponds to the special case of a blind strategy that chooses the control actions in advance. In particular, almost-sure weak and strong synchronization with blind strategies has been studied [12] and the main result is the undecidability of deciding the existence of a blind almost-sure winning strategy for weak synchronization, and the PSPACE-completeness of the emptiness problem for strong synchronization [11,13]. In contrast, for general strategies (which also correspond to input trees), we establish the PSPACE-completeness and PTIME-completeness of deciding almost-sure weak and strong synchronization respectively.

A typical application scenario is the design of a control program for a group of mobile robots running in a stochastic environment. The possible behaviors of the robots and the stochastic response of the environment (such as obstacle encounters) are represented by an MDP, and a synchronizing strategy corresponds to a control program that can be embedded in every robot to ensure that they meet (or synchronize) eventually once, infinitely often, or eventually forever.

2 Markov Decision Processes and Synchronization

We closely follow the definitions of [14]. A *probability distribution* over a finite set S is a function $d : S \to [0, 1]$ such that $\sum_{s\in S} d(s) = 1$. The *support* of d is the set $\mathsf{Supp}(d) = \{s \in S \mid d(s) > 0\}$. We denote by $\mathcal{D}(S)$ the set of all probability distributions over S. Given a set $T \subseteq S$, let $d(T) = \sum_{s\in T} d(s)$ and $\|d\|_T = max_{s\in T} d(s)$. For $T \neq \emptyset$, the *uniform distribution* on T assigns probability $\frac{1}{|T|}$ to every state in T. Given $s \in S$, the *Dirac distribution* on s assigns probability 1 to s, and by a slight abuse of notation, we denote it simply by s.

A *Markov decision process* (MDP) is a tuple $\mathcal{M} = \langle Q, \mathsf{A}, \delta\rangle$ where Q is a finite set of states, A is a finite set of actions, and $\delta : Q \times \mathsf{A} \to \mathcal{D}(Q)$ is a probabilistic transition function. A state q is *absorbing* if $\delta(q, a)$ is the Dirac distribution on q for all actions $a \in \mathsf{A}$.

Given state $q \in Q$ and action $a \in \mathsf{A}$, the successor state of q under action a is q' with probability $\delta(q, a)(q')$. Denote by $\mathsf{post}(q, a)$ the set $\mathsf{Supp}(\delta(q, a))$, and given $T \subseteq Q$ let $\mathsf{Pre}(T) = \{q \in Q \mid \exists a \in \mathsf{A} : \mathsf{post}(q, a) \subseteq T\}$ be the set of states from which there is an action to ensure that the successor state is in T. For $k > 0$, let $\mathsf{Pre}^k(T) = \mathsf{Pre}(\mathsf{Pre}^{k-1}(T))$ with $\mathsf{Pre}^0(T) = T$.

A *path* in $\mathcal{M}$ is an infinite sequence $\pi = q_0a_0q_1a_1\dots$ such that $q_{i+1} \in \mathsf{post}(q_i, a_i)$ for all $i \geq 0$. A finite prefix $\rho = q_0a_0q_1a_1\dots q_n$ of a path (or simply a finite path) has length $|\rho| = n$ and last state $\mathsf{Last}(\rho) = q_n$. We denote by $\mathsf{Play}(\mathcal{M})$ and $\mathsf{Pref}(\mathcal{M})$ the set of all paths and finite paths in $\mathcal{M}$ respectively.

Strategies. A *randomized strategy* for $\mathcal{M}$ (or simply a strategy) is a function $\alpha : \mathsf{Pref}(\mathcal{M}) \to \mathcal{D}(\mathsf{A})$ that, given a finite path ρ, returns a probability distribution $\alpha(\rho)$ over the action set, used to select a successor state q' of ρ with probability $\sum_{a\in\mathsf{A}} \alpha(\rho)(a) \cdot \delta(q, a)(q')$ where $q = \mathsf{Last}(\rho)$.

A strategy α is *pure* if for all $\rho \in \mathsf{Pref}(\mathcal{M})$, there exists an action $a \in \mathsf{A}$ such that $\alpha(\rho)(a) = 1$; and *memoryless* if $\alpha(\rho) = \alpha(\rho')$ for all ρ, ρ' such that $\mathsf{Last}(\rho) = \mathsf{Last}(\rho')$. We view pure strategies as functions $\alpha : \mathsf{Pref}(\mathcal{M}) \to \mathsf{A}$, and memoryless strategies as functions $\alpha : Q \to \mathcal{D}(\mathsf{A})$.

Finally, a strategy α uses *finite-memory* if it can be represented by a finite-state transducer $T = \langle \mathsf{Mem}, m_0, \alpha_u, \alpha_n\rangle$ where Mem is a finite set of modes (the memory of the strategy), $m_0 \in \mathsf{Mem}$ is the initial mode, $\alpha_u : \mathsf{Mem} \times (\mathsf{A} \times Q) \to \mathsf{Mem}$ is an update function, that given the current memory, last action and state updates the memory, and $\alpha_n : \mathsf{Mem} \times Q \to \mathcal{D}(\mathsf{A})$ is a next-move function that selects the probability distribution $\alpha_n(m, q)$ over actions when the current

	Eventually	Weakly	Strongly
Sure	PSPACE-C	**PSPACE-C**	**PTIME-C**
Almost-sure	PSPACE-C	**PSPACE-C**	**PTIME-C**
Limit-sure	PSPACE-C		

Table 2. Computational complexity of the membership problem (new results in boldface).

mode is m and the current state of $\mathcal{M}$ is q. For pure strategies, we assume that $\alpha_n : \mathsf{Mem} \times Q \to \mathsf{A}$. The *memory size* of the strategy is the number $|\mathsf{Mem}|$ of modes. For a finite-memory strategy α, let $\mathcal{M}(\alpha)$ be the Markov chain obtained as the product of $\mathcal{M}$ with the transducer defining α. We assume general knowledge of the reader about Markov chains, such as recurrent and transient states, periodicity, and stationary distributions [23].

Outcomes and winning modes. Given an initial distribution $d_0 \in \mathcal{D}(Q)$ and a strategy α in an MDP $\mathcal{M}$, a *path-outcome* is a path $\pi = q_0 a_0 q_1 a_1 \dots$ in $\mathcal{M}$ such that $q_0 \in \mathsf{Supp}(d_0)$ and $a_i \in \mathsf{Supp}(\alpha(q_0 a_0 \dots q_i))$ for all $i \geq 0$. The probability of a finite prefix $\rho = q_0 a_0 q_1 a_1 \dots q_n$ of π is

$$d_0(q_0) \cdot \prod_{j=0}^{n-1} \alpha(q_0 a_0 \dots q_j)(a_j) \cdot \delta(q_j, a_j)(q_{j+1}).$$

We denote by $\mathit{Outcomes}(d_0, \alpha)$ the set of all path-outcomes from d_0 under strategy α. An *event* $\Omega \subseteq \mathsf{Play}(\mathcal{M})$ is a measurable set of paths, and given an initial distribution d_0 and a strategy α, the probability $Pr^{\alpha}(\Omega)$ of Ω is uniquely defined [24]. We consider the following classical winning modes. Given an initial distribution d_0 and an event Ω, we say that $\mathcal{M}$ is: *sure winning* if there exists a strategy α such that $\mathit{Outcomes}(d_0, \alpha) \subseteq \Omega$; *almost-sure winning* if there exists a strategy α such that $\Pr^{\alpha}(\Omega) = 1$;

For example, given a set $T \subseteq Q$ of target states, and $k \in \mathbb{N}$, we denote by $\Box T = \{q_0 a_0 q_1 \dots \in \mathsf{Play}(\mathcal{M}) \mid \forall i : q_i \in T\}$ the safety event of always staying in T, by $\Diamond T = \{q_0 a_0 q_1 \dots \in \mathsf{Play}(\mathcal{M}) \mid \exists i : q_i \in T\}$ the event of reaching T, and by $\Diamond^k T = \{q_0 a_0 q_1 \dots \in \mathsf{Play}(\mathcal{M}) \mid q_k \in T\}$ the event of reaching T after exactly k steps. Let $\Diamond^{\leq k} T = \bigcup_{j \leq k} \Diamond^j T$. Hence, if $\Pr^{\alpha}(\Diamond T) = 1$ then almost-surely a state in T is reached under strategy α.

We consider a symbolic outcome of MDPs viewed as generators of sequences of probability distributions over states [21]. Given an initial distribution $d_0 \in \mathcal{D}(Q)$ and a strategy α in $\mathcal{M}$, the *symbolic outcome* of $\mathcal{M}$ from d_0 is the sequence $(\mathcal{M}^{\alpha}_n)_{n \in \mathbb{N}}$ of probability distributions defined by $\mathcal{M}^{\alpha}_k(q) = Pr^{\alpha}(\Diamond^k \{q\})$ for all $k \geq 0$ and $q \in Q$. Hence, $\mathcal{M}^{\alpha}_k$ is the probability distribution over states after k steps under strategy α. Note that $\mathcal{M}^{\alpha}_0 = d_0$ and the symbolic outcome is a deterministic sequence of distributions: each distribution $\mathcal{M}^{\alpha}_k$ has a unique (determinisitic) successor.

Informally, synchronizing objectives require that the probability of some state (or some group of states) tends to 1 in the sequence $(\mathcal{M}_n^\alpha)_{n\in\mathbb{N}}$, either once, infinitely often, or always after some point. Given a set $T \subseteq Q$, consider the functions $sum_T : \mathcal{D}(Q) \to [0,1]$ and $max_T : \mathcal{D}(Q) \to [0,1]$ that compute $sum_T(X) = \sum_{q\in T} X(q)$ and $max_T(X) = \max_{q\in T} X(q)$. For $f \in \{sum_T, max_T\}$ and $p \in [0,1]$, we say that a probability distribution X is p-synchronized according to f if $f(X) \geq p$, and that a sequence $\bar{X} = X_0X_1\dots$ of probability distributions is [11,14]:

(a) *event* (or *eventually*) *p-synchronizing* if X_i is p-synchronized for some $i \geq 0$;
(b) *weakly p-synchronizing* if X_i is p-synchronized for infinitely many i's;
(c) *strongly p-synchronizing* if X_i is p-synchronized for all but finitely many i's.

For $p = 1$, these definitions are analogous to the traditional reachability, Büchi, and coBüchi conditions [2], and the following winning modes can be considered [14]: given an initial distribution d_0 and a function $f \in \{sum_T, max_T\}$, we say that for the objective of {eventually, weak, strong} synchronization from d_0, $\mathcal{M}$ is:

- *sure winning* if there exists a strategy α such that the symbolic outcome of α from d_0 is {eventually, weakly, strongly} 1-synchronizing according to f;
- *almost-sure winning* if there exists a strategy α such that for all $\varepsilon > 0$ the symbolic outcome of α from d_0 is {eventually, weakly, strongly} $(1-\varepsilon)$-synchronizing according to f;
- *limit-sure winning* if for all $\varepsilon > 0$, there exists a strategy α such that the symbolic outcome of α from d_0 is {eventually, weakly, strongly} $(1-\varepsilon)$-synchronizing according to f;

Note that the winning modes for synchronization objectives differ from the classical winning modes in MDPs: they can be viewed as a specification of the set of sequences of distributions that are winning in a non-stochastic system (since the symbolic outcome is deterministic), while the traditional almost-sure and limit-sure winning modes for path-outcomes consider a probability measure over paths and specify the probability of a specific event (i.e., a set of paths). Thus for instance a strategy is almost-sure synchronizing if the (single) symbolic outcome it produces belongs to the corresponding winning set, whereas traditional almost-sure winning requires a certain event to occur with probability 1.

We often write $\|X\|_T$ instead of $max_T(X)$ (and we omit the subscript when $T = Q$) and $X(T)$ instead of $sum_T(X)$, as in Table 1 where the definitions of the various winning modes and synchronizing objectives for $f = sum_T$ are summarized.

Decision problems. For $f \in \{sum_T, max_T\}$ and $\lambda \in \{event, weakly, strongly\}$, the *winning region* $\langle\langle 1\rangle\rangle_{sure}^{\lambda}(f)$ is the set of initial distributions such that $\mathcal{M}$ is sure winning for λ-synchronizing (we assume that $\mathcal{M}$ is clear from the context). We define analogously the sets $\langle\langle 1\rangle\rangle_{almost}^{\lambda}(f)$ and $\langle\langle 1\rangle\rangle_{limit}^{\lambda}(f)$. For a singleton $T = \{q\}$ we have $sum_T = max_T$, and we simply write $\langle\langle 1\rangle\rangle_{\mu}^{\lambda}(q)$ (where

	Eventually	Weakly	Strongly	
			sum_T	max_T
Sure	exponential	**exponential**	**memoryless**	**linear**
Almost-sure	infinite	**infinite**	**memoryless**	**linear**
Limit-sure	unbounded			

Table 3. Memory requirement (new results in boldface).

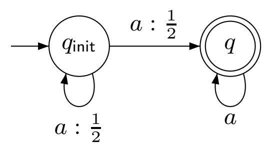

Fig. 1. An MDP $\mathcal{M}$ such that $\langle\langle 1\rangle\rangle^{\lambda}_{sure}(q) \neq \langle\langle 1\rangle\rangle^{\lambda}_{almost}(q)$ for $\lambda \in \{weakly, strongly\}$.

$\mu \in \{sure, almost, limit\}$). It follows from the definitions that $\langle\langle 1\rangle\rangle^{strongly}_{\mu}(f) \subseteq \langle\langle 1\rangle\rangle^{weakly}_{\mu}(f) \subseteq \langle\langle 1\rangle\rangle^{event}_{\mu}(f)$ and thus strong and weak synchronization are conservative approximations of eventually synchronization. It is easy to see that $\langle\langle 1\rangle\rangle^{\lambda}_{sure}(f) \subseteq \langle\langle 1\rangle\rangle^{\lambda}_{almost}(f) \subseteq \langle\langle 1\rangle\rangle^{\lambda}_{limit}(f)$, and for $\lambda = event$ the inclusions are strict [14]. In contrast, weak and strong synchronization are more robust as we show in this paper that the almost-sure and limit-sure winning modes coincide.

Lemma 1. *There exists an MDP $\mathcal{M}$ and state q such that $\langle\langle 1\rangle\rangle^{\lambda}_{sure}(q) \subsetneq \langle\langle 1\rangle\rangle^{\lambda}_{almost}(q)$ for $\lambda \in \{weakly, strongly\}$.*

Proof. Consider the MDP $\mathcal{M}$ with initial state q_{init} and action set $\{a\}$ as shown in Fig. 1. On action a in q_{init}, the successor is q_{init} or q with probability $\frac{1}{2}$, and q is an absorbing state.

We show that $q_{\mathsf{init}} \in \langle\langle 1\rangle\rangle^{strongly}_{almost}(q)$ and $q_{\mathsf{init}} \notin \langle\langle 1\rangle\rangle^{strongly}_{sure}(q)$. Since $\mathcal{M}$ has only a single action, so it is a Markov chain with a unique possible strategy α: always playing a. The outcome under α is such that the probability to be in q after k steps is $1 - \frac{1}{2^k}$ for all k, showing that $\mathcal{M}$ is almost-sure winning for the strongly synchronizing objective in $\{q\}$ (from q_{init}). On the other hand, $q_{\mathsf{init}} \notin \langle\langle 1\rangle\rangle^{strongly}_{sure}(q_1)$ because under α, the probability in q_{init} remains always positive, and thus in q we have $\mathcal{M}^{\alpha}_{n}(q) < 1$ for all $n \geq 0$, showing that $\mathcal{M}$ is not sure winning for the strongly synchronizing objective in $\{q\}$ (from q_{init}). The same argument holds for weakly synchronizing objective. □

The *membership problem* is to decide, given an initial probability distribution d_0, whether $d_0 \in \langle\langle 1\rangle\rangle^{\lambda}_{\mu}(f)$. It is sufficient to consider Dirac initial distributions (i.e., assuming that MDPs have a single initial state) because the answer to the general membership problem for an MDP $\mathcal{M}$ with initial distribution d_0 can be obtained by solving the membership problem for a copy of $\mathcal{M}$ with a new initial state from which the successor distribution on all actions is d_0.

For eventually synchronizing, the membership problem is PSPACE-complete for all winning modes [14]. In this paper, we show that the complexity of the membership problem is PSPACE-complete for weak synchronization, and even PTIME-complete for strong synchronization. The complexity results are summarized in Table 2, and we present the memory requirement for winning strategies in Table 3.

3 Weak Synchronization

We establish the complexity and memory requirement for weakly synchronizing objectives. We show that the membership problem is PSPACE-complete for sure and almost-sure winning, that exponential memory is necessary and sufficient for sure winning while infinite memory is necessary for almost-sure winning, and we show that limit-sure and almost-sure winning coincide.

3.1 Sure weak synchronization

The PSPACE upper bound of the membership problem for sure weak synchronization is obtained by the following characterization.

Lemma 2. *Let $\mathcal{M}$ be an MDP and T be a target set. For all states q_{init}, we have $q_{\mathsf{init}} \in \langle\langle 1 \rangle\rangle^{weakly}_{sure}(sum_T)$ if and only if there exists a set $S \subseteq T$ such that $q_{\mathsf{init}} \in \mathsf{Pre}^m(S)$ for some $m \geq 0$ and $S \subseteq \mathsf{Pre}^n(S)$ for some $n \geq 1$.*

Proof. First, if $q_{\mathsf{init}} \in \langle\langle 1 \rangle\rangle^{weakly}_{sure}(sum_T)$, then let α be a sure winning weakly synchronizing strategy. Then there are infinitely many positions n such that $\mathcal{M}^{\alpha}_n(T) = 1$, and since the state space is finite, there is a set S of states such that for infinitely many positions n we have $\mathsf{Supp}(\mathcal{M}^{\alpha}_n) = S$ and $\mathcal{M}^{\alpha}_n(T) = 1$, and thus $S \subseteq T$. By the result of [14, Lemma 4], it follows that $q_{\mathsf{init}} \in \mathsf{Pre}^m(S)$ for some $m \geq 0$, and by considering two positions $n_1 < n_2$ where $\mathsf{Supp}(\mathcal{M}^{\alpha}_{n_1}) = \mathsf{Supp}(\mathcal{M}^{\alpha}_{n_2}) = S$, it follows that $S \subseteq \mathsf{Pre}^n(S)$ for $n = n_2 - n_1 \geq 1$.

The reverse direction is straightforward by considering a strategy α that ensures $\mathcal{M}^{\alpha}_m(S) = 1$ for some $m \geq 0$, and then ensures that the probability mass from all states in S remains in S after every multiple of n steps where $n > 0$ is such that $S \subseteq \mathsf{Pre}^n(S)$, showing that α is a sure winning weakly synchronizing strategy in S (and thus in T) from q_{init}, thus $q_{\mathsf{init}} \in \langle\langle 1 \rangle\rangle^{weakly}_{sure}(sum_T)$. □

The PSPACE upper bound follows from the characterization in Lemma 2. A (N)PSPACE algorithm is to guess the set $S \subseteq T$, and the numbers m, n (with $m, n \leq 2^{|Q|}$ since the sequence $\mathsf{Pre}^n(S)$ of predecessors is ultimately periodic), and check that $q_{\mathsf{init}} \in \mathsf{Pre}^m(S)$ and $S \subseteq \mathsf{Pre}^n(S)$. The PSPACE lower bound follows from the PSPACE-completeness of the membership problem for sure eventually synchronization [14, Theorem 2].

Lemma 3. *The membership problem for $\langle\langle 1 \rangle\rangle^{weakly}_{sure}(sum_T)$ is PSPACE-hard even if T is a singleton.*

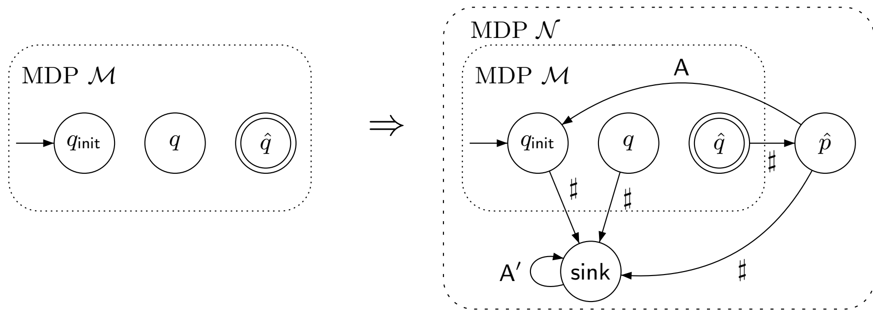

Fig. 2. The reduction sketch to show PSPACE-hardness of the emptiness problem for sure weak synchronization in MDPs.

Proof. The proof is by a reduction from the membership problem for $\langle\!\langle 1 \rangle\!\rangle^{event}_{sure}(sum_T)$ with a singleton T, which is PSPACE-complete [14, Theorem 2]. From an MDP $\mathcal{M} = \langle Q, \mathsf{A}, \delta \rangle$ with initial state q_{init} and target state $\hat{q}$, we construct another MDP $\mathcal{N} = \langle Q', \mathsf{A}', \delta' \rangle$ and a target state $\hat{p}$ such that $q_{\mathsf{init}} \in \langle\!\langle 1 \rangle\!\rangle^{event}_{sure}(\hat{q})$ in $\mathcal{M}$ if and only if $q_{\mathsf{init}} \in \langle\!\langle 1 \rangle\!\rangle^{weakly}_{sure}(\hat{p})$ in $\mathcal{N}$.

The MDP $\mathcal{N}$ is a copy of $\mathcal{M}$ with two new states $\hat{p}$ and sink that are reachable only by a new action $\sharp$ (see Fig. 2). Formally, $Q' = Q \cup \{\hat{p}, \mathsf{sink}\}$ and $\mathsf{A}' = \mathsf{A} \cup \{\sharp\}$. The transition function δ' is defined as follows: $\delta'(q, a) = \delta(q, a)$ for all states $q \in Q$ and $a \in \mathsf{A}$, $\delta(q, \sharp)(\mathsf{sink}) = 1$ for all $q \in Q' \setminus \{\hat{q}\}$ and $\delta(\hat{q}, \sharp)(\hat{p}) = 1$. The state sink is absorbing and from state $\hat{p}$ all other transitions lead to the initial state, i.e. $\delta(\mathsf{sink}, a)(\mathsf{sink}) = 1$ and $\delta(\hat{p}, a)(q_{\mathsf{init}}) = 1$ for all $a \in \mathsf{A}$.

We establish the correctness of the reduction as follows. First, if $q_{\mathsf{init}} \in \langle\!\langle 1 \rangle\!\rangle^{event}_{sure}(\hat{q})$ in $\mathcal{M}$, then let α be a sure winning strategy in $\mathcal{M}$ for eventually synchronization in $\{\hat{q}\}$. A sure winning strategy in $\mathcal{N}$ for weak synchronization in $\{\hat{p}\}$ is to play according to α until the whole probability mass is in $\hat{q}$, then play $\sharp$ followed by some $a \in \mathcal{A}$ to visit $\hat{p}$ and get back to the initial state q_{init}, and then repeat the same strategy from q_{init}. Hence $q_{\mathsf{init}} \in \langle\!\langle 1 \rangle\!\rangle^{weakly}_{sure}(\hat{p})$ in $\mathcal{N}$.

Second, if $q_{\mathsf{init}} \in \langle\!\langle 1 \rangle\!\rangle^{weakly}_{sure}(\hat{p})$ in $\mathcal{N}$, then consider a strategy α such that $\mathcal{N}^{\alpha}_{n}(\hat{p}) = 1$ for some $n \geq 0$. By construction of $\mathcal{N}$, it follows that $\mathcal{N}^{\alpha}_{n-1}(\hat{q}) = 1$, that is all path-outcomes of α of length $n-1$ reach $\hat{q}$, and α plays $\sharp$ in the next step. If α never plays $\sharp$ before position $n-1$, then α is a valid strategy in $\mathcal{M}$ up to step $n-1$ and it shows that $q_{\mathsf{init}} \in \langle\!\langle 1 \rangle\!\rangle^{event}_{sure}(\hat{q})$ is sure winning in $\mathcal{M}$ for eventually synchronization in $\{\hat{q}\}$. Otherwise let m be the largest number such that there is a finite path-outcome ρ of α of length $m < n-1$ such that $\sharp \in \mathsf{Supp}(\alpha(\rho))$. Note that the action $\sharp$ can be played by α only in the state $\hat{q}$, and thus the initial state is reached again after one more step. It follows that in some path-outcome ρ' of α of length $m+2$, we have $\mathsf{Last}(\rho') = q_{\mathsf{init}}$, and by the choice of m, the action $\sharp$ is not played by α until position $n-1$ where all the probability mass is in $\hat{q}$. Hence the strategy that plays like α from ρ' in $\mathcal{N}$ is a valid strategy from q_{init} in $\mathcal{M}$, and is a witness that $q_{\mathsf{init}} \in \langle\!\langle 1 \rangle\!\rangle^{event}_{sure}(\hat{q})$. □

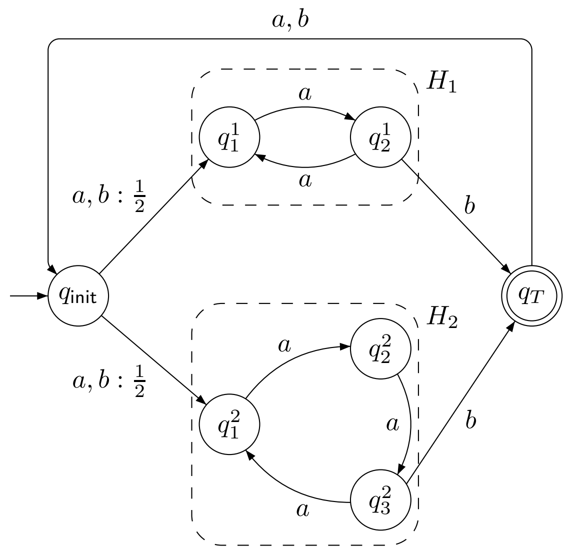

Fig. 3. The MDP $\mathcal{M}_2$.

The proof of Lemma 2 suggests an exponential-memory strategy for sure weakly synchronization that in $q \in \mathsf{Pre}^n(S)$ plays an action a such that $\mathsf{post}(q,a) \subseteq \mathsf{Pre}^{n-1}(S)$, which can be realized with exponential memory since $n \leq 2^{|Q|}$. It can be shown that exponential memory is necessary in general. The argument is very similar to the proof of exponential memory lower bound for sure eventually synchronization [14, Section 4.1]. For the sake of completeness, we present a family of MDPs $\mathcal{M}_n$ $(n \in \mathbb{N})$ over alphabet $\{a,b\}$ that are sure winning for weak synchronization, and where the sure winning strategies require exponential memory. The MDP $\mathcal{M}_2$ is shown in Fig. 3. The structure of $\mathcal{M}_n$ is an initial uniform probabilistic transition to n components $H_1,\dots,H_n$ where H_i is a cycle of length p_i the i-th prime number. On action a, the next state in the cycle is reached, and on action b the target state q_T is reached, only from the last state in the cycles. From other states, the action b leads to an absorbing sink state (transitions not depicted). A sure winning strategy from q_{init} for weak synchronization in $\{q_T\}$ is to play a in the first $p_n^{\#} = \prod_{i=1}^{n} p_i$ steps, and then play bb to reach q_{init} again, through q_T. This requires memory of size $p_n^{\#} > 2^n$ while the size of $\mathcal{M}_n$ is in $O(n^2 \log n)$ [4]. It can be proved that all winning strategies for weak synchronization need to be, from q_{init}, sure eventually synchronizing in $\{q_T\}$ (consider the last occurrence of q_{init} along a play before all the probability mass is in q_T) and this requires memory of size at least $p_n^{\#}$ by standard pumping arguments as in [14].

Theorem 1. *For sure weak synchronization in MDPs:*

1. *(Complexity). The membership problem is PSPACE-complete.*

2. (Memory). Exponential memory is necessary and sufficient for both pure and randomized strategies, and pure strategies are sufficient.

3.2 Almost-sure weak synchronization

We present a characterization of almost-sure weak synchronization that gives a PSPACE upper bound for the membership problem. Our characterization uses the limit-sure eventually synchronizing objectives *with exact support* [14]. This objective requires that the probability mass tends to 1 in a target set T, and moreover that after the same number of steps the support of the probability distribution is contained in a given set U. Formally, given an MDP $\mathcal{M}$, let $\langle\langle 1\rangle\rangle^{event}_{limit}(sum_T, U)$ for $T \subseteq U$ be the set of all initial distributions such that for all $\varepsilon > 0$ there exists a strategy α and $n \in \mathbb{N}$ such that $\mathcal{M}^{\alpha}_{n}(T) \geq 1 - \varepsilon$ and $\mathcal{M}^{\alpha}_{n}(U) = 1$.

We show that an MDP is almost-sure weakly synchronizing in target T if (and only if), for some set U, there is a sure eventually synchronizing strategy in target U, and from the probability distributions with support U there is a limit-sure winning strategy for eventually synchronizing in $\mathsf{Pre}(T)$ with support in $\mathsf{Pre}(U)$. This ensures that from the initial state we can have the whole probability mass in U, and from U have probability $1-\varepsilon$ in $\mathsf{Pre}(T)$ (and in T in the next step), while the whole probability mass is back in $\mathsf{Pre}(U)$ (and in U in the next step), allowing to repeat the strategy for $\varepsilon \to 0$, thus ensuring infinitely often probability at least $1-\varepsilon$ in T (for all $\varepsilon > 0$).

Lemma 4. *Let $\mathcal{M}$ be an MDP and T be a target set. For all states q_{init}, we have $q_{\mathsf{init}} \in \langle\langle 1\rangle\rangle^{weakly}_{almost}(sum_T)$ if and only if there exists a set U such that*

- *$q_{\mathsf{init}} \in \langle\langle 1\rangle\rangle^{event}_{sure}(sum_U)$, and*
- *$d_U \in \langle\langle 1\rangle\rangle^{event}_{limit}(sum_{\mathsf{Pre}(T)}, \mathsf{Pre}(U))$ where d_U is the uniform distribution over U.*

Proof. First, if $q_{\mathsf{init}} \in \langle\langle 1\rangle\rangle^{weakly}_{almost}(sum_T)$, then there exists a strategy α such that for all $i \geq 0$ there exists $n_i \in \mathbb{N}$ such that $\mathcal{M}^{\alpha}_{n_i}(T) \geq 1 - 2^{-i}$, and moreover $n_{i+1} > n_i$ for all $i \geq 0$. Let $s_i = \mathsf{Supp}(\mathcal{M}^{\alpha}_{n_i})$ be the support of $\mathcal{M}^{\alpha}_{n_i}$. Since the state space is finite, there is a set U that occurs infinitely often in the sequence $s_0 s_1 \dots$, thus for all $k > 0$ there exists $m_k \in \mathbb{N}$ such that $\mathcal{M}^{\alpha}_{m_k}(T) \geq 1-2^{-k}$ and $\mathcal{M}^{\alpha}_{m_k}(U) = 1$. It follows that α is sure eventually synchronizing in U from q_{init}, i.e. $q_{\mathsf{init}} \in \langle\langle 1\rangle\rangle^{event}_{sure}(sum_U)$. Moreover, we can assume that $m_{k+1} > m_k$ for all $k > 0$ and thus $\mathcal{M}$ is also limit-sure eventually synchronizing in $\mathsf{Pre}(T)$ with exact support in $\mathsf{Pre}(U)$ from the initial distribution $d_1 = \mathcal{M}^{\alpha}_{m_1}$. Since $\mathsf{Supp}(d_1) = U = \mathsf{Supp}(d_U)$ and since only the support of the initial probability distributions is relevant for the limit-sure eventually synchronizing objective [14, Corollary 1], it follows that $d_U \in \langle\langle 1\rangle\rangle^{event}_{limit}(sum_{\mathsf{Pre}(T)}, \mathsf{Pre}(U))$.

To establish the converse, note that since $d_U \in \langle\langle 1\rangle\rangle^{event}_{limit}(sum_{\mathsf{Pre}(T)}, \mathsf{Pre}(U))$, it follows from [14, Corollary 1] that from all initial distributions with support in U, for all $\varepsilon > 0$ there exists a strategy α_ε and a position n_ε such that

$\mathcal{M}^{\alpha_\varepsilon}_{n_\varepsilon}(T) \geq 1-\varepsilon$ and $\mathcal{M}^{\alpha_\varepsilon}_{n_\varepsilon}(U) = 1$. We construct an almost-sure weakly synchronizing strategy α as follows. Since $q_{\mathsf{init}} \in \langle\!\langle 1 \rangle\!\rangle^{event}_{sure}(sum_U)$, play according to a sure eventually synchronizing strategy from q_{init} until all the probability mass is in U. Then for $i = 1, 2, \dots$ and $\varepsilon_i = 2^{-i}$, repeat the following procedure: given the current probability distribution, select the corresponding strategy α_{ε_i} and play according to α_{ε_i} for n_{ε_i} steps, ensuring probability mass at least $1-2^{-i}$ in $\mathsf{Pre}(T)$ and support of the probability mass in $\mathsf{Pre}(U)$. Then from states in $\mathsf{Pre}(T)$, play an action to ensure reaching T in the next step, and from states in $\mathsf{Pre}(U)$ ensure reaching U. Continue playing according to $\alpha_{\varepsilon_{i+1}}$ for $n_{\varepsilon_{i+1}}$ steps, etc. Since $n_{\varepsilon_i} + 1 > 0$ for all $i \geq 0$, this strategy ensures that $\limsup_{n\to\infty} \mathcal{M}^{\alpha}_{n}(T) = 1$ from q_{init}, hence $q_{\mathsf{init}} \in \langle\!\langle 1 \rangle\!\rangle^{weak}_{almost}(sum_T)$.

□

Since the membership problems for sure eventually synchronizing and for limit-sure eventually synchronizing with exact support are PSPACE-complete ([14, Theorem 2 and 4]), the membership problem for almost-sure weak synchronization is in PSPACE by guessing the set U, and checking that $q_{\mathsf{init}} \in \langle\!\langle 1 \rangle\!\rangle^{event}_{sure}(sum_U)$, and that $d_U \in \langle\!\langle 1 \rangle\!\rangle^{event}_{limit}(sum_{\mathsf{Pre}(T)}, \mathsf{Pre}(U))$. We establish a matching PSPACE lower bound.

Lemma 5. *The membership problem for $\langle\!\langle 1 \rangle\!\rangle^{weakly}_{almost}(sum_T)$ is PSPACE-hard even if T is a singleton.*

Proof. The problem of deciding, given an MDP $\mathcal{M}$ and a singleton T, whether $\mathsf{Pre}^n_{\mathcal{M}}(T) \neq \emptyset$ for all $n \geq 0$ is PSPACE-complete [14, Lemma 3]. We present a reduction of this problem to the membership problem for almost-sure weak synchronization, very similar to the proof of PSPACE-hardness for limit-sure eventually synchronizing [14, Lemma 11].

The reduction is as follows (see also Fig. 4). Given an MDP $\mathcal{M} = \langle Q, \mathsf{A}, \delta \rangle$ and a singleton $T \subseteq Q$, we construct an MDP $\mathcal{N} = \langle Q', \mathsf{A}', \delta' \rangle$ with state space $Q' = Q \uplus \{q_{\mathsf{init}}\}$ such that $\mathsf{Pre}^n_{\mathcal{M}}(T) \neq \emptyset$ for all $n \geq 0$ if and only if q_{init} is almost-sure weakly synchronizing in T. The MDP $\mathcal{N}$ is essentially a copy of $\mathcal{M}$ with alphabet $\mathsf{A} \uplus \{\sharp\}$ and the transition function on action $\sharp$ is the uniform distribution on Q from q_{init}, and the Dirac distribution on q_{init} from the other states $q \in Q$. There are self-loops on q_{init} for all other actions $a \in \mathsf{A}$. Formally, the transition function δ' is defined as follows, for all $q \in Q$:

- $\delta'(q, a) = \delta(q, a)$ for all $a \in \mathsf{A}$ (copy of $\mathcal{M}$), and $\delta'(q, \sharp)(q_{\mathsf{init}}) = 1$;
- $\delta'(q_{\mathsf{init}}, a)(q_{\mathsf{init}}) = 1$ for all $a \in \mathsf{A}$, and $\delta'(q_{\mathsf{init}}, \sharp)(q) = \frac{1}{|Q|}$.

We establish the correctness of the reduction as follows. For the first direction, assume that $\mathsf{Pre}^n_{\mathcal{M}}(T) \neq \emptyset$ for all $n \geq 0$. It follows that there exist numbers $k_0, r \leq 2^{|Q|}$ such that $\mathsf{Pre}^{k_0+r}_{\mathcal{M}}(T) = \mathsf{Pre}^{k_0}_{\mathcal{M}}(T) \neq \emptyset$.

By Lemma 4 with $U = Q$, we need to show that (i) $q_{\mathsf{init}} \in \langle\!\langle 1 \rangle\!\rangle^{event}_{sure}(sum_Q)$, and (ii) $d_Q \in \langle\!\langle 1 \rangle\!\rangle^{event}_{limit}(sum_{\mathsf{Pre}(T)}, \mathsf{Pre}(Q))$ where d_Q is the uniform distribution over Q. To show (i), we can play $\sharp$ from q_{init} to get the probability mass synchronized in Q. To show (ii), since playing $\sharp$ from d_Q ensures to reach q_{init}, it

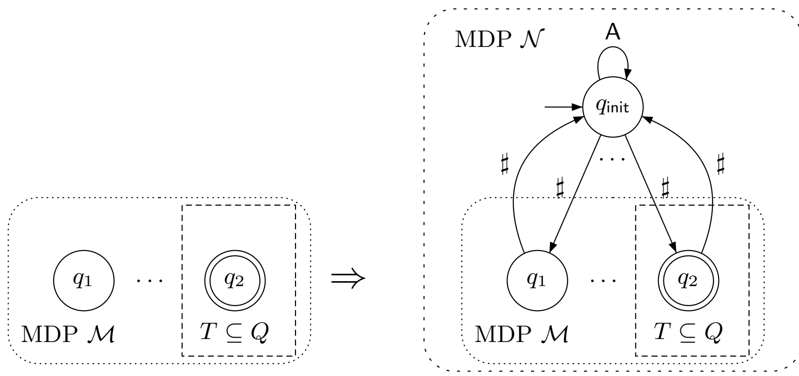

Fig. 4. Sketch of reduction to show PSPACE-hardness of the membership problem for almost-sure weak synchronization.

suffices to prove that $q_{\mathsf{init}} \in \langle\!\langle 1 \rangle\!\rangle^{event}_{limit}(sum_T, Q)$, and it is sufficient to prove this in $\mathcal{M}$ since $\mathcal{N}$ embeds a copy of $\mathcal{M}$ (note that the requirement that the exact support is in Q becomes trivial then). Using [14, Lemma 8] with $k = k_0$ and $R = \mathsf{Pre}^{k_0}_{\mathcal{M}}(T)$ (and $U = Z = Q$ is the trivial support), it is sufficient to prove that $q_{\mathsf{init}} \in \langle\!\langle 1 \rangle\!\rangle^{event}_{limit}(sum_R)$ to get $q_{\mathsf{init}} \in \langle\!\langle 1 \rangle\!\rangle^{event}_{limit}(sum_T)$. We show the stronger statement that q_{init} is actually almost-sure eventually synchronizing in R with the pure strategy α defined as follows, for all play prefixes ρ (let $m = |\rho| \bmod r$):

- if $\mathsf{Last}(\rho) = q_{\mathsf{init}}$, then $\alpha(\rho) = \sharp$;
- if $\mathsf{Last}(\rho) = q \in Q$, then
  - if $q \in \mathsf{Pre}^{r-m}_{\mathcal{M}}(R)$ for $0 \leq m < r$, then $\alpha(\rho) = a$ such that $\mathsf{post}(q, a) \subseteq \mathsf{Pre}^{r-m-1}_{\mathcal{M}}(R)$;
  - otherwise, $\alpha(\rho) = \sharp$.

Note that if $q \in R$, then $q \in \mathsf{Pre}^{r-m}_{\mathcal{M}}(R)$ for $m = 0$ since $\mathsf{Pre}^{r}_{\mathcal{M}}(R) = R$. The strategy α ensures that the probability mass that is not (yet) in the sequence of predecessors $\mathsf{Pre}^{n}_{\mathcal{M}}(R)$ goes to q_{init}, where by playing $\sharp$ at least a fraction $\frac{1}{|Q|}$ of it would reach the sequence of predecessors (at a synchronized position). It follows that after $2i$ steps, the probability mass in q_{init} is at most $(1 - \frac{1}{|Q|})^i$ and the probability mass in the sequence of predecessors is at least $1 - (1 - \frac{1}{|Q|})^i$. For $i \to \infty$, the probability in the sequence of predecessors tends to 1 and since $\mathsf{Pre}^{n}_{\mathcal{M}}(R) = R$ for all positions n that are a multiple of r, we get $\sup_n \mathcal{M}^{\alpha}_n(R) = 1$ and $q_{\mathsf{init}} \in \langle\!\langle 1 \rangle\!\rangle^{event}_{almost}(sum_R)$.

For the converse direction, assume that q_{init} is almost-sure weakly synchronizing in T, then q_{init} is also limit-sure eventually synchronizing in T. By [14, Lemma 8], either (1) q_{init} is limit-sure eventually synchronizing in $\mathsf{Pre}^{n}_{\mathcal{N}}(T)$ for all $n \geq 0$, and then it follows that $\mathsf{Pre}^{n}_{\mathcal{N}}(T) \neq \emptyset$ for all $n \geq 0$, or (2) q_{init} is sure eventually synchronizing in T, and then since only the action $\sharp$ leaves the state q_{init} (and $\mathsf{post}(q_{\mathsf{init}}, \sharp) = Q$), and since $q_{\mathsf{init}} \in \langle\!\langle 1 \rangle\!\rangle^{event}_{sure}(sum_T)$ if and only if there

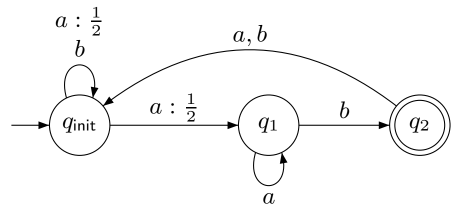

Fig. 5. An MDP where infinite memory is necessary for almost-sure weakly synchronizing strategies.

exists $k \geq 0$ such that $q_{\mathsf{init}} \in \mathsf{Pre}^k_{\mathcal{N}}(T)$ [14, Lemma 4], we have $Q \subseteq \mathsf{Pre}^{k-1}_{\mathcal{N}}(T)$. Moreover, since $Q \subseteq \mathsf{Pre}_{\mathcal{N}}(Q)$ and $\mathsf{Pre}_{\mathcal{N}}(\cdot)$ is a monotone operator, it follows that $Q \subseteq \mathsf{Pre}^n_{\mathcal{N}}(T)$ for all $n \geq k-1$ and thus $\mathsf{Pre}^n_{\mathcal{N}}(T) \neq \emptyset$ for all $n \geq 0$. We conclude the proof by noting that $\mathsf{Pre}^n_{\mathcal{M}}(T) = \mathsf{Pre}^n_{\mathcal{N}}(T) \cap Q$ and therefore $\mathsf{Pre}^n_{\mathcal{M}}(T) \neq \emptyset$ for all $n \geq 0$. □

Simple examples show that winning strategies require infinite memory for almost-sure weak synchronization.

Theorem 2. *For almost-sure weak synchronization in MDPs:*

1. *(Complexity). The membership problem is PSPACE-complete.*
2. *(Memory). Infinite memory is necessary in general for both pure and randomized strategies, and pure strategies are sufficient.*

Proof. The result on memory requirement is established by following example. The example and argument are analogous to the proof that infinite memory is necessary for almost-sure eventually synchronizing [14, Section 4.2]. Consider the MDP $\mathcal{M}$ shown in Fig. 5 with three states $q_{\mathsf{init}}, q_1, q_2$ and two actions a, b. The only probabilistic transition is in q_{init} on action a that has successors q_{init} and q_1 with probability $\frac{1}{2}$. The other transitions are deterministic. Let q_{init} be the initial state. We construct a strategy that is almost-sure weakly synchronizing in $\{q_2\}$, showing that $q_{\mathsf{init}} \in \langle\langle 1 \rangle\rangle^{weakly}_{almost}(q_2)$. First, observe that for all $\varepsilon > 0$ we can have probability at least $1-\varepsilon$ in q_2 after finitely many steps from q_{init}: playing n times a and then b leads to probability $1 - \frac{1}{2^n}$ in q_2. Note that after that, the current probability distribution has support $\{q_{\mathsf{init}}, q_2\}$ and that from such a distribution, we can as well ensure probability at least $1 - \varepsilon$ in q_2. Thus for a fixed ε, the MDP is $(1-\varepsilon)$-synchronizing in $\{q_2\}$ (after finitely many steps), and by taking a smaller value of ε, we can continue to play a strategy to have probability at least $1 - \varepsilon$ in q_2, and repeat this for $\varepsilon \to 0$. This strategy ensures almost-sure weak synchronization in $\{q_2\}$. Below, we show that infinite memory is necessary for almost-sure winning in this MDP.

Assume towards contradiction that there exists a finite-memory strategy α that is almost-sure weakly synchronizing in $\{q_2\}$. Consider the Markov chain $\mathcal{M}(\alpha)$ (the product of the MDP $\mathcal{M}$ with the finite-state transducer defining

α). A state (q, m) in $\mathcal{M}(\alpha)$ is called a *q-state*. Since α is almost-sure weakly synchronizing in $\{q_2\}$, there is a q_2-state in the recurrent states of $\mathcal{M}(\alpha)$. Since on all actions q_{init} is a successor of q_2, and q_{init} is a successor of itself, it follows that there is a recurrent q_{init}-state in $\mathcal{M}(\alpha)$, and that all periodic classes of recurrent states in $\mathcal{M}(\alpha)$ contain a q_{init}-state. Hence, in each stationary distribution there is a q_{init}-state with a positive probability, and therefore the probability mass in q_{init} is bounded away from zero. It follows that the probability mass in q_2 is bounded away from 1 thus α is not almost-sure weakly synchronizing in $\{q_2\}$, a contradiction. □

3.3 Limit-sure weak synchronization

We show that the winning regions for almost-sure and limit-sure weak synchronization coincide. The result is not intuitively obvious (recall that it does not hold for eventually synchronizing) and requires a careful analysis of the structure of limit-sure winning strategies to show that they always induce the existence of an almost-sure winning strategy. The construction of an almost-sure winning strategy from a family of limit-sure winning strategies is illustrated in the following example.

Consider the MDP $\mathcal{M}$ in Fig. 6 with initial state q_{init} and target set $T = \{q_4\}$. Note that there is a relevant strategic choice only in q_3, and that q_{init} is limit-sure winning for eventually synchronization in $\{q_4\}$ since we can inject a probability mass arbitrarily close to 1 in q_3 (by always playing a in q_3), and then switching to playing b in q_3 gets probability $1 - \varepsilon$ in T (for arbitrarily small ε). Moreover, the same holds from state q_4. These two facts are sufficient to show that q_{init} is limit-sure winning for weak synchronization in $\{q_4\}$: given $\varepsilon > 0$, play from q_{init} a strategy to ensure probability at least $p_1 = 1 - \frac{\varepsilon}{2}$ in q_4 (in finitely many steps), and then play according to a strategy that ensures from q_4 probability $p_2 = p_1 - \frac{\varepsilon}{4}$ in q_4 (in finitely many, and at least one step), and repeat this process using strategies that ensure, if the probability mass in q_4 is at least p_i, that the probability in q_4 is at least $p_{i+1} = p_i - \frac{\varepsilon}{2^{i+1}}$ (in at least one step). It follows that $p_i = 1 - \frac{\varepsilon}{2} - \frac{\varepsilon}{4} - \cdots - \frac{\varepsilon}{2^i} > 1 - \varepsilon$ for all $i \geq 1$, and thus $\limsup_{i\to\infty} p_i \geq 1 - \varepsilon$ showing that q_{init} is limit-sure weakly synchronizing in target $\{q_4\}$.

It follows from the result that we establish in this section (Theorem 3) that q_{init} is actually almost-sure weakly synchronizing in target $\{q_4\}$. To see this, consider the sequence $\mathsf{Pre}^i(T)$ for $i \geq 0$: $\{q_4\}, \{q_3\}, \{q_2\}, \{q_3\}, \dots$ is ultimately periodic with period $r = 2$ and $R = \{q_3\} = \mathsf{Pre}(T)$ is such that $R = \mathsf{Pre}^2(R)$. The period corresponds to the loop $q_2 q_3$ in the MDP. It turns out that *limit-sure* eventually synchronizing in T implies *almost-sure* eventually synchronizing in R (by the proof of [14, Lemma 9]), thus from q_{init} a *single* strategy ensures that the probability mass in R is 1, either in the limit or after finitely many steps. Note that in both cases since $R = \mathsf{Pre}^r(R)$ this even implies almost-sure weakly synchronizing in R. The same holds from state q_4.

Moreover, note that all distributions produced by an almost-sure weakly synchronizing strategy are themselves almost-sure weakly synchronizing. An almost-sure winning strategy for weak synchronization in $\{q_4\}$ consists in playing from

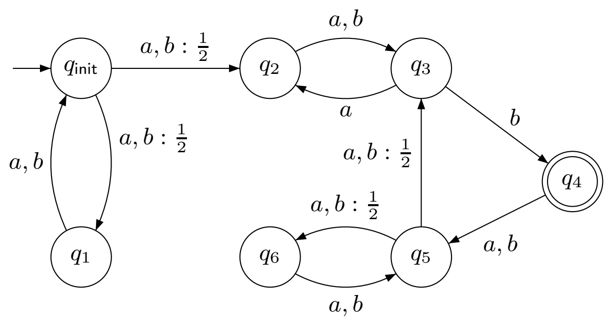

Fig. 6. An example to show $q_{\mathsf{init}} \in \langle\langle 1 \rangle\rangle^{weakly}_{limit}(q_4)$ implies $q_{\mathsf{init}} \in \langle\langle 1 \rangle\rangle^{weakly}_{almost}(q_4)$.

q_{init} an *almost-sure* eventually synchronizing strategy in target $R = \{q_3\}$, and considering a decreasing sequence ε_i such that $\lim_{i\to\infty} \varepsilon_i = 0$, when the probability mass in R is at least $1 - \varepsilon_i$, inject it in $T = \{q_4\}$. Then the remaining probability mass defines a distribution (with support $\{q_1, q_2\}$ in the example) that is still almost-sure eventually synchronizing in R, as well as the states in T. Note that in the example, (almost all) the probability mass in $T = \{q_4\}$ can move to q_3 in an even number of steps, while from $\{q_1, q_2\}$ an odd number of steps is required, resulting in a *shift* of the probability mass. However, by repeating the strategy two times from q_4 (injecting large probability mass in q_3, moving to q_4, and injecting in q_3 again), we can make up for the shift and reach q_3 from q_4 in an even number of steps, thus in synchronization with the probability mass from $\{q_1, q_2\}$. This idea is formalized in the rest of this section, and we prove that we can always make up for the shifts, which requires a carefully analysis of the allowed amounts of shifting.

The result is easier to prove when the target T is a singleton, as in the example. For an arbitrary target set T, we need to get rid of the states in T that do not contribute a significant (i.e., bounded away from 0) probability mass in the limit, that we call the 'vanishing' states. We show that they can be removed from T without changing the winning region for limit-sure winning. When the target set has no vanishing state, we can construct an almost-sure winning strategy as in the case of a singleton target set.

Given an MDP $\mathcal{M}$ with initial state $q_{\mathsf{init}} \in \langle\langle 1 \rangle\rangle^{weakly}_{limit}(sum_T)$ that is limit-sure winning for the weakly synchronizing objective in target set T, let $(\alpha_i)_{i\in\mathbb{N}}$ be a family of limit-sure winning strategies such that $\limsup_{n\to\infty} \mathcal{M}^{\alpha_i}_n(T) \geq 1 - \varepsilon_i$ where $\lim_{i\to\infty} \varepsilon_i = 0$. Hence by definition of lim sup, for all $i \geq 0$ there exists a strictly increasing sequence $k_{i,0} < k_{i,1} < \cdots$ of positions such that $\mathcal{M}^{\alpha_i}_{k_{i,j}}(T) \geq 1 - 2\varepsilon_i$ for all $j \geq 0$. A state $q \in T$ is *vanishing* if $\liminf_{i\to\infty} \liminf_{j\to\infty} \mathcal{M}^{\alpha_i}_{k_{i,j}}(q) = 0$ for some family of limit-sure weakly synchronizing strategies $(\alpha_i)_{i\in\mathbb{N}}$. Intuitively, the contribution of a vanishing state q to the probability in T tends to 0 and therefore $\mathcal{M}$ is also limit-sure winning for the weakly synchronizing objective in target set $T \setminus \{q\}$.

Lemma 6. *If an MDP $\mathcal{M}$ is limit-sure weakly synchronizing in target set T, then there exists a set $T' \subseteq T$ such that $\mathcal{M}$ is limit-sure weakly synchronizing in T' without vanishing states.*

Proof. If there is no vanishing state for $(\alpha_i)_{i\in\mathbb{N}}$, then take $T' = T$ and the proof is complete. Otherwise, let $(\alpha_i)_{i\in\mathbb{N}}$ be a family of limit-sure winning strategies such that $\limsup_{n\to\infty} \mathcal{M}_n^{\alpha_i}(T) \geq 1 - \varepsilon_i$ where $\lim_{i\to\infty} \varepsilon_i = 0$ and let q be a vanishing state for $(\alpha_i)_{i\in\mathbb{N}}$. We show that $(\alpha_i)_{i\in\mathbb{N}}$ is limit-sure weakly synchronizing in $T \setminus \{q\}$. For every $i \geq 0$ let $k_{i,0} < k_{i,1} < \cdots$ be a strictly increasing sequence such that (a) $\mathcal{M}_{k_{i,j}}^{\alpha_i}(T) \geq 1 - 2\varepsilon_i$ for all $i, j \geq 0$, and (b) $\liminf_{i\to\infty} \liminf_{j\to\infty} \mathcal{M}_{k_{i,j}}^{\alpha_i}(q) = 0$.

It follows from (b) that for all $\varepsilon > 0$ and all $x > 0$ there exists $i > x$ such that for all $y > 0$ there exists $j > y$ such that $\mathcal{M}_{k_{i,j}}^{\alpha_i}(q) < \varepsilon$, and thus

$$\mathcal{M}_{k_{i,j}}^{\alpha_i}(T \setminus \{q\}) \geq 1 - 2\varepsilon_i - \varepsilon$$

by (a). Since this holds for infinitely many i's, we can choose i such that $\varepsilon_i < \varepsilon$ and we have

$$\limsup_{j\to\infty} \mathcal{M}_{k_{i,j}}^{\alpha_i}(T \setminus \{q\}) \geq 1 - 3\varepsilon$$

and thus

$$\limsup_{n\to\infty} \mathcal{M}_{n}^{\alpha_i}(T \setminus \{q\}) \geq 1 - 3\varepsilon$$

since the sequence $(k_{i,j})_{j\in\mathbb{N}}$ is strictly increasing. This shows that $(\alpha_i)_{i\in\mathbb{N}}$ is limit-sure weakly synchronizing in $T \setminus \{q\}$.

By repeating this argument as long as there is a vanishing state (thus at most $|T| - 1$ times), we can construct the desired set $T' \subseteq T$ without vanishing state. ◻

For a limit-sure weakly synchronizing MDP in target set T (without vanishing states), we show that from a probability distribution with support T, a probability mass arbitrarily close to 1 can be injected synchronously back in T (in at least one step), that is $d_T \in \langle\langle 1 \rangle\rangle_{limit}^{event}(sum_{\mathsf{Pre}(T)})$. The same holds from the initial state q_{init} of the MDP. This property is the key to construct an almost-sure weakly synchronizing strategy.

Lemma 7. *If an MDP $\mathcal{M}$ with initial state q_{init} is limit-sure weakly synchronizing in a target set T without vanishing states, then $q_{\mathsf{init}} \in \langle\langle 1 \rangle\rangle_{limit}^{event}(sum_{\mathsf{Pre}(T)})$ and $d_T \in \langle\langle 1 \rangle\rangle_{limit}^{event}(sum_{\mathsf{Pre}(T)})$ where d_T is the uniform distribution over T.*

Proof. Since $q_{\mathsf{init}} \in \langle\langle 1 \rangle\rangle_{limit}^{weakly}(sum_T)$ and $\langle\langle 1 \rangle\rangle_{limit}^{weakly}(sum_T) \subseteq \langle\langle 1 \rangle\rangle_{limit}^{event}(sum_T)$, we have $q_{\mathsf{init}} \in \langle\langle 1 \rangle\rangle_{limit}^{event}(sum_T)$ and thus it suffices to prove that $d_T \in \langle\langle 1 \rangle\rangle_{limit}^{event}(sum_{\mathsf{Pre}(T)})$. This is because then from q_{init}, probability arbitrarily close to 1 can be injected in $\mathsf{Pre}(T)$ through a distribution with support in T (since by [14, Corollary 1] only the support of the initial probability distribution is important for limit-sure eventually synchronizing).

Let $(\alpha_i)_{i\in\mathbb{N}}$ be a family of limit-sure winning strategies such that $\limsup_{n\to\infty} \mathcal{M}_n^{\alpha_i}(T) \geq 1-\varepsilon_i$ where $\lim_{i\to\infty}\varepsilon_i = 0$, and such that there is no vanishing state. For every $i \geq 0$ let $k_{i,0} < k_{i,1} < \cdots$ be a strictly increasing sequence such that $\mathcal{M}_{k_{i,j}}^{\alpha_i}(T) \geq 1-2\varepsilon_i$ for all $i,j \geq 0$, and let $B = \min_{q\in T} \liminf_{i\to\infty} \liminf_{j\to\infty} \mathcal{M}_{k_{i,j}}^{\alpha_i}(q)$. Note that $B > 0$ since there is no vanishing state. It follows that there exists $x > 0$ such that for all $i > x$ there exists $y_i > 0$ such that for all $j > y_i$ and all $q \in T$ we have $\mathcal{M}_{k_{i,j}}^{\alpha_i}(q) \geq \frac{B}{2}$.

Given $\nu > 0$, let $i > x$ such that $\varepsilon_i < \frac{\nu B}{4}$, and for $j > y_i$, consider the positions $n_1 = k_{i,j}$ and $n_2 = k_{i,j+1}$. We have $n_1 < n_2$ and $\mathcal{M}_{n_1}^{\alpha_i}(T) \geq 1-2\varepsilon_i$ and $\mathcal{M}_{n_2}^{\alpha_i}(T) \geq 1-2\varepsilon_i$, and $\mathcal{M}_{n_1}^{\alpha_i}(q) \geq \frac{B}{2}$ for all $q \in T$. Consider the strategy β that plays like α_i plays from position n_1 and thus transforms the distribution $\mathcal{M}_{n_1}^{\alpha_i}$ into $\mathcal{M}_{n_2}^{\alpha_i}$. For all states $q \in T$, from the Dirac distribution on q under strategy β, the probability to reach $Q \setminus T$ in $n_2 - n_1$ steps is thus at most $\frac{\mathcal{M}_{n_2}^{\alpha_i}(Q\setminus T)}{\mathcal{M}_{n_1}^{\alpha_i}(q)} \leq \frac{2\varepsilon_i}{B/2} < \nu$.

Therefore, from an arbitrary probability distribution with support T we have $\mathcal{M}_{n_2-n_1}^{\beta}(T) > 1-\nu$, showing that d_T is limit-sure eventually synchronizing in T and thus in $\mathsf{Pre}(T)$ since $n_2 - n_1 > 0$ (it is easy to show that if the mass of probability in T is at least $1-\nu$, then the mass of probability in $\mathsf{Pre}(T)$ one step before is at least $1-\frac{\nu}{\eta}$ where η is the smallest positive probability in $\mathcal{M}$). $\square$

To show that limit-sure and almost-sure winning coincide for weakly synchronizing objectives, from a family of limit-sure winning strategies we construct an almost-sure winning strategy that uses the eventually synchronizing strategies of Lemma 7. The construction consists in using successively strategies that ensure probability mass $1-\varepsilon_i$ in the target T, for a decreasing sequence $\varepsilon_i \to 0$. Such strategies exist by Lemma 7, both from the initial state and from the set T. However, the mass of probability that can be guaranteed to be synchronized in T by the successive strategies is always smaller than 1, and therefore we need to argue that the remaining masses of probability (of size ε_i) can also get synchronized in T, and despite their possible shift with the main mass of probability.

Two main key arguments are needed to establish the correctness of the construction: (1) eventually synchronizing implies that a finite number of steps is sufficient to obtain a probability mass of $1-\varepsilon_i$ in T, and thus the construction of the strategy is well defined, and (2) by the finiteness of the period r (such that $R = \mathsf{Pre}^r(R)$ where $R = \mathsf{Pre}^k(T)$ for some k) we can ensure to eventually make up for the shifts, and every piece of the probability mass can contribute (synchronously) to the target infinitely often.

Theorem 3. $\langle\!\langle 1 \rangle\!\rangle_{limit}^{weakly}(sum_T) = \langle\!\langle 1 \rangle\!\rangle_{almost}^{weakly}(sum_T)$ *for all MDPs and target sets* T*.*

Proof. Since $\langle\!\langle 1 \rangle\!\rangle_{almost}^{weakly}(sum_T) \subseteq \langle\!\langle 1 \rangle\!\rangle_{limit}^{weakly}(sum_T)$ holds by the definition, it is sufficient to prove that $\langle\!\langle 1 \rangle\!\rangle_{limit}^{weakly}(sum_T) \subseteq \langle\!\langle 1 \rangle\!\rangle_{almost}^{weakly}(sum_T)$ and by Lemma 6 it is sufficient to prove that if $q_{\mathsf{init}} \in \langle\!\langle 1 \rangle\!\rangle_{limit}^{weakly}(sum_T)$ is limit-sure weakly synchronizing in T without vanishing state, then q_{init} is almost-sure weakly synchronizing in T. If T has vanishing states, then consider $T' \subseteq T$ as in Lemma 6 and it will follows that q_{init} is almost-sure weakly synchronizing in T' and thus also in T. We proceed with the proof that $q_{\mathsf{init}} \in \langle\langle 1 \rangle\rangle^{weakly}_{limit}(sum_T)$ implies $q_{\mathsf{init}} \in \langle\langle 1 \rangle\rangle^{weakly}_{almost}(sum_T)$.

For $i = 1, 2, \dots$ consider the sequence of predecessors $\mathsf{Pre}^i(T)$, which is ultimately periodic: let $1 \leq k, r \leq 2^{|Q|}$ such that $\mathsf{Pre}^k(T) = \mathsf{Pre}^{k+r}(T)$, and let $R = \mathsf{Pre}^k(T)$. Thus $R = \mathsf{Pre}^{k+r}(T) = \mathsf{Pre}^r(R)$.

Claim 1. We have $q_{\mathsf{init}} \in \langle\langle 1 \rangle\rangle^{event}_{almost}(sum_R)$ and $d_T \in \langle\langle 1 \rangle\rangle^{event}_{almost}(sum_R)$.

Proof of Claim 1. By Lemma 7, since there is no vanishing state in T we have $q_{\mathsf{init}} \in \langle\langle 1 \rangle\rangle^{event}_{limit}(sum_{\mathsf{Pre}(T)})$ and $d_T \in \langle\langle 1 \rangle\rangle^{event}_{limit}(sum_{\mathsf{Pre}(T)})$, and it follows from the characterization of [14, Lemma 8] and the proof of [14, Lemma 9] that:

either (1) $q_{\mathsf{init}} \in \langle\langle 1 \rangle\rangle^{event}_{sure}(sum_{\mathsf{Pre}(T)})$ or (2) $q_{\mathsf{init}} \in \langle\langle 1 \rangle\rangle^{event}_{almost}(sum_R)$, and
either (a) $d_T \in \langle\langle 1 \rangle\rangle^{event}_{sure}(sum_{\mathsf{Pre}(T)})$ or (b) $d_T \in \langle\langle 1 \rangle\rangle^{event}_{almost}(sum_R)$.

Note that (a) implies (b) (and thus (b) holds) since (a) implies $T \subseteq \mathsf{Pre}^i(T)$ for some $i \geq 1$ (by [14, Lemma 4]) and thus $T \subseteq \mathsf{Pre}^{n \cdot i}(T)$ for all $n \geq 0$ by monotonicity of $\mathsf{Pre}^i(\cdot)$, which entails for $n \cdot i \geq k$ that $T \subseteq \mathsf{Pre}^m(R)$ where $m = (n \cdot i - k) \mod r$ and thus d_T is sure (and almost-sure) winning for the eventually synchronizing objective in target R.

Note also that (1) implies (2) since by (1) we can play a sure-winning strategy from q_{init} to ensure in finitely many steps probability 1 in $\mathsf{Pre}(T)$ and in the next step probability 1 in T, and by (b) play an almost-sure winning strategy for eventually synchronizing in R. Hence $q_{\mathsf{init}} \in \langle\langle 1 \rangle\rangle^{event}_{almost}(sum_R)$ and thus $(2b)$ holds, which concludes the proof of Claim 1.

We now show that there exists an almost-sure winning strategy for the weakly synchronizing objective in target T.

Recall that $\mathsf{Pre}^r(R) = R$ and thus once some probability mass p is in R, it is possible to ensure that the probability mass in R after r steps is at least p, and thus that (with period r) the probability in R does not decrease. By the result of [14, Lemma 9], almost-sure winning for eventually synchronizing in R implies that there exists a strategy α such that the probability in R tends to 1 at periodic positions: for some $0 \leq h < r$ the strategy α is *almost-sure eventually synchronizing in R with shift h*, that is $\forall \varepsilon > 0 \cdot \exists N \cdot \forall n \geq N : n \equiv h \mod r \implies \mathcal{M}^{\alpha}_n(R) \geq 1 - \varepsilon$. We also say that the initial distribution $d_0 = \mathcal{M}^{\alpha}_0$ is almost-sure eventually synchronizing in R with shift h.

Claim 2.

($\star$) *If $\mathcal{M}^{\alpha}_0$ is almost-sure eventually synchronizing in R with some shift h, then $\mathcal{M}^{\alpha}_i$ is almost-sure eventually synchronizing in R with shift $h - i \mod r$.*

($\star\star$) *Let t such that d_T is almost-sure eventually synchronizing in R with shift t. If a distribution is almost-sure eventually synchronizing in R with some shift h, then it is also almost-sure eventually synchronizing in R with shift $h + k + t \mod r$ (where we chose k such that $R = \mathsf{Pre}^k(T)$).*

Proof of Claim 2. The result ($\star$) immediately follows from the definition of shift, and we prove ($\star\star$) as follows. We show that almost-sure eventually synchronizing in R with shift h implies almost-sure eventually synchronizing in R with shift $h + k + t \mod r$. Intuitively, the probability mass that is in R with shift h can be injected in T in k steps, and then from T we can play an almost-sure eventually synchronizing strategy in target R with shift t, thus a total shift of $h + k + t \mod r$. Precisely, an almost-sure winning strategy α is constructed as follows: given a finite prefix of play ρ, if there is no state $q \in R$ that occurs in ρ at a position $n \equiv h \mod r$, then play in ρ according to the almost-sure winning strategy α_h for eventually synchronizing in R with shift h. Otherwise, if there is no $q \in T$ that occurs in ρ at a position $n \equiv h + k \mod r$, then we play according to a sure winning strategy α_{sure} for eventually synchronizing in T, and otherwise we play according to an almost-sure winning strategy α_t from T for eventually synchronizing in R with shift t. To show that α is almost-sure eventually synchronizing in R with shift $h + k + t$, note that α_h ensures with probability 1 that R is reached at positions $n \equiv h \mod r$, and thus T is reached at positions $h + k \mod r$ by α_{sure}, and from the states in T the strategy α_t ensures with probability 1 that R is reached at positions $h + k + t \mod r$. This concludes the proof of Claim 2.

Construction of an almost-sure winning strategy. We construct strategies α_ε for $\varepsilon > 0$ that ensure, from a distribution that is almost-sure eventually synchronizing in R (with some shift h), that after finitely many steps, a distribution d' is reached such that $d'(T) \geq 1 - \varepsilon$ and d' is almost-sure eventually synchronizing in R (with some shift h'). Since q_{init} is almost-sure eventually synchronizing in R (with some shift h), it follows that the strategy α_{as} that plays successively the strategies (each for finitely many steps) $\alpha_{\frac{1}{2}}, \alpha_{\frac{1}{4}}, \alpha_{\frac{1}{8}}, \dots$ is almost-sure winning from q_{init} for the weakly synchronizing objective in target T.

We define the strategies α_ε as follows. Given an initial distribution that is almost-sure eventually synchronizing in R with a shift h and given $\varepsilon > 0$, let α_ε be the strategy that plays according to the almost-sure winning strategy α_h for eventually synchronizing in R with shift h for a number of steps $n \equiv h \mod r$ until a distribution d is reached such that $d(R) \geq 1 - \varepsilon$, and then from d it plays according to a sure winning strategy α_{sure} for eventually synchronizing in T from the states in R (for k steps), and keeps playing according to α_h from the states in $Q \setminus R$ (for k steps). The distribution d' reached from d after k steps is such that $d'(T) \geq 1 - \varepsilon$ and we claim that it is almost-sure eventually synchronizing in R with shift t. This holds by definition from the states in $\mathsf{Supp}(d') \cap T$, and by ($\star$) the states in $\mathsf{Supp}(d') \setminus T$ are almost-sure eventually synchronizing in R with shift $h - (h + k) \mod r$, and by ($\star\star$) with shift $h - (h + k) + k + t = t$.

It follows that the strategy α_{as} is well-defined and ensures, for all $\varepsilon > 0$, that the probability mass in T is infinitely often at least $1 - \varepsilon$, thus is almost-sure weakly synchronizing in T. This concludes the proof of Theorem 3. □

The complexity results of Theorem 1 and Theorem 2 hold for the membership problem with function max_T by the following lemma.

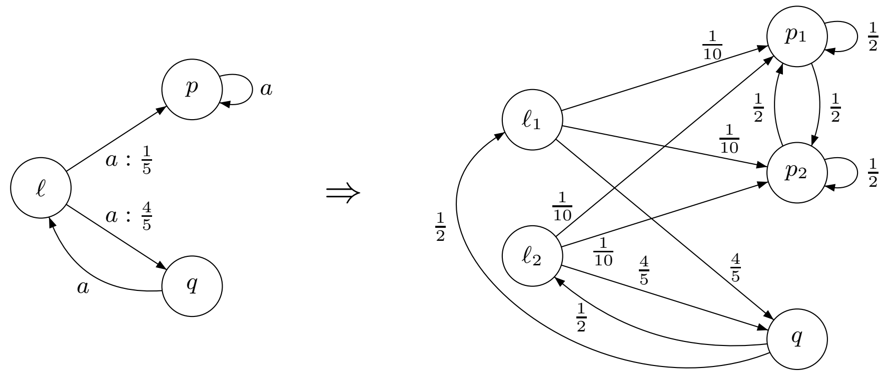

Fig. 7. State duplication ensures that the probability mass can never be accumulated in a single state except in q (we omit action a for readability).

Lemma 8. *For weak synchronization and each winning mode, the membership problems with functions max and max_T are polynomial-time equivalent to the membership problem with function $sum_{T'}$ with a singleton T'.*

Proof. First, for $\mu \in \{sure, almost, limit\}$, we have $\langle\langle 1 \rangle\rangle_{\mu}^{weakly}(max_T) = \bigcup_{q \in T} \langle\langle 1 \rangle\rangle_{\mu}^{weakly}(q)$, showing that the membership problems for max are polynomial-time reducible to the corresponding membership problem for sum_T with singleton T. The reverse reduction is as follows. Given an MDP $\mathcal{M}$, a state q and an initial distribution d_0, we can construct an MDP $\mathcal{M}'$ and initial distribution d_0' such that $d_0 \in \langle\langle 1 \rangle\rangle_{\mu}^{weakly}(q)$ iff $d_0' \in \langle\langle 1 \rangle\rangle_{\mu}^{weakly}(max_{Q'})$ where Q' is the state space of $\mathcal{M}'$ (thus $max_{Q'}$ is simply the function max). The idea is to construct $\mathcal{M}'$ and d_0' as a copy of $\mathcal{M}$ and d_0 where all states except q are duplicated, and the initial and transition probabilities are equally distributed between the copies (see Fig. 7). Therefore if the probability tends to 1 in some state, it has to be in q.

□

4 Strong Synchronization

In this section, we show that the membership problem for strongly synchronizing objectives can be solved in polynomial time, for all winning modes, and both with function max_T and function sum_T. We show that linear-size memory is necessary in general for max_T, and memoryless strategies are sufficient for sum_T. It follows from our results that the limit-sure and almost-sure winning modes coincide for strong synchronization.

4.1 Strong synchronization with function *max*

First, note that for strong synchronization the membership problem with function max_T reduces to the membership problem with function max_Q where Q is the entire state space, by a construction similar to the proof of Lemma 8: states in $Q \setminus T$ are duplicated, ensuring that only states in T are used to accumulate probability.

The strongly synchronizing objective with function *max* requires that from some point on, almost all the probability mass is at every step in a single state. The sequence of states that contain almost all the probability corresponds to a sequence of deterministic transitions in the MDP, and thus eventually to a cycle of deterministic transitions.

The *graph of deterministic transitions* of an MDP $\mathcal{M} = \langle Q, \mathsf{A}, \delta \rangle$ is the directed graph $G = \langle Q, E \rangle$ where $E = \{\langle q_1, q_2 \rangle \mid \exists a \in \mathsf{A} : \delta(q_1, a)(q_2) = 1\}$. For $\ell \geq 1$, a *deterministic cycle in* $\mathcal{M}$ of length ℓ is a finite path $\hat{q}_\ell \hat{q}_{\ell-1} \cdots \hat{q}_0$ in G (that is, $\langle \hat{q}_i, \hat{q}_{i-1} \rangle \in E$ for all $1 \leq i \leq \ell$) such that $\hat{q}_0 = \hat{q}_\ell$. The cycle is *simple* if $\hat{q}_i \neq \hat{q}_j$ for all $1 \leq i < j \leq \ell$.

We show that sure (resp., almost-sure and limit-sure) strong synchronization is equivalent to sure (resp., almost-sure and limit-sure) reachability to a state in such a cycle, with the requirement that it can be reached in a synchronized way, that is by finite paths whose lengths are congruent modulo the length ℓ of the cycle. To check this, we keep track of a modulo-ℓ counter along the play.

Define the MDP $\mathcal{M} \times [\ell] = \langle Q', \mathsf{A}, \delta' \rangle$ where $Q' = Q \times \{0, 1, \cdots, \ell - 1\}$ and $\delta'(\langle q, i \rangle, a)(\langle q', i - 1 \rangle) = \delta(q, a)(q')$ (where $i - 1$ is $\ell - 1$ for $i = 0$) for all states $q, q' \in Q$, actions $a \in \mathsf{A}$, and $0 \leq i \leq \ell - 1$.

Lemma 9. *Let η be the smallest positive probability in the transitions of $\mathcal{M}$, and let $\frac{1}{1+\eta} < p \leq 1$. There exists a strategy α such that $\liminf_{n\to\infty} \|\mathcal{M}_n^\alpha\| \geq p$ from an initial state q_{init} if and only if there exists a simple deterministic cycle $\hat{q}_\ell \hat{q}_{\ell-1} \cdots \hat{q}_0$ in $\mathcal{M}$ and a strategy β in $\mathcal{M} \times [\ell]$ such that $\Pr^\beta(\Diamond\{\langle \hat{q}_0, 0 \rangle\}) \geq p$ from $\langle q_{\mathsf{init}}, 0 \rangle$.*

Proof. First, assume that there exists a simple deterministic cycle $\hat{q}_\ell \hat{q}_{\ell-1} \cdots \hat{q}_0$ with length ℓ and a strategy β in $\mathcal{M} \times [\ell]$ that ensures the target set $\Diamond\{\langle \hat{q}_0, 0 \rangle\}$ is reached with probability at least p from the state $\langle q_{\mathsf{init}}, 0 \rangle$. Since randomization is not necessary for reachability objectives, we can assume that β is a pure strategy. We show that there exists a strategy α such that $\liminf_{n\to\infty} \|\mathcal{M}_n^\alpha\| \geq p$ from q_{init}. From β, we construct a pure strategy α in $\mathcal{M}$. Given a finite path $\rho = q_0 a_0 q_1 a_1 \ldots q_n$ in $\mathcal{M}$ (with $q_0 = q_{\mathsf{init}}$), there is a corresponding path $\rho' = \langle q_0, k_0 \rangle a_0 \langle q_1, k_1 \rangle a_1 \ldots \langle q_n, k_n \rangle$ in $\mathcal{M} \times [\ell]$ where $k_i = -i \; mod \; \ell$. Since the sequence $k_0 k_1 \ldots$ is uniquely determined from ρ, there is a clear bijection between the paths in $\mathcal{M}$ and the paths in $\mathcal{M} \times [\ell]$ that we often omit to apply and mention. For ρ, we define α as follows: if $q_n = \hat{q}_{k_n}$, then there exists an action a such that $\mathsf{post}(\hat{q}_{k_n}, a) = \{\hat{q}_{k_{n+1}}\} = \{\hat{q}_{n+1}\}$ and we define $\alpha(\rho) = a$, otherwise let $\alpha(\rho) = \beta(\rho')$. Thus α mimics β unless a state q is reached at step n such that $q = \hat{q}_k$ where $k = -n \; mod \; \ell$, and then α switches to always playing actions that keeps $\mathcal{M}$ in the simple deterministic cycle $\hat{q}_\ell \hat{q}_{\ell-1} \cdots \hat{q}_0$. Below, we prove that given $\varepsilon > 0$ there exists k such that for all $n \geq k$, we have $\|\mathcal{M}_n^\alpha\| \geq p - \varepsilon$. It follows that $\liminf_{n\to\infty} \|\mathcal{M}_n^\alpha\| \geq p$ from q_{init}. Since $\Pr^\beta(\Diamond\{\langle \hat{q}_0, 0\rangle\}) \geq p$, there exists k such that $\Pr^\beta(\Diamond^{\leq k}\{\langle \hat{q}_0, 0\rangle\}) \geq p - \varepsilon$. We assume w.l.o.g. that $k \textit{ mod } \ell = 0$. For $i = 0, 1, \ldots, \ell - 1$, let $R_i = \{\langle \hat{q}_i, i\rangle\}$. Then trivially $\Pr^\beta(\Diamond^{\leq k} \bigcup_{i=0}^{\ell-1} R_i) \geq p - \varepsilon$ and since α agrees with β on all finite paths that do not (yet) visit $\bigcup_{i=0}^{\ell-1} R_i$, given a path ρ that visits $\bigcup_{i=0}^{\ell-1} R_i$ (for the first time), only actions that keep $\mathcal{M}$ in the simple cycle $\hat{q}_\ell \hat{q}_{\ell-1} \cdots \hat{q}_0$ are played by α and thus all continuations of ρ in the outcome of α will visit $\hat{q}_0$ after k steps (in total). It follows that $\Pr^\beta(\Diamond^k\{\langle \hat{q}_0, 0\rangle\}) \geq p - \varepsilon$, that is $\mathcal{M}_k^\alpha(\hat{q}_0) \geq p - \varepsilon$. Thus, $\|\mathcal{M}_k^\alpha\| \geq p - \varepsilon$. Since next, α will always play actions that keeps $\mathcal{M}$ looping through the cycle $\hat{q}_\ell \hat{q}_{\ell-1} \cdots \hat{q}_0$, we have $\|\mathcal{M}_n^\alpha\| \geq p - \varepsilon$ for all $n \geq k$.

Second, assume that there exists a strategy α such that $\liminf_{n\to\infty} \|\mathcal{M}_n^\alpha\| \geq p$ from q_{init}. Thus, for all $\varepsilon > 0$ there exists $k \in \mathbb{N}$ such that for all $n \geq k$ we have $\|\mathcal{M}_n^\alpha\| \geq p - \varepsilon$. Fix $\varepsilon < p - \frac{1}{1+\eta}$. Let k be such that for all $n \geq k$, there exists a unique state $\hat{p}_n$ such that $\mathcal{M}_n^\alpha(\hat{p}_n) \geq p - \varepsilon$. Below, we prove that for all $n \geq k$, there exists some action $a \in \mathsf{A}$ such that $\mathsf{post}(\hat{p}_n, a) = \{\hat{p}_{n+1}\}$. Assume towards contradiction that there exists $j > k$ such that for all a there exists $q \neq \hat{p}_{j+1}$ such that $\{q, \hat{p}_{j+1}\} \subseteq \mathsf{post}(\hat{p}_j, a)$. Therefore, $\mathcal{M}_{j+1}^\alpha(q) \geq \mathcal{M}_j^\alpha(\hat{p}_j) \cdot \eta \geq (p - \varepsilon) \cdot \eta$. Hence,

$$\mathcal{M}_{j+1}^\alpha(\hat{p}_{j+1}) \leq 1 - \mathcal{M}_{j+1}^\alpha(q) \leq 1 - (p - \varepsilon) \cdot \eta.$$

Thus, $p - \varepsilon \leq \|\mathcal{M}_{j+1}^\alpha\| \leq 1 - (p - \varepsilon) \cdot \eta$ that gives $\varepsilon \geq p - \frac{1}{1+\eta}$, a contradiction. This argument proves that for all $n \geq k$, there exists an action $a \in \mathsf{A}$ such that $\mathsf{post}(\hat{p}_n, a) = \{\hat{p}_{n+1}\}$. The finiteness of the state space Q entails that in the sequence $\hat{p}_k \hat{p}_{k+1} \cdots$, some state and thus some simple deterministic cycle occur infinitely often. Let $\hat{q}_\ell \hat{q}_{\ell-1} \cdots \hat{q}_0$ be a cycle that occurs infinitely often in the sequence $\hat{p}_k \hat{p}_{k+1} \cdots$ (in the right order). For all j, let i_j be the position of $\hat{q}_0$ in all occurrences of the cycle $\hat{q}_\ell \hat{q}_{\ell-1} \cdots \hat{q}_0$ in the sequence $\hat{p}_k \hat{p}_{k+1} \cdots$; and let $t_j = i_j \textit{ mod } \ell$. In the sequence $t_0 t_1 \cdots$, there exists $0 \leq t < \ell$ that appears infinitely often. Let the cycle $r_\ell r_{\ell-1} \cdots r_0$ be such that $r_{(i+t) \textit{ mod } \ell} = \hat{q}_i$. Then, the cycle $r_\ell r_{\ell-1} \cdots r_0$ happens infinitely often in the sequence $\hat{p}_k \hat{p}_{k+1} \cdots$ such that the positions of r_0 are infinitely often 0 (modulo ℓ). Therefore, the probability of $\mathcal{M}$ to be in r_0 in positions (modulo ℓ) equals to 0, is infinitely often equal or greater than p. Hence, for a strategy β in $\mathcal{M} \times [\ell]$ that copies all the plays of the strategy α, we have $\Pr^\beta(\Diamond\{\langle r_0, 0\rangle\}) \geq p$ from $\langle q_{\mathsf{init}}, 0\rangle$. □

It follows directly from Lemma 9 with $p = 1$ that almost-sure strong synchronization is equivalent to almost-sure reachability to a deterministic cycle in $\mathcal{M} \times [\ell]$. The same equivalence holds for the sure and limit-sure winning modes.

Lemma 10. *A state q_{init} is sure (resp., almost-sure or limit-sure) winning for the strongly synchronizing objective (according to max_Q) if and only if there exists a simple deterministic cycle $\hat{q}_\ell \hat{q}_{\ell-1} \cdots \hat{q}_0$ such that $\langle q_{\mathsf{init}}, 0\rangle$ is sure (resp., almost-sure or limit-sure) winning for the reachability objective $\Diamond\{\langle \hat{q}_0, 0\rangle\}$ in $\mathcal{M} \times [\ell]$.*

Proof. The proof is organized in three sections:

(1) sure winning mode: First, assume that there exists a simple deterministic cycle $\hat{q}_\ell \hat{q}_{\ell-1} \cdots \hat{q}_0$ with length ℓ such that $\langle q_{\mathsf{init}}, 0\rangle$ is sure winning for the reachability objective $\Diamond\{\langle \hat{q}_0, 0\rangle\}$. Thus, there exists a pure memoryless strategy β such that $\mathit{Outcomes}(\langle q_{\mathsf{init}}, 0\rangle, \beta) \subseteq \Diamond\{\langle \hat{q}_0, 0\rangle\}$. Since β is memoryless, there must be $k \leq |Q| \times \ell$ such that $\mathit{Outcomes}(\langle q_{\mathsf{init}}, 0\rangle, \beta) \subseteq \Diamond^{\leq k}\{\langle \hat{q}_0, 0\rangle\}$ meaning that all infinite paths starting in $\langle q_{\mathsf{init}}, 0\rangle$ and following β reach $\langle \hat{q}_0, 0\rangle$ within k steps. From β, we construct a pure finite-memory strategy α in $\mathcal{M}$ that is represented by $T = \langle \mathsf{Mem}, i, \alpha_u, \alpha_n\rangle$ where $\mathsf{Mem} = \{0, \cdots, \ell-1\}$ is the set of modes. The idea is that α simulates what β plays in the state $\langle q, i\rangle$, in the state q of $\mathcal{M}$ and mode i of T (there is only one exception). Thus, the initial mode is 0. The update function only decreases modes by 1 ($\alpha_u(i, a, q) = (i-1)\ \mathit{mod}\ \ell$ for all states q and actions a) since by taking any transition the mode is decreased by 1. The next-move function $\alpha_n(i, q)$ is defined as follows: $\alpha_n(i, q) = \beta(\langle q, i\rangle)$ for all states q and modes $0 \leq i < \ell$, except when $q = \hat{q}_i$, in this case let $\alpha_n(i, q) = a$ where $\mathsf{post}(\hat{q}_i, a) = \{q_{i-1}\}$. Thus β mimics α unless a state q is reached at step n such that $q = \hat{q}_{-n\ \mathit{mod}\ \ell}$, and then α switches to always playing actions that keeps $\mathcal{M}$ in the simple deterministic cycle $\hat{q}_\ell \hat{q}_{\ell-1} \cdots \hat{q}_0$. Now we prove that q_{init} is sure winning for the strongly synchronizing objective according to max_Q. Let $j \geq k$ be such that $j\ \mathit{mod}\ \ell = 0$. Let $R = \{\langle \hat{q}_i, i\rangle \mid 0 \leq i < \ell\}$. Thus obviously $\mathit{Outcomes}(\langle q_{\mathsf{init}}, 0\rangle, \beta) \subseteq \Diamond R$. and since α agrees with β on all finite paths that do not (yet) visit R, given a path ρ that visits R (for the first time), only actions that keep $\mathcal{M}$ in the simple cycle $\hat{q}_\ell \hat{q}_{\ell-1} \cdots \hat{q}_0$ are played by α and thus all continuations of ρ in the outcome of α will visit $\hat{q}_0$ after j steps. It follows that $\Pr^\beta(\Diamond^j\{\langle \hat{q}_0, 0\rangle\}) = 1$, that is $\mathcal{M}^\alpha_j(q_0) = 1$. Thus, $\|\mathcal{M}^\alpha_j\| = 1$. Since next, α will always play actions that keeps $\mathcal{M}$ looping through the cycle $\hat{q}_\ell \hat{q}_{\ell-1} \cdots \hat{q}_0$, we have $\|\mathcal{M}^\alpha_n\| = 1$ for all $n \geq j$.

Second, assume that there exists a strategy α and k such that for all $n \geq k$ we have $\|\mathcal{M}^\alpha_n\| = 1$ from the initial state q_{init}. For all $n \geq k$, let $\hat{p}_n$ be a state such that $\mathcal{M}^\alpha_n(\hat{p}_n) = 1$. The finiteness of the state space Q entails that in the sequence $\hat{p}_k \hat{p}_{k+1} \cdots$, some state and thus some simple deterministic cycle occur infinitely often. Let $\hat{q}_\ell \hat{q}_{\ell-1} \cdots \hat{q}_0$ be a cycle that occurs infinitely often in the sequence $\hat{p}_k \hat{p}_{k+1} \cdots$ (in the right order). For all j, let i_j be the position of $\hat{q}_0$ in all occurrences of the cycle $\hat{q}_\ell \hat{q}_{\ell-1} \cdots \hat{q}_0$ in the sequence $\hat{p}_k \hat{p}_{k+1} \cdots$; and let $t_j = i_j\ \mathit{mod}\ \ell$. In the sequence $t_0 t_1 \cdots$, there exists $0 \leq t < \ell$ that appears infinitely often. Let the cycle $r_\ell r_{\ell-1} \cdots r_0$ be such that $r_{(i+t)\ \mathit{mod}\ \ell} = \hat{q}_i$. Then, the cycle $r_\ell r_{\ell-1} \cdots r_0$ happens infinitely often in the sequence $\hat{p}_k \hat{p}_{k+1} \cdots$ such that the positions of r_0 are infinitely often 0 (modulo ℓ). Hence, for a strategy β in $\mathcal{M} \times [\ell]$ that copies all the plays of the strategy α, we have $\mathit{Outcomes}(\langle q_{\mathsf{init}}, 0\rangle, \beta) \subseteq \Diamond\{\langle r_0, 0\rangle\}$ from the initial state $\langle q_{\mathsf{init}}, 0\rangle$.

(2) almost-sure winning mode: This case is an immediate result from Lemma 9, by taking $p = 1$.

(3) limit-sure winning mode: First, assume that there exists a simple deterministic cycle $\hat{q}_\ell \hat{q}_{\ell-1} \cdots \hat{q}_0$ with length ℓ such that $\langle q_{\mathsf{init}}, 0\rangle$ is limit-sure (and thus almost-sure) winning for the reachability objective $\Diamond\{\langle \hat{q}_0, 0\rangle\}$).

Since $\langle \hat{q}_{\mathsf{init}}, 0\rangle$ is almost-sure for reachability objective, then q_{init} is almost-sure (and thus limit-sure) for strongly synchronizing objective. Second, assume that q_{init} is limit-sure winning for the strongly synchronizing objective (according to max_Q). It means that for all i there exists a strategy α_i such that $\liminf_{n\to\infty} \|\mathcal{M}_n^{\alpha_i}\| \geq 1 - 2^{-i}$. Let k be such that $1 - 2^{-k} \geq \frac{1}{1+\eta}$. By Lemma 9, for all $i \geq k$ there exists a simple deterministic cycle $c_i = \hat{p}_{\ell_i}\hat{p}_{\ell_i-1}\cdots\hat{p}_0$ with length ℓ_i and a strategy β_i in $\mathcal{M} \times [\ell_i]$ such that $\Pr^{\beta_i}(\Diamond\{\langle \hat{q}_0, 0\rangle\}) \geq 1 - 2^{-i}$ from $\langle q_{\mathsf{init}}, 0\rangle$. Since the number of simple deterministic cycle is finite, there exists some simple cycle c that occurs infinitely often in the sequence $c_k c_{k+1} c_{k+2}\cdots$. We see that for the cycle $c = \hat{q}_\ell \hat{q}_{\ell-1}\cdots\hat{q}_0$, the states $\langle \hat{q}_{\mathsf{init}}, 0\rangle$ is limit-sure winning for the reachability objective $\Diamond\{\langle \hat{q}_0, 0\rangle\}$). ☐

Since the winning regions of almost-sure and limit-sure winning coincide for reachability objectives in MDPs [3], the next corollary follows from Lemma 10.

Corollary 1. $\langle\!\langle 1 \rangle\!\rangle_{limit}^{strongly}(max_T) = \langle\!\langle 1 \rangle\!\rangle_{almost}^{strongly}(max_T)$ *for all target sets* T.

If there exists a cycle c satisfying the condition in Lemma 10, then all cycles reachable from c in the graph G of deterministic transitions also satisfies the condition. Hence it is sufficient to check the condition for an arbitrary simple cycle in each strongly connected component (SCC) of G. It follows that strong synchronization can be decided in polynomial time (SCC decomposition can be computed in polynomial time, as well as sure, limit-sure, and almost-sure reachability in MDPs). The length of the cycle gives a linear bound on the memory needed to win, and the bound is tight.

Theorem 4. *For the three winning modes of strong synchronization according to* max_T *in MDPs:*

1. *(Complexity). The membership problem is PTIME-complete.*
2. *(Memory). Linear memory is necessary and sufficient for both pure and randomized strategies, and pure strategies are sufficient.*

Proof. First, we prove the PTIME upper bound. Given an MDP $\mathcal{M} = \langle Q, \mathsf{A}, \delta\rangle$ and a state q_{init}, we say a simple deterministic cycle $c = \hat{q}_\ell \hat{q}_{\ell-1}\cdots\hat{q}_0$ is sure winning (resp., almost-sure and limit-sure) for strong synchronization from q_{init} if $\langle q_{\mathsf{init}}, 0\rangle$ is sure winning (resp., almost-sure and limit-sure) for the reachability objective $\Diamond\{\langle \hat{q}_0, 0\rangle\}$ in $\mathcal{M} \times [\ell]$. We show that if c is sure winning (resp., almost-sure and limit-sure) for strong synchronization from q_{init}, then so are all simple cycles $c' = \hat{p}_{\ell'}\hat{p}_{\ell'-1}\cdots\hat{p}_0$ reachable from c in the deterministic digraph induced by $\mathcal{M}$.

(1) sure winning: Since c is sure winning for strong synchronization from q_{init}, $\mathcal{M}$ is 1-synchronized in $\hat{q}_0$. Since there is a path via deterministic transitions from $\hat{q}_0$ to $\hat{p}_0$, $\mathcal{M}$ is 1-synchronized in $\hat{p}_0$ too. So the cycle c' is sure wining for strong synchronization from q_{init}, too.

(2) limit-sure winning: Assume that c is limit-sure winning for strong synchronization from q_{init}. By definition, for all $i \in \mathbb{N}$, there exists n such that for all $j > n$ we have $\mathcal{M}$ is $1-2^{-i-j}$ in $\hat{q}_0$. It implies that for all i there exists n such that $\mathcal{M}$ is $1-2^{-2i}$-synchronized in $\hat{q}_0$. Since there is a path via deterministic transitions from $\hat{q}_0$ to $\hat{p}_0$, then $\mathcal{M}$ is $1-2^{-2i}$-synchronized in $\hat{p}_0$ for all i. So the cycle c' is limit-sure wining for strong synchronization from q_{init}, too.

(3) almost-sure winning: By corollary 1, since a cycle is almost-sure winning for strong synchronization from q_{init} if and only if it is limit-sure winning, the results follows.

The above arguments prove that if a simple deterministic cycle c is sure winning (resp., almost-sure and limit-sure) for strongly synchronizing objective from q_{init}, then all simple cycles reachable from c in the graph of deterministic transitions G induced by $\mathcal{M}$, are sure winning (resp., almost-sure and limit-sure). In particular, it holds for all simple cycles in the bottoms SCCs reachable from c in G. Therefore, to decide membership problem for strongly synchronizing objective, it suffices to only check whether one cycle in each bottom SCC of G is sure winning (resp., almost-sure and limit-sure). Since the SCC decomposition for a digraph is in PTIME, and since the number of bottom SCCs in a connected digraph is at most the size of the digraph (the number of states $|Q|$), the PTIME upper bound follows.

For the PTIME-hardness, for all $\mu \in \{sure, almost, limit\}$ the proof is by a reduction from monotone Boolean circuit problem (MBC). Given an MBC with an underlying binary tree, the value of leaves are either 0 or 1 (called 0 or 1-leaves), and the value of other vertices, labeled with $\wedge$ or $\vee$, are computed inductively. It is known that deciding whether the value of the *root* is 1 for a given MBC, is PTIME-complete [16]. From an MBC, we construct an MDP $\mathcal{M}$ where the states are the vertices of the tree with three new absorbing states sync, q_1 and q_2, and two actions L, R. On both actions L and R, the next successor of the 1-leaves is only sync, and the next successor of the 0-leaves is q_1 or q_2 with probability $\frac{1}{2}$. The next successor of a $\wedge$-state is each of their children with equal probability, on all actions. The next successor of a $\vee$-state is its left (resp., right) child by action L (resp., action R). We can see that $\mathcal{M}$ can be synchronized only in sync.

We call a subtree *complete* if (1) *root* is in subtree, (2) at least one child of all $\vee$-vertices is in the subtree, (3) both children of all $\wedge$-vertices are in the subtree. There is a bijection between a complete subtree and a strategy in $\mathcal{M}$. The value of *root* is 1 if and only if there is a complete subtree such that it has no 0-leaves (all leaves are 1-leaves). For such subtrees, all plays under the corresponding strategy reach some 1-leave and thus are synchronized in sync. It means that $root \in \langle\langle 1 \rangle\rangle_{\mu}^{strongly}(\mathsf{sync})$ if and only if the value of *root* is 1.

Finally, the result on memory requirement is established as follows. Since memoryless strategies are sufficient for reachability objectives in MDPs, it follows from the proof of Lemma 9 and Lemma 10 that the (memoryless) winning strategies in $\mathcal{M} \times [\ell]$ can be transferred to winning strategies with memory $\{0, 1, \cdots, \ell-1\}$ in $\mathcal{M}$. Since $\ell \leq |Q|$, linear-size memory is sufficient to win strongly synchronizing objectives. We present a family of MDPs $\mathcal{M}_n$ ($n \in \mathbb{N}$) that are sure winning for strongly synchronization (according to $\max_Q$), and

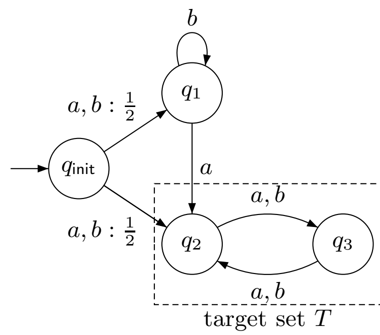

Fig. 8. An MDP that all strategies to win sure strongly synchronizing with function $max_{\{q_2,q_3\}}$ require memory.

where the sure winning strategies require linear memory. The MDP $\mathcal{M}_2$ is shown in Fig. 8, and the MDP $\mathcal{M}_n$ is obtained by replacing the cycle q_2q_3 of deterministic transitions by a simple cycle of length n. Note that only in q_1 there is a real strategic choice. Since q_1 and q_2 contain probability, we need to wait in q_1 (by playing b) until we play a when the probability in q_2 comes back in q_2 through the cycle. We need to play $n-1$ times b, and then a, thus linear memory is sufficient and it is easy to show that it is necessary to ensure strongly synchronization. ◻

4.2 Strong synchronization with function *sum*

The strongly synchronizing objective with function sum_T requires that eventually all the probability mass remains in T. We show that this is equivalent to a traditional reachability objective with target defined by the set of sure winning initial distributions for the safety objective $\Box T$.

It follows that almost-sure (and limit-sure) winning for strong synchronization is equivalent to almost-sure (or equivalently limit-sure) winning for the coBüchi objective $\Diamond\Box T = \{q_0 a_0 q_1 \cdots \in \mathsf{Play}(\mathcal{M}) \mid \exists j \cdot \forall i > j : q_i \in T\}$. However, sure strong synchronization is not equivalent to sure winning for the coBüchi objective: the MDP in Fig. 9 is sure winning for the coBüchi objective $\Diamond\Box\{q_{\mathsf{init}}, q_2\}$ from q_{init}, but not sure winning for the reachability objective $\Diamond S$ where $S = \{q_2\}$ is the winning region for the safety objective $\Box\{q_{\mathsf{init}}, q_2\}$ (and thus not sure strongly synchronizing). Note that this MDP is almost-sure strongly synchronizing in target $T = \{q_{\mathsf{init}}, q_2\}$ from q_{init}, and almost-sure winning for the coBüchi objective $\Diamond\Box T$, as well as almost-sure winning for the reachability objective $\Diamond S$.

Lemma 11. *Given a target set T, an MDP $\mathcal{M}$ is sure (resp., almost-sure or limit-sure) winning for the strongly synchronizing objective according to sum_T if*

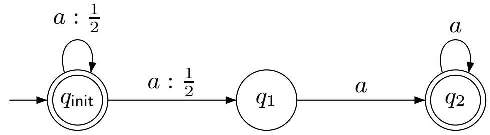

Fig. 9. An MDP such that q_{init} is sure-winning for coBüchi objective in $T = \{q_{\mathsf{init}}, q_2\}$ but not for strong synchronization according to sum_T.

and only if $\mathcal{M}$ is sure (resp., almost-sure or limit-sure) winning for the reachability objective $\Diamond S$ where S is the sure winning region for the safety objective $\Box T$.

Proof. First, assume that a state q_{init} of $\mathcal{M}$ is sure (resp., almost-sure or limit-sure) winning for the strongly synchronizing objective according to sum_T, and show that q_{init} is sure (resp., almost-sure or limit-sure) winning for the reachability objective $\Diamond S$.

(i) *Limit-sure winning.* For all $\varepsilon > 0$, let $\varepsilon' = \frac{\varepsilon}{|Q|} \cdot \eta^{|Q|}$ where η is the smallest positive probability in the transitions of $\mathcal{M}$. By the assumption, from q_{init} there exists a strategy α and $N \in \mathbb{N}$ such that for all $n \geq N$, we have $\mathcal{M}_n^\alpha(T) \geq 1 - \varepsilon'$. We claim that at step N, all non-safe states have probability at most $\frac{\varepsilon}{|Q|}$, that is $\mathcal{M}_N^\alpha(q) \leq \frac{\varepsilon}{|Q|}$ for all $q \in Q \setminus S$. Towards contradiction, assume that $\mathcal{M}_N^\alpha(q) > \frac{\varepsilon}{|Q|}$ for some non-safe state $q \in Q \setminus S$. Since $q \notin S$ is not safe, there is a path of length $\ell \leq |Q|$ from q to a state in $Q \setminus T$, thus with probability at least $\eta^{|Q|}$. It follows that after $N + \ell$ steps we have $\mathcal{M}_{N+\ell}^\alpha(Q \setminus T) > \frac{\varepsilon}{|Q|} \cdot \eta^{|Q|} = \varepsilon'$, in contradiction with the fact $\mathcal{M}_n^\alpha(T) \geq 1 - \varepsilon'$ for all $n \geq N$. Now, since all non-safe states have probability at most $\frac{\varepsilon}{|Q|}$ at step N, it follows that $\mathcal{M}_N^\alpha(Q \setminus S) \leq \frac{\varepsilon}{|Q|} \cdot |Q| = \varepsilon$ and thus $\Pr^\alpha(\Diamond S) \geq 1 - \varepsilon$. Therefore $\mathcal{M}$ is limit-sure winning for the reachability objective $\Diamond S$ from q_{init}.

(ii) *Almost-sure winning.* Since almost-sure strongly synchronizing implies limit-sure strongly synchronizing, it follows from (i) that $\mathcal{M}$ is limit-sure (and thus also almost-sure) winning for the reachability objective $\Diamond S$, as limit-sure and almost-sure reachability coincide for MDPs [3].

(iii) *Sure winning.* From q_{init} there exists a strategy α and $N \in \mathbb{N}$ such that for all $n \geq N$, we have $\mathcal{M}_n^\alpha(T) = 1$. Hence α is sure winning for the reachability objective $\Diamond \mathsf{Supp}(\mathcal{M}_N^\alpha)$, and from all states in $\mathsf{Supp}(\mathcal{M}_N^\alpha)$ the strategy α ensures that only states in T are visited. It follows that $\mathsf{Supp}(\mathcal{M}_N^\alpha) \subseteq S$ is sure winning for the safety objective $\Box T$, and thus α is sure winning for the reachability objective $\Diamond S$ from q_{init}.

For the converse direction of the lemma, assume that a state q_{init} is sure (resp., almost-sure or limit-sure) winning for the reachability objective $\Diamond S$. We construct a winning strategy for strong synchronization in T as follows: play according to a sure (resp., almost-sure or limit-sure) winning strategy for the reachability objective $\Diamond S$, and whenever a state of S is reached along the play, then switch to a winning strategy for the safety objective $\Box T$. The constructed strategy is sure (resp., almost-sure or limit-sure) winning for strong synchronization according to sum_T because for sure winning, after finitely many steps all paths from q_{init} end up in $S \subseteq T$ and stay in S forever, and for almost-sure (or equivalently limit-sure) winning, for all $\varepsilon > 0$, after sufficiently many steps, the set S is reached with probability at least $1 - \varepsilon$, showing that the outcome is strongly $(1 - \varepsilon)$-synchronizing in $S \subseteq T$, thus the strategy is almost-sure (and also limit-sure) strongly synchronizing. □

Corollary 2. $\langle\langle 1 \rangle\rangle^{strongly}_{limit}(sum_T) = \langle\langle 1 \rangle\rangle^{strongly}_{almost}(sum_T)$ *for all target sets* T.

The following result follows from Lemma 11 and the fact that the winning region for sure safety, sure reachability, and almost-sure reachability can be computed in polynomial time for MDPs [3]. Moreover, memoryless strategies are sufficient for these objectives.

Theorem 5. *For the three winning modes of strong synchronization according to* sum_T *in MDPs:*

1. *(Complexity). The membership problem is PTIME-complete.*
2. *(Memory). Pure memoryless strategies are sufficient.*

References

1. Agrawal, M., Akshay, S., Genest, B., Thiagarajan, P.S.: Approximate verification of the symbolic dynamics of Markov chains. In: Proc. of LICS. pp. 55–64. IEEE (2012)
2. de Alfaro, L., Henzinger, T.A.: Concurrent omega-regular games. In: Proc. of LICS. pp. 141–154 (2000)
3. de Alfaro, L., Henzinger, T.A., Kupferman, O.: Concurrent reachability games. Theor. Comput. Sci. 386(3), 188–217 (2007)
4. Bach, E., Shallit, J.: Algorithmic Number Theory, Vol. 1: Efficient Algorithms. MIT Press (1996)
5. Baier, C., Bertrand, N., Größer, M.: On decision problems for probabilistic Büchi automata. In: Proc. of FoSSaCS. pp. 287–301. LNCS 4962, Springer (2008)
6. Baldoni, R., Bonnet, F., Milani, A., Raynal, M.: On the solvability of anonymous partial grids exploration by mobile robots. In: Proc. of OPODIS. pp. 428–445. LNCS 5401, Springer (2008)
7. Burkhard, H.D.: Zum längenproblem homogener experimente an determinierten und nicht-deterministischen automaten. Elektronische Informationsverarbeitung und Kybernetik 12(6), 301–306 (1976)
8. Cerný, J.: Poznámka k. homogénnym experimentom s konecnymi automatmi. In: Matematicko-fyzikálny časopis. vol. 14(3), pp. 208–216 (1964)
9. Chadha, R., Korthikanti, V.A., Viswanathan, M., Agha, G., Kwon, Y.: Model checking MDPs with a unique compact invariant set of distributions. In: Proc. of QEST. pp. 121–130. IEEE Computer Society (2011)
10. Courcoubetis, C., Yannakakis, M.: The complexity of probabilistic verification. J. ACM 42(4), 857–907 (1995)
11. Doyen, L., Massart, T., Shirmohammadi, M.: Infinite synchronizing words for probabilistic automata. In: Proc. of MFCS. pp. 278–289. LNCS 6907, Springer (2011)

12. Doyen, L., Massart, T., Shirmohammadi, M.: Synchronizing objectives for Markov decision processes. In: Proc. of iWIGP: Interactions, Games and Protocols. pp. 61–75. EPTCS 50 (2011)
13. Doyen, L., Massart, T., Shirmohammadi, M.: Infinite synchronizing words for probabilistic automata (Erratum). CoRR abs/1206.0995 (2012)
14. Doyen, L., Massart, T., Shirmohammadi, M.: Limit synchronization in Markov decision processes. CoRR abs/1310.2935 (2013), to appear in Proc. of FoSSaCS 2014.
15. Filar, J., Vrieze, K.: Competitive Markov Decision Processes. Springer (1997)
16. Gibbons, A., Rytter, W.: Efficient parallel algorithms. Cambridge University Press (1988)
17. Gimbert, H., Oualhadj, Y.: Probabilistic automata on finite words: Decidable and undecidable problems. In: Proc. of ICALP (2). pp. 527–538. LNCS 6199, Springer (2010)
18. Henzinger, T.A., Mateescu, M., Wolf, V.: Sliding window abstraction for infinite Markov chains. In: Proc. of CAV. pp. 337–352. LNCS 5643, Springer (2009)
19. Imreh, B., Steinby, M.: Directable nondeterministic automata. Acta Cybern. 14(1), 105–115 (1999)
20. Kfoury, D.: Synchronizing sequences for probabilistic automata. Studies in Applied Mathematics 29, 101–103 (1970)
21. Korthikanti, V.A., Viswanathan, M., Agha, G., Kwon, Y.: Reasoning about MDPs as transformers of probability distributions. In: Proc. of QEST. pp. 199–208. IEEE Computer Society (2010)
22. Martyugin, P.: Computational complexity of certain problems related to carefully synchronizing words for partial automata and directing words for nondeterministic automata. Theory Comput. Syst. 54(2), 293–304 (2014)
23. Pinsky, M.A., Karlin, S.: An Introduction to Stochastic Modeling. Academic Press (2011)
24. Vardi, M.Y.: Automatic verification of probabilistic concurrent finite-state programs. In: Proc. of FOCS. pp. 327–338. IEEE Computer Society (1985)
25. Volkov, M.V.: Synchronizing automata and the Cerny conjecture. In: Proc. of LATA. pp. 11–27. LNCS 5196, Springer (2008)