

Taming denumerable Markov decision processes with decisiveness

Nathalie Bertrand¹, Patricia Bouyer², Thomas Brihaye³, and Paulin Fournier⁴

¹ Univ. Rennes, Inria, CNRS, IRISA – Rennes (France)

² LSV, CNRS & ENS Paris-Saclay, Univ. Paris-Saclay – Cachan (France)

³ Université Mons – Mons (Belgium)

⁴ CLEARSY Systems Engineering – Aix en Provence (France)

Abstract. Decisiveness has proven to be an elegant concept for denumerable Markov chains: it is general enough to encompass several natural classes of denumerable Markov chains, and is a sufficient condition for simple qualitative and approximate quantitative model checking algorithms to exist. In this paper, we explore how to extend the notion of decisiveness to Markov decision processes. Compared to Markov chains, the extra non-determinism can be resolved in an adversarial or cooperative way, yielding two natural notions of decisiveness. We then explore whether these notions yield model checking procedures concerning the infimum and supremum probabilities of reachability properties.

Keywords: Verification · Markov decision processes · Approximation scheme

1 Introduction

Formal methods for systems with random or unknown behaviours call for models with probabilistic aspects, and appropriate automated verification techniques. Perhaps one of the simplest classes of probabilistic models is the one of Markov chains. The verification of finite-state Markov chains has been quite thoroughly studied, several algorithms for their verification appeared in the literature, and are implemented in mature tools such as PRISM [20] and STORM [12].

Denumerable Markov chains. For some systems however, *finite* Markov chains fall short at providing an appropriate modelling formalism, and *infinite* Markov chains must be considered. There are two general directions for the model checking of infinite-state Markov chains. One option is to focus on Markov chains generated in a specific way: for instance when the underlying transition systems is the configuration graph of a lossy channel system [2], a pushdown automaton [19], or a 1-counter system [13]. In this case, *ad hoc* model checking techniques have been developed for the qualitative and quantitative analysis. The second option is to establish general criteria on denumerable Markov chains that are sufficient conditions for the qualitative and quantitative model checking to be feasible. Abdulla *et al.* explored this direction, and proposed the elegant

notion of *decisive* Markov chains [1]. Intuitively, decisive denumerable Markov chains enjoy some nice properties of finite-state Markov chains. For instance, if a state is continuously reachable, then it will almost-surely be reached. Precisely, a Markov chain is decisive (with respect to a target state s_{\odot} , from a given initial state s_0) if almost all runs from s_0 either reach s_{\odot} , or end in states from which s_{\odot} is no longer reachable. Assuming decisiveness, the qualitative model checking of reachability properties reduces –as in the finite case– to simple graph analyses. Moreover, one can approximate the probability of reachability properties up to any desired error. A stronger property for denumerable Markov chains is the existence of a finite attractor, *i.e.* a finite set of states that is reached almost-surely from any state of the Markov chain. Sufficient conditions for this finite attractor property are given in [4].

Markov decision processes. Purely probabilistic models are too limited to represent features such as, *e.g.* the lack of any assumption regarding scheduling policies or relative speeds (in concurrent systems), or the lack of any information regarding values that have been abstracted away (in abstract models), or the latitude left for later implementation decisions (in early designs). In such situations, it is not desirable to assume the choices to be resolved probabilistically, and nondeterminism is needed. Markov decision processes (MDPs) do encompass nondeterminism and probabilities. They can be seen as an extension of Markov chains with nondeterministic choices. In MDPs, the nondeterminism is resolved by a *scheduler*, which can either be angelic or demonic, so that for a given property it is relevant to consider both the infimum and supremum probabilities, when ranging over all schedulers.

Similarly to the purely probabilistic case, one can either opt for *ad hoc* model checking algorithms for classes of infinite-state MDPs, or derive generic results under appropriate assumptions. In the first scenario, one can mention MDPs which are generated by lossy channel systems [2,5], with nondeterministic action choices and probabilistic message losses. Up to our knowledge, only qualitative verification algorithms –based on the finite-attractor property– have been developed. In particular, the existence of a scheduler that ensures a reachability property with probability 1 (resp. with positive probability) is decidable [5]. However, the existence of a scheduler ensuring a Büchi objective with positive probability is undecidable [2]. As for the second direction, general denumerable MDPs have been considered recently, with the aim of identifying the memory requirements for optimal (or ε -optimal) schedulers [17,16]. Up to our knowledge, there are however no generic approaches to provide quantitative model checking algorithms. This is the purpose of this paper.

Stochastic games. MDPs can be seen as a particular case of stochastic turn-based games, with a single player. Stochastic turn-based games were mostly studied for finite arenas, due to algorithmic concerns. There are notable exceptions of games with underlying tractable model, for which decidability result exist: recursive concurrent stochastic games [11,14], 1-counter stochastic games [9,10] or lossy channel systems [8,3]. For infinite arenas, general results mostly concern purely

non-algorithmical, aspects, such as determinacy. Still, there are a few algorithmical results for general stochastic turn-based two-player games with reachability objectives [18,11]. Rephrased in terms of MDPs (with a single player), under appropriate assumptions, one can show the existence of pure and positional optimal strategies. Also, the optimal probabilities can be characterized as being the least fixpoint of some natural functional. We recall these results in the paper.

Contributions. In this paper, we precisely address the design of generic algorithms for the quantitative model checking of denumerable MDPs. To do so, first, we build on the seminal work on decisive Markov chains [1] and explore how the notion of decisiveness can be extended to Markov decision processes. We propose two notions of decisiveness, called inf-decisiveness and sup-decisiveness, which differ on whether the resolution of nondeterminism is angelic or demonic. These two notions are both conservative extensions of decisiveness for Markov chains. Second, we provide natural approximation schemes for the optimum (infimum and supremum) probabilities of reachability properties. These schemes provide a nondecreasing sequence of lower bounds, as well as a nonincreasing sequence of upper bounds, for the probability one wishes to compute. We then identify sufficient conditions for the two sequences to converge towards the same limit, thus yielding an approximation algorithm. Third, we show that the decisiveness notions we introduced are sufficient conditions for termination of our approximation schemes. As a consequence, for inf-decisive MDPs, one can approximate the infimum reachability probability up to any error, and for sup-decisive MDPs, one can approximate the supremum reachability up to any error.

2 Preliminaries on Markov decision processes

2.1 Markov decision processes

Definition 1. A Markov decision process (MDP) is a tuple $\mathcal{M} = (S, \text{Act}, P)$ where S is a denumerable set of states, Act is a denumerable set of actions, $P: S \times \text{Act} \times S \rightarrow [0, 1] \cap \mathbb{Q}$ is a probabilistic transition function satisfying $\sum_{s' \in S} P(s, a, s') \in \{0, 1\}$ for all $(s, a) \in S \times \text{Act}$.

The MDP \mathcal{M} is *finite* if S is finite. Given $(s, a) \in S \times \text{Act}$, we say that a is *enabled* at s whenever $\sum_{s' \in S} P(s, a, s') = 1$; otherwise it is said *disabled*. We write $\text{En}(s)$ for the set of actions enabled at s . The MDP \mathcal{M} is said *finitely action-branching* if the number of for every $s \in S$, $\text{En}(s)$ is finite. It is *finitely proba-branching* if for every (s, a) , the support of $P(s, a, -)$ is finite. It is *finitely branching* if it is both finitely action-branching and finitely proba-branching.

A *history* (resp. *path*) in \mathcal{M} is an element $s_0 s_1 s_2 \dots$ of S^+ (resp. S^ω) such that for every relevant i , there is $a_i \in \text{Act}$ such that $P(s_i, a_i, s_{i+1}) > 0$ (in particular a_i is enabled at s_i). We write $\text{Hist}(\mathcal{M})$ for the set of histories in \mathcal{M} and $\text{Paths}(\mathcal{M})$ for the set of paths in \mathcal{M} . We define the length of a history $h = s_0 s_1 \dots s_k$ as $\text{length}(h) = k$, and its last state $\text{last}(h) = s_k$. We sometimes write $h \cdot s$ for a history, to emphasize its last state.

We consider the σ -algebra generated by cylinders in $\text{Paths}(\mathcal{M})$: for $h \in \text{Hist}(\mathcal{M})$ a history, the cylinder generated by h is

$$\text{Cyl}(h) = \{\rho \in \text{Paths}(\mathcal{M}) \mid h \text{ is a prefix of } \rho\}$$

Definition 2. A scheduler in \mathcal{M} is a function $\sigma: \text{Hist}(\mathcal{M}) \rightarrow \text{Dist}(\text{Act})$ which assigns a probability distribution over actions to any history, with the constraint that for every $h \in \text{Hist}(\mathcal{M})$, the support of $\sigma(h)$ is included in the set of enabled actions at $\text{last}(h)$. We write $\text{Sched}(\mathcal{M})$ for the set of schedulers in \mathcal{M} .

We fix a scheduler σ . We say that σ enables a after the history $h = s_0 \cdots s_k$ whenever $\sigma(h)(a) > 0$; we then write $a \in \text{Enabled}_\sigma(h)$. If σ does not depend on histories, i.e. if $\text{last}(h) = \text{last}(h')$ implies $\sigma(h) = \sigma(h')$, then it is called *positional*. If for every h , $\sigma(h)$ is a Dirac probability measure, it is said *pure*. A pure and positional scheduler can alternatively be described as a function $\sigma: S \rightarrow \text{Act}$. We write $\text{Sched}_{\text{pp}}(\mathcal{M})$ for the set of pure positional schedulers in \mathcal{M} , and $\text{Sched}_{\text{ph}}(\mathcal{M})$ for the set of pure (*a priori* not positional, that is, history-dependent) schedulers.

Given a scheduler σ in \mathcal{M} and an initial state $s_0 \in S$, one can define a probability measure $\mathbb{P}_{\mathcal{M}, s_0}^\sigma$ on $\text{Paths}(\mathcal{M})$ inductively as follows:

- $\mathbb{P}_{\mathcal{M}, s_0}^\sigma(\text{Cyl}(s_0)) = 1$;
- if $h = s_0 \cdots s_k \in \text{Hist}(\mathcal{M})$ and $h \cdot s_{k+1} \in \text{Hist}(\mathcal{M})$, then:

$$\mathbb{P}_{\mathcal{M}, s_0}^\sigma(\text{Cyl}(h \cdot s_{k+1})) = \mathbb{P}_{\mathcal{M}, s_0}^\sigma(\text{Cyl}(h)) \cdot \sum_{a \in \text{Enabled}_\sigma(h)} \sigma(h)(a) \cdot \mathbf{P}(s_k, a, s_{k+1})$$

Equivalently, it is the probability measure in the (infinite-state) Markov chain \mathcal{M}_σ induced by scheduler σ on \mathcal{M} .

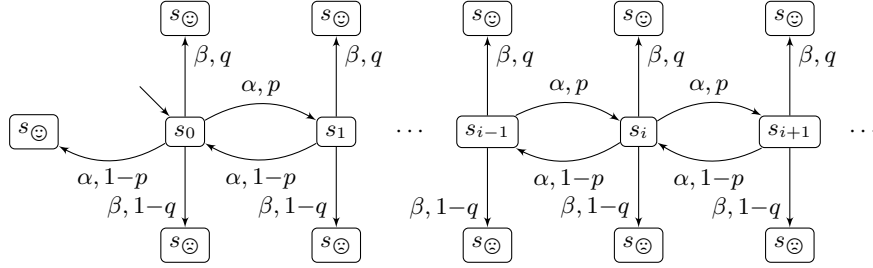


Fig. 1. Example of a finitely branching MDP with infinite state space. For readability, in this picture the absorbing states s_\ominus and s_\ominus are duplicated.

Figure 1 presents an example of MDP, which is finitely branching. Under a scheduler which always selects α , this yields a random walk, which will be diverging if $p > \frac{1}{2}$, that is, the probability not to visit s_\ominus will be positive (say λ_p). In particular, in this case, the infimum probability of reaching s_\ominus will depend on the relative values of q and λ_p . We discuss this example again on page 16.

2.2 Optimum reachability probabilities

Depending on the application, the non-determinism in Markov decision processes can be thought of as adversarial or as cooperative. For the probability of a given property, it thus makes sense to consider on the one hand the infimum and supremum probabilities when ranging over all schedulers.

We describe path properties using the standard LTL operators \mathbf{F} and \mathbf{G} , and their step-bounded variants $\mathbf{F}_{\leq n}$ and $\mathbf{G}_{\leq n}$. Let $\rho = s_0 s_1 \dots \in \text{Paths}(\mathcal{M})$ be a path in \mathcal{M} . If ψ is a state property, the path property $\mathbf{F}\psi$ holds on ρ if there is some index $k \in \mathbb{N}$ such that s_k satisfies ψ . Given $n \in \mathbb{N}$, $\mathbf{F}_{\leq n}$ holds on ρ if there is some index $k \leq n$ such that s_k satisfies ψ . Dually, ρ satisfies $\mathbf{G}\psi$ if all indices $k \in \mathbb{N}$ are such that s_k satisfies ψ ; and ρ satisfies $\mathbf{G}_{\leq n}\psi$ if for all indices $k \leq n$, s_k satisfies ψ . Now, given a path property ϕ , we write $\llbracket \phi \rrbracket_{\mathcal{M}, s_0}$ for the set of paths from s_0 in \mathcal{M} that satisfy ϕ .

In this paper, we focus on the optimization of reachability properties, and thus aim at computing or approximating the following values: given \mathcal{M} an MDP, s_0 an initial state for \mathcal{M} , $T \subseteq S$ a set of target states, and $\text{opt} \in \{\inf, \sup\}$

$$\mathbb{P}_{\mathcal{M}, s_0}^{\text{opt}}(\mathbf{F} T) \stackrel{\text{def}}{=} \text{opt}_{\sigma \in \text{Sched}(\mathcal{M})} \mathbb{P}_{\mathcal{M}, s_0}^{\sigma}(\mathbf{F} T) .$$

Without loss of generality, one can assume that T consists of a single absorbing state (*i.e.* with no enabled actions), that we denote s_{\odot} in the sequel.

For finite MDPs, the computation of these values for $\text{opt} = \inf$ and $\text{opt} = \sup$ is well-known (see *e.g.* [6, Chap. 10]). It reduces to solving a linear program (of linear size), resulting in a polynomial time algorithm. Moreover, the infimum and supremum values are attained by pure and positional schedulers.

Alternatively to solving a linear program, value iteration techniques can also be used and often turn out to be more efficient in practice, see [15]. They rely on a fixed-point characterization of the values $\text{val}_{\mathcal{M}}^{\text{opt}}(s) \stackrel{\text{def}}{=} \mathbb{P}_{\mathcal{M}, s}^{\text{opt}}(\mathbf{F} s_{\odot})$, where $\text{opt} \in \{\inf, \sup\}$. This characterization in fact also holds for finitely action-branching denumerable MDPs, and can even be extended to stochastic turn-based two-player games with reachability objectives [18, 11]. The Bellman functional $\Gamma^{\text{opt}}: [0, 1]^S \rightarrow [0, 1]^S$ is defined as follows. For every $\nu \in [0, 1]^S$ and every $s \in S$:

$$\Gamma^{\text{opt}}(\nu)(s) = \begin{cases} 1 & \text{if } s = s_{\odot} \\ \text{opt}_{\delta \in \text{Dist}(\text{En}(s))} \sum_{a \in \text{Act}} \sum_{s' \in S} \delta(a)(s') \cdot \nu(s') & \text{otherwise} \end{cases}$$

Lemma 1 ([11]). *Let \mathcal{M} be a finitely action-branching MDP. Then, $(\text{val}_{\mathcal{M}}^{\text{opt}}(s))_{s \in S}$ is the least fixed-point of the Bellman equations defined by Γ^{opt} .*

The proof is quite technical and is detailed in [11].

As for the existence of simple optimal schedulers, one can also rephrase for MDPs general results on stochastic reachability games.

Lemma 2 (adapted from [11, 18]). *Let \mathcal{M} be a finitely action-branching MDP, s_0 an initial state, and s_{\odot} a target state.*

1. *there exists $\sigma \in \text{Sched}_{\text{pp}}(\mathcal{M})$ s.t. for all $s \in S$, $\mathbb{P}_{\mathcal{M},s}^\sigma(\mathbf{F} s_\odot) = \mathbb{P}_{\mathcal{M},s}^{\text{inf}}(\mathbf{F} s_\odot)$;*
2. *for $\text{opt} \in \{\text{inf}, \text{sup}\}$, for every $n \in \mathbb{N}$, there exists $\sigma_n \in \text{Sched}_{\text{ph}}(\mathcal{M})$ s.t. $\mathbb{P}_{\mathcal{M},s_0}^\sigma(\mathbf{F}_{\leq n} s_\odot) = \mathbb{P}_{\mathcal{M},s_0}^{\text{opt}}(\mathbf{F}_{\leq n} s_\odot)$.*

Note that the first item of Lemma 2, *i.e.* the existence of optimal pure positional schedulers does not hold for supremum reachability probabilities. Moreover, in case of infimum reachability probabilities, the assumption of finite action-branching is required.

For the second item, the finite action-branching assumption is crucial. Indeed, consider the infinitely action-branching MDP $\mathcal{M} = (\{s, s_\odot, s_\ominus\}, \{\alpha_i \mid i \in \mathbb{N}\}, \mathbf{P})$ with $\mathbf{P}(s, \alpha_i, s_\odot) = 1 - 2^{-k}$ and $\mathbf{P}(s, \alpha_i, s_\ominus) = 2^{-k}$. It is such that there is no pure optimal scheduler that achieves $\mathbb{P}_{\mathcal{M},s}^{\text{sup}}(\mathbf{F}_{\leq 2} s_\odot) = 1$. Also an optimal scheduler for step-bounded reachability can require memory.

Approximation algorithms. Even if characterizations of the values exist (recall the Bellman equations), for infinite MDPs no general algorithm is known to compute $\mathbb{P}_{\mathcal{M},s_0}^{\text{inf}}(\mathbf{F} s_\odot)$ and $\mathbb{P}_{\mathcal{M},s_0}^{\text{sup}}(\mathbf{F} s_\odot)$, or to decide whether these values exceed a threshold. Of course such algorithms would very much depend on the representation of denumerable MDPs.

We first establish an undecidability result for the supremum probability of MDPs given as non-deterministic and probabilistic lossy channel system (NPLCS). These models are built on channel systems, they have probabilistic message losses, and non-deterministic choices between possible read/write actions [7,2].

Theorem 1. *The following decision problem is undecidable:*

- **Input:** $\mathcal{M} = (S, \text{Act}, \mathbf{P})$ an MDP defined by a NPLCS, with $s_0, s_\odot \in S$
- **Output:** *yes iff $\mathbb{P}_{\mathcal{M},s_0}^{\text{sup}}(\mathbf{F} s_\odot) = 1$.*

Whether the supremum equals 1 is also known as the value-1 problem. This undecidability contrasts with the fact that the existence of a scheduler ensuring almost-surely a reachability objective is decidable for NPLCS [5]. We sketch the undecidability proof, which is given with more details in Appendix A.

Proof (Sketch of proof). We reduce the boundedness problem, which is undecidable for lossy channel systems (LCSs) [21]. Given an LCS \mathcal{L} , consider the finitely action-branching MDP \mathcal{M} as represented in Figure 2.

To define \mathcal{M} , we proceed in two steps. First, the semantics of \mathcal{L} can be seen as an MDP $\mathcal{M}(\mathcal{L})$, with state space $S_{\mathcal{L}}$ is the set of *configurations* of \mathcal{L} . A configuration is a pair (q, w) where q is a control state and w describes the channel contents. As for the probabilistic transition function, from every configuration (q, w) , finitely many actions (sendings and receptions) are available. The next configuration depends on the message losses, and we assume the probability distribution for a letter on a channel to be lost to be uniform.

We then embed $\mathcal{M}(\mathcal{L})$ into a bigger MDP \mathcal{M} , whose state space is $S_{\mathcal{L}} \cup \{s_\odot, s_\ominus\}$. The two states s_\odot and s_\ominus are sink states. From any state of $\mathcal{M}(\mathcal{L})$, two extra actions are enabled: **try** and **restart**. Action **restart** is deterministic and

leads to (q_0, ε) the initial configuration of \mathcal{L} . From (q, w) , action **try** leads with probability $\frac{1}{2^{|w|}}$ to s_\ominus and remaining probability $1 - \frac{1}{2^{|w|}}$ to s_\ominus .

We claim that under this construction, \mathcal{M} enjoys the following properties:

- if \mathcal{L} is bounded, there exists $p > 0$ such that $\mathbb{P}_{\mathcal{M}, s_0}^{\sup}(\mathbf{F} s_\ominus) \leq 1 - p < 1$;
- if \mathcal{L} is unbounded, $\mathbb{P}_{\mathcal{M}, s_0}^{\sup}(\mathbf{F} s_\ominus) = 1$.

As such, the constructed MDP is not directly derived from an NPLCS, but it can easily be turned to such an MDP, which allows to derive the expected result. \square

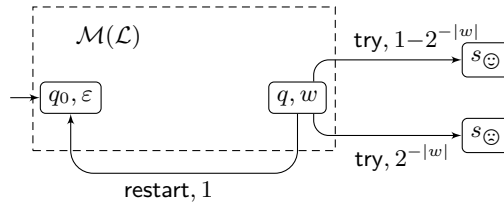


Fig. 2. Undecidability of the value-1 problem.

In this paper, we aim at providing generic *approximation algorithms* for infimum and supremum reachability probabilities in denumerable MDPs.

Definition 3. An approximation algorithm takes as input an MDP \mathcal{M} , an initial state s_0 , a target state s_\ominus , an optimization criterion $\text{opt} \in \{\inf, \sup\}$, and a precision $\varepsilon > 0$, and returns a value v such that $|v - \mathbb{P}_{\mathcal{M}, s_0}^{\text{opt}}(\mathbf{F} s_\ominus)| \leq \varepsilon$.

Our approach is to provide generic *approximation schemes*, defined by two sequences $(r_n^-)_n$ and $(r_n^+)_n$ such that for every n , $r_n^- \leq p \leq r_n^+$, where p is the desired probability. An approximation scheme can be turned into an approximation algorithm for p if for every precision $\varepsilon > 0$, there exists n such that $|r_n^+ - r_n^-| \leq \varepsilon$, by picking any v in the interval $[r_n^-, r_n^+]$.

To prove the correctness or termination of our approximation schemes, we often make use of the following general observation:

Remark 1. Let \mathcal{M} be an MDP, s a state of \mathcal{M} , $\text{opt} \in \{\inf, \sup\}$, and ϕ_1 and ϕ_2 be two path properties. For every scheduler σ , $\mathbb{P}_{\mathcal{M}, s}^\sigma(\phi_1 \vee \phi_2) \leq \mathbb{P}_{\mathcal{M}, s}^\sigma(\phi_1) + \mathbb{P}_{\mathcal{M}, s}^{\sup}(\phi_2)$. Thus:

$$\mathbb{P}_{\mathcal{M}, s}^{\text{opt}}(\phi_1 \vee \phi_2) \leq \mathbb{P}_{\mathcal{M}, s}^{\text{opt}}(\phi_1) + \mathbb{P}_{\mathcal{M}, s}^{\sup}(\phi_2) .$$

2.3 Decisiveness for MDPs

Fix an MDP $\mathcal{M} = (S, \text{Act}, \text{P})$ in the rest of this section and $s_\ominus \in S$ be a target state.

Avoid sets. For Markov chains, Abdulla *et al.* introduced the set of states from which one can no longer reach s_\ominus (denoted $\widetilde{s_\ominus}$) [1]. We generalize this notion to MDPs by defining avoid sets in two flavors, depending on whether one considers infimum or supremum over schedulers. For $\text{opt} \in \{\inf, \sup\}$, we let:

$$\text{Avoid}_{\mathcal{M}}^{\text{opt}}(s_\ominus) = \{s \in S \mid \text{opt}_{\sigma \in \text{Sched}(\mathcal{M})} \mathbb{P}_{\mathcal{M},s}^\sigma(\mathbf{F} s_\ominus) = 0\} .$$

Clearly, $\text{Avoid}_{\mathcal{M}}^{\sup}(s_\ominus) \subseteq \text{Avoid}_{\mathcal{M}}^{\inf}(s_\ominus)$. Note that $\sup_{\sigma \in \text{Sched}(\mathcal{M})} \mathbb{P}_{\mathcal{M},s}^\sigma(\mathbf{F} s_\ominus) = 0$ iff for all $\sigma \in \text{Sched}(\mathcal{M})$, $\mathbb{P}_{\mathcal{M},s}^\sigma(\mathbf{F} s_\ominus) = 0$. In contrast, it may happen that $\inf_{\sigma \in \text{Sched}(\mathcal{M})} \mathbb{P}_{\mathcal{M},s}^\sigma(\mathbf{F} s_\ominus) = 0$, yet there is no $\sigma \in \text{Sched}(\mathcal{M})$ such that $\mathbb{P}_{\mathcal{M},s}^\sigma(\mathbf{F} s_\ominus) = 0$. For instance on the MDP of Figure 3, when choosing action α_i from s_0 , the probability of $\mathbf{F} s_\ominus$ is $\frac{1}{2^i}$. Recall that thanks to Lemma 2, this behaviour requires infinite action-branching.

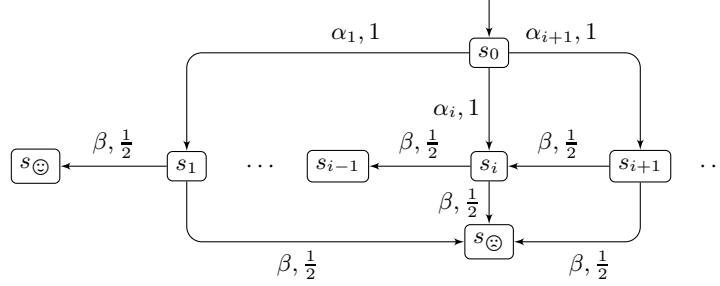


Fig. 3. MDP for which $\mathbb{P}^{\inf}(\mathbf{F} s_\ominus) = 0$, yet for every scheduler σ , $\mathbb{P}^\sigma(\mathbf{F} s_\ominus) > 0$.

Decisiveness properties. We now define two notions of decisiveness for MDPs, that are both conservative extensions of the decisiveness for Markov chains [1].

Definition 4. Let $\mathcal{M} = (S, \text{Act}, \mathbf{P})$ be an MDP, $s_\ominus \in S$ a target state, $s \in S$ an initial state, and $\text{opt} \in \{\inf, \sup\}$. \mathcal{M} is opt -decisive w.r.t. s_\ominus from s whenever

$$\forall \sigma \in \text{Sched}_{\text{pp}}(\mathcal{M}), \mathbb{P}_{\mathcal{M},s}^\sigma(\mathbf{F} s_\ominus \vee \mathbf{F} \text{Avoid}_{\mathcal{M}}^{\text{opt}}(s_\ominus)) = 1 .$$

Since $\text{Avoid}_{\mathcal{M}}^{\sup}(s_\ominus) \subseteq \text{Avoid}_{\mathcal{M}}^{\inf}(s_\ominus)$, sup-decisiveness is a stronger condition than inf-decisiveness.

Example 1. The two MDPs of Figure 4, where the ε_i 's satisfy $\prod_i (1 - \varepsilon_i) > 0$, illustrate the notions of avoid sets and decisiveness. The MDP on the left \mathcal{M}^L is such that $\text{Avoid}_{\mathcal{M}^L}^{\sup}(s_\ominus) = \{s_\ominus\}$, but $\text{Avoid}_{\mathcal{M}^L}^{\inf}(s_\ominus) = S \setminus \{s_\ominus\}$, since from states s_i 's, only the scheduler which always selects α avoids s_\ominus . Thus, \mathcal{M}^L is inf-decisive, but not sup-decisive, from s_0 w.r.t. s_\ominus . The MDP on the right \mathcal{M}^R is such that $\text{Avoid}_{\mathcal{M}^R}^{\inf}(s_\ominus) = \{s_\ominus\}$ since it is not possible to avoid s_\ominus (except from s_\ominus). The MDP \mathcal{M}^R is not inf-decisive (and thus not sup-decisive) from

s_0 w.r.t. s_\ominus , since the scheduler which always selects α avoids s_\ominus and s_\ominus with positive probability $\prod_i (1 - \varepsilon_i)$. We will discuss this example again in Section 5.2.

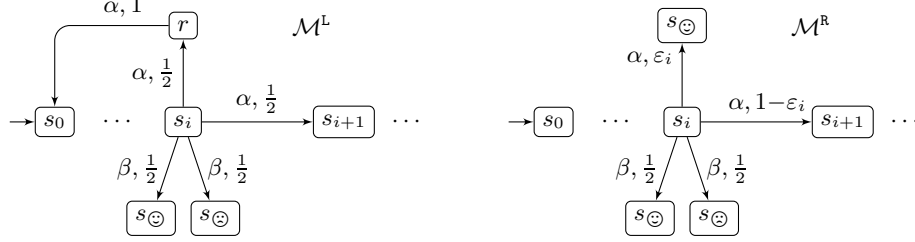


Fig. 4. Left: MDP \mathcal{M}^L with $\text{Avoid}_{\mathcal{M}^L}^{\text{sup}}(s_\ominus) = \{s_\ominus\}$, $\text{Avoid}_{\mathcal{M}^L}^{\text{inf}}(s_\ominus) = S \setminus \{s_\ominus\}$; \mathcal{M}^L is inf-decisive, but not sup-decisive, from s_0 w.r.t. s_\ominus . Right: MDP \mathcal{M}^R with $\text{Avoid}_{\mathcal{M}^R}^{\text{inf}}(s_\ominus) = \{s_\ominus\}$; \mathcal{M}^R is not inf-decisive (and thus not sup-decisive) from s_0 w.r.t. s_\ominus .

Remark that, in particular, all finite MDPs are inf-decisive. Changing the definition by requiring that the condition holds for all history-dependent schedulers, would imply that some finite MDPs would not be inf-decisive.

With our definition however, not all finite MDPs are sup-decisive. For instance, the 2-state MDP with $P(s_0, \alpha, s_0) = 1$ and $P(s_0, \beta, s_\ominus)$, is not sup-decisive from s_0 w.r.t. s_\ominus . For finite MDPs, a characterization of sup-decisiveness is that all end components are terminal; in particular they are atomic, *i.e.* they admit no sub end component.

3 Generic approximation schemes

The objective of this section is to provide generic approximations schemes for optimum reachability probabilities, and to state sufficient conditions under which termination and correctness are guaranteed (thus yielding approximation algorithms). The section starts with a construction that collapses avoid sets.

3.1 Collapsing avoid sets and first approximation scheme

For $\text{opt} \in \{\text{inf}, \text{sup}\}$, from \mathcal{M} we build a new MDP $\mathcal{M}^{\text{opt}} = (S^{\text{opt}}, \text{Act}, P^{\text{opt}})$ by merging states in $\text{Avoid}_{\mathcal{M}}^{\text{opt}}(s_\ominus)$ into s_\ominus^{opt} a fresh absorbing state (see the formal construction in Appendix C.1).

Formally, $\mathcal{M}^{\text{opt}} = (S^{\text{opt}}, \text{Act}, P^{\text{opt}})$ with

- $S^{\text{opt}} = (S \setminus \text{Avoid}_{\mathcal{M}}^{\text{opt}}(s_\ominus)) \cup \{s_\ominus^{\text{opt}}\}$;
- for every $s, s' \in S^{\text{opt}} \setminus \{s_\ominus^{\text{opt}}\}$, for every $a \in \text{Act}$, $P^{\text{opt}}(s, a, s') = P(s, a, s')$;

- for every $s \in S^{\text{opt}} \setminus \{s_{\ominus}^{\text{opt}}\}$, $\mathbb{P}^{\text{opt}}(s, a, s_{\ominus}^{\text{opt}}) = \sum_{s' \in \text{Avoid}_{\mathcal{M}}^{\text{opt}}(s_{\ominus})} \mathbb{P}(s, a, s')$.
- for every $a \in \text{Act}$, $\mathbb{P}^{\text{opt}}(s_{\ominus}^{\text{opt}}, a, s_{\ominus}^{\text{opt}}) = 1$;

In both cases (inf and sup), remark that $s_{\ominus} \in S \cap S^{\text{opt}}$. W.l.o.g. we assume that the initial state is preserved in the collapsed MDP (*i.e.* $s_0 \in S^{\text{opt}} \cap S$); otherwise, by definition of $\text{Avoid}_{\mathcal{M}}^{\text{opt}}(s_{\ominus})$, $\text{opt}_{\sigma} \mathbb{P}_{\mathcal{M},s}^{\sigma}(\mathbf{F} s_{\ominus}) = 0$ and the value to be computed is trivially 0.

Note the following two properties:

- for every $s \in S^{\text{inf}} \setminus \{s_{\ominus}^{\text{inf}}\}$, for every $\sigma \in \text{Sched}(\mathcal{M}^{\text{inf}})$, $\mathbb{P}_{\mathcal{M}^{\text{inf}},s}^{\sigma}(\mathbf{F} s_{\ominus}) > 0$;
- for every $s \in S^{\text{sup}} \setminus \{s_{\ominus}^{\text{sup}}\}$, there is $\sigma \in \text{Sched}(\mathcal{M}^{\text{sup}})$ s.t. $\mathbb{P}_{\mathcal{M}^{\text{sup}},s}^{\sigma}(\mathbf{F} s_{\ominus}) > 0$.

The above constructions collapsing avoid sets preserve optimum probabilities:

Lemma 3. $\mathbb{P}_{\mathcal{M},s_0}^{\text{opt}}(\mathbf{F} s_{\ominus}) = \mathbb{P}_{\mathcal{M}^{\text{opt}},s_0}^{\text{opt}}(\mathbf{F} s_{\ominus})$.

Note that this lemma holds with no prior assumption on \mathcal{M} (neither decisiveness, nor finite-branching).

Proof. We first consider the case $\text{opt} = \text{inf}$. In general, any scheduler σ in \mathcal{M} straightforwardly yields a scheduler σ' in \mathcal{M}^{opt} by mimicking σ until s_{\ominus}^{opt} is reached (if so happens) and then behaves arbitrarily. Clearly enough $\mathbb{P}_{\mathcal{M}^{\text{opt}},s_0}^{\sigma'}(\mathbf{F} s_{\ominus}) \leq \mathbb{P}_{\mathcal{M}^{\text{opt}},s_0}^{\sigma}(\mathbf{F} s_{\ominus})$. Indeed,

$$\begin{aligned} \mathbb{P}_{\mathcal{M}^{\text{opt}},s_0}^{\sigma'}(\mathbf{F} s_{\ominus}) &= \mathbb{P}_{\mathcal{M}^{\text{opt}},s_0}^{\sigma'}(\mathbf{F} s_{\ominus} \wedge \mathbf{G} \neg s_{\ominus}^{\text{opt}}) \\ &= \mathbb{P}_{\mathcal{M}^{\text{opt}},s_0}^{\sigma}(\mathbf{F} s_{\ominus} \wedge \mathbf{G} \neg \text{Avoid}_{\mathcal{M}}^{\text{opt}}(s_{\ominus})) \\ &\leq \mathbb{P}_{\mathcal{M}^{\text{opt}},s_0}^{\sigma}(\mathbf{F} s_{\ominus}) . \end{aligned}$$

Thus, $\mathbb{P}_{\mathcal{M}^{\text{inf}},s_0}^{\text{inf}}(\mathbf{F} s_{\ominus}) \leq \mathbb{P}_{\mathcal{M},s_0}^{\text{inf}}(\mathbf{F} s_{\ominus})$.

To prove the other inequality, for every scheduler σ in \mathcal{M}^{inf} , and every $\varepsilon > 0$, we show that there exists a scheduler σ' in \mathcal{M} with $\mathbb{P}_{\mathcal{M},s_0}^{\sigma'}(\mathbf{F} s_{\ominus}) \leq \mathbb{P}_{\mathcal{M}^{\text{inf}},s_0}^{\sigma}(\mathbf{F} s_{\ominus}) + \varepsilon$. To do so, for every $\varepsilon > 0$ and every state $s \in \text{Avoid}_{\mathcal{M}}^{\text{inf}}(s_{\ominus})$, let σ_{ε}^s be a scheduler achieving at most ε for $\mathbb{P}_{\mathcal{M},s}^{\sigma_{\varepsilon}^s}(\mathbf{F} s_{\ominus})$. This is possible by definition of $\text{Avoid}_{\mathcal{M}}^{\text{inf}}(s_{\ominus})$. Construct scheduler σ' which plays according to σ until the first visit to $\text{Avoid}_{\mathcal{M}}^{\text{inf}}(s_{\ominus})$, and then switches to σ_{ε}^s (where s is the first visited state in $\text{Avoid}_{\mathcal{M}}^{\text{inf}}(s_{\ominus})$). Then, $\mathbb{P}_{\mathcal{M},s_0}^{\sigma'}(\mathbf{F} s_{\ominus}) \leq \mathbb{P}_{\mathcal{M}^{\text{inf}},s_0}^{\sigma}(\mathbf{F} s_{\ominus}) + \varepsilon$ and thus $\mathbb{P}_{\mathcal{M},s_0}^{\text{inf}}(\mathbf{F} s_{\ominus}) \leq \mathbb{P}_{\mathcal{M}^{\text{inf}},s_0}^{\text{inf}}(\mathbf{F} s_{\ominus}) + \varepsilon$. Since this holds for every ε , we conclude that $\mathbb{P}_{\mathcal{M},s_0}^{\text{inf}}(\mathbf{F} s_{\ominus}) \leq \mathbb{P}_{\mathcal{M}^{\text{inf}},s_0}^{\text{inf}}(\mathbf{F} s_{\ominus})$.

Let us now consider the case $\text{opt} = \text{sup}$. Observe that any path in \mathcal{M} that reaches s_{\ominus} satisfies $\mathbf{G} \neg \text{Avoid}_{\mathcal{M}}^{\text{sup}}(s_{\ominus})$. Indeed, by definition of $\text{Avoid}_{\mathcal{M}}^{\text{sup}}(s_{\ominus})$, $\mathbf{F} \text{Avoid}_{\mathcal{M}}^{\text{sup}}(s_{\ominus})$ implies $\mathbf{G} \neg s_{\ominus}$. Therefore for any scheduler σ in \mathcal{M} , $\mathbb{P}_{\mathcal{M}}^{\sigma}(\mathbf{F} s_{\ominus}) = \mathbb{P}_{\mathcal{M}}^{\sigma}(\mathbf{F} s_{\ominus} \wedge \mathbf{G} \neg \text{Avoid}_{\mathcal{M}}^{\text{sup}}(s_{\ominus}))$. We use this fact to prove both inequalities on supremum probability values.

Pick a scheduler σ in \mathcal{M}^{sup} , and extend it to \mathcal{M} into a scheduler σ' that mimicks σ , unless $\text{Avoid}_{\mathcal{M}}^{\text{sup}}(s_{\ominus})$ is reached in which case it behaves arbitrarily. We have:

$$\begin{aligned}\mathbb{P}_{\mathcal{M},s_0}^{\sigma'}(\mathbf{F} s_{\ominus}) &= \mathbb{P}_{\mathcal{M},s_0}^{\sigma'}(\mathbf{F} s_{\ominus} \wedge \mathbf{G} \neg \text{Avoid}_{\mathcal{M}}^{\text{sup}}(s_{\ominus})) \\ &= \mathbb{P}_{\mathcal{M}^{\text{sup}},s_0}^{\sigma}(\mathbf{F} s_{\ominus} \wedge \mathbf{G} \neg s_{\ominus}^{\text{sup}}) \\ &= \mathbb{P}_{\mathcal{M}^{\text{sup}},s_0}^{\sigma}(\mathbf{F} s_{\ominus}) .\end{aligned}$$

Therefore $\mathbb{P}_{\mathcal{M}^{\text{sup}},s_0}^{\text{sup}}(\mathbf{F} s_{\ominus}) \leq \mathbb{P}_{\mathcal{M},s_0}^{\text{sup}}(\mathbf{F} s_{\ominus})$.

Conversely, with the same construction as above, for any scheduler σ in \mathcal{M} , we build σ' a scheduler in \mathcal{M}^{sup} that mimicks σ until s_{\ominus}^{opt} is reached. We observe that

$$\begin{aligned}\mathbb{P}_{\mathcal{M}}^{\sigma}(\mathbf{F} s_{\ominus}) &= \mathbb{P}_{\mathcal{M}}^{\sigma}(\mathbf{F} s_{\ominus} \wedge \mathbf{G} \neg \text{Avoid}_{\mathcal{M}}^{\text{sup}}(s_{\ominus})) \\ &= \mathbb{P}_{\mathcal{M}^{\text{sup}}}^{\sigma'}(\mathbf{F} s_{\ominus} \wedge \mathbf{G} \neg s_{\ominus}^{\text{sup}}) \\ &= \mathbb{P}_{\mathcal{M}^{\text{sup}}}^{\sigma'}(\mathbf{F} s_{\ominus}) .\end{aligned}$$

Thus $\mathbb{P}_{\mathcal{M}^{\text{sup}},s_0}^{\text{sup}}(\mathbf{F} s_{\ominus}) \geq \mathbb{P}_{\mathcal{M},s_0}^{\text{sup}}(\mathbf{F} s_{\ominus})$, which allows us to conclude. \square

According to Lemma 3 computing the supremum probability (resp. infimum probability) in \mathcal{M} can equivalently be done in \mathcal{M}^{sup} (resp. \mathcal{M}^{inf}).

For every integer n , we define the following path properties in \mathcal{M}^{opt} :

$$\begin{cases} R_n = \mathbf{F}_{\leq n} s_{\ominus} \\ H_n^{\text{opt}} = \mathbf{G}_{\leq n} (\neg s_{\ominus} \wedge \neg s_{\ominus}^{\text{opt}}) . \end{cases}$$

In words, R_n expresses that the target is *reached* within n steps, and H_n^{opt} denotes that the target has not been reached within n steps, but there is still some *hope* to succeed. Hope here depends on the optimization objective (and thus on the collapsed MDP). In \mathcal{M}^{inf} (resp. \mathcal{M}^{sup}) it means that the infimum probability (resp. supremum probability) to reach s_{\ominus} from that point is positive. Note that $R_n \vee H_n^{\text{opt}} = \mathbf{F}_{\leq n} s_{\ominus} \vee \mathbf{G}_{\leq n} \neg s_{\ominus}^{\text{opt}}$.

We first establish a series of inequalities on probabilities in \mathcal{M}^{opt} :

Lemma 4. *For every initial state $s_0 \in S$ and every $n \in \mathbb{N}$*

$$\mathbb{P}_{\mathcal{M}^{\text{opt}},s_0}^{\text{opt}}(R_n) \leq \mathbb{P}_{\mathcal{M}^{\text{opt}},s_0}^{\text{opt}}(\mathbf{F} s_{\ominus}) \leq \mathbb{P}_{\mathcal{M}^{\text{opt}},s_0}^{\text{opt}}(R_n \vee H_n^{\text{opt}}) \leq \mathbb{P}_{\mathcal{M}^{\text{opt}},s_0}^{\text{opt}}(R_n) + \mathbb{P}_{\mathcal{M}^{\text{opt}},s_0}^{\text{sup}}(H_n^{\text{opt}}) .$$

Proof. Let us show that for every n ,

$$\llbracket R_n \rrbracket_{\mathcal{M}^{\text{opt}},s_0} \subseteq \llbracket \mathbf{F} s_{\ominus} \rrbracket_{\mathcal{M}^{\text{opt}},s_0} \subseteq \llbracket R_n \vee H_n^{\text{opt}} \rrbracket_{\mathcal{M}^{\text{opt}},s_0} .$$

The first inclusion is obvious. To prove the second, observe that the complement of $\llbracket R_n \vee H_n^{\text{opt}} \rrbracket_{\mathcal{M}^{\text{opt}},s_0}$ is $\llbracket (\mathbf{G}_{\leq n} \neg s_{\ominus}) \wedge (\mathbf{F}_{\leq n} s_{\ominus}^{\text{opt}}) \rrbracket_{\mathcal{M}^{\text{opt}},s_0} = \llbracket \mathbf{F}_{\leq n} s_{\ominus}^{\text{opt}} \rrbracket_{\mathcal{M}^{\text{opt}},s_0}$. Any path in that set thus reaches the sink losing state s_{\ominus}^{opt} in the first n steps, and therefore belongs to $\llbracket \neg \mathbf{F} s_{\ominus} \rrbracket_{\mathcal{M}^{\text{opt}},s_0}$. Finally, the last inequality is a direct application of Remark 1. \square

Thanks to Lemma 4, it is natural to define an approximation scheme with $\mathbb{P}_{\mathcal{M}^{\text{opt}}, s_0}^{\text{opt}}(R_n)$ as lower bound, and $\mathbb{P}_{\mathcal{M}^{\text{opt}}, s_0}^{\text{opt}}(R_n \vee H_n^{\text{opt}})$ as upper bound. It is formalised in Algorithm 1.

```

input : An MDP  $\mathcal{M}$ ,  $s_0, s_{\odot} \in S$ ,  $\varepsilon \in (0, 1)$ 
output: A value  $v \in [0, 1]$ 

 $n := 0$ ;
repeat
   $n := n + 1$ 
   $p_n^{\text{opt}, -} := \mathbb{P}_{\mathcal{M}^{\text{opt}}, s_0}^{\text{opt}}(\mathbf{F}_{\leq n} s_{\odot})$ ;
   $p_n^{\text{opt}, +} := \mathbb{P}_{\mathcal{M}^{\text{opt}}, s_0}^{\text{opt}}(\mathbf{F}_{\leq n} s_{\odot} \vee \mathbf{G}_{\leq n} (\neg s_{\odot} \wedge \neg s_{\odot}^{\text{opt}}))$ 
until  $|p_n^{\text{opt}, +} - p_n^{\text{opt}, -}| \leq \varepsilon$ ;
return  $p_n^{\text{opt}, -}$ 

```

Algorithm 1: $\text{Approx_Scheme}_1^{\text{opt}}$

Thanks to the rightmost inequality of Lemma 4, if we prove that, uniformly over all schedulers, H_n^{opt} becomes negligible, then this will ensure termination of the algorithm and help proving that this is actually an approximation algorithm.

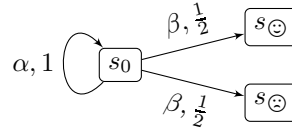
Theorem 2. *Let \mathcal{M} be an MDP such that $\lim_{n \rightarrow +\infty} \mathbb{P}_{\mathcal{M}^{\text{opt}}, s_0}^{\text{sup}}(H_n^{\text{opt}}) = 0$. Then $\text{Approx_Scheme}_1^{\text{opt}}$ provides an approximation algorithm for $\mathbb{P}_{\mathcal{M}, s_0}^{\text{opt}}(\mathbf{F} s_{\odot})$.*

Proof. The sequence $(p_n^{\text{opt}, -})_n$ is nondecreasing and $(p_n^{\text{opt}, +})_n$ is nonincreasing. Assuming they converge to the same value, which is the case, thanks to Lemma 4, when $\lim_{n \rightarrow +\infty} \mathbb{P}_{\mathcal{M}^{\text{opt}}, s_0}^{\text{sup}}(H_n^{\text{opt}}) = 0$, then $\text{Approx_Scheme}_1^{\text{opt}}$ terminates.

When it terminates, thanks to Lemmas 3 and 4, $\text{Approx_Scheme}_1^{\text{opt}}$ returns an ε -approximation of $\mathbb{P}_{\mathcal{M}, s_0}^{\text{opt}}(\mathbf{F} s_{\odot})$.

Algorithm $\text{Approx_Scheme}_1^{\text{opt}}$ is based on unfoldings of the MDP to deeper and deeper depths. Precisely, the lower bound $p_n^{\text{opt}, -}$ is the probability in the unfolding up to depth n of paths that reached s_{\odot} ; $p_n^{\text{opt}, +}$ is the probability in the same unfolding of paths that either reached s_{\odot} or end in a state from which there is a path in \mathcal{M} to s_{\odot} .

When $\text{opt} = \text{sup}$, $\text{Approx_Scheme}_1^{\text{opt}}$ has the following drawback: it may not terminate, even on finite MDPs. For instance the 3-state MDP on the right satisfies $\text{Avoid}^{\text{sup}}(s_{\odot}) = \{s_{\odot}\}$, and for every $n \in \mathbb{N}$, $p_n^{\text{sup}, -} = \frac{1}{2}$ and $p_n^{\text{sup}, +} = 1$.



3.2 Sliced MDP and second approximation scheme

To overcome the above-mentioned shortcoming of $\text{Approx_Scheme}_1^{\text{opt}}$ (in case $\text{opt} = \text{sup}$), we propose a refined approximation scheme. Intuitively, instead of unfolding the MDP up to depth n , as implicitly done in $\text{Approx_Scheme}_1^{\text{opt}}$, we consider slices of the MDP consisting of the restrictions to all states that are

reachable from s_0 within a fixed number of steps. Doing so, the termination on finite MDPs will be ensured.

Let $\mathcal{M} = (S, \text{Act}, \mathbb{P})$ be an MDP, $\text{opt} \in \{\inf, \sup\}$, and \mathcal{M}^{opt} as defined in Subsection 3.1. For every $n \in \mathbb{N}$, we define $\mathcal{M}_n^{\text{opt}}$, inductively on n , as the sub-MDP of \mathcal{M}^{opt} , restricted to states that can be reached within n steps. As illustrated in Figure 5, actions that exclusively lead to states unreachable within n steps are removed; for other actions, transitions leading out of states reachable within n steps are redirected to the fresh state s_\perp^n .

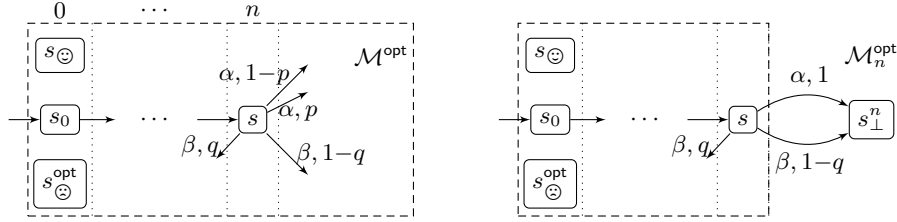


Fig. 5. Construction of the sliced MDP $\mathcal{M}_n^{\text{opt}}$ (right) from \mathcal{M}^{opt} (left).

The formal construction of $\mathcal{M}_n^{\text{opt}}$ is given in Appendix C.2. Writing S_n^{opt} for the set of states reachable from s_0 in at most n steps, the state spaces of \mathcal{M}^{opt} and $\mathcal{M}_n^{\text{opt}}$ coincide on $S_n^{\text{opt}} \setminus \{s_\perp^n\}$, and all outgoing transitions from S_n^{opt} in \mathcal{M}^{opt} are directed to s_\perp^n in $\mathcal{M}_n^{\text{opt}}$. Any path in \mathcal{M}^{opt} induces a unique path in $\mathcal{M}_n^{\text{opt}}$ which either stays in the common state space $S_n^{\text{opt}} \setminus \{s_\perp^n\}$, or reaches s_\perp^n . Moreover, any path in $\mathcal{M}_n^{\text{opt}}$ that reaches s_\ominus , can be seen as a path in \mathcal{M}^{opt} that also reaches s_\ominus . In the sequel, we use transparently the correspondence between paths in \mathcal{M} that only visit states reachable within n steps, and paths in $\mathcal{M}_n^{\text{opt}}$ that avoid s_\perp^n .

The sliced MDP $\mathcal{M}_n^{\text{opt}}$ enjoys the following inequalities:

- Lemma 5.** 1. $\mathbb{P}_{\mathcal{M}_n^{\text{opt}}, s_0}^{\text{opt}}(\mathbf{F}_{\leq n} s_\ominus) \leq \mathbb{P}_{\mathcal{M}_n^{\text{opt}}, s_0}^{\text{opt}}(\mathbf{F} s_\ominus)$
 2. $\mathbb{P}_{\mathcal{M}_n^{\text{opt}}, s_0}^{\text{opt}}(\mathbf{F} s_\ominus) \leq \mathbb{P}_{\mathcal{M}^{\text{opt}}, s_0}^{\text{opt}}(\mathbf{F} s_\ominus) \leq \mathbb{P}_{\mathcal{M}_n^{\text{opt}}, s_0}^{\text{opt}}(\mathbf{F}(s_\ominus \vee s_\perp^n))$
 3. $\mathbb{P}_{\mathcal{M}_n^{\text{opt}}, s_0}^{\text{opt}}(\mathbf{F}(s_\ominus \vee s_\perp^n)) \leq \mathbb{P}_{\mathcal{M}^{\text{opt}}, s_0}^{\text{opt}}(\mathbf{F}_{\leq n} s_\ominus \vee \mathbf{G}_{\leq n}(\neg s_\ominus \wedge \neg s_\oplus^{\text{opt}}))$
 4. $\mathbb{P}_{\mathcal{M}_n^{\text{opt}}, s_0}^{\text{opt}}(\mathbf{F}(s_\ominus \vee s_\perp^n)) \leq \mathbb{P}_{\mathcal{M}_n^{\text{opt}}, s_0}^{\text{opt}}(\mathbf{F} s_\ominus) + \mathbb{P}_{\mathcal{M}_n^{\text{opt}}, s_0}^{\text{sup}}(\mathbf{F} s_\perp^n)$

Proof. The proof of the first three items is by inclusion of events:

- Every path in \mathcal{M}^{opt} which reaches s_\ominus in at most n steps also reaches s_\ominus in $\mathcal{M}_n^{\text{opt}}$ (avoiding s_\perp^n)
- Conversely, every path which reaches s_\ominus in $\mathcal{M}_n^{\text{opt}}$ also reaches s_\ominus in \mathcal{M}^{opt} . Also, every path in \mathcal{M}^{opt} which reaches s_\ominus can be partly read in \mathcal{M}^{opt} : it either reaches s_\ominus in \mathcal{M}^{opt} , or ends up in s_\perp^n .
- Every path which either reaches s_\ominus or hits s_\perp^n in \mathcal{M}^{opt} , either hits s_\ominus in at most n steps, or does not visit s_\ominus and s_\oplus^{opt} in \mathcal{M}^{opt} during the n first steps.

Finally, the last item is a direct application of Remark 1. \square

We now define a second approximation scheme as formalised in Algorithm 2.

```

input : An MDP  $\mathcal{M}$ ,  $s_0, s_\odot \in S$ ,  $\varepsilon \in (0, 1)$ 
output: A value  $v \in [0, 1]$ 

 $n := 0$ ;
repeat
   $n := n + 1$ 
   $q_n^{\text{opt}, -} := \mathbb{P}_{\mathcal{M}_n^{\text{opt}}, s_0}^{\text{opt}}(\mathbf{F} s_\odot)$ ;
   $q_n^{\text{opt}, +} := \mathbb{P}_{\mathcal{M}_n^{\text{opt}}, s_0}^{\text{opt}}(\mathbf{F}(s_\odot \vee s_\perp^n))$ 
until  $|q_n^{\text{opt}, +} - q_n^{\text{opt}, -}| \leq \varepsilon$ ;
return  $q_n^{\text{opt}, -}$ 

```

Algorithm 2: $\text{Approx_Scheme}_2^{\text{opt}}$

Using the sequences defined in our two approximation schemes, from Lemma 5 (first and third items) we learn that for every n , $p_n^{\text{opt}, -} \leq q_n^{\text{opt}, -}$, and $q_n^{\text{opt}, +} \leq p_n^{\text{opt}, +}$. Thus, $\text{Approx_Scheme}_2^{\text{opt}}$ is a refinement of $\text{Approx_Scheme}_1^{\text{opt}}$. In particular, under the hypotheses of Theorem 3, it terminates. We give below a refined⁵ criterion for ensuring that $\text{Approx_Scheme}_2^{\text{opt}}$ is an approximation scheme.

Theorem 3. *Let \mathcal{M} be an MDP such that $\lim_{n \rightarrow +\infty} \mathbb{P}_{\mathcal{M}_n^{\text{opt}}, s_0}^{\text{sup}}(\mathbf{F} s_\perp^n) = 0$. Then $\text{Approx_Scheme}_2^{\text{opt}}$ provides an approximation algorithm for $\mathbb{P}_{\mathcal{M}, s_0}^{\text{opt}}(\mathbf{F} s_\odot)$.*

Proof. The sequence $(q_n^{\text{opt}, -})_n$ (resp. $(q_n^{\text{opt}, +})_n$) is non-decreasing (resp. non-increasing). When they converge to the same limit, $\text{Approx_Scheme}_2^{\text{opt}}$ terminates. Moreover, thanks to Lemma 5, for every $n \in \mathbb{N}$, $q_n^{\text{opt}, -} \leq \mathbb{P}_{\mathcal{M}_n^{\text{opt}}, s_0}^{\text{opt}}(\mathbf{F} s_\odot) \leq q_n^{\text{opt}, +}$ so that upon termination, $\text{Approx_Scheme}_2^{\text{opt}}$ returns an ε -approximation of $\mathbb{P}_{\mathcal{M}, s_0}^{\text{opt}}(\mathbf{F} s_\odot)$.

Under the assumption that $\lim_{n \rightarrow +\infty} \mathbb{P}_{\mathcal{M}_n^{\text{opt}}, s_0}^{\text{sup}}(\mathbf{F} s_\perp^n) = 0$ the last item of Lemma 5 implies that $\lim_{n \rightarrow +\infty} \mathbb{P}_{\mathcal{M}_n^{\text{opt}}, s_0}^{\text{opt}}(\mathbf{F}(s_\odot \vee s_\perp^n)) = \lim_{n \rightarrow +\infty} \mathbb{P}_{\mathcal{M}_n^{\text{opt}}, s_0}^{\text{opt}}(\mathbf{F} s_\odot)$. Hence the two sequences $(q_n^{\text{opt}, -})_n$ and $(q_n^{\text{opt}, +})_n$ converge towards $\mathbb{P}_{\mathcal{M}_n^{\text{opt}}, s_0}^{\text{opt}}(\mathbf{F} s_\odot)$ and $\text{Approx_Scheme}_2^{\text{opt}}$ terminates. \square

The approximation scheme $\text{Approx_Scheme}_2^{\text{opt}}$ terminates for finite MDPs. In contrast, recall that approximating the supremum reachability property with $\text{Approx_Scheme}_1^{\text{sup}}$ may not terminate on some finite MDPs (see page 12).

In the remainder of the paper, we relate decisiveness to our approximation schemes. More precisely, we explicit which decisiveness hypotheses are sufficient to ensure termination of each of our approximation schemes. These sufficient conditions depend on the optimization objective, so that in Section 4 we focus on infimum probability, and in Section 5 we treat supremum probability.

⁵ It is easy to check that $\mathbb{P}_{\mathcal{M}_n^{\text{opt}}, s_0}^{\text{opt}}(\mathbf{F} s_\perp^n) \leq \mathbb{P}_{\mathcal{M}_n^{\text{opt}}, s_0}^{\text{opt}}(H_n^{\text{opt}})$.

4 Infimum reachability probability

Focusing on the infimum probability, we give conditions related to decisiveness which ensure termination and correctness of our the approximation schemes.

4.1 Finite action-branching and inf-decisiveness to approximate infimum reachability probability

Theorem 4. *Let $\mathcal{M} = (S, \text{Act}, \mathbb{P})$ be an MDP, $s_0 \in S$ an initial state and s_\odot a target state. Assume that \mathcal{M} is finitely action-branching and inf-decisive w.r.t. s_\odot from s_0 . Then $\text{Approx_Scheme}_1^{\text{inf}}$ and $\text{Approx_Scheme}_2^{\text{inf}}$ terminate and are correct for \mathcal{M} from s_0 .*

The rest of this section is devoted to the proof of this result. We first prove that the decisiveness property of \mathcal{M} transfers to \mathcal{M}^{inf} :

Lemma 6. *Let \mathcal{M} be a finitely action-branching MDP. If \mathcal{M} is inf-decisive w.r.t. s_\odot from s_0 , then so is \mathcal{M}^{inf} .*

Proof. Since \mathcal{M} is finitely action-branching, Lemma 2 applies: pick $\sigma_0 \in \text{Sched}_{\text{pp}}(\mathcal{M})$ a pure positional scheduler defined on $\text{Avoid}_{\mathcal{M}}^{\text{inf}}(s_\odot)$ such that for every $s \in \text{Avoid}_{\mathcal{M}}^{\text{inf}}(s_\odot)$, $\mathbb{P}_{\mathcal{M},s}^{\sigma_0}(\mathbf{F} s_\odot) = 0$. Note that, starting from $\text{Avoid}_{\mathcal{M}}^{\text{inf}}(s_\odot)$, paths under scheduler σ_0 only visit states in $\text{Avoid}_{\mathcal{M}}^{\text{inf}}(s_\odot)$.

Towards a contradiction, assume that \mathcal{M}^{inf} is not inf-decisive. Let $\sigma \in \text{Sched}_{\text{pp}}(\mathcal{M}^{\text{inf}})$ be a pure positional scheduler in \mathcal{M}^{inf} such that $\mathbb{P}_{\mathcal{M}^{\text{inf}},s_0}^{\sigma}(\mathbf{F} s_\odot \vee \mathbf{F} \text{Avoid}_{\mathcal{M}^{\text{inf}}}^{\text{inf}}(s_\odot)) < 1$. Remark that by construction of \mathcal{M}^{inf} , $\text{Avoid}_{\mathcal{M}^{\text{inf}}}^{\text{inf}}(s_\odot) = \{s_\odot^{\text{inf}}\}$. Thus, $\mathbb{P}_{\mathcal{M}^{\text{inf}},s_0}^{\sigma}(\mathbf{F} s_\odot \vee \mathbf{F} s_\odot^{\text{inf}}) < 1$, and for every $s \in S \setminus \text{Avoid}_{\mathcal{M}^{\text{inf}}}^{\text{inf}}(s_\odot)$, $\mathbb{P}_{\mathcal{M}^{\text{inf}},s}^{\sigma}(\mathbf{F} s_\odot) > 0$.

We define σ' , scheduler in \mathcal{M} that mimicks σ as long as $\text{Avoid}_{\mathcal{M}^{\text{inf}}}^{\text{inf}}(s_\odot)$ is not reached, and if so behaves as σ_0 . Gluing these two pure and positional partial schedulers, σ' is a pure and positional scheduler that satisfies $\mathbb{P}_{\mathcal{M},s_0}^{\sigma'}(\mathbf{F} s_\odot \vee \mathbf{F} \text{Avoid}_{\mathcal{M}}^{\text{inf}}(s_\odot)) = \mathbb{P}_{\mathcal{M}^{\text{inf}},s_0}^{\sigma}(\mathbf{F} s_\odot \vee \mathbf{F} s_\odot^{\text{inf}}) < 1$. This contradicts the fact that \mathcal{M} is inf-decisive w.r.t. s_\odot from s_0 . \square

It is now sufficient to prove the hypotheses of Proposition 2 to show the correctness and termination of the approximation scheme $\text{Approx_Scheme}_1^{\text{inf}}$ when \mathcal{M} is finitely action-branching and inf-decisive. Let us show that H_n^{inf} is uniformly negligible:

Lemma 7. *If \mathcal{M} is finitely action-branching and inf-decisive w.r.t. s_\odot from s_0 , then*

$$\lim_{n \rightarrow +\infty} \mathbb{P}_{\mathcal{M}^{\text{inf}},s_0}^{\text{sup}}(H_n^{\text{inf}}) = 0 .$$

Proof (sketch). We give here the proof assuming \mathcal{M} is finitely branching; the proof under the weaker assumption of finite action-branching is given in Appendix. Towards a contradiction, assume that there exists $\varepsilon > 0$ and a sequence $N = \{n_0 < n_1 < \dots\}$ of integers such that for every $n \in N$, $\mathbb{P}_{\mathcal{M}^{\text{inf}},s_0}^{\text{sup}}(H_n^{\text{inf}}) > \varepsilon$.

The property H_n^{inf} is a step-bounded safety property, thus using Lemma 2, for any $n \in N$ there exists a pure (history-dependent) scheduler $\sigma_n \in \text{Sched}_{\text{ph}}(\mathcal{M}^{\text{inf}})$ such that: $\mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma_n}(H_n^{\text{inf}}) > \varepsilon$. Using the schedulers $(\sigma_n)_{n \in N}$, we extract a scheduler $\sigma^* \in \text{Sched}_{\text{ph}}(\mathcal{M}^{\text{inf}})$ as follows. Since \mathcal{M}^{inf} is finitely action-branching, N admits an infinite subset $N_0 \subseteq N$ such that the first decision $\sigma_n(s_0)$ is the same for every $n \in N_0$; we define $\sigma^*(s_0)$ to be this decision. Let K_1 be the set of length-2 histories of σ^* : this set is finite since \mathcal{M}^{inf} is finitely proba-branching as well. Let $N_1 \subseteq N_0$ be an infinite subset of N_0 such that for every $h \in K_1$, all $\sigma_n(h)$ coincide for $n \in N_1$; we define $\sigma^*(h)$ to be this uniform decision. Iterating this process, the resulting scheduler is pure (and history-dependent). Furthermore, it ensures, for all k , $\mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma^*}(H_k^{\text{inf}}) \geq \varepsilon$. In particular,

$$\mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma^*}(\mathbf{G}(\neg s_{\ominus} \wedge \neg s_{\ominus}^{\text{inf}})) = \lim_{k \rightarrow \infty} \mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma^*}(H_k^{\text{inf}}) \geq \varepsilon.$$

Now, for the safety objective $\mathbf{G}(\neg s_{\ominus} \wedge \neg s_{\ominus}^{\text{inf}})$, applying Lemma 2 since \mathcal{M}^{inf} is finitely action-branching, there is a pure and positional optimal strategy σ^{**} such that

$$\mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma^{**}}(\mathbf{G}(\neg s_{\ominus} \wedge \neg s_{\ominus}^{\text{inf}})) \geq \mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma^*}(\mathbf{G}(\neg s_{\ominus} \wedge \neg s_{\ominus}^{\text{inf}})) \geq \varepsilon$$

However, \mathcal{M} is inf-decisive, hence so is \mathcal{M}^{inf} (by Lemma 6), and $\text{Avoid}_{\mathcal{M}^{\text{inf}}}(s_{\ominus}) = \{s_{\ominus}^{\text{inf}}\}$:

$$\mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma^{**}}(\mathbf{F} s_{\ominus} \vee \mathbf{F} s_{\ominus}^{\text{inf}}) = 1$$

which contradicts the above inequality. Hence the result. \square

4.2 Limits and example of applicability of the approach

The inf-decisiveness assumption is required to ensure termination of the algorithms $\text{Approx_Scheme}_1^{\text{inf}}$ and $\text{Approx_Scheme}_2^{\text{inf}}$. Consider indeed the MDP \mathcal{M} from Figure 1, which has $\text{Avoid}_{\mathcal{M}}(s_{\ominus}) = \{s_{\ominus}\}$, so that $\mathcal{M}^{\text{inf}} = \mathcal{M}$. In case $p > \frac{1}{2}$, \mathcal{M} is not inf-decisive from s_0 w.r.t. s_{\ominus} since the pure positional scheduler σ that always picks action α has a positive probability, say λ_p , to never reach s_{\ominus} nor s_{\ominus}^{inf} . Now, one can set values for p ($> \frac{1}{2}$) and q such that, for every $n \in \mathbb{N}$, $q_n^{\text{opt}, -} \leq 1 - \lambda_p < q$ and $q_n^{\text{opt}, +} = q$. Thus, $\text{Approx_Scheme}_2^{\text{inf}}$ does not terminate on that non inf-decisive example.

In [1], several classes of Markov chains are given as examples of so-called decisive Markov chains. For infinite Markov decision processes, we focus on non-deterministic and probabilistic lossy channel system (NPLCS), already mentioned page 6. For an NPLCS, write \mathcal{M} its MDP semantics. For every pure positional scheduler $\sigma \in \text{Sched}_{\text{pp}}(\mathcal{M})$, one can argue that \mathcal{M}_{σ} has a finite attractor (the set of configurations with empty channels). Hence, one can conclude that the Markov chain \mathcal{M}_{σ} is decisive –in the sense of [1]– from initial configuration (q_0, ε) w.r.t. any set T defined by control states only. Writing \tilde{T} for the set of configurations that can no longer reach T , decisiveness of \mathcal{M}_{σ} means that $\mathbb{P}_{\mathcal{M}, (q_0, \varepsilon)}^{\sigma}(\mathbf{F} T \vee \mathbf{F} \tilde{T}) = 1$. Hence \mathcal{M} is inf-decisive as well w.r.t.

T from (q_0, ε) , since $\text{Avoid}_{\mathcal{M}}^{\text{inf}}(T) \subseteq \tilde{T}$. The two schemes $\text{Approx_Scheme}_1^{\text{inf}}$ and $\text{Approx_Scheme}_2^{\text{inf}}$ are therefore approximation schemes for computing the infimum reachability probability of T in \mathcal{M} . It remains to discuss the effectiveness of the schemes. Assuming finite action-branching, thanks to Lemma 2, computing $\text{Avoid}_{\mathcal{M}}^{\text{inf}}(T)$ amounts to computing states from which one can almost-surely avoid T , which amounts to computing states from which one can (surely) avoid T ; this can be computed in \mathcal{M} , since reachability is decidable for lossy channel systems.

5 Supremum reachability probability

We now turn to supremum probability, and give sufficient conditions (related to decisiveness) to turn our generic schemes into approximation algorithms.

5.1 Finite action-branching and sup-decisiveness to approximate supremum reachability probability

Similarly to inf-decisiveness to approximate the infimum reachability probability, sup-decisiveness together with finite action branching is sufficient to obtain approximation algorithms for the supremum probability. Recall that sup-decisiveness is a more restrictive condition than inf-decisiveness.

Theorem 5. *Let $\mathcal{M} = (S, \text{Act}, \mathbb{P})$ be a denumerable MDP, $s_0 \in S$ the initial state and s_{\odot} a target state. Assume that \mathcal{M} is finitely action-branching and sup-decisive w.r.t. s_{\odot} from s_0 . Then $\text{Approx_Scheme}_1^{\text{sup}}$ and $\text{Approx_Scheme}_2^{\text{sup}}$ are correct, and terminate for \mathcal{M} from s_0 .*

To prove Theorem 5, we show that the hypothesis of Theorem 2 is satisfied, implying the correctness and termination of $\text{Approx_Scheme}_1^{\text{sup}}$ and, because $\text{Approx_Scheme}_2^{\text{sup}}$ refines $\text{Approx_Scheme}_1^{\text{sup}}$, we also obtain its correctness and termination.

Lemma 8. *If \mathcal{M} is sup-decisive w.r.t. s_{\odot} from s_0 , then so is \mathcal{M}^{sup} .*

Proof. Notice that by construction of \mathcal{M}^{sup} , $\text{Avoid}_{\mathcal{M}^{\text{sup}}}^{\text{sup}}(s_{\odot}) = \{s_{\odot}^{\text{sup}}\}$. Towards a contradiction, assume that \mathcal{M}^{sup} is not sup-decisive. Then, there is a pure positional scheduler $\sigma \in \text{Sched}_{\text{pp}}(\mathcal{M}^{\text{sup}})$ such that $\mathbb{P}_{\mathcal{M}^{\text{sup}}, s_0}^{\sigma}(\mathbf{F} s_{\odot} \vee \mathbf{F} s_{\odot}^{\text{sup}}) < 1$. We view σ as a partial scheduler in \mathcal{M} , and extend it to $\sigma' \in \text{Sched}_{\text{pp}}(\mathcal{M})$ in an arbitrary pure positional way from states in $\text{Avoid}_{\mathcal{M}}^{\text{sup}}(s_{\odot})$. Then $\mathbb{P}_{\mathcal{M}, s_0}^{\sigma'}(\mathbf{F} s_{\odot} \vee \mathbf{F} \text{Avoid}_{\mathcal{M}}^{\text{sup}}(s_{\odot})) < 1$. This contradicts the fact that \mathcal{M} is sup-decisive. \square

Notice that, in contrast to the case of infimum (see Lemma 6), the proof of Lemma 8 does not require finite action-branching.

Lemma 9. *If \mathcal{M} is finitely action-branching and sup-decisive w.r.t. s_{\odot} from s_0 , then*

$$\lim_{n \rightarrow +\infty} \mathbb{P}_{\mathcal{M}^{\text{sup}}, s_0}^{\text{sup}}(H_n^{\text{sup}}) = 0 .$$

Proof. Towards a contradiction, assume that there exists $\varepsilon > 0$ and a sequence $N = \{n_0 < n_1 < \dots\}$ of integers such that for every $n \in N$, $\mathbb{P}_{\mathcal{M}^{\text{sup}}, s_0}^{\text{sup}}(H_n^{\text{sup}}) > \varepsilon$.

For every $n \in N$, fix $\sigma_n \in \text{Sched}(\mathcal{M}^{\text{sup}})$ such that $\mathbb{P}_{\mathcal{M}^{\text{sup}}, s_0}^{\sigma_n}(H_n^{\text{sup}}) > \varepsilon$. By Lemma 2, since H_n^{sup} is a step-bounded safety property, σ_n can be assumed to be pure. Then, following the same reasoning as in the proof of Lemma 7, we extract from the sequence $(\sigma_n)_{n \in N}$ a pure scheduler σ^* such that

$$\mathbb{P}_{\mathcal{M}^{\text{sup}}, s_0}^{\sigma^*}(\mathbf{G}(\neg s_{\odot} \wedge \neg s_{\odot}^{\text{sup}})) \geq \varepsilon'$$

where $\varepsilon' = \varepsilon - (1 - \prod_{j=1}^{\infty} (1 - \varepsilon_j))$.

Now, for the safety objective $\mathbf{G}(\neg s_{\odot} \wedge \neg s_{\odot}^{\text{sup}})$ since \mathcal{M}^{sup} is finitely action-branching, there is a pure and positional optimal strategy σ^{**} such that

$$\mathbb{P}_{\mathcal{M}^{\text{sup}}, s_0}^{\sigma^{**}}(\mathbf{G}(\neg s_{\odot} \wedge \neg s_{\odot}^{\text{sup}})) \geq \mathbb{P}_{\mathcal{M}^{\text{sup}}, s_0}^{\sigma^*}(\mathbf{G}(\neg s_{\odot} \wedge \neg s_{\odot}^{\text{sup}})) \geq \varepsilon'$$

Indeed, this is a consequence of Lemma 2 (first item), applied to the dual property $\mathbf{F}(s_{\odot} \vee s_{\odot}^{\text{sup}})$ in the finitely action-branching MDP \mathcal{M}^{sup} . It contradicts the fact that $\mathbb{P}_{\mathcal{M}^{\text{sup}}, s_0}^{\sigma^{**}}(\mathbf{F} s_{\odot} \vee \mathbf{F} s_{\odot}^{\text{sup}}) = 1$, because \mathcal{M}^{sup} is sup-decisive. \square

5.2 Limits of applicability of the approach

The sup-decisiveness property turns out to be a necessary condition for the scheme $\text{Approx_Scheme}_2^{\text{sup}}$ to terminate. Consider again the MDP \mathcal{M}^{R} on the right of Figure 4. For every $n \in \mathbb{N}$, $q_n^{\text{sup},+} = 1$, and yet $\mathbb{P}_{\mathcal{M}^{\text{R}}, s_1}^{\text{sup}}(\mathbf{F} s_{\odot}) < 1$. Thus, $\text{Approx_Scheme}_2^{\text{sup}}$ does not terminate for precision $\varepsilon \leq \frac{1}{2} (1 - \mathbb{P}_{\mathcal{M}^{\text{R}}, s_1}^{\text{sup}}(\mathbf{F} s_{\odot}))$. This MDP is not sup-decisive (nor inf-decisive) from s_0 w.r.t. s_{\odot} . In fact, already the Markov chain obtained by restricting to action α is not decisive w.r.t. s_{\odot} in the sense of [1], and $\text{Approx_Scheme}_2^{\text{sup}}$ would not terminate.

6 Conclusion

In this paper, we studied how to extend decisiveness from Markov chains to Markov decision processes, and how to use this property to derive approximation algorithms for optimum reachability probabilities. The notion of sup-decisiveness is quite strong, limiting the applicability of our approach to approximate supremum reachability probabilities. The notion of inf-decisiveness is weaker, increasing the potential applicability of our approach to approximate infimum probabilities.

As future work we plan to consider richer properties, for instance, in the case of infimum probabilities, study repeated reachability or quantitative payoff functions. We also aim at considering classes of denumerable MDPs defined by high-level models such as lossy channel systems or VASS, to see whether for these instances, some of our hypotheses can be relaxed. Finally, we would like to clarify the decidability frontier for quantitative model checking of denumerable MDPs.

References

1. P. A. Abdulla, N. Ben Henda, and R. Mayr. Decisive Markov chains. *Logical Methods in Computer Science*, 3(4:7):1–32, 2007.
2. P. A. Abdulla, N. Bertrand, A. Rabinovich, and Ph. Schnoebelen. Verification of probabilistic systems with faulty communication. *Information and Computation*, 202(2):141–165, 2005.
3. P. A. Abdulla, L. Clemente, R. Mayr, and S. Sandberg. Stochastic parity games on lossy channel systems. *Logical Methods in Computer Science*, 10(4), 2014.
4. C. Baier, N. Bertrand, and Ph. Schnoebelen. A note on the attractor-property of infinite-state Markov chains. *Information Processing Letters*, 97(2):58–63, Jan. 2006.
5. C. Baier, N. Bertrand, and Ph. Schnoebelen. Verifying nondeterministic probabilistic channel systems against ω -regular linear-time properties. *ACM Transactions on Computational Logic*, 9(1), 2007.
6. C. Baier and J.-P. Katoen. *Principles of Model Checking*. MIT Press, 2008.
7. N. Bertrand and Ph. Schnoebelen. Model checking lossy channels systems is probably decidable. In *Proc. 6th International Conference on Foundations of Software Science and Computation Structures (FoSSaCS’03)*, volume 2620 of *Lecture Notes in Computer Science*, pages 120–135. Springer, 2003.
8. N. Bertrand and Ph. Schnoebelen. Solving stochastic Büchi games on infinite arenas with a finite attractor. In *Proc. 11th International Workshop on Quantitative Aspects of Programming Languages and Systems (QAPL’13)*, volume 117 of *EPTCS*, pages 116–131, 2013.
9. T. Brázdil, V. Brozek, and K. Etessami. One-counter stochastic games. In *Proc. IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS’10)*, volume 8 of *LIPIcs*, pages 108–119. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2010.
10. T. Brázdil, V. Brozek, K. Etessami, and A. Kucera. Approximating the termination value of one-counter mdps and stochastic games. In *Proc. 38th International Colloquium on Automata, Languages and Programming (ICALP’11)*, volume 6756 of *Lecture Notes in Computer Science*, pages 332–343. Springer, 2011.
11. T. Brázdil, V. Brozek, A. Kucera, and J. Obdržálek. Qualitative reachability in stochastic BPA games. *Information and Computation*, 209(8):1160–1183, 2011.
12. C. Dehnert, S. Junges, J. Katoen, and M. Volk. A storm is coming: A modern probabilistic model checker. In *Proc. 29th International Conference on Computer Aided Verification (CAV’17)*, volume 10427 of *Lecture Notes in Computer Science*, pages 592–600. Springer, 2017.
13. K. Etessami, D. Wojtczak, and M. Yannakakis. Quasi-birth-death processes, tree-like qbds, probabilistic 1-counter automata, and pushdown systems. *Perform. Eval.*, 67(9):837–857, 2010.
14. K. Etessami and M. Yannakakis. Recursive concurrent stochastic games. *Logical Methods in Computer Science*, 4(4), 2008.
15. S. Haddad and B. Monmege. Interval iteration algorithm for mdps and imdps. *Theor. Comput. Sci.*, 735:111–131, 2018.
16. S. Kiefer, R. Mayr, M. Shirmohammadi, and P. Totzke. Büchi objectives in countable mdps. In *Proc. 46th International Colloquium on Automata, Languages, and Programming, (ICALP’19)*, volume 132 of *LIPIcs*, pages 119:1–119:14. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2019.

17. S. Kiefer, R. Mayr, M. Shirmohammadi, and D. Wojtczak. Parity objectives in countable MDPs. In *Proc. 32th Annual Symposium on Logic in Computer Science (LICS'17)*, pages 1–11. IEEE Computer Society Press, 2017.
18. A. Kucera. *Lectures in Game Theory for Computer Scientists*, chapter Turn-Based Stochastic Games. Cambridge University Press, 2011.
19. A. Kucera, J. Esparza, and R. Mayr. Model checking probabilistic pushdown automata. *Logical Methods in Computer Science*, 2(1), 2006.
20. M. Kwiatkowska, G. Norman, and D. Parker. PRISM 4.0: Verification of probabilistic real-time systems. In *Proc. 23rd International Conference on Computer Aided Verification (CAV'11)*, volume 6806 of *Lecture Notes in Computer Science*, pages 585–591. Springer, 2011.
21. R. Mayr. Undecidable problems in unreliable computations. *Theoretical Computer Science*, 297(1-3):337–354, 2003.

Appendix

In this appendix, we provide complement proofs that cannot appear in the core of the paper due to space constraints.

A Undecidability of the value problem (Theorem 1)

Theorem 1. *The following decision problem is undecidable:*

- **Input:** $\mathcal{M} = (S, \text{Act}, \mathbf{P})$ an MDP defined by a NPLCS, with $s_0, s_\ominus \in S$
- **Output:** *yes* iff $\mathbb{P}_{\mathcal{M}, s_0}^{\sup}(\mathbf{F} s_\ominus) = 1$.

Proof. We reduce the boundedness problem, which is undecidable for lossy channel systems (LCSs) [21]. Given an LCS \mathcal{L} , consider the finitely action-branching MDP \mathcal{M} as represented in Figure 6.

To define \mathcal{M} , we proceed in two steps. First, the semantics of \mathcal{L} can be seen as an MDP $\mathcal{M}(\mathcal{L})$, whose state space $S_{\mathcal{L}}$ is the set of *configurations* of \mathcal{L} . A configuration is a pair (q, w) where q is a control state and w describes the channel contents. As for the probabilistic transition function, from every configuration (q, w) , finitely many actions (sendings and receptions) are available. The next configuration depends on the message losses, and we assume the probability distribution for a letter on a channel to be lost to be uniform.

We then embed $\mathcal{M}(\mathcal{L})$ into a bigger MDP \mathcal{M} , whose state space is $S_{\mathcal{L}} \cup \{s_\ominus, s_\oplus\}$. The two states s_\ominus and s_\oplus are sink states. From any state of $\mathcal{M}(\mathcal{L})$, two extra actions are enabled: **try** and **restart**. Action **restart** is deterministic and leads to (q_0, ε) the initial configuration of \mathcal{L} . From (q, w) , action **try** leads with probability $\frac{1}{2^{|w|}}$ to s_\ominus and remaining probability $1 - \frac{1}{2^{|w|}}$ to s_\oplus .

We claim that under this construction, \mathcal{M} enjoys the following properties:

- if \mathcal{L} is bounded, there exists $p > 0$ such that $\mathbb{P}_{\mathcal{M}, s_0}^{\sup}(\mathbf{F} s_\ominus) \leq 1 - p < 1$;
- if \mathcal{L} is unbounded, $\mathbb{P}_{\mathcal{M}, s_0}^{\sup}(\mathbf{F} s_\ominus) = 1$.

Assume first that \mathcal{L} is bounded. In order to reach s_\ominus , a scheduler must take the action **try** at some point. Let p be the minimum probability to move from some reachable configuration in \mathcal{L} to s_\ominus when taking the **try**-transition. Since \mathcal{L} is bounded, p is positive (and equal to $2^{-\ell}$ where ℓ is the maximal channel contents length in \mathcal{L}). Whenever the scheduler chooses action **try**, the probability is at least p to move to s_\ominus . All in all, under any scheduler σ , $\mathbb{P}_{\mathcal{M}}^{\sigma}(\mathbf{F} s_\ominus) \leq 1 - p$, and thus $\mathbb{P}_{\mathcal{M}}^{\sup}(\mathbf{F} s_\ominus) \leq 1 - p$. Note that p only depends on the LCS \mathcal{L} .

Assume now that \mathcal{L} is unbounded, and fix $\eta > 0$. Consider a configuration (q, w) such that $2^{-|w|} \leq \eta$. Thus from configuration (q, w) , when **try** is played, the probability to move to s_\ominus is at most η . We define a scheduler σ as follows. The objective of σ is to reach configuration (q, w) and then play **try**. Since (q, w) is reachable, there is a sequence of actions that allows to reach it with positive probability from the initial configuration (q_0, ε) . If this path is not realised (due to unwanted message losses), σ changes mode and restarts the simulation with

action **restart**. So defined, σ almost surely eventually succeeds in reaching (q, w) , and thus ensures $\mathbb{P}_{\mathcal{M}}^{\sigma}(\mathbf{F} s_{\ominus}) = 1 - 2^{-|w|} \geq 1 - \eta$. Therefore $\mathbb{P}_{\mathcal{M}, s_0}^{\sup}(\mathbf{F} s_{\ominus}) = 1$.

As such, this is not an MDP directly derived from an NPLCS, but it can easily be turned to such an MDP. First notice that, instead of $2^{-|w|}$, one only needs a decreasing function $f(|w|)$ which converges to 0 when the length of w diverges; at some step, the probability to lose all messages is such a function. Hence:

- action **try** can be replaced by the deterministic writing of a fresh action δ , followed by a reading of δ (this requires all messages in w to be lost, hence this happens with probability $f(|w|)$) leading to s_{\ominus} , or any other reading leading to s_{\ominus} ;
- action **restart** can be replaced by a writing of a fresh letter δ , followed by readings of letter until reading of δ (the channel is then empty, as expected). \square

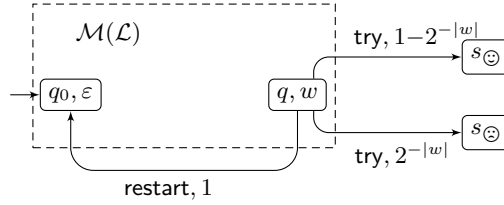


Fig. 6. Undecidability of the value-1 problem.

B Proof of Lemma 7

Lemma 7. *If \mathcal{M} is finitely action-branching and inf-decisive w.r.t. s_{\ominus} from s_0 , then*

$$\lim_{n \rightarrow +\infty} \mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sup}(H_n^{\text{inf}}) = 0 .$$

Proof. For every $k > 0$, we let $\varepsilon_k > 0$ be such that $\prod_{j=1}^{\infty} (1 - \varepsilon_j) > 1 - \varepsilon$. We note $\varepsilon' = \varepsilon - (1 - \prod_{j=1}^{\infty} (1 - \varepsilon_j))$.

We define a new scheduler $\sigma^* \in \text{Sched}_{\text{ph}}(\mathcal{M}^{\text{inf}})$ using the above family $(\sigma_n)_{n \in N}$, inductively on the length of histories. We explicit the induction hypothesis for $k \geq 1$: there is an infinite subset $N_k \subseteq N$ and a finite set K_k of histories of length k where:

- for every h , which is a strict prefix of some $h' \in K_k$, $\sigma^*(h) = \sigma_n(h)$ for all $n \in N_k$;
- $\mathbb{P}_{\mathcal{M}, s_0}^{\sigma^*}(\text{Cyl}(K_k)) \geq \prod_{j=1}^k (1 - \varepsilon_j)$.

Note: in this proof, we will write K_k instead of $\text{Cyl}(K_k)$ to ease readability.

We first consider the case $k = 1$. There is a unique length-0 history $h = s_0$. Since \mathcal{M}^{inf} is finitely action-branching, there is an infinite subset $N_1 \subseteq N$ such that $\sigma_n(s_0)$ is the same for every $n \in N_1$; we define $\sigma^*(s_0)$ as that precise value. Unless \mathcal{M}^{inf} is finite proba-branching, there might be infinitely many outcomes of σ^* of length 1. We select a finite number of such outcomes K_1 which has probability at least $1 - \varepsilon_1$: formally $\mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma^*}(K_1) \geq 1 - \varepsilon_1$. This shows that the induction hypothesis initially holds.

We now assume that the induction hypothesis holds for $k \geq 1$. We let $N_{k+1} \subseteq N_k$ be an infinite set such that for every $h \in K_k$ (which have length k), all $\sigma_n(h)$ coincide for any $n \in N_{k+1}$; We set $\sigma^*(h)$ the uniform value. As in the initial case, there might be infinitely many outcomes of σ^* of length $k+1$ whose prefixes are in K_k . We select a finite portion of such outcomes, with a large relative probability; formally let $K_{k+1} \subseteq K_k$ be a finite subset of K_k such that $\mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma^*}(K_{k+1} \mid K_k) \geq 1 - \varepsilon_{k+1}$. Using the induction hypothesis and Bayes theorem, we get the expected condition $\mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma^*}(K_{k+1}) \geq \prod_{j=1}^{k+1} (1 - \varepsilon_j)$. This closes the induction step.

Now, fix $k \in \mathbb{N}$ and fix some $n \in N_k$ such that $n \geq k$. Then we know that $\mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma_n}(H_n^{\text{inf}}) \geq \varepsilon$, hence that $\mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma_n}(H_k^{\text{inf}}) \geq \varepsilon$. We also know that σ^* and σ_n coincide on K_k . We compute:

$$\begin{aligned}
\mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma^*}(H_k^{\text{inf}}) &\geq \mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma^*}(H_k^{\text{inf}} \cap K_k) \\
&= \mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma_n}(H_k^{\text{inf}} \cap K_k) \\
&= \mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma_n}(H_k^{\text{inf}}) - \mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma_n}(H_k^{\text{inf}} \cap \neg K_k) \\
&\geq \mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma_n}(H_k^{\text{inf}}) - \mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma_n}(\neg K_k) \\
&\geq \mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma_n}(H_k^{\text{inf}}) - \left(1 - \prod_{j=1}^k (1 - \varepsilon_j)\right) \\
&\geq \mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma_n}(H_k^{\text{inf}}) - \left(1 - \prod_{j=1}^{\infty} (1 - \varepsilon_j)\right) \\
&\geq \varepsilon - \left(1 - \prod_{j=1}^{\infty} (1 - \varepsilon_j)\right) \\
&= \varepsilon'
\end{aligned}$$

The remainder of the proof is as in the core of the paper. For completeness, we recall it here.

In particular,

$$\mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma^*}(\mathbf{G}(\neg s_{\odot} \wedge \neg s_{\odot}^{\text{inf}})) = \lim_{k \rightarrow \infty} \mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma^*}(H_k^{\text{inf}}) \geq \varepsilon'.$$

Now, for the safety objective $\mathbf{G}(\neg s_{\odot} \wedge \neg s_{\odot}^{\text{inf}})$, applying Lemma 2 (first item) since \mathcal{M}^{inf} is finitely action-branching, there is a pure and positional optimal

strategy σ^{**} such that

$$\mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma^{**}}(\mathbf{G}(\neg s_{\ominus} \wedge \neg s_{\ominus}^{\text{inf}})) \geq \mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma^*}(\mathbf{G}(\neg s_{\ominus} \wedge \neg s_{\ominus}^{\text{inf}})) \geq \varepsilon'$$

However, \mathcal{M} is inf-decisive, hence so is \mathcal{M}^{inf} (by Lemma 6), and $\text{Avoid}_{\mathcal{M}^{\text{inf}}}(s_{\ominus}) = \{s_{\ominus}^{\text{inf}}\}$:

$$\mathbb{P}_{\mathcal{M}^{\text{inf}}, s_0}^{\sigma^{**}}(\mathbf{F} s_{\ominus} \vee \mathbf{F} s_{\ominus}^{\text{inf}}) = 1$$

which contradicts the above inequality. Hence the result. \square

C Formal constructions of Section 3

C.1 Construction of the MDP \mathcal{M}^{opt} (Subsection 3.1)

Formally, $\mathcal{M}^{\text{opt}} = (S^{\text{opt}}, \text{Act}, \mathbf{P}^{\text{opt}})$ with

- $S^{\text{opt}} = (S \setminus \text{Avoid}_{\mathcal{M}}^{\text{opt}}(s_{\ominus})) \cup \{s_{\ominus}^{\text{opt}}\}$;
- for every $s, s' \in S^{\text{opt}} \setminus \{s_{\ominus}^{\text{opt}}\}$, for every $a \in \text{Act}$, $\mathbf{P}^{\text{opt}}(s, a, s') = \mathbf{P}(s, a, s')$;
- for every $s \in S^{\text{opt}} \setminus \{s_{\ominus}^{\text{opt}}\}$, $\mathbf{P}^{\text{opt}}(s, a, s_{\ominus}^{\text{opt}}) = \sum_{s' \in \text{Avoid}_{\mathcal{M}}^{\text{opt}}(s_{\ominus})} \mathbf{P}(s, a, s')$.
- for every $a \in \text{Act}$, $\mathbf{P}^{\text{opt}}(s_{\ominus}^{\text{opt}}, a, s_{\ominus}^{\text{opt}}) = 1$;

C.2 Construction of the sliced MDPs (Subsection 3.2)

Formally, for $n = 0$, $\mathcal{M}_0^{\text{opt}} = (S_0^{\text{opt}}, \text{Act}, \mathbf{P}_0^{\text{opt}})$ is defined by

- $S_0^{\text{opt}} = \{s_0, s_{\ominus}, s_{\ominus}^{\text{opt}}, s_{\perp}^0\}$;
- let $a \in \text{En}(s_0)$ be such that $\mathbf{P}^{\text{opt}}(s_0, a, s_0) > 0$, then for every $s \in \{s_0, s_{\ominus}, s_{\ominus}^{\text{opt}}\}$, we set $\mathbf{P}_0^{\text{opt}}(s_0, a, s) = \mathbf{P}^{\text{opt}}(s_0, a, s)$ and $\mathbf{P}_0^{\text{opt}}(s_0, a, s_{\perp}^0) = \sum_{s \notin \{s_0, s_{\ominus}, s_{\ominus}^{\text{opt}}\}} \mathbf{P}^{\text{opt}}(s_0, a, s)$;
- let $a \in \text{En}(s_0)$ be such that $\mathbf{P}^{\text{opt}}(s_0, a, s_0) = 0$: then we set $\mathbf{P}_0^{\text{opt}}(s_0, a, s_{\perp}^0) = 1$.

Notice that all transitions leaving S_0^{opt} are directed to s_{\perp}^0 .

For $n \geq 1$, the MDP $\mathcal{M}_n^{\text{opt}} = (S_n^{\text{opt}}, \text{Act}, \mathbf{P}_n^{\text{opt}})$ is defined inductively as follows:

- $S_n^{\text{opt}} = (S_{n-1}^{\text{opt}} \setminus \{s_{\perp}^{n-1}\}) \cup \{s' \in S \mid \exists s \in S_{n-1}^{\text{opt}} \exists a \in \text{Act} \text{ s.t. } \mathbf{P}^{\text{opt}}(s, a, s') > 0\} \cup \{s_{\perp}^n\}$;
- for all $s, s' \in S_{n-1}^{\text{opt}} \cap S_n^{\text{opt}}$, for all $a \in \text{Act}$, $\mathbf{P}_n^{\text{opt}}(s, a, s') = \mathbf{P}_{n-1}^{\text{opt}}(s, a, s')$;
- for all $s \in S_{n-1}^{\text{opt}} \cap S_n^{\text{opt}}$, for all $s' \in S_n^{\text{opt}} \setminus (S_{n-1}^{\text{opt}} \cup \{s_{\perp}^n\})$, $\mathbf{P}_n^{\text{opt}}(s, a, s') = \mathbf{P}^{\text{opt}}(s, a, s')$;
- for all $s \in S_n^{\text{opt}} \setminus S_{n-1}^{\text{opt}}$, if $\mathbf{P}^{\text{opt}}(s, a, (S_n^{\text{opt}} \setminus \{s_{\perp}^n\})) > 0$, then for every $s' \in S_n^{\text{opt}} \setminus \{s_{\perp}^n\}$, $\mathbf{P}_n^{\text{opt}}(s, a, s') = \mathbf{P}^{\text{opt}}(s, a, s')$ and $\mathbf{P}_n^{\text{opt}}(s, a, s_{\perp}^n) = \sum_{s' \notin S_n^{\text{opt}}} \mathbf{P}^{\text{opt}}(s, a, s')$;
- for all $s \in S_n^{\text{opt}} \setminus S_{n-1}^{\text{opt}}$, if $a \in \text{En}(s)$ is such that $\mathbf{P}^{\text{opt}}(s, a, (S_n^{\text{opt}} \setminus \{s_{\perp}^n\})) = 0$, then, $\mathbf{P}_n^{\text{opt}}(s, a, s_{\perp}^n) = 1$.