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-ε)-synchronizing sequence; (*iii*) *limit-sure* winning if for all $\varepsilon > 0$, there is a strategy that produces a (1-ε)-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 [22, 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

⋆ 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).

⋆⋆ Fuller version: [1].

represents the possible choices of the controller, and the stochasticity represents the uncertainties about the system response. The controller synthesis problem is 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, 4].

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 [6, 14, 17], and recently it has been used in several works on design and verification of reactive systems [2, 9, 20]. While the verification of MDPs may yield undecidability, both with traditional specifications [5, 16], and symbolic specifications [20, 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 [12, 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 $\bar{X} = X_0 X_1 \dots$ 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

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

if a given MDP is winning by providing tight (matching) upper and lower bounds: 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 [23] 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, 18, 19, 12], leading to results of PSPACE-completeness [21], or even undecidability [19, 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 $\pi = q_0 a_0 q_1 a_1 \dots$ such that $q_{i+1} \in \mathsf{post}(q_i, a_i)$ for all $i \geq 0$. A finite prefix $\rho = q_0 a_0 q_1 a_1 \dots q_n$ 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 $\mathsf{Last}(\rho) = \mathsf{Last}(\rho')$. Finally, a strategy α uses *finite-memory* if there exists a right congruence $\approx$ over $\mathsf{Pref}(\mathcal{M})$ (i.e., if $\rho \approx \rho'$, then $\rho \cdot a \cdot q \approx \rho' \cdot a \cdot q$ for all $\rho, \rho' \in \mathsf{Pref}(\mathcal{M})$ and $a \in \mathsf{A}$, $q \in Q$) of finite index such that $\rho \approx \rho'$ implies $\alpha(\rho) = \alpha(\rho')$. The index of $\approx$ is the *memory size* of the strategy.

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 $Outcomes(d_0, \alpha)$ 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 [22]. 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 $Outcomes(d_0, \alpha) \subseteq \Omega$; *almost-sure winning* if there exists a strategy α such that $\Pr^{\alpha}(\Omega) = 1$; and *limit-sure winning* if $\sup_{\alpha} \Pr^{\alpha}(\Omega) = 1$.

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. 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 [20]. 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 $(\mathcal{M}_n^{\alpha})_{n \in \mathbb{N}}$ of probability distributions defined by $\mathcal{M}_k^{\alpha}(q) = Pr^{\alpha}(\Diamond^k \{q\})$ for all $k \geq 0$ and $q \in Q$. Hence, $\mathcal{M}_k^{\alpha}$ is the probability distribution over states after k steps under strategy α. Note that $\mathcal{M}_0^{\alpha} = d_0$ and the symbolic outcome is a deterministic sequence of distributions: each distribution $\mathcal{M}_k^{\alpha}$ has a unique (deterministic) successor.

Informally, synchronizing objectives require that the probability of a given 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 $\bar{X} = X_0 X_1 \dots$ of probability distributions is [12, 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 [3], 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;

<table>
<tr><td></td><td>Eventually</td><td>Weakly</td><td>Strongly</td></tr>
<tr><td>Sure</td><td>PSPACE-C [14]</td><td>PSPACE-C</td><td>PTIME-C</td></tr>
<tr><td>Almost-sure</td><td>PSPACE-C [14]</td><td rowspan="2">PSPACE-C</td><td rowspan="2">PTIME-C</td></tr>
<tr><td>Limit-sure</td><td>PSPACE-C [14]</td></tr>
</table>

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

- *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* $\langle\langle 1\rangle\rangle^{\lambda}_{sure}(f)$ 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 $\langle\langle 1\rangle\rangle^{\lambda}_{almost}(f)$ and $\langle\langle 1\rangle\rangle^{\lambda}_{limit}(f)$. For a singleton $T = \{q\}$ we have $sum_T = max_T$, and we simply write $\langle\langle 1\rangle\rangle^{\lambda}_{\mu}(q)$ (where $\mu \in \{sure, almost, limit\}$). 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 $\lambda \in \{weakly, strongly\}$.*

	Eventually	Weakly	Strongly	
			sum_T	max_T
Sure	exponential [14]	**exponential**	**memoryless**	**linear**
Almost-sure	infinite [14]	**infinite**	**memoryless**	**linear**
Limit-sure	unbounded [14]			

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

The *membership problem* is to decide, given an initial probability distribution d_0, whether $d_0 \in \langle\langle 1 \rangle\rangle_{\mu}^{\lambda}(f)$. 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 $q_{\mathsf{init}} \in \langle\langle 1 \rangle\rangle_{sure}^{weakly}(sum_T)$ 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 $n \geq 1$.*

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.*

The proof of Lemma 2 suggests an exponential-memory strategy for sure weak 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.

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*.*

Since the membership problems for sure eventually synchronizing and for limit-sure eventually synchronizing with exact support are PSPACE-complete ([14, Theorem 2 and Theorem 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.*

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.*

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. 1 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 $i \geq 0$: $\{q_4\}, \{q_3\}, \{q_2\}, \{q_3\}, \ldots$ is ultimately

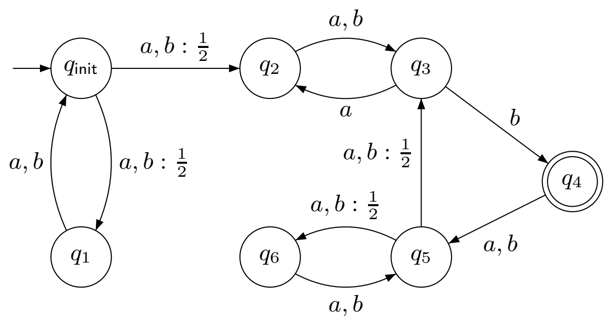

Fig. 1. An example to show $q_{\mathsf{init}} \in \langle\langle 1 \rangle\rangle^{weakly}_{limit}(q_4)$ implies $q_{\mathsf{init}} \in \langle\langle 1 \rangle\rangle^{weakly}_{almost}(q_4)$.

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 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 $\limsup$, 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.*

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

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

Finally, we note that the complexity results of Theorem 1 and Theorem 2 hold for the membership problem with functions *max* and max_T by the lemma below. 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.Therefore if the probability mass tends to 1 in some state, it has to be in q.

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'.*

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 $\langle q_{\mathsf{init}}, 0 \rangle$.*

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 $\mathcal{M} \times [\ell]$.*

Since the winning regions of almost-sure and limit-sure winning coincide for reachability objectives in MDPs [4], the next corollary follows from Lemma 10.

Corollary 1. *$\langle\!\langle 1 \rangle\!\rangle^{strongly}_{limit}(max_T) = \langle\!\langle 1 \rangle\!\rangle^{strongly}_{almost}(max_T)$ 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 satisfy 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.*

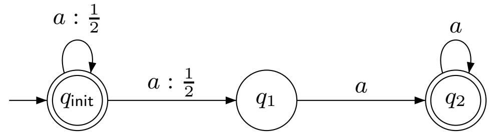

Fig. 2. An MDP such that q_{init} is sure-winning for coBüchi objective in $T = \{q_{\mathsf{init}}, q_2\}$ but not for strong synchronization according to sum_T.

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. 2 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 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 $\Box T$.*

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 sure winning region for safety and reachability, and the almost-sure winning region for reachability can be computed in polynomial time for MDPs [4]. 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. Full version. In: ArXiv CoRR: http://arxiv.org/abs/1402.2840 (2014)
2. Agrawal, M., Akshay, S., Genest, B., Thiagarajan, P.S.: Approximate verification of the symbolic dynamics of Markov chains. In: LICS. pp. 55–64. IEEE (2012)
3. de Alfaro, L., Henzinger, T.A.: Concurrent omega-regular games. In: Proc. of LICS. pp. 141–154 (2000)
4. de Alfaro, L., Henzinger, T.A., Kupferman, O.: Concurrent reachability games. Theor. Comput. Sci. 386(3), 188–217 (2007)
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: QEST. pp. 121–130. IEEE (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. 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. In: Proc. of FoSSaCS. pp. 58–72. LNCS 8412, Springer (2014)
15. Filar, J., Vrieze, K.: Competitive Markov Decision Processes. Springer (1997)
16. Gimbert, H., Oualhadj, Y.: Probabilistic automata on finite words: Decidable and undecidable problems. In: ICALP (2). pp. 527–538. LNCS 6199, Springer (2010)
17. 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)
18. Imreh, B., Steinby, M.: Directable nondeterministic automata. Acta Cybern. 14(1), 105–115 (1999)
19. Kfoury, D.: Synchronizing sequences for probabilistic automata. Studies in Applied Mathematics 29, 101–103 (1970)
20. Korthikanti, V.A., Viswanathan, M., Agha, G., Kwon, Y.: Reasoning about MDPs as transformers of probability distributions. In: QEST. pp. 199–208. IEEE (2010)
21. 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)
22. Vardi, M.Y.: Automatic verification of probabilistic concurrent finite-state programs. In: Proc. of FOCS. pp. 327–338. IEEE Computer Society (1985)
23. Volkov, M.V.: Synchronizing automata and the Cerny conjecture. In: Proc. of LATA. pp. 11–27. LNCS 5196, Springer (2008)