Statistic Analysis for Probabilistic Processes

Michel de Rougemont, Mathieu Tracol
University Paris-South, LRI-CNRS
Orsay, France
mdr@lri.fr, tracol@lri.fr

***Abstract*—We associate a statistical vector to a trace and a geometrical embedding to a Markov Decision Process, based on a distance on words, and study basic Membership and Equivalence problems. The Membership problem for a trace w and a Markov Decision Process S decides if there exists a strategy on S which generates with high probability traces close to w. We prove that Membership of a trace is *testable* and Equivalence of MDPs is polynomial time approximable. For Probabilistic Automata, Membership is not testable, and approximate Equivalence is undecidable. We give a class of properties, based on results concerning the structure of the tail sigma-field of a finite Markov chain, which characterizes equivalent Markov Decision Processes in this context.**

Keywords: Markov Decision Processes, Probabilistic Automata, State Action frequency, tail σ field, Property Testing, Approximation.

I. INTRODUCTION

We consider probabilistic systems with both non deterministic and probabilistic transitions, and basic questions concerning their traces such as *Statistical membership* for a given system and *Equivalence* of two systems which are known to be hard in this context [3], [16]. We study their approximation, based on Property testing, with a natural distance dist on words, and show that *Statistical membership* becomes *testable* and *Equivalence* polynomial time computable in the size of the system, but remain hard for *Probabilistic Automata*.

Property testing [22], [12] is a classical method to approximate decision problems, given a distance between two inputs. An *ε-tester* for a property P on words, is a randomized algorithm $\mathscr{A}$ which takes a word w_n of size n as input, and distinguishes with high probability between w_n satisfies P and w_n ε-far from P. A property P is *testable* if there exists a randomized algorithm such that for every $\varepsilon > 0$, $\mathscr{A}(\varepsilon)$ is an ε-tester for P whose time complexity only depends on ε, *i.e.* is independent of the size n.

Let $\mathscr{S}$ be an MDP (Markov Decision Process) of size m, $\lambda \leq 1$ a threshold value and $0 \leq \varepsilon \leq 1$. Given an input w_n of size n, the *Statistical membership* decides if there exists a strategy σ, which assigns decisions on each non deterministic state, such that $Prob_\sigma[\mathsf{dist}(r_n, w_n) \leq \varepsilon] \geq \lambda$, i.e. the probability to observe a trace r_n which is ε-close to w_n is greater than λ. Although this problem is PSPACE-hard in $Max(m,n)$, we will show that it is testable, i.e. can be approximated in time independent of n. We present a method which generalizes the approach introduced in [11] for non deterministic systems and considers the statistical behavior of a system. It associates a statistical vector x to w_n and a convex set $\mathscr{H}(\mathscr{S})$ of vectors to $\mathscr{S}$ in such a way that the geometrical distance between x and $\mathscr{H}(\mathscr{S})$ is close to the distance between w_n and the set of traces of $\mathscr{S}$. We define a distance between two MDPs $\mathscr{S}_1$ and $\mathscr{S}_2$ as the geometrical distance between $\mathscr{H}(\mathscr{S}_1)$ and $\mathscr{H}(\mathscr{S}_2)$ and show how to approximate it in polynomial time. A motivation for this statistical analysis is to decide if there are runs with some statistic constraints, such as the proportion of action a greater than ten percent.

In [26], Tzeng studied the equivalence between two probabilistic automata and proved that the exact equivalence is in PTIME. This result was extended in [10], where the authors study labeled Markov chains in a context close to ours. Tzeng defined the *approximate equivalence* in a natural way: two probabilistic automata are ε-close if for all words w, the probabilities to be accepted are ε-close. The undecidability of this problem is proved in [16]. In this context, given a word w_n, the *membership* simply decides if $Prob_A[w_n \text{ is accepted}] \geq \lambda$. We show that this property is not testable.

There are many other approaches which associate distances to probabilistic systems. In [27], [9], distances generalize the classical probabilistic bisimulation between two states and in [18], [24] the $\bar{D}$ generalizes the Trace Equivalence between two Markov chains. The distance introduced differentiates systems which have a different long term behavior, and is most relevant for systems which are not supposed to stop. Our main results are: a generalization of Derman's theorem to higher order statistics **(Theorem 1)**. The *Statistical membership* on MDPs is testable **(Theorem 2)**, whereas it is not testable **(Theorem 3)** for Probabilistic Automata. Approximate Equivalence is polynomial time computable for MDPs **(Proposition 4)** whereas it is undecidable for Probabilistic Automata. *Ultimate properties* characterize the MDPs and the Markov chains at distance 0 **(Theorem 5,6,7).**

In section 2, we review the main definitions for Testers, MDPs and State-Action Frequencies, Probabilistic Automata. In section 3, we generalize known results on MDPs and statistics to higher order statistics. In section 4 we define the *Statistical membership* and *Equivalence* problems, and prove positive results for MDPs. In section 5 we prove negative results for Probabilistic Automata. In section 6 we present a class of properties (ultimate properties), which characterizes exactly the equivalence and simulation relations between MDPs induced by the distance.

II. Preliminaries

A. Testers and Statistics on words

An *elementary operation* on a word w_n of size n on an alphabet Σ is an insertion, a deletion, a substitution of a single letter, or the *move* of a whole subword of w to another position. The *edit distance with moves* $\mathsf{dist}(w,w')$ between w and w' is the minimal number of elementary operations performed on w to obtain w', divided by $max\{|w|,|w'|\}$ as we only consider relative distances. The distance between w and a language L, noted $\mathsf{dist}(w,L)$ is the minimum distance $\mathsf{dist}(w,w')$ for $w' \in L$.

Definition 1 (Property Tester [22], [12])**.** *Let $\varepsilon > 0$. An ε-tester for a language L is a randomized algorithm $\mathscr{A}$ such that, for all words w as input:*
(1) If $w \in L$, then $\mathscr{A}$ accepts with probability at least $2/3$,
(2) If w is ε-far from L, then $\mathscr{A}$ rejects with probability at least $2/3$.

Applying a polynomial number of times the algorithm $\mathscr{A}$ on the same input w, and using a majority vote, the $2/3$ bound could be replaced by any real number $\rho > 1/2$. A *query* asks for the value of $w[i]$ for some i. The *query complexity* is the number of queries made to the word, and the *time complexity* is the usual definition, where we assume that arithmetic operations, a uniformly random choice of an integer from any finite range not larger than the input size, and a query to the input, take constant time. A language L is *testable*, if there exists a randomized algorithm $\mathscr{A}$ such that, for every real $\varepsilon > 0$ as input, $\mathscr{A}(\varepsilon)$ is an ε-tester of L, and the query and time complexities of $\mathscr{A}$ depend only on ε.

The $\mathsf{ustat}_k(w_n)$ vector of w_n, of dimension $|\Sigma|^k$, also called the k-gram of w_n is a vector whose u component, $\mathsf{ustat}_k(w_n)[u]$ for u a word of size k, is the number of different occurrences of u in w_n divided by $n-k+1$, the number of blocks of size k in w_n. It is also the probability to find u in a uniform random block. For instance, for $k=2$ and $\Sigma = \{0,1\}$ there are 4 possible words u of length k, which we take in lexicographic order. For $w_6 = 101101 \in \Sigma^*$ we get $ustat_2(w_6) = (0, 2/5, 2/5, 1/5)$. We will use the result of [11], which relates dist to the L_1 distance between ustat vectors:

Proposition 1. *For large enough words $w,w' \in \Sigma^*$, $\forall \delta > 0$, for large enough k:*

- *if $\mathsf{dist}(w,w') \le \delta^2$, then $||ustat_k(w) - ustat_k(w')||_1 \le 7 \cdot \delta$*
- *if $||ustat_k(w) - ustat_k(w')||_1 \le \delta$, then $\mathsf{dist}(w,w') \le 7 \cdot \delta$*

B. Markov Decision Processes and Probabilistic Automata

All the MDPs and automata are on a finite alphabet Σ. If S is finite set, we write $\Delta(S)$ for the set of distributions on S.

Definition 2. *A Markov Decision Process (MDP) is a triple $\mathscr{S} = (S, \Sigma, P)$ where S is a finite set of states, Σ is a set of* actions*, and $P : S \times \Sigma \times S \to [0;1]$ is the transition relation. $P(s,a,t)$, also written $P(t|s,a)$, is the probability to arrive in t in one step when the current state is s and action $a \in \Sigma$ is chosen for the transition.*

If action a is not allowed from state s, $P(t|s,a) = 0$ for all $t \in S$. The initial state of the system is chosen randomly according to an initial probability distribution α on its state space. A *history*, or *run*, on $\mathscr{S}$ is a finite or infinite alternating sequence of states and actions, which begins with a state and ends with a state when finite. We write Ω^* for the set of finite histories, Ω for the set of infinite histories on $\mathscr{S}$. If $n \in \mathbb{N}$ and $r \in \Omega$ we write $r_{|n}$ for the sequence of the first $n-1$ state action couples in r and the n-th state in r. The *trace* $Tr(r)$ of a run r is the sequence of actions. If $n \in \mathbb{N}$, X_n and Y_n are the random variables on Ω which associate to a run r its n-th state and its n-th action. A *policy* on $\mathscr{S}$, see [28], [23], is a function $\sigma : \Omega^* \to \Delta(\Sigma)$. A policy resolves the non determinism of the system by choosing a distribution on the set of available actions from the last state of the given history. We write *HR* for the set of history dependent and randomized policies. A policy is *deterministic* when for all history $h = (s_1, a_1, \ldots, a_{i-1}, s_i)$ on $\mathscr{S}$, $\sigma(h) \in \Sigma$. We write *HD* for the set of history dependent deterministic policies.

If $k \in \mathbb{N}$, a policy σ is said to have *memory* k if for any history $h = (s_1, a_1, \ldots, a_{i-1}, s_i)$ of length at least k we have $\sigma((s_1, a_1, \ldots, a_{i-1}, s_i)) = \sigma((s_{i-k}, a_{i-k}, \ldots, a_{i-1}, s_i))$. We write $MR(k)$ for the set of randomized policies with memory less than k. A policy is *stationary*, or *memoryless*, if it has memory 0, i.e. for any history $h = (s_1, a_1, \ldots, a_{i-1}, s_i)$ we have $\sigma(h) = \sigma(s_i)$. We write *SR* for the set of stationary randomized policies, and we write *SD* for the set of stationary deterministic policies.

A policy σ and an initial distribution α induce a probability distribution $\mathbb{P}^{\sigma,\alpha}$ on the σ-field $\mathscr{F}$ of Ω generated by the cones $C_\rho = \{r \in \Omega \mid r_{|\rho|} = \rho\}$, for $\rho \in \Omega^*$,(see [5], [28]). If the initial distribution α is concentrated on a state, that is if there exists $s \in S$ such that $\alpha(s) = 1$, we may write $\mathbb{P}^{\sigma,s}$ instead of $\mathbb{P}^{\sigma,\alpha}$.

Let $\mathscr{S}$ be an MDP and s be a state of $\mathscr{S}$. The set $Leave(s) \subseteq \Omega$ is the set of runs which do not cross s after a finite number of steps. That is,

$$Leave(s) = \{r \in \Omega \mid \exists k \in \mathbb{N} \ s.t. \ \forall l \ge k \ \ X_l(r) \ne s\}$$

Given a policy σ on $\mathscr{S}$, the state s is said to be *transient under* σ if $\mathbb{P}^{\sigma,s}[Leave(s)] = 1$. That is, s is transient for σ

if with probability one, after a finite number of steps, if the system is initiated on s, the runs do not cross s.

As for Markov chains, the communication properties between the states of an MDP is important. An MDP is *weakly communicating*, [20], if the set of states can be partitioned into a set of states that are accessible from each other (i.e., for any two states s and s' in that set, there exists a policy under which there is a positive probability to reach s' from s), and a set S_0 of states which are transient under *all* policies. An MDP is *communicating* if this decomposition can be done with an empty set S_0.

For a general MDP, there is a decomposition [6] into maximal disjoint end components (MECs) $S_1,...,S_l$ of the state space such that $S = S_0 \cup S_1 \cup ... \cup S_l$, where S_0 is the set of states which are transient for any policy on $\mathscr{S}$. A MEC is a maximal closed subset S_i of states such that the underlying graph is strongly connected, and such that once entered, there exists a policy σ which keeps the associated run in S_i forever. If $T \subseteq S$, $Reach(T)$ is the event: $\{r \in \Omega | \exists k \in \mathbb{N}\ s.t.\ X_k(r) \in T\}$, which is a measurable event [5], [28]. The maximal reachability problem [5], [7] asks, given a set $T \subset S$ of *destination states*, and an initial distribution α on S, for:

$$MaxReach_{\mathscr{S}}(\alpha, T) = Sup_{\sigma \in HR(\mathscr{S})} \mathbb{P}^{\alpha,\sigma}(Reach(T))$$

In [7], a polynomial time algorithm for $MaxReach_{\mathscr{S}}(\alpha, T)$ is given, and the optimal policy σ is deterministic, and computable in polynomial time as well.

A *probabilistic automaton (PA)*, [21], is an MDP $\mathscr{A}$ with a marked initial state s_i and an extra set of final states $F \subseteq S$, usually given with a probability threshold $\lambda \in [0;1]$ for the acceptance condition. Given an input word $w_n \in \Sigma^n$, $\mathbb{P}_{\mathscr{A}}(w_n)$ is the probability to reach a state in F after reading w_n, when the system is initiated on s_i.

1) State-Action Frequencies: Statistics on runs for an MDP, have been introduced in [17], [8], [20], [24] as random variables for the empirical state-action frequency vectors. We consider a run on an MDP with state space S as a sequence of couples in S and Σ. The statistics of a run will be the statistics taken on this alphabet.

Definition 3 (Expected state action frequency vector)**.** *Let σ be a policy on $\mathscr{S}$, $k \in \mathbb{N}$ and $T \geq 0$. $\hat{x}_k^T$ is the random variable on the set of histories Ω, which associates to all $r \in \Omega$ the k-gram of its prefix of length T. That is, $\hat{x}_k^T = ustat_k(r_{|T})$. Given an initial distribution α, the Expected state action frequency vector $x_{\sigma,\alpha,k}^T$ is $\mathbb{E}_{\sigma,\alpha}[\hat{x}_k^T]$, i.e. the expectation of $\hat{x}_k^T$.*

We may forget the k in the notations when $k = 1$. $x_{\sigma,\alpha,1}^T$ is a vector in $[0;1]^{S\times\Sigma}$ whose components are non-negative and sum to one, and $x_{\sigma,\alpha,1}^T[(s,a)]$ is the expected frequency, up to time T, of taking state-action (s,a), given the initial distribution α and the non determinism resolved by σ. If σ is a policy on $\mathscr{S}$, then $x_{\sigma,\alpha,k}^\infty$ is the empty set if $x_{\sigma,\alpha,k}^T$ does not converge as $T \to +\infty$, and the limit point if $x_{\sigma,\alpha,k}^T$ converges. If K is a class of policies ($K = SD, HR...$), we define:

$$H_k^K(\alpha) = \bigcup_{\sigma \in K} x_{\sigma,\alpha,k}^\infty$$

2) The polytope $\mathscr{H}$ for communicating MDPs: For $k = 1$, let $\mathscr{H}(\mathscr{S})$ be the set of vectors $x \in \Delta(S \times \Sigma)$ which satisfy for all $s' \in S$ the linear equations:

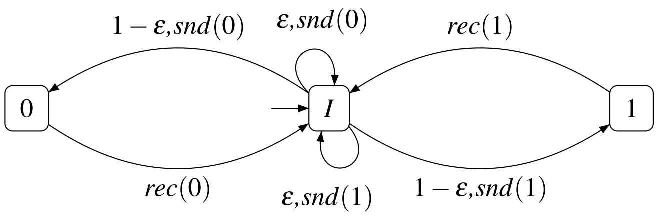
$$\sum_{s\in S}\sum_{a\in\Sigma} P(s'|s,a)\cdot x(s,a) = \sum_{a'\in\Sigma} x(s',a') \qquad (1)$$

If $H \subseteq \Delta(S \times \Sigma)$, let $\overline{H}$ be the convex closure of H. The following proposition is an improvement by [17], [20], [13] of a first result of [8].

Proposition 2. *Let $\mathscr{S}$ be a weakly communicating MDP. Then for all distribution α on $\mathscr{S}$,*

$$\mathscr{H}(\mathscr{S}) = H^{HR}(\alpha) = \overline{(H^{SD}(\alpha))},$$

Example 1 (A lossy communication channel)**.**



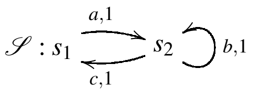
This lossy communication channel is communicating. We will be interested in the actions appearing in the runs, in $\Sigma = \{snd(1), snd(0), rec(1), rec(0)\}$. In order to get vectors of reasonable size, we take the statistics of the traces. Non determinism is only present on state I, from which the system chooses betwen $snd(0)$ and $snd(1)$. It induces two possible stationary and deterministic policies. One chooses $snd(0)$ and leads to the limit statistic vector of order one on Σ: $y_1 = (snd(0) : \frac{1}{2-\varepsilon}, rec(0) : \frac{1-\varepsilon}{2-\varepsilon}, snd(1) : 0, rec(1) : 0)$. The other chooses $snd(1)$ in I, and gives the symmetric point $y_2 = (snd(0) : 0, rec(0) : 0, snd(1) : \frac{1}{2-\varepsilon}, rec(1) : \frac{1-\varepsilon}{2-\varepsilon})$. As the system is weakly communicating, the projection of $\mathscr{H}(\mathscr{S})$ in $\mathbb{R}^{\Sigma^k}$ is the segment between y_1 and y_2.

III. State-action frequencies and MDPs of higher order.

A. Higher order statistics

We first generalize the results of [17], [8], [20] to higher order statistics. In this section k is a natural number greater than 0. We fix an MDP $\mathscr{S}$ and α an initial distribution on $\mathscr{S}$. The analogous of Derman's theorem ([8], chapter 7), is not true any more when we consider statistics of higher orders: if $k \geq 2$, in general $H_k^{HR}(\alpha)$ is not the convex hull of $H_k^{SR}(\alpha)$. Still, we will see that in that case we have $H_k^{HR}(\alpha) = \overline{(H_k^{MR(k)}(\alpha))}$.

Example 2. *Consider the following MDP:*



s_1 is the initial state. On $\mathscr{S}$, consider the policy $\sigma \in HR$ such that the choice on state s_2 depends on the history: if the system was in s_2, then σ chooses action c with probability one. If the system was in s_1 and just arrives in s_2, σ chooses action b with probability one. Then, $x^T_{\sigma,\alpha,2}$ converges to a point $x \in \Delta((S\times\Sigma)^2)$, such that $x[s_1as_2b] = x[s_2bs_2c] = x[s_2cs_1a] = 1/3$, and all the other coordinates are zero. If $x \in H_2^{SR}(\alpha)$, $x[s_1as_2b] > 0$ and $x[s_2cs_1a] > 0$ implies $x[s_1as_2c] > 0$. This proves that $x \notin H_2^{SR}(\alpha)$, and in fact we have $d_{L_1}(x, H_2^{SR}(\alpha)) \geq 1/6$, which proves that we do not have $H_2^{HR}(\alpha) = \overline{H_2^{SR}(\alpha)}$. We will prove that $H_2^{HR}(\alpha) = \overline{H_2^{MR(2)}(\alpha)}$.

We describe now the construction of the k-th self product of an MDP. Our goal is the following: the state-action frequency vectors on the k-th self product $\mathscr{S}^k$ of $\mathscr{S}$ should correspond to the order k state-action frequency vectors on $\mathscr{S}$.

Definition 4. *The k self product of an MDP $\mathscr{S} = (S,\Sigma,P)$ is the MDP $\mathscr{S}^k = (S',\Sigma,P')$ where:*
$S' = (\prod_{i=1}^{k-1} S\times\Sigma)\times S$,
If $t' = (t_1,b_1\ldots,b_{k-1},t_k)$ and $s' = (s_1,a_1\ldots,a_{k-1},s_k)$ are in S' and $a \in \Sigma$,

$$P'(t'|s',a) = \begin{cases} P(t_k|s_k,a) \text{ if } (t_1,b_1\ldots,b_{k-2},t_{k-1}) = \\ \qquad (s_2,a_2\ldots,a_{k-1},s_k) \text{ and } a = b_{k-1} \\ 0 \text{ otherwise} \end{cases}$$

If $\mathscr{S}$ is a communicating MDP, then $\mathscr{S}^k$ is communicating as well. A run on $\mathscr{S}^k$ is a sequence of couples in S' and Σ, which can be seen as a sequence of couples in S and Σ, i.e. a run on $\mathscr{S}$. Also, given a policy σ on $\mathscr{S}$, there is a naturally associated policy σ' on $\mathscr{S}^k$, which takes the same actions given the same histories. However, several policies on $\mathscr{S}$ may be associated to the same policy σ' on $\mathscr{S}^k$. Indeed when we choose a policy on $\mathscr{S}^k$ we lose the first k steps of σ. We can use the notion of state-action statistic vector for the MDP $\mathscr{S}^k$. If σ' is a policy on $\mathscr{S}^k$, $T \geq 1$, and α' is an initial distribution on $\mathscr{S}^k$, then $x^T_{\sigma',\alpha',1}(\mathscr{S}^k) \in \mathbb{R}^{(S'\times\Sigma)} = \mathbb{R}^{(S\times\Sigma)^k}$.

An initial distribution α and a policy σ on $\mathscr{S}$ induce a unique initial distribution $\alpha'(\sigma,\alpha)$ on $\mathscr{S}^k$, such that for all state $(s_1,a_1\ldots,s_k)$ in $\mathscr{S}^k$,

$$\alpha'(\sigma,\alpha)((s_1,a_1\ldots,s_k)) = \mathbb{P}^{\sigma,\alpha}(C_{s_1,a_1\ldots,s_k}).$$

Lemma 1. *Let $\mathscr{S}$ be an MDP, α an initial distribution, $k \geq 1$, and σ a policy on $\mathscr{S}$. We write σ' for the policy on $\mathscr{S}^k$ associated to σ, and $\alpha' = \alpha'(\sigma,\alpha)$ for the initial distribution on $\mathscr{S}^k$ associated to σ and α. Then for all $T \geq 1$, we have:*

$$x^{T+k-1}_{\sigma,\alpha,k}(\mathscr{S}) = x^T_{\sigma',\alpha',1}(\mathscr{S}^k),$$

Both are vectors in $\mathbb{R}^{(S\times\Sigma)^k}$. The following theorem and corollary generalize proposition 2 to the context of higher order statistics.

Theorem 1. *Let $\mathscr{S}$ be a weakly communicating MDP and $k \geq 1$. Then $H_1^{HR}(\mathscr{S}^k) = H_k^{HR}(\mathscr{S})$, and for all initial distribution α on $\mathscr{S}$,*

$$\bigcup_{\sigma\in MR(k)(\mathscr{S})} H^{SR}(\mathscr{S}^k)(\alpha'(\sigma,\alpha)) = H_k^{MR(k)}(\mathscr{S})(\alpha)$$

Corollary 1. *If $\mathscr{S}$ is weakly communicating,*

$$\mathscr{H}(\mathscr{S}^k) = H_1^{HR}(\mathscr{S}^k) = H_k^{HR}(\mathscr{S}) = \overline{[H_1^{SR}(\mathscr{S}^k)]} = \overline{[H_k^{MR(k)}(\mathscr{S})]}.$$

So far the polytope that we have associated to an MDP lies in a vector space whose dimension depends on the state space of the considered system. We eliminate this dependence, in order to be able to compare systems with very different state spaces. For this we introduce the linear projection $\pi: \mathbb{R}^{(S\times\Sigma)^k} \to \mathbb{R}^{\Sigma^k}$ such that if $x \in \mathbb{R}^{(S\times\Sigma)^k}$, on a component $v \in \Sigma^k$ we have

$$\pi(x)[v] = \sum_{u\in(S\times\Sigma)^k \text{ s.t. } Tr(u)=v} x[u] \qquad (2)$$

In the future, if $i,k \in \mathbb{N}$, we write $H_k^i(\mathscr{S})$ for $H_k^{MR(i)}(\mathscr{S})$, and $\Pi_k^i(\mathscr{S})$ for $\pi(H_k^i(\mathscr{S}))$.

B. Distances

In this paragraph we compute the distance between a statistic vector and a polytope $\Pi_k^k(\mathscr{S})$, and define a distance between MDPs.

Definition 5 (The distance $d_k(x,\mathscr{S})$)**.** *If $\mathscr{S}$ is a weakly communicating MDP, $k \in \mathbb{N}$, and $x \in \mathbb{R}^{\Sigma^k}$, let*

$$d_k(x,\mathscr{S}) = Inf_{y\in\Pi_k^k(\mathscr{S})} ||x-y||_1$$

The distance $d_k(x,\mathscr{S})$ can be computed in time polynomial in $(|S|\cdot|\Sigma|)^k$. We get non exponential bounds because the polytope $H_k^k(\mathscr{S})$ is characterized by a number of linear equations polynomial in $|\mathscr{S}|$ (see equation 1). Thus, since we are considering the L_1 norm, we can use a linear program of size polynomial in the size of $\mathscr{S}$ to compute $d_k(x,\mathscr{S})$. Notice that the polytope $H_k^k(\mathscr{S})$ may have an exponential number of extremal points (which correspond to the exponential number of possible stationary deterministic policies).

Definition 6 (d_k between weakly-communicating MDPs)**.** *If $\mathscr{S}_1$ and $\mathscr{S}_2$ are two weakly communicating MDPs let $d_k(\mathscr{S}_1,\mathscr{S}_2)$ be the Hausdorff distance (with respect to the norm L_1) between their polytopes for k statistics:*

$$d_k(\mathscr{S}_1,\mathscr{S}_2) = \frac{Sup_{x\in\Pi_k^k(\mathscr{S}_1)} Inf_{y\in\Pi_k^k(\mathscr{S}_2)} ||x-y||_1}{2} + \frac{Sup_{x\in\Pi_k^k(\mathscr{S}_2)} Inf_{y\in\Pi_k^k(\mathscr{S}_1)} ||x-y||_1}{2}.$$

For instance, with two lossy channels $\mathscr{S}_1, \mathscr{S}_2$ with respective parameters ε_1 and ε_2, it is not difficult to see that the distance between the order one polytopes is $|\varepsilon_1 - \varepsilon_2|/2$. It is $|\varepsilon_1 - \varepsilon_2|$ for the DGJP-metric of [9], for instance, and we will see that in general our distance is far from the DGJP-metric. Unfortunately, this distance is difficult to compute, and hard to approximate to any ratio smaller than the dimension $|\Sigma|^k$. In fact, [14] proves that even the L_1-diameter of a polytope is not computable in PTIME, and that it is not well approximable. We use the fact that the Hausdorff distance with the norm L_∞ is computable using a linear program of polynomial size, to approximate the L_1 Hausdorff distance within a factor $|\Sigma|^k$.

Proposition 3. *Suppose $\mathscr{S}_1$ and $\mathscr{S}_2$ weakly communicating. Then we can compute the distance $d_k(\mathscr{S}_1, \mathscr{S}_2)$ within a factor $|\Sigma|^k$ in PTIME$(((|S_1| + |S_2|) \cdot |\Sigma|)^k)$.*

In general, for a non weakly communicating MDP, the set of the limit statistics is a union of polytopes. Given two general MDPs $\mathscr{S}_1, \alpha_1$ and $\mathscr{S}_2, \alpha_2$, we partition their state spaces into MECs, as in [6]. Write $S_j = S_j^0 \cup S_j^1 \cup ... \cup S_j^{l_j}$, for $j \in \{1,2\}$. For all $j \in \{1,2\}$ and $i \in [1; l_j]$, we write $\mathscr{S}_j^i$ for the MDP associated to the MEC. If $x \in \mathbb{R}^{\Sigma^k}$, $\delta \geq 0$ and $j \in \{1,2\}$, we write $V_j(x, \delta)$ for the union of the state spaces of the MECs whose associated polytope is δ-close to x. That is:

$$V_j(x, \delta) = \bigcup_i \{S_i^j | d_k(x, \mathscr{S}_j^i) \leq \delta\}$$

Now, the distance between the two systems may have two parameters: one for the distance of the respective points on the polytopes associated to the MECs, the other for the maximal probability to reach these MECs. If $\varepsilon, \delta \in [0;1]$, $\mathscr{S}_1, \alpha_1$ will be said to be (ε, δ)-simulated by $\mathscr{S}_2, \alpha_2$ for order k statistics, written $\mathscr{S}_1, \alpha_1 \prec^k_{(\varepsilon,\delta)} \mathscr{S}_2, \alpha_2$, if for all $x \in \mathbb{R}^{\Sigma^k}$,

$$MaxReach_1(\alpha_1, V_1(x, 0)) \leq MaxReach_2(\alpha_2, V_2(x, \varepsilon)) + \delta$$

That is, given a statistic vector x, the maximal probability to reach on $\mathscr{S}_2$ a MEC whose polytope is ε-close to x is at least equal to the maximal probability to reach on $\mathscr{S}_1$ a MEC whose polytope contains x, minus δ.

Taking $\delta = \varepsilon$, this notion induces a quasi metric $d_k^{\prec}$, and a pseudo metric d_k, which extend d_k on the set of general MDPs.

Definition 7 (Pseudometrics d_k and $d_k^{\prec}$)**.** *Given $\mathscr{S}_1, \alpha_1$ and $\mathscr{S}_2, \alpha_2$ two MDPs, let*

$$d_k^{\prec}(\mathscr{S}_1, \mathscr{S}_2) = \text{Inf}\{\varepsilon > 0 | \mathscr{S}_1, \alpha_1 \prec_{(\varepsilon,\varepsilon)} \mathscr{S}_2, \alpha_2\}$$

$$d_k(\mathscr{S}_1, \mathscr{S}_2) = Min(d_k^{\prec}(\mathscr{S}_1, \mathscr{S}_2), d_k^{\prec}(\mathscr{S}_2, \mathscr{S}_1))$$

Notice that d_k is symmetric, whereas $d_k^{\prec}$ is not in general. In the case of weakly communicating MDPs, the two definitions for $d_k(\mathscr{S}_1, \mathscr{S}_2)$ coincide. Both $d_k^{\prec}(\mathscr{S}_1, \mathscr{S}_2)$ and $d_k(\mathscr{S}_1, \mathscr{S}_2)$ can be approximated in time polynomial in $((|S_1| \cdot |\Sigma|^k + |S_2| \cdot |\Sigma|)^k)$, within a factor $|\Sigma|^k$.

Our distances between MDPs cannot be compared to the metrics defined in [27], [9]. These metrics are in fact metrics between states of a given system, which may induce metrics between systems, by taking their initial states. In our approach, we do not rely on states, and the distance between systems does not depend on the initial distributions if the systems are weakly communicating.

IV. Problems on MDPs

We consider two classes of problems on MDPs. First ”membership” type problems, where an MDP is fixed and the input consists in a word or a statistic vector, and second “comparison” type problems, where the input consists in two MDPs that we want to compare. For the following, if $\varepsilon > 0$, $n \in \mathbb{N}$, $k \in \mathbb{N}$ and $x \in \mathbb{R}^{\Sigma^k}$, let

$$B_n(x, \varepsilon) = \{r \in \Omega | \|ustat_k(r_{|n}) - x\| \leq \varepsilon\}$$

As well, if $w \in \Sigma^n$, let $B_n(w, \varepsilon) = \{r \in \Omega | \|(ustat_k(r_{|n}) - ustat_k(w)\| \leq \varepsilon\}$.

A. Membership problems

Most of the standard problems in the context of MDPs optimization, (see [4], [19]), can be presented as follows. We are given an MDP $\mathscr{S}$, a length $n \in \mathbb{N}$, a probability threshold $\lambda \in [0;1]$, and an objective x. The question is to find a policy σ on $\mathscr{S}$ such that the probability of the set of runs of length n which satisfy the objective is greater than λ. In our context, a natural objective is, given a word $w \in \Sigma^*$, to decide if there exists a policy such that with high probability the traces are close to w. Since by proposition 1 two words are close iff their k-grams are close, we can formulate our problem as follows.

For this subsection, we fix an MDP $\mathscr{S}$ with initial distribution α, a threshold $\lambda \in [0;1]$, and a radius parameter $\delta \in [0;1]$.

Problem 1 (Statistical (λ, δ)-membership)**.**
Input: $w \in \Sigma^n$.
Question: *$\exists \sigma \in HR(\mathscr{S})$ such that $\mathbb{P}^\sigma(B_n(w, \delta)) \geq \lambda$?*

Consider the associated language:

$$\mathscr{L}_\delta^\lambda = \{w \in \Sigma^* | \exists \sigma \in HR(\mathscr{S}) \ s.t. \ \mathbb{P}^\sigma(B_n(w, \delta)) \geq \lambda\}$$

Proposition 4. *Given an MDP $\mathscr{S}$ of size n and a word w of length not greater than n, it is PSPACE-hard to decide if w satisfies the requirements of the (λ, δ)-membership problem on $\mathscr{S}$.*

If $\varepsilon \in [0;1]$, we say that a word $w \in \Sigma^n$ is ε-close to $\mathscr{L}_\delta^\lambda$ if there exists $w' \in \Sigma^n$ such that $d(w, w') \leq \varepsilon$, and $w' \in \mathscr{L}_{\delta \cdot (1+\varepsilon)}^{\lambda \cdot (1-\varepsilon)}$. We prove that $\mathscr{L}_\delta^\lambda$ is constant time testable: For all $\varepsilon \in]0;1]$, there exists a tester $\mathscr{T}_\varepsilon$ such that for long enough input $w \in \Sigma^*$:

- If $w \in \mathscr{L}_\delta^\lambda$, then $\mathscr{T}_\varepsilon$ answers YES with probability at least $2/3$.
- If w is ε-far from $\mathscr{L}_\delta^\lambda$, $\mathscr{T}_\varepsilon$ answers NO with probability at least $2/3$.

Fix $\varepsilon > 0$ and $k = \ulcorner 1/\varepsilon \urcorner$. The construction of $\mathscr{T}_\varepsilon$ is as follows: Let $S = S_0 \cup S_1 \cup \ldots \cup S_l$ be the decomposition of $\mathscr{S}$ into MECs. Let $\mathscr{S}_i$, $i \in [1;l]$, be the communicating MDPs associated to the MECs S_i, $i \in [1;l]$. For all $i \in [1;l]$, compute the set of linear equations which give $\Pi_k^k(\mathscr{S}_i)$. As before, if $x \in \mathbb{R}^{\Sigma^k}$, $V(x,\delta) = \bigcup\{S_i | d_k(x,\mathscr{S}_i) \leq \delta\}$.

Algorithm 1 (The tester $\mathscr{T}_\varepsilon(w)$)**.** ***Input:*** $w \in \Sigma^*$.

- *Sample w to obtain $\hat{x}$, an approximation of $ustat_k(w)$, in constant time.*
- *Compute $p = MaxReach(\alpha, V(\hat{x}, \delta \cdot (1+\varepsilon)))$.*
- *If $p \geq \lambda \cdot (1-\varepsilon/2)$, then $\mathscr{T}_\varepsilon$ answers YES, if not, $\mathscr{T}_\varepsilon$ answers NO.*

The following lemma, proved in [11], and based on a Chernoff bound, proves that we can sample $\hat{x}$ efficiently.

Lemma 2. *There exists a probabilistic algorithm which works in constant time on inputs $w \in \Sigma^*$ and which produces a vector $\hat{x} \in \mathbb{R}^{\Sigma^k}$ such that $\mathbb{P}(\|ustat_k(w) - \hat{x}\| < \delta \cdot \varepsilon/14) \geq 2/3$.*

To prove the correctness of the tester, we present a generalization of the results in [25] and [17] to the context of general MDPs. By theorem 2 of [25], and theorem 5.1. of [17], we know that for all $i \in [1;l]$, we can find two constants C_1^i, C_2^i, such that for all $x \in \mathbb{R}^{\Sigma^k}$, $n \in \mathbb{N}$, $\alpha_i \in \Delta(S_i)$, and $\varepsilon > 0$ we have:

- If $d_k(x,\mathscr{S}_i) \leq \delta \cdot (1+\varepsilon/2)$, then $\exists \sigma_i \in HR(\mathscr{S}_i)$ s.t. $\mathbb{P}^{\sigma_i,\alpha_i}(B_n(x,\delta \cdot (1+\varepsilon))) \geq 1 - C_1^i \cdot e^{-C_2^i \cdot n \cdot \varepsilon^2}$
- If $d_k(x,\mathscr{S}_i) > \delta \cdot (1+\varepsilon)$, then $\forall \sigma_i \in HR(\mathscr{S}_i)$, $\mathbb{P}^{\sigma_i,\alpha_i}(B_n(x,\delta \cdot (1+\varepsilon/2))) \leq C_1^i \cdot e^{-C_2^i \cdot n \cdot \varepsilon^2}$

Moreover, we can find two constants C_1^0, C_2^0, such that for any policy σ on $\mathscr{S}$, all $\alpha \in \Delta(S)$ and all $n \in \mathbb{N}$, $\mathbb{P}^{\sigma,\alpha}(X_n \in S_0) \leq C_1^0 \cdot e^{-C_1^0 \cdot n}$.

In the following, $q, N \in \mathbb{N}$ are such that

$$C_1^0 \cdot e^{-C_1^0 \cdot q} < \frac{\lambda \cdot \varepsilon}{4};\ \sum_{i=1}^{l} C_1^i \cdot e^{-C_2^i \cdot (N-q)\varepsilon^2 1} < \frac{\lambda \cdot \varepsilon}{4}$$

Lemma 3. *Let $x \in \mathbb{R}^{\Sigma^k}, n \geq N$.*

1) *If $MaxReach(\alpha, V(x,\delta \cdot (1+\varepsilon))) \leq \lambda \cdot (1-\varepsilon/2)$, then $\forall \sigma \in HR(\mathscr{S})$, $\mathbb{P}^{\sigma,\alpha}(B_n(x,\delta \cdot (1+\varepsilon/2)) \leq \lambda$.*
2) *If $MaxReach(\alpha, V(x,\delta \cdot (1+\varepsilon))) \geq \lambda \cdot (1-\varepsilon/2)$, then $\exists \sigma \in HR(\mathscr{S})$ s.t. $\mathbb{P}^{\sigma,\alpha}(B_n(x,\delta \cdot (1+\varepsilon)) \geq \lambda \cdot (1-\varepsilon)$.*

Theorem 2 (Correctness of $\mathscr{T}_\varepsilon$)**.** *For all $\varepsilon \in]0;1[$, $\mathscr{T}_\varepsilon$ is an ε-tester for $\mathscr{L}_\delta^\lambda$.*

Proof:

- Suppose $w \in \mathscr{L}_\delta^\lambda$, $|w| = n$. Let $x = ustat_k(w)$. Then, $\exists \sigma \in HR(\mathscr{S})$ such that $\mathbb{P}^{\sigma,\alpha}(B_n(w,\delta)) \geq \lambda$. By lemma 2, with probability at least $2/3$, $\|x - \hat{x}\| \leq \delta \cdot \varepsilon/2$, so with probability at least $2/3$, $\exists \sigma \in HR(\mathscr{S})$ such that $\mathbb{P}^{\sigma,\alpha}(B_n(\hat{x},\delta \cdot (1+\varepsilon/2))) > \lambda$. Using point 1 of lemma 3, with probability at least $2/3$,
$MaxReach(\alpha, V(\hat{x},\delta \cdot (1+\varepsilon))) \geq \lambda \cdot (1-\varepsilon/2)$.
I.e. $\mathscr{T}_\varepsilon$ answers YES with probability at least $2/3$
- Conversely, suppose $d(w, \mathscr{L}_{\delta \cdot (1+\varepsilon)}^{\lambda \cdot (1-\varepsilon)}) > \varepsilon$. Let $w' \in \Sigma^n$, and $y = ustat_k(w')$. By contraposition, suppose $MaxReach(\alpha, V(y,\delta \cdot (1+\varepsilon))) \geq \lambda \cdot (1-\varepsilon/2)$. Then, by the point 2 of lemma 3, $\exists \sigma \in HR(\mathscr{S})$ such that $\mathbb{P}^{\sigma,\alpha}(B_n(y,\delta \cdot (1+\varepsilon))) \geq \lambda \cdot (1-\varepsilon)$. This implies $w' \in \mathscr{L}_{\delta \cdot (1+\varepsilon)}^{\lambda \cdot (1-\varepsilon)}$, and by hypothesis, $\mathsf{dist}(w,w') \geq \varepsilon$. Using proposition 1, this implies $\|x - y\| \geq \varepsilon/7$. Since with probability at least $2/3$, $\|x - \hat{x}\| \leq \varepsilon/14$, with probability at least $2/3$ we have $MaxReach(\alpha, V(\hat{x},\delta \cdot (1+\varepsilon))) < \lambda \cdot (1-\varepsilon/2)$. Hence, $\mathscr{T}_\varepsilon$ answers NO with probability at least $2/3$.

∎

Using proposition 3 and the results of [7], $\mathscr{T}_\varepsilon$ works in time polynomial in $|S|^2 \cdot (|\Sigma| \cdot |S|)^k$, which is independent of $|w|$.

B. Comparison problems

Two MDPs should be close if they can generate close words with close probabilities. An MDP $\mathscr{S}_1$ should be approximately simulated by an MDP $\mathscr{S}_2$ if $\mathscr{S}_2$ can generate words close to the words generated by $\mathscr{S}_1$, with higher probability.

Problem 2 (MDP ε-simulation)**.**
Input: *$\mathscr{S}_1, \mathscr{S}_2$ with initial distributions α_1 and α_2.*
For $w \in \Sigma^n$, let

$$\lambda_1(w) = Sup_{\sigma_1 \in HR(\mathscr{S}_1)} \mathbb{P}^{\sigma_1}(B_n(w,\varepsilon)),$$
$$\lambda_2(w) = Sup_{\sigma_2 \in HR(\mathscr{S}_2)} \mathbb{P}^{\sigma_2}(B_n(w,\varepsilon)).$$

Question: *$\forall w \in \Sigma^*$ large enough, $\lambda_1(w) \leq \lambda_2(w)$?*

The distance $d_k^{\prec}$ can be seen as a distance of simulation. Given two MDPs $\mathscr{S}_1$ and $\mathscr{S}_2$ with respective initial distributions α_1 and α_2, let N_1, N_2 be as for lemma 3, for $\mathscr{S}_1$ and $\mathscr{S}_2$ respectively. If $i \in \{1,2\}$, $w \in \Sigma^n$ and $\varepsilon > 0$, let $\lambda_i(w,\varepsilon) = Sup_{\sigma_i \in HR(\mathscr{S}_i)} \mathbb{P}_i^{\sigma_i,\alpha_i}(B_n(w,\varepsilon))$. The following proposition is the analogous of the tester 1 for comparing MDPs. It gives an approximate solution for problem 2.

Proposition 5. *Let $n \geq Max(N_1, N_2)$.*

- *If $d_k^{\prec}(\mathscr{S}_1, \mathscr{S}_2) \leq \varepsilon/2$, then for all $w \in \Sigma^*$, $\lambda_1(w,0) \leq \lambda_2(w, 2 \cdot \varepsilon) + 2 \cdot \varepsilon$.*
- *Conversely, suppose $d_k^{\prec}(\mathscr{S}_1, \mathscr{S}_2) > 3 \cdot \varepsilon/2$. Then we can find $w \in \Sigma^*$ such that $\lambda_1(w,\varepsilon/2) > \lambda_2(w,\varepsilon) + \varepsilon/4$*

Using the approximation algorithms to compute $d_k^{\prec}(\mathscr{S}_1, \mathscr{S}_2)$, we get a procedure to compute approximately the simulation relation between MDPs. A bisimulation relation can be defined in the same way by taking the symmetrization of the simulation relation, and can be computed approximately by using the approximation algorithms for $d_k(\mathscr{S}_1, \mathscr{S}_2)$.

V. Problems related to Probabilistic Automata

In this section we study some problems on Probabilistic Automata, analogous to the problems studied in the previous section concerning MDPs. We will see that these are much more difficult to solve, even approximately. We will deal with three problems concerning PAs: the language emptiness problem, the membership problem of a given word to the language of a PA, and the comparison problem between two PAs.

If $\mathscr{A}$ is a probabilistic automaton and $w \in \Sigma^*$, $\mathbb{P}_{\mathscr{A}}(w)$ is the probability to arrive in an accepting state when w is read on $\mathscr{A}$.

Recall first the undecidability of the *emptiness problem for a PA*. Given a PA $\mathscr{A}$ for which one of the two cases hold, it is undecidable to decide if there exists $w \in \Sigma^*$ such that $\mathbb{P}_{\mathscr{A}}(w) > 1 - \varepsilon$, or for all $w \in \Sigma^*$, $\mathbb{P}_{\mathscr{A}}(w) < \varepsilon$ (Corollary 3.4 of [16]). If the length of the word is fixed, the problem is NP-complete:

Problem 3 (n-Emptyness)**.**
Input*:* $\mathscr{A}$*,* $\varepsilon \in [0;1]$*,* $n \in \mathbb{N}$*.*
Question*: Decide if there is a word* $w \in \Sigma^n$ *such that* $\mathbb{P}_{\mathscr{A}}(w) \geq 1 - \varepsilon$*, or if for all word* $w \in \Sigma^n$*,* $\mathbb{P}_{\mathscr{A}}(w) \leq \varepsilon$*.*

Proposition 6. *Problem 3 is NP-complete.*

A. Membership problems

We consider now membership problems. An automaton $\mathscr{A}$, and a threshold $\lambda \in]0;1[$, are fixed.

Problem 4 (Membership problem)**.**
Input: $w \in \Sigma^*$*.*
Question: *Do we have* $\mathbb{P}_{\mathscr{A}}(w) \geq \lambda$*?*

The associated language is:

$$\mathscr{L}^\lambda = \{w \in \Sigma^* | \mathbb{P}_{\mathscr{A}}(w) \geq \lambda\}$$

Problem 4 is clearly computable in time $O(|\mathscr{A}| \cdot |w|)$. We prove that the language $\mathscr{L}^\lambda$ is not constant time testable, which contrasts with the results of [11] and our results in the context of MDPs. In [11], the authors consider the membership problem for non-deterministic automata. Using a geometric construction, they prove that this problem is testable in time independent of w.

Lemma 4. *Let* $\mathscr{T}$ *be a randomized* $O(1)$*-algorithm with inputs in* Σ^* *which works in time* $N \in \mathbb{N}$ *and samples subwords of length at most* $k \in \mathbb{N}$*. Let* $w, w' \in \Sigma^*$ *be such that* $\|ustat_k(w) - ustat_k(w')\| \leq \frac{\varepsilon}{2 \cdot N \cdot |\Sigma|^{k \cdot N}}$*. Then*

$$|\mathbb{P}_{\mathscr{T}}(YES|w) - \mathbb{P}_{\mathscr{T}}(YES|w')| + \\ |\mathbb{P}_{\mathscr{T}}(NO|w) - \mathbb{P}_{\mathscr{T}}(NO|w')| \leq \varepsilon.$$

Thus, a constant time tester gives close results on inputs with close statistics, *independently of the size of the input*. It implies that if a property can be tested in constant time, it should be testable by considering inputs of bounded size. In order to test the language $\mathscr{L}^\lambda$, as we did in the context of MDPs, we may allow a relaxation on the threshold: if $\varepsilon \geq 0$, a word $w \in \Sigma^n$ is ε-close to $\mathscr{L}^\lambda$ if there exists $w' \in \Sigma^n$ such that $d(w, w') \leq \varepsilon$, and $w' \in \mathscr{L}^{\lambda \cdot (1-\varepsilon)}$. We prove that there exists some PAs for which we cannot construct a constant time tester for the associated membership problem, even with the relaxation on the threshold.

Reading a word on a PA is the same as following a purely time dependent policy on the associated MDP. A general policy is far more flexible, and this explains the differences betweens the complexities of the problems considered on PAs or on MDPs.

Theorem 3. *The membership problem for PAs is not constant time testable.*

B. Comparison problems

In [26], the author proves that we can decide whether two PAs accept the same words with the same probabilities in PTIME. However, the following problem, which can be seen as a relaxation of the equivalence problem, is undecidable.

Problem 5 (Approximate equivalence)**.**
Input*: Two PAs* $\mathscr{A}_1$ *and* $\mathscr{A}_2$*,* $\varepsilon \in [0;1]$*.*
Question*: for all* $w \in \Sigma^*$*,* $|\mathbb{P}_1(w) - \mathbb{P}_2(w)| \leq \varepsilon$*?*

Even if we introduce a relaxation on the input, as in [11], the problem is still undecidable.

Problem 6 (Approximate simulation on close inputs)**.**
Input*: Two PAs* $\mathscr{A}_1$ *and* $\mathscr{A}_2$*,* $\varepsilon \in [0;1]$*.*
Question*: for all* $w \in \Sigma^*$*, there exists a word* $w' \in \Sigma^*$ *such that* $d(w, w') \leq \varepsilon$ *and* $|\mathbb{P}_1(w) - \mathbb{P}_2(w')| \leq \varepsilon$*?*

We reduce the undecidable emptiness problem for a PA to each of these problems. Let $\mathscr{A}_\emptyset$ be a probabilistic automaton with no accepting state, which accepts no word with probability greater than zero. A given automaton $\mathscr{A}$ will be ε-close to $\mathscr{A}_\emptyset$ (for problems 5) iff $\mathscr{A}$ does not accept any word with probability greater than ε. Considering problem 6, $\mathscr{A}$ is ε-simulated by $\mathscr{A}_\emptyset$ iff it does not accept any word with probability greater to ε. Since it is undecidable to decide if all the finite words are accepted by $\mathscr{A}$ with probability at most ε, this proves that problem 5 and 6 are undecidable.

VI. Ultimate properties

In this section, we present a class of properties on infinite words which satisfy two conditions:

- The properties are *ultimate*, that is if w is a trace and w' is obtained by changing only a finite number of letters in w, then w and w' should satisfy the same properties.

- They are not sensitive to time translation. That is, if w' is a suffix of w, then w and w' satisfy the same properties.

Two purely probabilistic processes will be *ultimately equivalent* if they satisfy the same ultimate properties with same probabilities. Two MDPs will be ultimately equivalent if for any policy on one of the MDPs there exists a policy on the second MDP such that the induced purely probabilistic processes are ultimately equivalent.

First we will prove that two probabilistic processes are ultimately equivalent iff they are *trace equivalent*, iff they have the same statistic (Theorem 5, 6). Second, we will prove that two weakly communicating MDPs are ultimately equivalent iff their statistic polytopes coincide, that is iff their distance is 0 (Theorem 7).

In order to compare different MDPs, we only consider the traces of their runs. We consider the probability space $(\Sigma^\omega, \mathscr{F}, \mathbb{P}^{\sigma,\alpha})$ where $\mathscr{F}$ is now the σ-field generated on Σ^ω by the cones $C_{a_1a_2...a_l} = \{w = b_1, b_2... \in \Sigma^\omega \mid (b_1,...,b_l) = (a_1,...,a_l)\}$, $a_1...a_l \in \Sigma^l$. We write $\mathbb{P}^{\sigma,\alpha}$ indifferently for the probability distribution on Ω, as in the previous section, and for the probability distribution induced by Tr on Σ^ω: If $\Gamma \in \mathscr{F}$, $\mathbb{P}^{\sigma,\alpha}(\Gamma)$ is $\mathbb{P}^{\sigma,\alpha}(Tr^{-1}(\Gamma))$. Recall that X_n is the random variable on Ω which associates to a run its state at time n, and Y_n associates to a run or a trace its action label at time n.

A *property* is a set $\Gamma \in \mathscr{F}$. Let $T : \Sigma^\omega \longrightarrow \Sigma^\omega$ be defined as $T(w_0w_1...) = (w_1w_2...)$. An *ultimate property* is a property Γ such that $T^{-1}(\Gamma) = \Gamma$. We write $\mathscr{G}$ for the class of ultimate properties. $\mathscr{G}$ is a σ-field, sometimes called the *invariant σ-field* ([1]). Clearly, $\mathscr{G}$ satisfies the two requirements formulated above.

For $n \geq 0$, let $\mathscr{F}_n = \mathscr{B}(Y_n, Y_{n+1}...)$ be the smallest σ-field on Σ^ω with respect to which all the $Y_i, i \geq n$ are measurable. Let $\mathscr{F}_\infty = \bigcap_{n\in\mathbb{N}} \mathscr{F}_n$, called the *tail σ-field* of $Y_n, n \geq 0$. Intuitively, an event Γ is in $\mathscr{F}_\infty$ iff changing a finite number of letters of an outcome $w \in \Sigma^\omega$ does not affect the occurrence of the event $r \in \Gamma$. Notice that $\mathscr{G}$ and $\mathscr{F}_\infty$, as σ-fields, are closed under union and intersection. The following result of [1] shows that ultimate properties form a particular class of properties of the tail σ-field.

Proposition 7. *T maps $\mathscr{F}_\infty$ one-to-one onto itself, and*

$$\mathscr{G} = \{\Gamma \in \mathscr{F}_\infty | T \cdot \Gamma = \Gamma\}$$

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, and let $\Gamma \in \mathscr{F}$. We say that Γ is a $\mathbb{P}$-*atomic* set of $\mathscr{F}$ if $\mathbb{P}(\Gamma) > 0$, and for all $\Gamma' \in \mathscr{F}$ such that $\Gamma' \subseteq \Gamma$ we have $\mathbb{P}(\Gamma') = 0$ or $\mathbb{P}(\Gamma') = \mathbb{P}(\Gamma)$. If $\mathscr{F}$ can be decomposed as a finite union of $\mathbb{P}$-atomic sets for $\mathscr{F}$ and a $\mathbb{P}$-negligible set, we shall say that $\mathscr{F}$ is *finite*. The following theorem is a generalization of the classical Kolmogorov's $0-1$ law to the context of Markov chains.

Theorem 4. *If σ has finite memory $i \in \mathbb{N}$, then the tail σ-field of $Y_n, n \geq 0$, is finite, and the number of $\mathbb{P}^{\sigma,\alpha}$-atomic sets of the invariant σ-field $\mathscr{G}$ does not exceed $|S|^i$.*

Theorem 4 is a generalization of results of [2]. Notice that if σ has infinite memory, $\mathscr{F}_\infty$ is not finite in general, and may contain no atomic subsets.

A. Markov chains with labelled transitions

In this subsection we consider ultimate properties associated to homogeneous Markov chains. A memoryless policy σ on an MDP $\mathscr{S} = (S, \Sigma, P)$ induces on the state space $S \times \Sigma$ a Labelled Transition Markov Chain $\mathscr{S}^\sigma$ (LTMC). This chain $(X_n, Y_n)_{n\in\mathbb{N}}$, with $(X_n, Y_n) \in S \times \Sigma$ for all $n \in \mathbb{N}$, has transition probabilities:

$$\mathbb{P}[(X_{n+1}, Y_{n+1}) = (x', y') | (X_n, Y_n) = (x, y)] = \\ P(x'|x, y) \cdot \sigma(x')(y').$$

Let $\mathscr{S} = (X_n, Y_n)_{n\in\mathbb{N}}$ be an irreducible LTMC, with initial distribution α. Let $k \in \mathbb{N}$. Then, since the chain is irreducible, by the law of large numbers, there exists a vector in $\mathbb{R}^{(S\times\Sigma)^k}$, which we call $ustat_k(\mathscr{S})$, such that with probability one the k-gram of the trace of a run on $\mathscr{S}$ converges to $ustat_k(\mathscr{S})$. Moreover, $ustat_k(\mathscr{S})$ is independent of the initial distribution α ([24], chapter 1). In the following, $\mathscr{S}_1, \alpha_1$ and $\mathscr{S}_2, \alpha_2$ are two LTMCs. We write $\mathbb{P}_1^{\alpha_1}$ and $\mathbb{P}_2^{\alpha_2}$ for the associated probability distributions on Σ^ω.

Definition 8 (Ultimate equivalence)**.** *$\mathscr{S}_1, \alpha_1$ and $\mathscr{S}_2, \alpha_2$ are said to be* ultimately equivalent*, written*
$\mathscr{S}_1, \alpha_1 \sim_u \mathscr{S}_2, \alpha_2$, if for all $\Gamma \in \mathscr{G}$, $\mathbb{P}_1^{\alpha_1}(\Gamma) = \mathbb{P}_2^{\alpha_2}(\Gamma)$

Proposition 8. *If $\mathscr{S}_1, \alpha_1 \sim_u \mathscr{S}_2, \alpha_2$, then for any $\Gamma \in \mathscr{G}$, Γ is $\mathbb{P}_1$-atomic iff Γ is $\mathbb{P}_2$-atomic.*

If $\mathscr{S}$ is an irreducible LTMC with initial distribution α, then for all $\Gamma \in \mathscr{G}$, $\mathbb{P}^\alpha(\Gamma) \in \{0, 1\}$, and $\mathbb{P}^\alpha(\Gamma) \in \{0, 1\}$ is independent of α.

As a consequence of the previous proposition, we can use the notation $\mathscr{S}_1 \sim_u \mathscr{S}_2$ to say that the LTMCs are ultimately equivalent.

Definition 9 (Trace equivalence for LTMCs)**.** *$\mathscr{S}_1$ and $\mathscr{S}_2$ are* trace equivalent *if there exists two initial distributions α_1 and α_2, on $\mathscr{S}_1$ and $\mathscr{S}_2$ respectively, such that:*

For all $w \in \Sigma^*$, $\mathbb{P}_1^{\alpha_1}(C_w) = \mathbb{P}_2^{\alpha_2}(C_w)$.

The following theorem shows that it is enough to know the k-grams of an irreducible chain for a bounded number of k's, to characterize completely the ultimate properties of the chain.

Theorem 5. *Let $\mathscr{S}_1$ and $\mathscr{S}_2$ be two irreducible LTMCs. Then the following are equivalent:*

1) $\forall k \in [1; (|S_1| + |S_2|)^2]$, $ustat_k(\mathscr{S}_1) = ustat_k(\mathscr{S}_2)$
2) *For all $k \in \mathbb{N}$, $ustat_k(\mathscr{S}_1) = ustat_k(\mathscr{S}_2)$*
3) $\mathscr{S}_1 \sim_u \mathscr{S}_2$
4) *$\mathscr{S}_1$ and $\mathscr{S}_2$ are trace equivalent.*

Given $\mathscr{S},\alpha$ a general LTMC on state space S, let $S=S_0\cup S_1\cup...\cup S_l$ be its decomposition into irreducible components: S_0 is the set of transient states, and the $S_i, i\in[1;l]$ are the irreducible components of the chain. Each S_i gives an irreducible LTMC $\mathscr{S}_i$. Given $i\in[1;l]$, let $Reach(S_i)$ be the set of infinite runs on $\mathscr{S}$ which enter S_i eventually (and then never leave it), and $p_i=\mathbb{P}^{\alpha}(Reach(S_i))$. Clearly, the p_i sum to one. Let $\lambda_{S_i}:\mathscr{G}\to\{0,1\}$ be such that for $\Gamma\in\mathscr{G}$, $\lambda_{S_i}(\Gamma)$ is the probability that a run executed on $\mathscr{S}_i$ is in Γ. The λ_{S_i} are well defined and take values in $\{0,1\}$, by the irreducibility of the $\mathscr{S}_i$.

Lemma 5. *Let $\mathscr{S},\alpha$ be an LTMC and $\Gamma\in\mathscr{G}$. Then $\mathbb{P}^{\alpha}(\Gamma)=\sum_{i=1}^{l}\mathbb{P}^{\alpha}(Reach(S_i))\cdot\lambda_{S_i}(\Gamma)$.*

Let $\mathscr{S}_1,\alpha_1$ and $\mathscr{S}_2,\alpha_2$ be two LTMCs on state spaces S_1 and S_2. Write $S_1=S_0^1\cup S_1^1\cup...\cup S_{l_1}^1$ and $S_2=S_0^2\cup S_1^2\cup...\cup S_{l_2}^2$ for the decompositions into irreducible components of the chains, and $Reach_j(S_i^j)$ for the set of infinite runs on $\mathscr{S}_j$ which enter S_i^j eventually (and then never leave it). Let $p_i^j=\mathbb{P}_j(Reach_j(S_i^j))$. The irreducible components S_i^j can be seen as irreducible LTMC $\mathscr{S}_i^j$, and $\sim_u$ is an equivalence relation on $\{\mathscr{S}_i^j, j\in\{0,1\}, i\in[1;l_j]\}$. Write $\{T_1,...,T_l\}$ for the equivalence classes of $\sim_u$. If $i\in[1;l]$, T_i is a union of S_j^1 and $S_{j'}^2$. The next theorem summarizes our results on LTMCs.

Theorem 6. *Let $\mathscr{S}_1,\alpha_1$ and $\mathscr{S}_2,\alpha_2$ be two LTMCs. Then the following are equivalent:*

1) $\mathscr{S}_1,\alpha_1\sim_u\mathscr{S}_2,\alpha_2$
2) $\forall i\in[1;l]\ \mathbb{P}_1^{\alpha_1}(Reach_1(T_i))=\mathbb{P}_2^{\alpha_2}(Reach_2(T_i))$

B. Ultimate simulation and equivalence for MDPs

In this section we compare the long term behaviors of MDPs, by comparing the Markovian processes induced on their state spaces by policies. Given a memoryless policy σ on an MDP $\mathscr{S}=(S,\Sigma,P)$, it induces an LTMC $\mathscr{S}^{\sigma}$ on S. If σ has memory $i\in\mathbb{N}$, we can see σ as a memoryless policy on $\mathscr{S}^i$, and we write also $\mathscr{S}^{\sigma}$ for the LTMC it induces on $\mathscr{S}^i$.

The class of weakly communicating MDPs will play the role of the irreducible Markov chains of the last subsection. If we consider only i-memory policies, the ultimate simulation relation between two weakly communicating MDPs would depend on the initial distributions of the MDPs. To tackle this problem, if $i\in\mathbb{N}$ we define the class $UR(i)(\mathscr{S})$ of the *ultimately memory i policies* on $\mathscr{S}$, introduced in [15] in the case $i=0$.

Definition 10 (Policies with ultimate finite memory)**.** *A policy σ is in the class $UR(i)(\mathscr{S})$ of the* ultimately memory i policies *on $\mathscr{S}$ if there exists a policy $\sigma^{\infty}\in MR(i)(\mathscr{S})$, called the* tail *of σ, and a random stopping time τ on Ω, called the* switching time *of σ, such that:*

- *If $r\in\Omega$, $\forall n\geq\tau(r)$ we have $\sigma(r_{|n})=\sigma^{\infty}(r_{|n})$.*
- $\mathbb{P}^{\sigma}(\{r|\tau(r)<\infty\})=1.$

In other words, σ is in $UR(i)(\mathscr{S})$ if with probability one, after a finite number of steps, σ behaves as a policy of memory at most i. We can prove a generalization of theorem 1 for policies in $UR(i)$:

Proposition 9. *If $\mathscr{S}$ is weakly communicating, then for all $k,i\in\mathbb{N}$ and all initial distribution α on $\mathscr{S}$,*

$$H_k^{UR(i)}(\alpha)(\mathscr{S})=\overline{H_k^{MR(i)}(\mathscr{S})}$$

We write $\mathscr{S}^{\sigma},\alpha$ for the probabilistic process with labelled transitions (which is not any more a Markov chain), induced by σ and the initial distribution α on $\mathscr{S}$. Two processes $\mathscr{S}_1^{\sigma_1},\alpha_1$ and $\mathscr{S}_2^{\sigma_2},\alpha_2$ are said to be *ultimately equivalent*, written $\mathscr{S}_1^{\sigma_1},\alpha_1\sim_u\mathscr{S}_2^{\sigma_2},\alpha_2$, if they give the same probabilities to the same ultimate properties. That is, if for all $\Gamma\in\mathscr{G}$ we have $\mathbb{P}^{\sigma_1,\alpha_1}(\Gamma)=\mathbb{P}^{\sigma_2,\alpha_2}(\Gamma)$.

The analogous of theorem 6 holds: $\mathscr{S}_1^{\sigma_1},\alpha_1$ and $\mathscr{S}_2^{\sigma_2},\alpha_2$ are ultimately equivalent iff for all $i\in[1;l]$, $\mathbb{P}_1^{\sigma_1,\alpha_1}(Reach_1(T_i))=\mathbb{P}_2^{\sigma_2,\alpha_2}(Reach_2(T_i))$. Here the T_i are equivalence classes on the set of irreducible components of $\mathscr{S}_1^{\sigma_1^{\infty}}$ and $\mathscr{S}_2^{\sigma_2^{\infty}}$

Definition 11 (Simulation between MDPs)**.** *$\mathscr{S}_1,\alpha_1$ is said to be* i-memory ultimately simulated *by $\mathscr{S}_2,\alpha_2$, written $\mathscr{S}_1,\alpha_1\prec_u^i\mathscr{S}_2,\alpha_2$, if for all $\sigma_1\in UR(i)(\mathscr{S}_1)$, there exists $\sigma_2\in UR(i)(\mathscr{S}_2)$ s.t. $\mathscr{S}_1^{\sigma_1},\alpha_1\sim_u\mathscr{S}_2^{\sigma_2},\alpha_2$.*

We say that $\mathscr{S}_1,\alpha_1$ and $\mathscr{S}_2,\alpha_2$ are *i-memory ultimately equivalent*, written $\mathscr{S}_1,\alpha_1\sim_u^i\mathscr{S}_2,\alpha_2$, if $\mathscr{S}_1,\alpha_1\prec_u^i\mathscr{S}_2,\alpha_2$ and $\mathscr{S}_2,\alpha_2\prec_u^i\mathscr{S}_1,\alpha_1$. As for irreducible LTMC, the simulation relation between weakly communicating MDPs does not depend on the initial distributions of the systems. This allows the notation $\mathscr{S}_1\prec_u^i\mathscr{S}_2$ if $\mathscr{S}_1$ and $\mathscr{S}_2$ are weakly communicating. The following theorem resumes the different notions presented in this paper: polytopes, distance, ultimate properties.

Theorem 7. *Let $\mathscr{S}_1$ and $\mathscr{S}_2$ be two weakly communicating MDPs. Then the following are equivalent:*

- $\mathscr{S}_1\prec_u^i\mathscr{S}_2$
- *For all $k\in\mathbb{N}$, $\Pi_k^i(\mathscr{S}_1)\subseteq\Pi_k^i(\mathscr{S}_2)$.*
- *For all $k\in[1;(|S_1|+|S_2|)^i]$, $\Pi_k^i(\mathscr{S}_1)\subseteq\Pi_k^i(\mathscr{S}_2)$.*
- $d^{\prec}_{(|S_1|+|S_2|)^i}(\mathscr{S}_1,\mathscr{S}_2)=0.$

Two weakly communicating MDPs are equivalent according to the relation $\sim_u^i$ induced by $\prec_u^i$ iff their polytopes coincide. The study of the ultimate properties of a general MDP can be done by studying the ultimate properties of its maximal end components, and the probabilities to reach these end components. The maximal end components play the role for MDPs that the irreducible components play for LTMCs.

VII. Conclusion

We introduced Property and Equivalence Testing as a method to approximate classical hard problems on the long term behavior of MDPs, and characterized Equivalent systems with the class of ultimate properties. These methods do not generalize to Probabilistic Automata. Potential applications are the approximate verification of quantitative properties of large probabilistic systems and future work will study how these methods may work with compact representations and with partially observed MDPs.

VIII. Acknowledgment

This research was funded by the French program on Computer Security **ANR-07-SESU-013**

References

[1] A.F. Abrahamse. The tail Field of a Markov Chain. *Ann. Math. Stat*, 1972.

[2] D. Blackwell and D. Freedman. The tail σ-field of a Markov chain and a theorem of Orey. *Ann. Math. Stat*, 35:1291–1295, 1964.

[3] V.D. Blondel. Undecidable Problems for Probabilistic Automata of Fixed Dimension. *Theory of Computing Systems*, 36(3):231–245, 2008.

[4] V.D. Blondel and J.N. Tsitsiklis. A survey of computational complexity results in systems and control. *Automatica-Oxford-*, 36:1249–1274, 2000.

[5] C. Courcoubetis and M. Yannakakis. The complexity of probabilistic verification. *JACM*, 42(4):857–907, 1995.

[6] L. de Alfaro. Formal Verification of Probabilistic Systems. *Technical Report: CS-TR-98-1601, Stanford University*, 1998.

[7] L. de Alfaro. Computing Minimum and Maximum Reachability Times in Probabilistic Systems. *LNCS*, pages 66–81, 1999.

[8] C. Derman. *Finite State Markovian Decision Processes*. Academic Press, Inc. Orlando, FL, USA, 1970.

[9] J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Metrics for labelled Markov processes. *Theoretical Computer Science*, 318(3):323–354, 2004.

[10] L. Doyen, T.A. Henzinger, J. Raskin, T. Harju, and J. Karhumaki. Equivalence of Labeled Markov Chains. *International Journal of Foundations of Computer Science*, 19(3):549–563, 2008.

[11] E. Fischer, F. Magniez, and M. de Rougemont. Approximate Satisfiability and Equivalence. In *LICS 2006*, pages 421–430, 2006.

[12] O. Goldreich, S. Goldwasser, and D. Ron. Property testing and its connection to learning and approximation. *JACM*, 45(4):653–750, 1998.

[13] LCM Kallenberg. Linear programming and finite Markovian control problems. 1983.

[14] S. Khot and A. Naor. Linear Equations Modulo 2 and the L1 Diameter of Convex Bodies. In *FOCS, 2007*, pages 318–328, 2007.

[15] D. Krass and OJ Vrieze. Achieving Target State-Action Frequencies in Multichain Average-Reward Markov Decision Processes. *Mathematics of Operations Research*, 27(3):545, 2002.

[16] O. Madani, S. Hanks, and A. Condon. On the Undecidability of Probabilistic Planning and Infinite-Horizon Partially Observable Markov Decision Problems. In *Proceedings of the National Conference on Artificial Intelligence*, pages 541–548, 1999.

[17] S. Mannor and J.N. Tsitsiklis. On the Empirical State-Action Frequencies in Markov Decision Processes Under General Policies. *Mathematics of Operations Research*, 30(3):545, 2005.

[18] D. Ornstein and B. Weiss. How sampling reveals a process. *The Annals of Probability*, 18(3):905–930, 1990.

[19] CH Papadimitriou and JN Tsitsiklis. The complexity of Markov decision processes. *Mathematics of operations research*, 12(3):441–450, 1987.

[20] M.L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc. New York, NY, USA, 1994.

[21] M.O. Rabin. Probabilistic Automata. *Sequential Machines: Selected Papers*, 1964.

[22] R. Rubinfeld and M. Sudan. Robust characterizations of polynomials with applications to program testing. *SIAM Journal on Computing*, 25:23–32, 1996.

[23] R. Segala and N. Lynch. Probabilistic simulations for probabilistic processes. *Nordic Journal of Computing*, 2(2):250–273, 1995.

[24] P. Shields. *The ergodic theory of discrete sample paths*. The American Mathematical Society, Graduate Studies in Mathematics, volume GS-13, 1996.

[25] M. Tracol. Fast convergence to state-action frequency polytopes for MDPs. *Operations Research Letters*, 2009.

[26] W.G. Tzeng. A Polynomial-Time Algorithm for the Equivalence of Probabilistic Automata. *SIAM Journal on Computing*, 21:216, 1992.

[27] F. van Breugel and J. Worrell. A behavioural pseudometric for probabilistic transition systems. *Theoretical Computer Science*, 331(1):115–142, 2005.

[28] M.Y. Vardi. Automatic verification of probabilistic concurrent finite state programs. In *FOCS 1984*, pages 327–338, 1985.