Minimising the Probabilistic Bisimilarity Distance

Stefan Kiefer
Department of Computer Science, University of Oxford, UK

Qiyi Tang
Department of Computer Science, University of Liverpool, UK

Abstract

A labelled Markov decision process (MDP) is a labelled Markov chain with nondeterminism; i.e., together with a strategy a labelled MDP induces a labelled Markov chain. The model is related to interval Markov chains. Motivated by applications to the verification of probabilistic noninterference in security, we study problems of minimising probabilistic bisimilarity distances of labelled MDPs, in particular, whether there exist strategies such that the probabilistic bisimilarity distance between the induced labelled Markov chains is less than a given rational number, both for memoryless strategies and general strategies. We show that the distance minimisation problem is $\exists\mathbb{R}$-complete for memoryless strategies and undecidable for general strategies. We also study the computational complexity of the qualitative problem about making the distance less than one. This problem is known to be NP-complete for memoryless strategies. We show that it is EXPTIME-complete for general strategies.

2012 ACM Subject Classification Theory of computation → Program verification; Theory of computation → Models of computation; Mathematics of computing → Probability and statistics

Keywords and phrases Markov decision processes, Markov chains

Digital Object Identifier 10.4230/LIPIcs.CONCUR.2024.32

Related Version *Full Version*: `https://arxiv.org/abs/2406.19830` [18]

Funding This work has been supported in part by the Engineering and Physical Sciences Research Council (EPSRC) through grant EP/X042596/1.

Acknowledgements We thank the referees for their constructive feedback.

1 Introduction

Given a model of computation (e.g., finite automata), and two instances of it, are they semantically equivalent (e.g., do the automata accept the same language)? Such *equivalence* problems can be viewed as a fundamental question for almost any model of computation. As such, they permeate computer science, in particular, theoretical computer science.

In *labelled Markov chains (LMCs)*, which are Markov chains whose states (or, equivalently, transitions) are labelled with an observable letter, there are two natural and very well-studied versions of equivalence, namely *trace (or language) equivalence* and *probabilistic bisimilarity*.

The *trace equivalence* problem has a long history, going back to Schützenberger [28] and Paz [21] who studied *weighted* and *probabilistic* automata, respectively. Those models generalise LMCs, but the respective equivalence problems are essentially the same. For LMCs, trace equivalence asks if the same label sequences have the same probabilities in the two LMCs. It can be extracted from [28] that equivalence is decidable in polynomial time, using a technique based on linear algebra; see also [32, 9].

Probabilistic bisimilarity is an equivalence that was introduced by Larsen and Skou [20]. It is finer than trace equivalence, i.e., probabilistic bisimilarity implies trace equivalence. A similar notion for Markov chains, called *lumpability*, can be traced back at least to the classical text by Kemeny and Snell [15]. Probabilistic bisimilarity can also be computed in

polynomial time [1, 7, 33]. Indeed, in practice, computing the bisimilarity quotient is fast and has become a backbone for highly efficient tools for probabilistic verification such as Prism [19] and Storm [12].

Numerous quantitative generalisations of this behavioural equivalence have been proposed, the probabilistic bisimilarity distance due to Desharnais et al. [8] being the most notable one. This distance can be at most 1, and is 0 if and only if the LMCs are probabilistic bisimilar. It was shown in [5] that the distance can be computed in polynomial time.

In this paper, we study distance minimisation problems for *(labelled) Markov decision processes (MDPs)*, which are LMCs plus nondeterminism; i.e., each state may have several *actions* (or "moves") one of which is chosen by a controller, potentially randomly. An MDP and a controller *strategy* together induce an LMC (potentially with infinite state space, depending on the complexity of the strategy). We consider both general strategies and the more restricted memoryless ones. There are good reasons to consider memoryless strategies, particularly their naturalness and simplicity in implementations, and their connection to *interval Markov chains* (see, e.g., [14, 6]) and *parametric MDPs* (see, e.g., [11, 35]). There are also good reasons to consider general unrestricted strategies, primarily their naturalness (in their definition for MDPs) and their generality. The latter is important particularly for security applications, see below, where general strategies can make programs more secure, in a precise, quantitative sense.

Let us elaborate on the connection to security. *Noninterference* refers to an information-flow property of a program, stipulating that information about *high* data (i.e., data with high confidentiality) may not leak to *low* (i.e., observable) data, or, quoting [25], "that a program is secure whenever varying the initial values of high variables cannot change the low-observable (observable by the attacker) behaviour of the program". It was proposed in [25] to reason about *probabilistic* noninterference in probabilistic multi-threaded programs by proving probabilistic bisimilarity; see also [29, 22]. More precisely, probabilistic noninterference is established if it can be shown that any two states that differ only in high data are probabilistic bisimilar, as then an attacker who only observes the low part of a state learns nothing about the high part. The observable behaviour of a multi-threaded program depends strongly on the *scheduler*, which in this context amounts to a strategy in the corresponding MDP.

Nevertheless, ensuring perfect (probabilistic) noninterference proves challenging, and a certain degree of information leakage may be acceptable [13, 24]. In such scenarios, where (probabilistic) bisimilarity might not hold under any scheduler, turning to bisimilarity distances allows us to estimate the security degree of a system under different schedulers. The smaller the distance, the more secure the system. Therefore, we would like to devise schedulers that minimise the probabilistic bisimilarity distances.

Some qualitative problems have already been studied in previous work. Concerning memoryless strategies, it was shown in [16] that the bisimilarity equivalence problem, i.e., whether strategies exist to make the distance 0, is NP-complete. Similarly, it was also shown in [16] that the problem whether memoryless strategies exist to make the distance less than one is NP-complete; cf. Table 1. The bisimilarity *in*equivalence problem, i.e., whether strategies exist to make the distance greater than 0, can be decided in polynomial time for memoryless strategies [16].

Concerning general strategies, the bisimilarity equivalence and inequivalence problems were studied in [17]. It was shown there that these problems are EXPTIME-complete and in P, respectively.

It remained open whether the existence of strategies to make the distance less than one is decidable for general strategies. We show that the distance less than one problem for general strategies is decidable. In fact, it is EXPTIME-complete, and therefore the problem has the

Table 1 Summary of results on distance minimisation problems.

Problem	Memoryless Strategy	General Strategy
distance $= 0$	NP-complete [16]	EXPTIME-complete [17]
distance < 1	NP-complete [16]	EXPTIME-complete (Section 6)
distance $< \theta$	$\exists\mathbb{R}$-complete (Section 4)	undecidable (Section 5)

same complexity as the bisimilarity equivalence problem for general strategies. To obtain this result, we prove a tight connection between the distance less than one problem and the bisimilarity equivalence problem: loosely speaking, whenever there are general strategies for two states to have distance less than one, the two states can reach a pair of states whose distance can be made 0, thus probabilistic bisimilar. This connection is natural and known for *finite* labelled Markov chains, but nontrivial to establish in general.

We also study *quantitative* distance minimisation problems: do there exist memoryless (resp. general) strategies for two given MDPs such that the induced LMCs have distance less than a given threshold? We show that the distance minimisation problem is $\exists\mathbb{R}$-complete for memoryless strategies and undecidable for general strategies. Here, $\exists\mathbb{R}$ refers to the class of problems that are many-one reducible to the existential theory of the reals; it is known that $\mathsf{NP} \subseteq \exists\mathbb{R} \subseteq \mathsf{PSPACE}$.

The rest of the paper is organised as follows. We give preliminaries in Section 2. In Section 3 we discuss probabilistic noninterference. In Sections 4 and 5 we prove our results on the quantitative distance minimisation problems for general strategies and memoryless strategies, respectively. We study the distance less than one problem for general strategies in Section 6. We conclude in Section 7. Missing proofs can be found in the full version of this paper [18].

2 Preliminaries

We write $\mathbb{N}$ for the set of nonnegative integers. Let S be a finite set. We denote by $\mathrm{Distr}(S)$ the set of probability distributions on S. For a distribution $\mu \in \mathrm{Distr}(S)$ we write $\mathrm{support}(\mu) = \{s \in S \mid \mu(s) > 0\}$ for its support. We denote the Dirac distribution concentrated on an element $s \in S$ by $\mathbf{1}_s$, that is, $\mathbf{1}_s(s) = 1$ and $\mathbf{1}_s(t) = 0$ for all $t \neq s$. We denote by $\rho(i)$ the i-th element of a sequence ρ. We denote the least fixed point of a function f by $\boldsymbol{\mu}.f$.

A *labelled Markov chain* (LMC) is a quadruple $\langle S, L, \tau, \ell\rangle$ consisting of a nonempty countable set S of states, a nonempty finite set L of labels, a transition function $\tau : S \to \mathrm{Distr}(S)$, and a labelling function $\ell : S \to L$. We denote by $\tau(s)(t)$ the transition probability from s to t. Similarly, we denote by $\tau(s)(E) = \sum_{t\in E} \tau(s)(t)$ the transition probability from s to $E \subseteq S$. We require the LMCs to be finitely branching, that is, $|\mathrm{support}(\tau(s))|$ is finite for every $s \in S$.

An equivalence relation $R \subseteq S \times S$ is a *probabilistic bisimulation* if for all $(s,t) \in R$, $\ell(s) = \ell(t)$ and $\tau(s)(E) = \tau(t)(E)$ for each R-equivalence class E. *Probabilistic bisimilarity*, denoted by $\sim_{\mathcal{M}}$ (or $\sim$ when $\mathcal{M}$ is clear), is the largest probabilistic bisimulation.

The probabilistic bisimilarity distance, a pseudometric on LMCs, was first defined by Desharnais, Gupta, Jagadeesan and Panangaden in [8]. Their definition is based on a real-valued modal logic. This logic can be viewed as a function which maps a formula f of the logic and a state s of the LMC to a real number $f(s) \in [0,1]$. The distance $d(s,t)$ between two states s, t is defined as $\sup_f |f(s) - f(t)|$. Later, Van Breugel and Worrell [34]

defined probabilistic bisimilarity distances on LMCs as a fixed point of a function. They showed that their pseudometric coincides with the one defined in [8]. In this paper, we use the definition from [34]. The *probabilistic bisimilarity distance*, denoted by $d_{\mathcal{M}}$ (or d when $\mathcal{M}$ is clear), is a function from $S \times S$ to $[0,1]$, that is, an element of $[0,1]^{S\times S}$. It is the least fixed point of the following function:

$$\Delta(e)(s,t) = \begin{cases} 1 & \text{if } \ell(s) \neq \ell(t) \\ \min\limits_{\omega \in \Omega(\tau(s),\tau(t))} \sum\limits_{u,v\in S} \omega(u,v)\, e(u,v) & \text{otherwise} \end{cases}$$

where the set $\Omega(\mu,\nu)$ of *couplings* of $\mu,\nu \in \mathrm{Distr}(S)$ is defined as $\Omega(\mu,\nu) = \left\{ \omega \in \mathrm{Distr}(S\times S) \;\middle|\; \sum_{t\in S} \omega(s,t) = \mu(s) \wedge \sum_{s\in S} \omega(s,t) = \nu(t) \right\}$. Note that a coupling $\omega \in \Omega(\mu,\nu)$ is a joint probability distribution with marginals μ and ν (see, e.g., [2, page 260-262]). For all $s,t \in S$, $s \sim t$ if and only if s and t has probabilistic bisimilarity distance zero [8, Theorem 1].

A *(labelled) Markov decision process* (MDP) is a tuple $\langle S, Act, L, \varphi, \ell\rangle$ consisting of a finite set S of states, a finite set Act of actions, a finite set L of labels, a partial function $\varphi : S \times Act \nrightarrow \mathrm{Distr}(S)$ denoting the probabilistic transition, and a labelling function $\ell : S \to L$. The set of available actions in a state s is $Act(s) = \{\mathsf{m} \in Act \mid \varphi(s,\mathsf{m}) \text{ is defined}\}$.

A path is a sequence $\rho = s_0\mathsf{m}_1 s_1 \cdots \mathsf{m}_n s_n$ such that $\varphi(s_i,\mathsf{m}_{i+1})$ is defined and $\varphi(s_i,\mathsf{m}_{i+1})(s_{i+1}) > 0$ for all $0 \leq i < n$. The last state of ρ is $\mathsf{last}(\rho) = s_n$. Let $\mathsf{Paths}(\mathcal{D})$ denote the set of paths in $\mathcal{D}$.

A (general) strategy for an MDP is a function $\alpha : \mathsf{Paths}(\mathcal{D}) \to \mathrm{Distr}(Act)$ that given a path ρ, returns a probability distribution on the available actions at the last state of ρ, $\mathsf{last}(\rho)$. A memoryless strategy depends only on $\mathsf{last}(\rho)$; so we can identify a memoryless strategy with a function $\alpha : S \to \mathrm{Distr}(Act)$ that given a state s, returns a probability distribution on the available actions at that state.

A general strategy α for $\mathcal{D}$ induces an LMC $\mathcal{D}(\alpha) = \langle \mathcal{P}, L, \tau, \ell'\rangle$, where $\mathcal{P} \subseteq \mathsf{Paths}(\mathcal{D})$. For $\rho \in \mathcal{P}$, we have $\tau(\rho)(\rho\mathsf{m}t) = \alpha(\rho)(\mathsf{m})\varphi(s,\mathsf{m})(t)$ and $\ell'(\rho) = \ell(s)$ where $s = \mathsf{last}(\rho)$ and $\mathsf{m} \in Act(s)$.

3 Probabilistic Noninterference

In this section we provide examples that show some challenges in distance minimisation and illustrate the relation between distance minimisation and probabilistic noninterference in security. As described in the introduction, we are interested in schedulers that minimise the information leakage.

▶ **Example 1.** We borrow an example from [25, Section 4] and [17, Section 3]. Consider the following simple program composed of two threads, involving a *high* boolean variable h (high confidentiality) and a *low* boolean variable l (observable):

$$l := h \quad | \quad l := \neg h$$

The vertical bar | separates two threads. The order in which the threads are executed is determined by a scheduler. We assume that assignments to the value of variable l are visible. One may model the program as the following MDP in Figure 1. Here, s_0 and s_1 correspond to initial states with $h = 0$ and $h = 1$, respectively. The two actions in the MDP, m_0 and m_1, correspond to the two possible orders of execution: action m_0 models the choice of executing $l := h$ first, followed by $l := \neg h$, while m_1 models the reverse order. The different colours

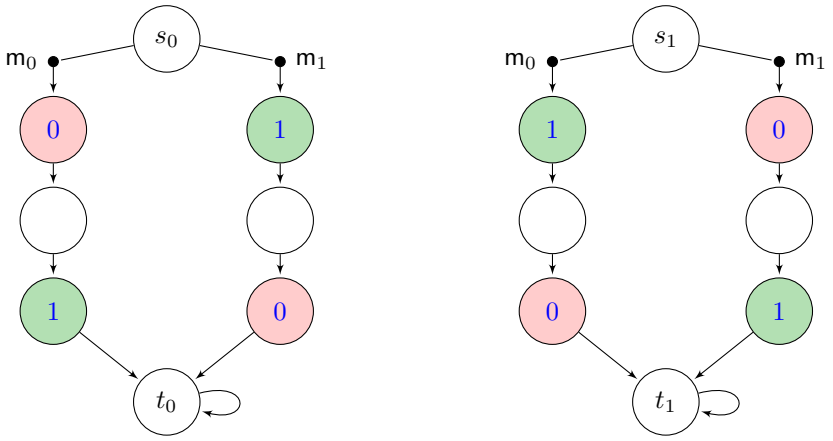


Figure 1 The program from Example 1 as an MDP. The states s_0 and s_1 have two available actions, m_0 and m_1. The default action m for the other states is omitted. Different colours (state labels) indicate the distinct values of the low data. Throughout the paper, transition probabilities out of each action are one unless explicitly specified.

represent the distinct values of the low, observable data. For instance, in state s_0, if the scheduler selects m_0 (the left branch of s_0), then l becomes 0 after executing $l := h$ and 1 after executing $l := \neg h$. All transitions are with probability one. A memoryless strategy that chooses actions $\mathsf{m}_0, \mathsf{m}_1$ uniformly at random (i.e., with probability 0.5 each) makes s_0, s_1 probabilistic bisimilar; i.e., $d(s_0, s_1) = 0$ under this strategy. ⌟

▶ **Example 2.** Consider the following variant of Example 1.

repeat
$l := h \quad | \quad l := \neg h$
until $coin(p) \vee h$

Here, $coin(p)$, for a fixed parameter $p \in [0, 1]$, models a biased coin that returns *true* with probability p and *false* with probability $1 - p$. One may model the program as the MDP in Figure 2, *except* that t_0, t_1 are sinks, as in Example 1. The value of h influences the termination condition of the loop and therefore "leaks" (with probability $1 - p$). As a result, under the optimal (in terms of minimising the distance) strategy, which is the same as in Example 1, we have now $d(s_0, s_1) = 1 - p$. The smaller p, the "worse" the leak. ⌟

The following example shows that general strategies may be needed for optimal security.

▶ **Example 3.** In order to mitigate the leak from Example 2, one might extend the program as follows, so that the scheduler is given an opportunity to disguise the fact that the program with $h = 1$ tends to terminate earlier than the program with $h = 0$:

repeat
$l := h \quad | \quad l := \neg h$
until $coin(p) \vee h$
repeat forever
$l := 0 \quad \oplus \quad l := 1$

Here, $\oplus$ stands for a nondeterministic choice, to be made by the scheduler, where exactly one of the instructions $l := 0$ and $l := 1$ is executed. In Figure 2, this corresponds to taking actions m_2 and m_3, respectively. One can show that the optimal *memoryless* strategy

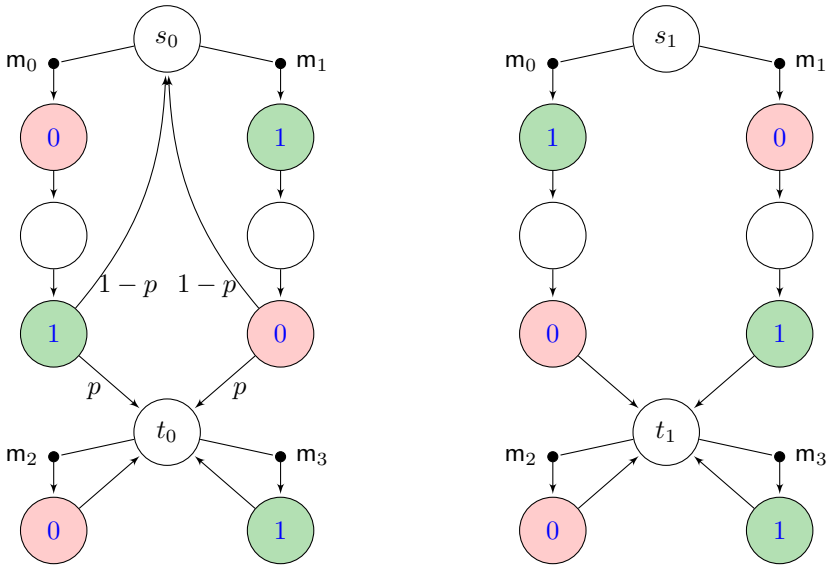


Figure 2 The program from Example 3 as an MDP. The states s_0 and s_1 have two available actions, m_0 and m_1. The states t_0 and t_1 also have two available actions, m_2 and m_3. Different colours (state labels) indicate the distinct values of the low data.

chooses between m_0 and m_1 uniformly at random (as before), and also chooses between m_2 and m_3 uniformly at random. Under this strategy we have $d(s_0,t_1) = 0.5+0.5(1-p)d(s_0,t_1)$, implying $d(s_0,t_1) = \frac{1}{1+p}$, and thus $d(s_0,s_1) = (1-p)d(s_0,t_1) = \frac{1-p}{1+p}$, which is, for $p \in (0,1)$, smaller (i.e., better) than the distance achievable in Example 2.

However, there is a general strategy α, not memoryless, that perfectly disguises when the first loop is exited. This strategy α chooses between m_0 and m_1 uniformly at random (as before). When the execution path visits t_0 or t_1 for the ith time, $i \geq 1$, then, if i is odd, α chooses between m_2 and m_3 uniformly at random, and if i is even, α chooses the action that was not taken upon the $(i-1)$th visit of t_0 or t_1. Under this strategy α we have $d(s_0,s_1) = 0$, i.e., s_0 and s_1 are probabilistic bisimilar. ⌟

4 Memoryless Strategies: Distance Minimisation

In this section we consider the *memoryless distance minimisation problem* which, given an MDP, two states s_1, s_2 of the MDP, and a rational number θ, asks whether there is a memoryless strategy α such that $d(s_1,s_2) < \theta$ holds in the LMC induced by α.

We show that the memoryless distance minimisation problem is $\exists\mathbb{R}$-complete. We prove the lower and upper bound in Theorems 7 and 8, respectively.

The *existential theory of the reals,* ETR, is the set of valid formulas of the form

$$\exists x_1 \ldots \exists x_n \; R(x_1,\ldots,x_n),$$

where R is a Boolean combination of comparisons of the form $p(x_1,\ldots,x_n) \sim 0$, in which $p(x_1,\ldots,x_n)$ is a multivariate polynomial (with rational coefficients) and $\sim \in \{<,>,\leq,\geq,=,\neq\}$. The complexity class $\exists\mathbb{R}$ [27] consists of those problems that are many-one reducible to ETR in polynomial time. Since ETR is NP-hard and in PSPACE [4, 23], we have NP $\subseteq \exists\mathbb{R} \subseteq$ PSPACE.

To prove that the memoryless distance minimisation problem is $\exists\mathbb{R}$-hard (Theorem 7), we proceed via a sequence of reductions, represented by the following lemmas, Lemmas 4–6.

▶ **Lemma 4.** *The following problem is $\exists\mathbb{R}$-complete: given a multivariate polynomial $p : \mathbb{R}^n \to \mathbb{R}$ of (total) degree at most 6, does there exist $x \in \mathbb{R}^n$ with $p(x) < 0$? The problem remains $\exists\mathbb{R}$-complete under the promise that if there is x with $p(x) < 0$ then there is x' with $p(x') < 0$ and $\|x'\| < 1$ (where $\|\cdot\|$ denotes the Euclidean norm).*

Proof. Membership in $\exists\mathbb{R}$ is clear. It remains to prove $\exists\mathbb{R}$-hardness. It is shown in [26, Lemma 3.9] that the following problem is $\exists\mathbb{R}$-complete: given multivariate polynomials $f_1, \dots, f_s : \mathbb{R}^n \to \mathbb{R}$, each of degree at most 2, does there exist $x \in \mathbb{R}^n$ with $\|x\| < 1$ such that $\bigwedge_{i=1}^{s}(f_i(x) = 0)$? It follows from the proof that the problem remains $\exists\mathbb{R}$-complete under the promise that $\bigwedge_i f_i(x) = 0$ implies $\|x\| < 1$. We reduce from this promise problem. Let $f_1, \dots, f_s : \mathbb{R}^n \to \mathbb{R}$, each of degree at most 2, such that for all $x \in \mathbb{R}^n$ we have that $\bigwedge_i f_i(x) = 0$ implies $\|x\| < 1$. Define the polynomial $q : \mathbb{R}^n \to \mathbb{R}$ by $q(x) := \sum_{i=1}^{s} f_i(x)^2$. Clearly, $q(x) \geq 0$ always holds, and we have $q(x) = 0$ if and only if $\bigwedge_i f_i(x) = 0$. Consider the two sets $\{(q(x), x) \in \mathbb{R}^{n+1} \mid \|x\| \leq 1\}$ and $\{(0, x) \in \mathbb{R}^{n+1} \mid \|x\| \leq 1\}$. If q has a root x, then the two sets overlap in the point $(0, x)$; otherwise, by [27, Corollary 3.4], they have distance at least $2^{2^{-k}}$, where k is a natural number whose unary representation can be computed in polynomial time. It follows that if $\|x\| \leq 1$ and $q(x) < 2^{2^{-k}}$ then there exists x' such that $q(x') = 0$.

In the following let us use real-valued variables $x_1, \dots, x_n, y_1, \dots, y_k$ and write $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_k)$. Define the polynomial $r : \mathbb{R}^{n+k} \to \mathbb{R}$ (of degree at most 6) by

$$r(x, y) \;:=\; (y_1 - 4)^2 + (y_2 - y_1^2)^2 + \cdots + (y_k - y_{k-1}^2)^2 + y_k^2 q(x) + \|x\|^2 - 1\,.$$

Let us also use a real-valued variable z. Define the polynomial $p : \mathbb{R}^{n+k+1}$ (of degree at most 6) by

$$p(x, y, z) \;:=\; z^6 r\left(\frac{x_1}{z}, \dots, \frac{x_n}{z}, \frac{y_1}{z}, \dots, \frac{y_k}{z}\right).$$

Suppose there is $x \in \mathbb{R}^n$ with $\bigwedge_i f_i(x) = 0$. Then $q(x) = 0$. For $1 \leq i \leq k$, set $y_i := 2^{2^i}$. Then $r(x, y) = \|x\|^2 - 1 < 0$. Set $z > 0$ small enough so that $z^2\left(\|x\|^2 + \|y\|^2 + 1\right) < 1$. For $1 \leq i \leq n$, set $x'_i := x_i z$. For $1 \leq i \leq k$, set $y'_i := y_i z$. Then $p(x', y', z) = z^6 r(x, y) < 0$ and $\|x'\|^2 + \|y'\|^2 + z^2 = z^2\left(\|x\|^2 + \|y\|^2 + 1\right) < 1$.

Towards the other direction, suppose there is $(x', y', z) \in \mathbb{R}^{n+k+1}$ with $p(x', y', z) < 0$. Since p is a polynomial, it is continuous. So we can assume without loss of generality that $z \neq 0$. For $1 \leq i \leq n$, set $x_i := x'_i/z$. For $1 \leq i \leq k$, set $y_i := y'_i/z$. Then $r(x, y) = p(x', y', z)/z^6 < 0$. This implies $y_k^2 q(x) < 1$ and $\|x\| < 1$. Using $r(x, y) < 0$, we show by induction that $y_i \geq 2^{2^{i-1}} + 1$ holds for all $i \in \{1, \dots, k\}$. For the induction base ($i = 1$) we have $(y_1 - 4)^2 \leq 1$. Thus, $y_1 - 4 \geq -1$, and so $y_1 \geq 3 = 2^{2^{1-1}} + 1$. For the step ($1 \leq i \leq k - 1$), suppose that $y_i \geq 2^{2^{i-1}} + 1$. Since $r(x, y) < 0$, we have $(y_{i+1} - y_i^2)^2 \leq 1$, and so

$$y_{i+1} \;\geq\; y_i^2 - 1 \;\geq\; (2^{2^{i-1}} + 1)^2 - 1 \;=\; 2^{2^i} + 2 \cdot 2^{2^{i-1}} \;\geq\; 2^{2^i} + 1\,.$$

Hence, we have shown that $y_k \geq 2^{2^{k-1}} + 1 > 2^{2^{k-1}}$. It follows that $q(x) < 1/y_k^2 < 2^{-2^k}$. Since $\|x\| < 1$, it follows from the argument at the beginning that there exists x' such that $q(x') = 0$ and so $\bigwedge_i f_i(x') = 0$.

This completes the hardness proof. Note that by combining the two directions, it follows that if there is $w \in \mathbb{R}^{n+k+1}$ with $p(w) < 0$, then there is $w' \in \mathbb{R}^{n+k+1}$ with $p(w') < 0$ and $\|w'\| < 1$, showing also $\exists\mathbb{R}$-hardness of the promise version of the problem. ◀

▶ **Lemma 5.** *The following problem is $\exists\mathbb{R}$-complete: given a multivariate polynomial $p : \mathbb{R}^n \to \mathbb{R}$ of degree at most 6, does there exist $x \in [0,1]^n$ with $p(x) > 0$?*

Proof. Membership in $\exists\mathbb{R}$ is clear. For hardness we reduce from the promise problem from the previous lemma. Let $p : \mathbb{R}^n \to \mathbb{R}$ be a multivariate polynomial of degree at most 6 such that if there is $x \in \mathbb{R}^n$ with $p(x) < 0$ then there is $x' \in \mathbb{R}^n$ with $p(x') < 0$ and $\|x'\| < 1$. Define the polynomial $q : \mathbb{R}^{2n} \to \mathbb{R}$ by $q(y_1, \ldots, y_n, z_1, \ldots, z_n) := -p(y_1 - z_1, \ldots, y_n - z_n)$. The degree of q is at most 6. We have to show that there is $x \in \mathbb{R}^n$ with $p(x) < 0$ if and only if there are $y_1, \ldots, y_n, z_1, \ldots, z_n \in [0,1]$ with $q(y_1, \ldots, y_n, z_1, \ldots, z_n) > 0$.

Suppose there are $x_1, \ldots, x_n \in \mathbb{R}$ with $p(x_1, \ldots, x_n) < 0$. By the property of p we can assume that $x_1^2 + \cdots + x_n^2 < 1$. It follows that $x_i \in [-1, 1]$ holds for all i. For all i with $x_i \geq 0$ define $y_i := x_i$ and $z_i := 0$. For all i with $x_i < 0$ define $y_i := 0$ and $z_i := -x_i$. Then we have $x_i = y_i - z_i$ and $y_i, z_i \in [0,1]$ for all i. Further,

$$q(y_1, \ldots, y_n, z_1, \ldots, z_n) \;=\; -p(y_1 - z_1, \ldots, y_n - z_n) \;=\; -p(x_1, \ldots, x_n) \;>\; 0\,.$$

Towards the other direction, suppose that there are $y_1, \ldots, y_n, z_1, \ldots, z_n \in [0,1]$ with $q(y_1, \ldots, y_n, z_1, \ldots, z_n) > 0$. For all i define $x_i := y_i - z_i$. Then we have

$$p(x_1, \ldots, x_n) \;=\; p(y_1 - z_1, \ldots, y_n - z_n) \;=\; -q(y_1, \ldots, y_n, z_1, \ldots, z_n) \;<\; 0\,,$$

as required. ◀

▶ **Lemma 6.** *The following problem is $\exists\mathbb{R}$-complete: given a rational number $\theta \geq 0$ and a multivariate (degree-6) polynomial $p : \mathbb{R}^n \to \mathbb{R}$ of the form $p(x) = \sum_{j=1}^{k} f_j(x)$ where each $f_j(x_1, \ldots, x_n)$ is a product of a nonnegative coefficient and 6 terms of the form x_i or $(1 - x_i)$, does there exist $x \in [0,1]^n$ with $p(x) > \theta$?*

Proof. Membership in $\exists\mathbb{R}$ is clear. Towards hardness, suppose $m : \mathbb{R}^n \to \mathbb{R}$ is a monomial with a negative coefficient, i.e.,

$$m(x_1, \ldots, x_n) \;=\; -c \prod_{j=1}^{d} x_{i_j} \qquad \text{for some } c > 0 \text{ and } i_1, \ldots, i_d \in \{1, \ldots, n\}\,.$$

Then we have

$$\begin{aligned} m(x_1, \ldots, x_n) \;&=\; -c \prod_{j=1}^{d} x_{i_j} \;=\; c(1 - x_{i_1}) \prod_{j=2}^{d} x_{i_j} - c \prod_{j=2}^{d} x_{i_j} \;=\; \ldots \\ &=\; -c + \sum_{k=1}^{d} c(1 - x_{i_k}) \prod_{j=k+1}^{d} x_{i_j}\,. \end{aligned}$$

We reduce from the problem from Lemma 5. Let $p : \mathbb{R}^n \to \mathbb{R}$ be a multivariate polynomial of degree at most 6. By rewriting each monomial of p that has a negative coefficient using the pattern above, we can write $p(x) = -\theta + q(x)$ for some $\theta \geq 0$ and some $q : \mathbb{R}^n \to \mathbb{R}$ of the form $q(x) = \sum_{j=1}^{k} f_j(x)$ where each $f_j(x)$ is a product of a nonnegative coefficient and at most 6 terms of the form x_i or $(1 - x_i)$. As long as there is an $f_j(x_1, \ldots, x_n)$ of degree less than 6, we can replace it by the two summands $x_1 f_j(x_1, \ldots, x_n)$ and $(1 - x_1) f_j(x_1, \ldots, x_n)$. So we can assume that every $f_j(x)$ has the required form. For all $x \in \mathbb{R}^n$ we have that $p(x) > 0$ if and only if $q(x) > \theta$, as required. ◀

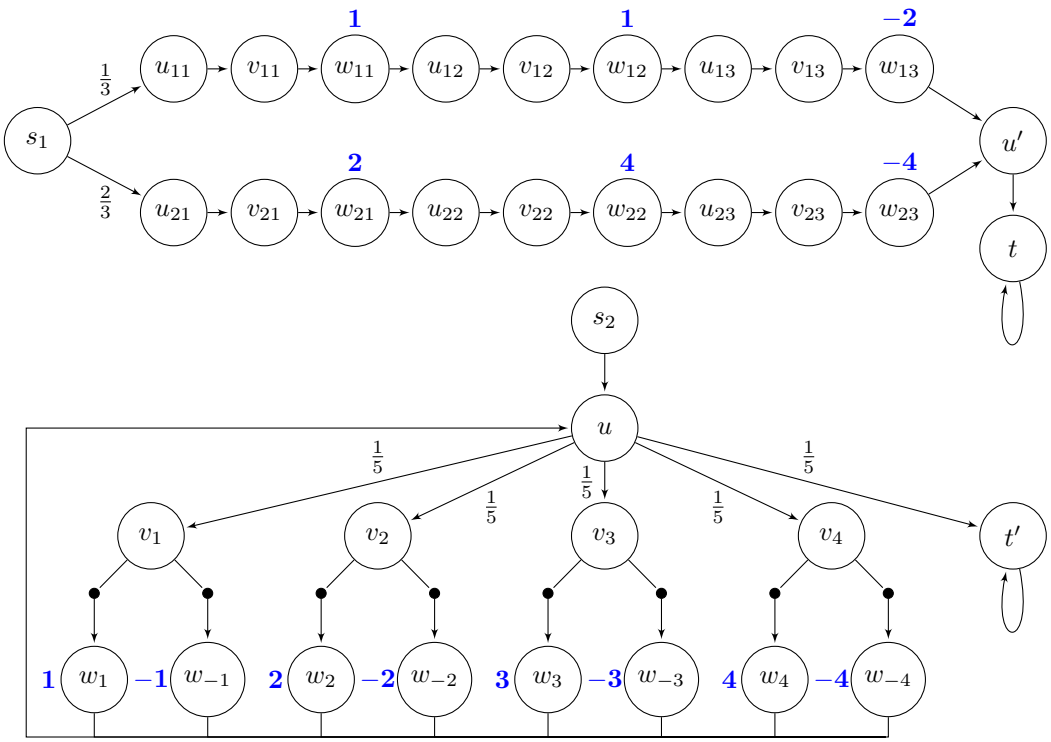


Figure 3 An illustration of the proof of Theorem 7. Consider the polynomial p with $p(x_1, x_2, x_3, x_4) = \frac{1}{3}x_1^2(1 - x_2) + \frac{2}{3}x_2x_4(1 - x_4)$. This example polynomial has degree 3 (instead of degree 6 in the proof) to allow for a more succinct picture. The analogous construction from the reduction yields the shown MDP. The labels are written next to the states in blue, unlike the other figures in this paper where we usually use different colours to indicate different state labels. We omit label 0. There is a one-to-one correspondence between an assignment $x \in [0, 1]^4$ and a memoryless strategy $\alpha(x)$ in the MDP. It is such that $d(s_1, s_2) = 1 - \frac{p(x)}{5^4}$, establishing a connection between an evaluation of p and the distance.

To show that the memoryless distance minimisation problem is $\exists\mathbb{R}$-hard, we reduce from the problem in Lemma 6. We give a brief outline of the reduction. Given a multivariate polynomial $p : \mathbb{R}^n \to \mathbb{R}$ of the form as in Lemma 6, we construct an MDP with initial states s_1 and s_2 such that each assignment $x \in [0, 1]^n$ corresponds to a memoryless strategy $\alpha(x)$ of the MDP. The distance of s_1 and s_2 in the LMC induced by the memoryless strategy $\alpha(x)$ is $1 - c \cdot p(x)$ where c is a constant. Therefore, there exists $x \in [0, 1]^n$ with $p(x) > \theta$ if and only if there exists a memoryless strategy $\alpha(x)$ such that the distance of s_1 and s_2 is less than $1 - c \cdot \theta$.

▶ **Theorem 7.** *The memoryless distance minimisation problem is $\exists\mathbb{R}$-hard.*

Proof. We reduce from the problem from Lemma 6. Let $\theta \geq 0$ and let $p : \mathbb{R}^n \to \mathbb{R}$ be a multivariate polynomial of the form $p(x) = \sum_{j=1}^{m} f_j(x)$ where each $f_j(x_1, \ldots, x_n)$ is a product of a nonnegative coefficient and 6 terms of the form x_i or $(1 - x_i)$. Let us write $f_j(x_1, \ldots, x_n) = c_j \prod_{k=1}^{6} x_{\ell(j,k)}$ where each $c_j \geq 0$ and each $\ell(j, k) \in \{-n, \ldots, -1, 1, \ldots, n\}$ and we use the notation x_{-i} for $i > 0$ to mean $1 - x_i$. We can assume that $\sum_{j=1}^{m} c_j = 1$ (otherwise, divide θ and each c_j by $\sum_{j=1}^{m} c_j$).

Construct an MDP which consists of two disjoint parts as follows; see Figure 3 for an illustration. The first part is an LMC. Include states $u_{j,k}, v_{j,k}, w_{j,k}$ for each $j \in \{1, \ldots, m\}$ and each $k \in \{1, \ldots, 6\}$. Each $u_{j,k}, v_{j,k}$ has label 0, and each $w_{j,k}$ has label $\ell(j, k)$. Each

$u_{j,k}$ transitions with probability 1 to $v_{j,k}$. Each $v_{j,k}$ transitions with probability 1 to $w_{j,k}$. Each $w_{j,k}$, except those with $k = 6$, transitions with probability 1 to $u_{j,k+1}$. Include also states s_1, u' and t with label 0. State s_1 transitions with probability c_j to $u_{j,1}$, for each j. State u' transitions with probability 1 to t. State t is a sink state, that is, it transitions with probability 1 to itself. Also each $w_{j,6}$ transitions with probability 1 to u'.

The second part is an MDP. Include states s_2, u, t' with label 0. State s_2 transitions with probability 1 to u. Include also states $v_1, \ldots, v_n$, each with label 0. State u transitions to each v_i and t' with probability $\frac{1}{n+1}$. State t' is a sink state. Include also states $w_{-n}, \ldots, w_{-1}, w_1, \ldots, w_n$, where each w_i has label i. Each v_i has two actions, one of which leads with probability 1 to w_i, the other one with probability 1 to w_{-i}. Each w_i transitions with probability 1 to u.

Each assignment $x \in [0,1]^n$ corresponds to a memoryless strategy $\alpha(x)$ such that in state v_i the memoryless strategy $\alpha(x)$ takes with probability x_i the action that leads to w_i, and $\alpha(x)$ takes with probability $1 - x_i$ the action that leads to w_{-i}. In fact, this mapping α (from an assignment to a memoryless strategy) is a bijection. Fix an arbitrary $x \in [0,1]^n$ and consider the distances in the LMC induced by $\alpha(x)$. For notational convenience, for any states s, s' let us write $\overline{d}(s,s') := 1 - d(s,s')$. Further, let us write $u_{j,7}$ for u'.

Let $j \in \{1, \ldots, m\}$ and $k \in \{1, \ldots, 6\}$. Then we have

$$\overline{d}(u_{j,k}, u) \;=\; \frac{1}{n+1}\overline{d}(v_{j,k}, v_{|\ell(j,k)|}) \;=\; \frac{1}{n+1}x_{\ell(j,k)}\overline{d}(w_{j,k}, w_{\ell(j,k)}) \;=\; \frac{1}{n+1}x_{\ell(j,k)}\overline{d}(u_{j,k+1}, u)\,.$$

Since $\overline{d}(u_{j,7}, u) = \overline{d}(u', u) = \frac{1}{n+1}$, it follows

$$\overline{d}(u_{j,1}, u) \;=\; \left(\frac{1}{n+1}\right)^7 \prod_{k=1}^{6} x_{\ell(j,k)}\,.$$

Hence,

$$\begin{aligned}\overline{d}(s_1, s_2) \;&=\; \sum_{j=1}^{m} c_j \overline{d}(u_{j,1}, u) \;=\; \sum_{j=1}^{m} c_j \left(\frac{1}{n+1}\right)^7 \prod_{k=1}^{6} x_{\ell(j,k)} \;=\; \left(\frac{1}{n+1}\right)^7 \sum_{j=1}^{m} f_j(x) \\ &=\; \frac{p(x)}{(n+1)^7}\,.\end{aligned}$$

Thus, we have $p(x) > \theta$ if and only if $\overline{d}(s_1, s_2) > \frac{\theta}{(n+1)^7}$ if and only if $d(s_1, s_2) < 1 - \frac{\theta}{(n+1)^7}$. This completes the hardness proof. ◀

The following theorem, proved in [18, A.1], provides a matching upper bound.

▶ **Theorem 8.** *The memoryless distance minimisation problem is in* $\exists\mathbb{R}$.

Together with Theorem 7 we obtain:

▶ **Corollary 9.** *The memoryless distance minimisation problem is* $\exists\mathbb{R}$*-complete.*

5 General Strategies: Distance Minimisation

In this section we consider the *general distance minimisation problem* which, given an MDP, two states s_1, s_2 of the MDP, and a rational number θ, asks whether there is a general strategy α such that $d(s_1, s_2) < \theta$ holds in the LMC induced by α.

To show that the general distance minimisation problem is undecidable, we establish a reduction from the emptiness problem for probabilistic automata.

A probabilistic automaton is a tuple $\mathcal{A} = \langle Q, q_0, L, \delta, F\rangle$ consisting of a finite set Q of states, an initial state $q_0 \in Q$, a finite set L of letters, a transition function $\delta : Q \times L \to \mathrm{Distr}(Q)$ assigning to every state and letter a distribution over states, and a set F of final states. We also extend δ to words, by letting $\delta(q_0, \varepsilon) = \mathbf{1}_{q_0}$ and $\delta(q_0, \sigma w) = \sum_{q\in Q} \delta(q_0, \sigma)(q)\delta(q, w)$ for $\sigma \in L$ and $w \in L^*$. For a state $q \in Q$, $\mathcal{A}_q$ is the probabilistic automaton obtained from $\mathcal{A}$ by making q the initial state.

We write $\Pr_{\mathcal{A}}(w) = \sum_{q\in F} \delta(q_0, w)(q)$ to denote the probability that $\mathcal{A}$ accepts a word w. The emptiness problem asks, given a probabilistic automaton $\mathcal{A}$, whether there exists a word w such that $\Pr_{\mathcal{A}}(w) > \frac{1}{2}$ holds. The probabilistic automaton $\mathcal{A}$ is called empty if no such word exists. This problem is known to be undecidable [10, 21], even for probabilistic automata with only two letters [3][1].

Let $\mathcal{A} = \langle Q, q_0, L, \delta, F\rangle$ be a probabilistic automaton; without loss of generality we assume that $q_0 \notin F$ and $L = \{a, b\}$. We construct an MDP $\mathcal{D}$ with states s_1 and s_2 and a number θ such that $\mathcal{A}$ is nonempty if and only if there is a general strategy such that $d(s_1, s_2) < \theta$ in the induced LMC.

Let us first outline the idea of the construction. Our MDP includes the part shown in Figure 4, where after a random word $w \in L^*$ is produced, the strategy must choose between taking the transition to x or to y. Lemma 10 below characterises the distance of s_1 and s_2 under strategy α in terms of α and $\Pr_{\mathcal{A}}$. It follows from Lemma 10 that the following strategy minimises the distance: if the random word w satisfies $\Pr_{\mathcal{A}}(w) \le \frac{1}{2}$, choose the transition to x; otherwise choose the transition to y. Setting θ as the distance under the strategy that always chooses the transition to x, we obtain that the distance can be made less than θ if and only if there is a word w with $\Pr_{\mathcal{A}}(w) > \frac{1}{2}$.

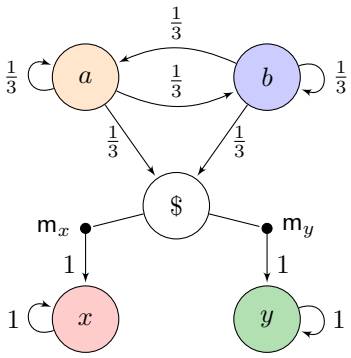


Figure 4 The first part of the MDP $\mathcal{D}$. The \$ state is the only one that has nondeterministic choices: it has two available actions, m_x and m_y. The default action m for the other states is omitted. Different colours indicate different state labels.

We now give the details of the construction. The MDP $\mathcal{D} = \langle S, Act, L', \varphi, \ell\rangle$ consists of two disjoint parts as follows; see Figure 4 and Figure 5. The set of actions is $Act = \{\mathsf{m}, \mathsf{m}_x, \mathsf{m}_y\}$. The set of labels is $L' = \{a, b, \$, x, y\}$.

The first part is an MDP shown in Figure 4. Its set of states is $\{a, b, \$, x, y\}$. The state s_1 is defined to be a. The transitions φ are defined as follows:

[1] It is stated in [3, Theorem 2.1] that the emptiness problem with unfixed threshold λ, i.e., whether there exists a word w such that $\Pr_{\mathcal{A}}(w) > \lambda$, is undecidable for probabilistic automata with only two letters. It is easy to adapt the proof to show undecidability of the emptiness problem with fixed threshold $\frac{1}{2}$.

- The state a (resp. b) transitions with uniform probability to its three successors a, b and \$, that is, $\varphi(s,\mathsf{m})(a) = \varphi(s,\mathsf{m})(b) = \varphi(s,\mathsf{m})(\$) = \frac{1}{3}$ for $s \in \{a,b\}$.
- The state \$ has two actions m_x and m_y; the action m_x goes with probability 1 to x and the action m_y goes with probability 1 to y. That is, $\varphi(\$,\mathsf{m}_x)(x) = \varphi(\$,\mathsf{m}_y)(y) = 1$.
- The states x and y are sink states, that is, $\varphi(s,\mathsf{m})(s) = 1$ for $s \in \{x,y\}$.

Each of the states is labelled with its name, that is, $\ell(s) = s$ for $s \in \{a,b,\$,x,y\}$. This sub-MDP "is almost" an MC, in the sense that a strategy α does not influence its behaviour until eventually a transition to x or y is taken. Since a, b, x and y have only one available action, we may omit the default action m in the paths that contain m only. For example, we may write $s_1ab\$$ to represent the path $s_1\mathsf{m}a\mathsf{m}b\mathsf{m}\$$.

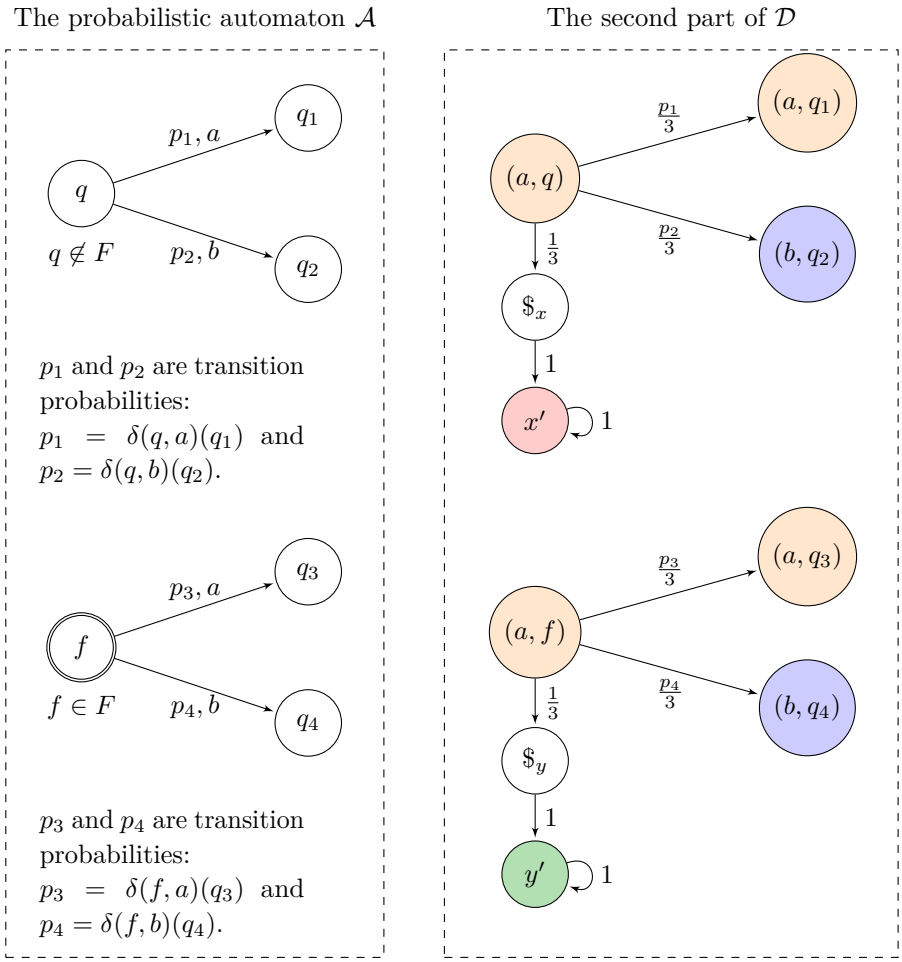


Figure 5 The second part of the MDP $\mathcal{D}$ is an LMC, constructed from the probabilistic automaton $\mathcal{A}$. The default deterministic action m for all states is omitted. The state (b,q) in the MDP $\mathcal{D}$, where $q \in Q$, has the same transitions as the state (a,q); it is labelled with b.

The other part of $\mathcal{D}$ is an LMC constructed from $\mathcal{A}$ as follows; see Figure 5. The set of states is $(L \times Q) \cup \{\$_x, \$_y, x', y'\}$. The state s_2 is defined to be (a, q_0).

We describe the transitions of the LMC using the transition function δ of $\mathcal{A}$. Consider a letter $\sigma \in L$ and a state $q \in Q$. The state (σ,q) with probability $\frac{1}{3}$ simulates the probabilistic automaton $\mathcal{A}$ reading the letter a, and with probability $\frac{1}{3}$ simulates the probabilistic automaton $\mathcal{A}$ reading the letter b. That is, $\varphi\big((\sigma,q),\mathsf{m}\big)\big((a,q')\big) = \frac{1}{3}\delta(q,a)(q')$ and $\varphi\big((\sigma,q),\mathsf{m}\big)\big((b,q')\big) = \frac{1}{3}\delta(q,b)(q')$.

For the remaining probability of $\frac{1}{3}$, we distinguish the following two cases:

- If $q \notin F$, the state (σ, q) transitions to $\$_x$ with probability $\frac{1}{3}$, that is, $\varphi\big((\sigma,q),\mathsf{m}\big)(\$_x) = \frac{1}{3}$.
- Otherwise, if $q \in F$, the state (σ, q) transitions to $\$_y$ with probability $\frac{1}{3}$, that is, $\varphi\big((\sigma,q),\mathsf{m}\big)(\$_y) = \frac{1}{3}$.

The state $\$_x$ (resp. $\$_y$) transitions with probability one to the sink state x' (resp. y'). That is, $\varphi(\$_x,\mathsf{m})(x') = \varphi(\$_y,\mathsf{m})(y') = \varphi(x',\mathsf{m})(x') = \varphi(y',\mathsf{m})(y') = 1$.

A state $(\sigma, q) \in L \times Q$ is labelled with σ. The states $\$_x$ and $\$_y$ are labelled with \$. The states x' and y' are labelled with x and y, respectively.

Given a general strategy α, the next lemma expresses the distance between s_1 and s_2 in terms of α and $\Pr_{\mathcal{A}}$. The proof is technical and can be found in [18, A.2].

▶ **Lemma 10.** *For any general strategy α, we have*

$$d_\alpha(s_1,s_2) = \sum_{w\in L^*} \frac{1}{3^{|w|+1}}\big((1-\Pr_{\mathcal{A}}(w))\alpha(s_1w\$)(\mathsf{m}_y) + \Pr_{\mathcal{A}}(w)\alpha(s_1w\$)(\mathsf{m}_x)\big).$$

Using Lemma 10, we prove the main theorem of this section:

▶ **Theorem 11.** *The general distance minimisation problem is undecidable.*

Proof. We reduce from the emptiness problem for probabilistic automata. Let $\mathcal{A} = \langle Q, q_0, L, \delta, F\rangle$ be a probabilistic automaton; without loss of generality we assume that $q_0 \notin F$ and $L = \{a, b\}$. Let $\mathcal{D}$ be the MDP constructed from $\mathcal{A}$ shown in Figures 4 and 5.

Let α_x be the memoryless strategy that chooses the action m_x whenever it is in state \$, that is, $\alpha_x(s_1w\$) = \mathbf{1}_{\mathsf{m}_x}$ for all $w \in L^*$. Let θ be the distance between s_1 and s_2 in the LMC $\mathcal{D}(\alpha_x)$. It can be computed in polynomial time [5]. We show in [18, A.3]that there is a word $w \in L^*$ such that $\Pr_{\mathcal{A}}(w) > \frac{1}{2}$ ($\mathcal{A}$ is nonempty) if and only if there is a general strategy α such that $d_\alpha(s_1, s_2) < \theta$ in the induced LMC. ◀

6 General Strategies: Distance Less Than One

In this section, we consider the distance less than one problem which, given an MDP and two states, asks whether there is a general strategy such that the two states have probabilistic bisimilarity distance less than one in the LMC induced by the general strategy. The challenge here is that general strategies induce, in general, LMCs with infinitely many states.

We show that the distance less than one problem is EXPTIME-complete. We prove the upper and lower bound in Sections 6.1 and 6.2, respectively.

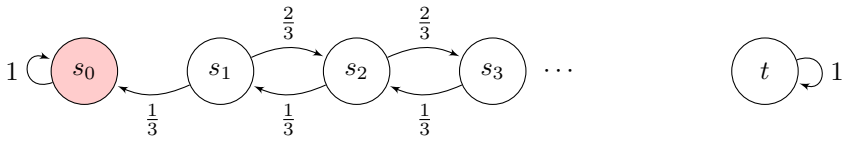
6.1 Membership in EXPTIME

Let $\mathcal{M} = \langle S, L, \tau, \ell\rangle$ be a (possibly infinite) LMC. We partition the set S^2 of state pairs into

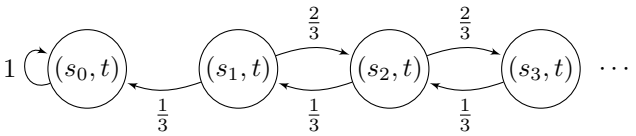
$$\begin{aligned} S^2_0 &= \{\,(s,t) \in S^2 \mid s \sim t\,\}\\ S^2_1 &= \{\,(s,t) \in S^2 \mid \ell(s) \neq \ell(t)\,\}\\ S^2_? &= S^2 \setminus (S^2_0 \cup S^2_1)\,. \end{aligned}$$

We call $T : S^2_? \to \mathrm{Distr}(S^2)$ a *policy* for the LMC if for all $(s,t) \in S^2_?$ we have $T(s,t) \in \Omega(\tau(s),\tau(t))$. We write $\mathcal{T}$ for the set of policies. Given a policy $T \in \mathcal{T}$, the Markov chain $\mathcal{C}^T_{\mathcal{M}} = \langle S^2, \tau'\rangle$ induced by T is defined by

$$\begin{aligned} \tau'\big((u,v)\big)\big((u,v)\big) &= 1 && \text{if } (u,v) \in S^2_0 \cup S^2_1;\\ \tau'\big((u,v)\big)\big((x,y)\big) &= T(u,v)(x,y) && \text{otherwise.} \end{aligned}$$



(a) An infinite LMC $\mathcal{M}$.



(b) The Markov chain $\mathcal{C}^T_{\mathcal{M}}$.

Figure 6 (a) An infinite state LMC $\mathcal{M}$ with an infinite state space $S = \{s_i \mid i \in \{0,1,2,\ldots\}\} \cup \{t\}$. All states have the same label except s_0. The states s_0 and t are sink states, that is, $\tau(s_0)(s_0) = \tau(t)(t) = 1$. Each s_i where $i \in \{1,2,\ldots\}$ transitions to s_{i-1} with probability $\frac{1}{3}$ and s_{i+1} with probability $\frac{2}{3}$. (b) The Markov chain $\mathcal{C}^T_{\mathcal{M}}$ induced by an arbitrary policy T in which only the states reachable from (s_1,t) are shown. The shown part of $\mathcal{C}^T_{\mathcal{M}}$ is the same for every policy T.

For $(s,t) \in S^2$ and a set of state pairs $Z \subseteq S^2$ we write $\mathcal{R}^T_{\mathcal{M}}((s,t),Z) \in [0,1]$ for the probability that in the Markov chain $\mathcal{C}^T_{\mathcal{M}}$ the state (s,t) reaches a state $(u,v) \in Z$.

By [31, Theorem 4, Proposition 5], the following proposition holds.

▶ **Proposition 12.** *Let $\mathcal{M} = \langle S, L, \tau, \ell\rangle$ be a* finite *LMC and $s,t \in S$. We have $d(s,t) < 1$ if and only if there exists a policy T such that $\mathcal{R}^T_{\mathcal{M}}((s,t), S^2_0) > 0$.*

The "only if" direction of Proposition 12 does not generally hold for LMCs with infinite state space, as the following example shows.

▶ **Example 13.** Consider the LMC $\mathcal{M}$ in Figure 6a. Let T be an arbitrary policy for $\mathcal{M}$. We have $T(s_i,t)(s_{i-1},t) = \frac{1}{3}$ and $T(s_i,t)(s_{i+1},t) = \frac{2}{3}$ for all $i \in \{1,2,\ldots\}$. The Markov chain $\mathcal{C}^T_{\mathcal{M}}$ induced by T is shown in Figure 6b; we only show the states that are reachable from (s_1,t). The shown part of $\mathcal{C}^T_{\mathcal{M}}$ is the same for every policy.

We have $d(s_i,t) = \frac{1}{2^i}$ for all $i \in \{0,1,2,\ldots\}$. In the Markov chain $\mathcal{C}^T_{\mathcal{M}}$, all state pairs that (s_1,t) can reach have distances greater than zero: for all $i \in \{1,2,\ldots\}$ the pair (s_1,t) can reach (s_i,t) and we have $d(s_i,t) = \frac{1}{2^i} > 0$. ⌟

The following theorem follows from [30, Theorem 6.1.7] for LMCs with finite state space. The same proof, see [18, A.4], works for LMCs with infinite state space.

▶ **Theorem 14.** *Let $\mathcal{M} = \langle S, L, \tau, \ell\rangle$ be an LMC. There is a policy $T \in \mathcal{T}$ such that we have*

$$d(s,t) \;=\; \mathcal{R}^T_{\mathcal{M}}((s,t),S^2_1) \;\leq\; \mathcal{R}^{T'}_{\mathcal{M}}((s,t),S^2_1) \quad \text{for all } (s,t) \in S^2 \text{ and all } T' \in \mathcal{T}.$$

In short, $d = \min_{T \in \mathcal{T}} \mathcal{R}^T_{\mathcal{M}}(\cdot, S^2_1)$.

The following corollary of Theorem 14 is similar to Proposition 12 but holds even for infinite-state LMCs.

▶ **Corollary 15.** *Let $\mathcal{M} = \langle S, L, \tau, \ell\rangle$ be an LMC and $s,t \in S$. We have $d(s,t) < 1$ if and only if there exists a policy T such that $\mathcal{R}^T_{\mathcal{M}}((s,t), S^2_1) < 1$. In particular, if there is a policy T with $\mathcal{R}^T_{\mathcal{M}}((s,t), S^2_0) > 0$ then $d(s,t) < 1$.*

Corollary 15 falls short of an "if and only if" connection between distance less than one and bisimilarity. Indeed, as we have seen, Proposition 12 does not always hold in infinite-state LMCs. However, the key technical insight of this section is that a version of Proposition 12 holds for (finite-state) MDPs and general strategies. More precisely, the following proposition characterises the existence of a strategy such that the distance is less than one.

▶ **Proposition 16.** *Let $\mathcal{D} = \langle S, Act, L, \varphi, \ell\rangle$ be an MDP, and let $s, t \in S$. There exists a strategy α'' with $d_{\mathcal{D}(\alpha'')}(s,t) < 1$ if and only if there are strategies α, α', a policy T for the LMC $\mathcal{D}(\alpha)$, two states $u, v \in S$ and two paths $\rho_1, \rho_2 \in \mathsf{Paths}(\mathcal{D})$ with $u = \mathsf{last}(\rho_1)$ and $v = \mathsf{last}(\rho_2)$, such that $\mathcal{R}^T_{\mathcal{D}(\alpha)}((s,t), \{(\rho_1,\rho_2)\}) > 0$ and u and v are probabilistically bisimilar in the LMC $\mathcal{D}(\alpha')$.*

The more difficult direction of the proof is the "only if" direction. It is based on Lévy's zero-one law, several applications of the Bolzano-Weierstrass theorem, and a characterisation of probabilistic bisimilarity in MDPs in terms of an "attacker-defender" game defined [17, Section 3.1].

The starting point of the proof of Proposition 16 is the following statement, which follows from Theorem 14 and Corollary 15 using Lévy's zero-one law.

▶ **Corollary 17.** *Let $\mathcal{M} = \langle S, L, \tau, \ell\rangle$ be an LMC and $s, t \in S$ with $d(s,t) < 1$. There exists a policy T such that for all $\varepsilon > 0$ there is $(u,v) \in S^2$ with $d(u,v) \le \varepsilon$ and $\mathcal{R}^T_{\mathcal{M}}((s,t), \{(u,v)\}) > 0$.*

Proof. Let $\mathcal{M} = \langle S, L, \tau, \ell\rangle$ be an LMC and $s, t \in S$ with $d(s,t) < 1$. By Corollary 15 there exists a policy T such that $\mathcal{R}^T_{\mathcal{M}}((s,t), S^2_1) < 1$. By Lévy's zero-one law, the probability in $\mathcal{C}^T_{\mathcal{M}}$ is one that a random run $(s_0,t_0)(s_1,t_1)\dots$ started from $(s_0,t_0) = (s,t)$ satisfies one of the following conditions:

1. the sequence $\mathcal{R}^T_{\mathcal{M}}((s_0,t_0), S^2_1), \mathcal{R}^T_{\mathcal{M}}((s_1,t_1), S^2_1), \dots$ converges to 1 and S^2_1 is reached;
2. the sequence $\mathcal{R}^T_{\mathcal{M}}((s_0,t_0), S^2_1), \mathcal{R}^T_{\mathcal{M}}((s_1,t_1), S^2_1), \dots$ converges to 0 and S^2_1 is not reached.

Event 1 can be equivalently characterised by saying that S^2_1 is reached. Since $\mathcal{R}^T_{\mathcal{M}}((s,t), S^2_1) < 1$, Event 2 happens with a positive probability. It follows that in $\mathcal{C}^T_{\mathcal{M}}$ there exists a run $(s_0,t_0)(s_1,t_1)\dots$ started from $(s_0,t_0) = (s,t)$ such that $\mathcal{R}^T_{\mathcal{M}}((s_0,t_0), S^2_1), \mathcal{R}^T_{\mathcal{M}}((s_1,t_1), S^2_1), \dots$ converges to 0. Let $\varepsilon > 0$. Then there exists $(u,v) \in S^2$ such that $\mathcal{R}^T_{\mathcal{M}}((u,v), S^2_1) \le \varepsilon$ and $\mathcal{R}^T_{\mathcal{M}}((s,t), \{(u,v)\}) > 0$. By Theorem 14 it follows that $d(u,v) \le \varepsilon$. ◀

▶ **Example 18.** Consider again Example 13. We have $d(s_1,t) = \frac{1}{2}$. Corollary 17 asserts that there is a policy T such that for all $\varepsilon > 0$, in $\mathcal{C}^T_{\mathcal{M}}$ the pair (s_1,t) can reach $(u,v) \in S^2$ with $d(u,v) \le \varepsilon$. Indeed, take an arbitrary policy T. Given any $\varepsilon > 0$ choose i with $\frac{1}{2^i} \le \varepsilon$. Then (s_1,t) can reach (s_i,t) and $d(s_i,t) = \frac{1}{2^i} \le \varepsilon$. ⌟

See [18, A.5] for the rest of the proof of Proposition 16. Proposition 16 is the key to proving the following result.

▶ **Theorem 19.** *The distance less than one problem is in EXPTIME.*

Proof. Let $\langle S, Act, L, \varphi, \ell\rangle$ be an MDP. Abusing the notation from the beginning of Section 6.1, let us define

$$\begin{aligned} S^2_0 &= \{\,(s,t) \in S^2 \mid \exists\, \alpha' \text{ such that } s,t \text{ are probabilistically bisimilar in } \mathcal{D}(\alpha')\,\} \\ S^2_1 &= \{\,(s,t) \in S^2 \mid \ell(s) \neq \ell(t)\,\} \\ S^2_? &= S^2 \setminus (S^2_0 \cup S^2_1)\,. \end{aligned}$$

By [17, Theorem 7] the set S^2_0 can be computed in exponential time. Consider the elements of S^2 as vertices of a directed graph with set of edges

$$E := \{(z,z) \mid z \in S_0^2 \cup S_1^2\} \cup \left\{((s_1,s_2),(t_1,t_2)) \in S_?^2 \times S^2 \mid \forall\, i \in \{1,2\}\ \exists\, \mathsf{m}_i \in Act(s_i) : \mathrm{support}(\varphi(s_i,\mathsf{m}_i)) \ni t_i\right\}.$$

After S_0^2 has been computed (in exponential time), the directed graph $G := (S^2, E)$ can be computed in polynomial time, and given two states $s, t \in S$, it can be checked in polynomial time if S_0^2 can be reached from (s,t) in G. It follows from Proposition 16 that this is the case if and only if there exists a strategy α'' with $d_{\mathcal{D}(\alpha'')}(s,t) < 1$. ◀

6.2 EXPTIME-Hardness

Given an MDP and two (initial) states, the *bisimilarity problem* asks whether there is a general strategy such that the two states are probabilistically bisimilar in the induced LMC. The bisimilarity problem was shown EXPTIME-complete in [17, Theorem 7]. We show in [18, A.6] that it can be reduced to the distance less than one problem. This gives us the following theorem.

▶ **Theorem 20.** *The distance less than one problem is EXPTIME-hard.*

Together with Theorem 19 we obtain:

▶ **Corollary 21.** *The distance less than one problem is EXPTIME-complete.*

7 Conclusion

Motivated by probabilistic noninterference, a security notion, we have settled the decidability and complexity of the most natural bisimilarity distance minimisation problems of MDPs under memoryless and general strategies.

Specifically, we have proved that the distance minimisation problem for memoryless strategies is $\exists\mathbb{R}$-complete (which implies, in particular, that it is NP-hard and in PSPACE). In contrast, we have shown that the distance minimisation problem for general strategies is undecidable, reducing from the emptiness problem for probabilistic automata.

We have also shown that it is EXPTIME-complete to decide if there are general strategies to make the probabilistic bisimilarity distance less than one. This extends a result from [17] that the bisimilarity equivalence problem under general strategies is EXPTIME-complete. The key technical link we need here is natural but nontrivial to establish under general strategies: if there are general strategies such that two states have distance less than one, these two states can reach another pair of states which can be made probabilistic bisimilar.

Distance *maximisation* problems also relate to probabilistic noninterference, but in terms of antagonistic schedulers wanting to maximise the information leakage. The decidability and complexity of several distance maximisation problems in MDPs is still open, including the distance equals one problem for general strategies.

References

1 Christel Baier. Polynomial time algorithms for testing probabilistic bisimulation and simulation. In Rajeev Alur and Thomas A. Henzinger, editors, *Computer Aided Verification*, pages 50–61, Berlin, Heidelberg, 1996. Springer Berlin Heidelberg.

2 Patrick Billingsley. *Probability and measure*. Wiley Series in Probability and Statistics. Wiley, New York, NY, USA, 3rd edition, 1995.

3 Vincent D. Blondel and Vincent Canterini. Undecidable problems for probabilistic automata of fixed dimension. *Theory Comput. Syst.*, 36(3):231–245, 2003. doi:10.1007/S00224-003-1061-2.

4 John Canny. Some algebraic and geometric computations in PSPACE. In *STOC*, pages 460–467, 1988.

5 Di Chen, Franck van Breugel, and James Worrell. On the complexity of computing probabilistic bisimilarity. In Lars Birkedal, editor, *Foundations of Software Science and Computational Structures - 15th International Conference, FOSSACS 2012, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2012, Tallinn, Estonia, March 24 - April 1, 2012. Proceedings*, volume 7213 of *Lecture Notes in Computer Science*, pages 437–451. Springer, 2012. doi:10.1007/978-3-642-28729-9_29.

6 Benoît Delahaye. Consistency for parametric interval Markov chains. In Étienne André and Goran Frehse, editors, *2nd International Workshop on Synthesis of Complex Parameters, SynCoP 2015, April 11, 2015, London, United Kingdom*, volume 44 of *OASICS*, pages 17–32. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2015.

7 Salem Derisavi, Holger Hermanns, and William H. Sanders. Optimal state-space lumping in Markov chains. *Inf. Process. Lett.*, 87(6):309–315, 2003.

8 Josée Desharnais, Vineet Gupta, Radha Jagadeesan, and Prakash Panangaden. Metrics for labeled Markov systems. In Jos Baeten and Sjouke Mauw, editors, *Proceedings of the 10th International Conference on Concurrency Theory*, volume 1664 of *Lecture Notes in Computer Science*, pages 258–273, Eindhoven, The Netherlands, August 1999. Springer-Verlag.

9 Laurent Doyen, Thomas A. Henzinger, and Jean-François Raskin. Equivalence of labeled Markov chains. *Int. J. Found. Comput. Sci.*, 19(3):549–563, 2008. doi:10.1142/S0129054108005814.

10 Nathanaël Fijalkow. Undecidability results for probabilistic automata. *ACM SIGLOG News*, 4(4):10–17, 2017. doi:10.1145/3157831.3157833.

11 Ernst Moritz Hahn, Holger Hermanns, and Lijun Zhang. Probabilistic reachability for parametric Markov models. *Int. J. Softw. Tools Technol. Transf.*, 13(1):3–19, 2011.

12 Christian Hensel, Sebastian Junges, Joost-Pieter Katoen, Tim Quatmann, and Matthias Volk. The probabilistic model checker Storm, 2020. arXiv:arXiv:2002.07080.

13 James W. Gray III. Probabilistic interference. In *Proceedings of the 1990 IEEE Symposium on Security and Privacy, Oakland, California, USA, May 7-9, 1990*, pages 170–179. IEEE Computer Society, 1990. doi:10.1109/RISP.1990.63848.

14 Bengt Jonsson and Kim Guldstrand Larsen. Specification and refinement of probabilistic processes. In *Proceedings of the Sixth Annual Symposium on Logic in Computer Science (LICS '91), Amsterdam, The Netherlands, July 15-18, 1991*, pages 266–277. IEEE Computer Society, 1991.

15 John G. Kemeny and J. Laurie Snell. *Finite Markov Chains*. Van Nostrand, 1960.

16 Stefan Kiefer and Qiyi Tang. Comparing labelled Markov decision processes. In Nitin Saxena and Sunil Simon, editors, *40th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS*, volume 182 of *LIPIcs*, pages 49:1–49:16. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.FSTTCS.2020.49.

17 Stefan Kiefer and Qiyi Tang. Strategies for MDP Bisimilarity Equivalence and Inequivalence. In Bartek Klin, Sławomir Lasota, and Anca Muscholl, editors, *33rd International Conference on Concurrency Theory (CONCUR 2022)*, volume 243 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 32:1–32:22, Dagstuhl, Germany, 2022. Schloss Dagstuhl – Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.CONCUR.2022.32.

18 Stefan Kiefer and Qiyi Tang. Minimising the probabilistic bisimilarity distance, 2024. arXiv:2406.19830.

19 Marta Kwiatkowska, Gethin Norman, and David Parker. PRISM 4.0: Verification of probabilistic real-time systems. In G. Gopalakrishnan and S. Qadeer, editors, *Proc. 23rd International Conference on Computer Aided Verification (CAV'11)*, volume 6806 of *LNCS*, pages 585–591. Springer, 2011.

20 Kim Guldstrand Larsen and Arne Skou. Bisimulation through probabilistic testing. *Inf. Comput.*, 94(1):1–28, 1991.

21 Azaria Paz. *Introduction to probabilistic automata*. Academic Press, 2014.

22 Andrei Popescu, Johannes Hölzl, and Tobias Nipkow. Formalizing probabilistic noninterference. In Georges Gonthier and Michael Norrish, editors, *Certified Programs and Proofs - Third International Conference*, volume 8307 of *Lecture Notes in Computer Science*, pages 259–275. Springer, 2013. doi:10.1007/978-3-319-03545-1_17.

23 James Renegar. On the computational complexity and geometry of the first-order theory of the reals. Parts I–III. *Journal of Symbolic Computation*, 13(3):255–352, 1992.

24 Peter Y. A. Ryan, John D. McLean, Jonathan K. Millen, and Virgil D. Gligor. Non-interference: Who needs it? In *14th IEEE Computer Security Foundations Workshop (CSFW-14 2001), 11-13 June 2001, Cape Breton, Nova Scotia, Canada*, pages 237–238. IEEE Computer Society, 2001. doi:10.1109/CSFW.2001.930149.

25 Andrei Sabelfeld and David Sands. Probabilistic noninterference for multi-threaded programs. In *Proceedings of the 13th IEEE Computer Security Foundations Workshop*, pages 200–214. IEEE Computer Society, 2000. doi:10.1109/CSFW.2000.856937.

26 Marcus Schaefer. Realizability of graphs and linkages. In *Thirty Essays on Geometric Graph Theory*, pages 461–482. Springer, 2012.

27 Marcus Schaefer and Daniel Stefankovic. Fixed points, Nash equilibria, and the existential theory of the reals. *Theory Comput. Syst.*, 60(2):172–193, 2017. doi:10.1007/s00224-015-9662-0.

28 Marcel-Paul Schützenberger. On the definition of a family of automata. *Information and Control*, 4:245–270, 1961.

29 Geoffrey Smith. Probabilistic noninterference through weak probabilistic bisimulation. In *16th IEEE Computer Security Foundations Workshop (CSFW-16 2003)*, pages 3–13. IEEE Computer Society, 2003. doi:10.1109/CSFW.2003.1212701.

30 Qiyi Tang. *Computing Probabilistic Bisimilarity Distances*. Phd thesis, York University, Toronto, September 2018. Available at https://yorkspace.library.yorku.ca/items/7640b6ad-edb3-4e60-8f09-3db33c817061.

31 Qiyi Tang and Franck van Breugel. Deciding probabilistic bisimilarity distance one for labelled Markov chains. In Hana Chockler and Georg Weissenbacher, editors, *Proceedings of the 30th International Conference on Computer Aided Verification*, volume 10981 of *Lecture Notes in Computer Science*, pages 681–699, Oxford, UK, July 2018. Springer-Verlag. doi:10.1007/978-3-319-96145-3_39.

32 Wen-Guey Tzeng. A polynomial-time algorithm for the equivalence of probabilistic automata. *SIAM Journal on Computing*, 21(2):216–227, 1992.

33 Antti Valmari and Giuliana Franceschinis. Simple $O(m \log n)$ time Markov chain lumping. In Javier Esparza and Rupak Majumdar, editors, *Tools and Algorithms for the Construction and Analysis of Systems, 16th International Conference, TACAS 2010, Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2010, Paphos, Cyprus, March 20-28, 2010. Proceedings*, volume 6015 of *Lecture Notes in Computer Science*, pages 38–52. Springer, 2010.

34 Franck van Breugel and James Worrell. Towards quantitative verification of probabilistic transition systems. In Fernando Orejas, Paul G. Spirakis, and Jan van Leeuwen, editors, *Automata, Languages and Programming, 28th International Colloquium, ICALP 2001, Crete, Greece, July 8-12, 2001, Proceedings*, volume 2076 of *Lecture Notes in Computer Science*, pages 421–432. Springer, 2001. doi:10.1007/3-540-48224-5_35.

35 Tobias Winkler, Sebastian Junges, Guillermo A. Pérez, and Joost-Pieter Katoen. On the complexity of reachability in parametric Markov decision processes. In Wan J. Fokkink and Rob van Glabbeek, editors, *30th International Conference on Concurrency Theory, CONCUR 2019, August 27-30, 2019, Amsterdam, the Netherlands*, volume 140 of *LIPIcs*, pages 14:1–14:17. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPIcs.CONCUR.2019.14.