Formal Methods in System Design (2022) 60:33–86
https://doi.org/10.1007/s10703-021-00364-6

Markov automata with multiple objectives

Tim Quatmann[1] · Sebastian Junges[2] · Joost-Pieter Katoen[1]

Received: 18 September 2019 / Accepted: 10 February 2021 / Published online: 29 March 2021

Abstract
Markov automata combine probabilistic branching, exponentially distributed delays and nondeterminism. This compositional variant of continuous-time Markov decision processes is used in reliability engineering, performance evaluation and stochastic scheduling. Their verification so far focused on single objectives such as (timed) reachability, and expected costs. In practice, often the objectives are mutually dependent and the aim is to reveal trade-offs. We present algorithms to analyze several objectives simultaneously and approximate Pareto curves. This includes, e.g., several (timed) reachability objectives, or various expected cost objectives. We also consider combinations thereof, such as on-time-within-budget objectives—which policies guarantee reaching a goal state within a deadline with at least probability p while keeping the allowed average costs below a threshold? We adopt existing approaches for classical Markov decision processes. The main challenge is to treat policies exploiting state residence times, even for *un*timed objectives. Experimental results show the feasibility and scalability of our approach.

Keywords Markov automata · Decision support · Continuous-time Markov decision processes · Multi-objective · Probabilistic model checking

1 Introduction

Markov automata [24,26] extend labeled transition systems with probabilistic branching and exponentially distributed delays. They are a compositional variant of continuous-time Markov decision processes (CTMDPs), in a similar vein as Segala's probabilistic automata extend classical MDPs. Transitions of a Markov automaton (MA) lead from states to probability distributions over states, and are either labeled with actions (allowing for interaction) or

✉ Sebastian Junges
sjunges@berkeley.edu

Tim Quatmann
tim.quatmann@cs.rwth-aachen.de

Joost-Pieter Katoen
katoen@cs.rwth-aachen.de

1 RWTH Aachen University, Aachen, Germany

2 University of California, Berkeley, CA, USA

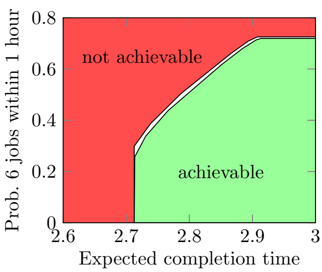
real numbers (rates of exponential distributions). MAs are used in reliability engineering [8, 51], hardware design [19], data-flow computation [38], dependability [9] and performance evaluation [25], as MAs are a natural semantic framework for modeling formalisms such as AADL, dynamic fault trees, stochastic Petri nets, stochastic activity networks, SADF etc. The verification of MAs so far focused on single objectives such as reachability, timed reachability, expected costs, and long-run averages [2,14,16,30,31,34,35]. These analyses cannot treat objectives that are mutually influencing each other. The aim of this paper is to analyze *multiple* objectives on MAs at once, facilitating *trade-off analysis* by approximating Pareto curves.

Consider the stochastic job scheduling problem of [12]: perform n jobs with exponential service times on k identical processors allowing pre-emptive scheduling. Once a job finishes, all k processors can be assigned any of the m remaining jobs. When $n-m$ jobs are finished, this yields $\binom{m}{k}$ non-deterministic choices. The largest-expected-service-time-first-policy is optimal to minimize the expected time to complete all jobs [12]. It is unclear how to schedule when imposing *extra* constraints, e.g., requiring a high probability to finish a batch of c jobs within a tight deadline (to accelerate their post-processing), or having a low average waiting time. These *multiple objectives* involve non-trivial *trade-offs*. Our algorithms analyze such trade-offs. Figure 1, e.g., shows the obtained result for 12 jobs and 3 processors. It approximates the set of points (p_1, p_2) for schedules achieving that (1) the expected time to complete all jobs is at most p_1 and (2) the probability to finish half of the jobs within an hour is at least p_2.

This paper extends an earlier paper [44] and presents techniques to verify MAs with multiple objectives. The contribution is threefold. First and foremost, we generalize the analysis of (classical, discrete-time) MDPs with multiple objectives to MA. Moreover, we remove some restrictions present in the classical analysis of MDPs with multiple objectives. Third, we provide an efficient implementation. We detail the contributions below.

Regarding the analysis of MA, we consider multiple (un)timed reachability and expected reward objectives as well as their combinations. Put shortly, we reduce all these problems to instances of multi-objective verification problems on classical MDPs. For multi-objective queries involving (combinations of) untimed reachability and expected reward objectives, corresponding algorithms on the *underlying* MDP can be used. In this case, the MDP is simply obtained by ignoring the timing information, see Fig. 2b. The crux is in relating MA schedulers—that can exploit state sojourn times to optimize their decisions—to MDP schedulers. For multiple timed reachability objectives, *digitization* [30,35] is employed to obtain an MDP, see Fig. 2c. The key is to mimic sojourn times by self-loops with appropriate probabilities. This provides a sound arbitrary close approximation of the timed behavior and also allows to combine timed reachability objectives with other types of objectives. The main

Fig. 1 Approx. Pareto curve for stochastic job scheduling



contribution is to show that digitization is sound for *all* possible MA schedulers. This requires a new proof strategy as the existing ones are tailored to single objectives.

We extend approaches for classical multi-objective MDP to support the multi-objective analysis of digitization MDPs. The extension embeds single-objective MA techniques from [35] into the multi-objective framework of [28]. Furthermore, we allow the simultaneous analysis of minimizing and maximizing expected reward objectives, lifting a restriction imposed in [28] by applying additional preprocessing steps that eliminate problematic end components.

Experiments on instances of four MA benchmarks show encouraging results. Multiple untimed reachability and expected reward objectives can be efficiently treated for models with millions of states. As for single objectives [30], timed reachability is more expensive. Our implementation is competitive to `PRISM` for multi-objective MDPs [28,39] and to `IMCA` [30], which implements [35], for single-objective MAs.

Compared with the version in [44], we (1) provide all technical ideas regarding the analysis of Markov automata, and provide formal proofs that previously were omitted. Furthermore, (2) the description of the existing multi-objective MDP model checking makes the paper self-contained, and allows us to give (3) the necessary extensions that allow, e.g., mixing upper and lower bounds on the different objectives. While such extensions may be considered minor, they are practically relevant and often intricate.

Related work Multi-objective decision making for MDPs with discounting and long-run objectives has been well investigated; for a recent survey, see [47]. Etessami et al. [27] consider verifying finite MDPs with multiple ω-regular objectives. Other multiple objectives include expected rewards under worst-case reachability [13,29], reward-bounded reachability [33], quantiles and conditional probabilities [5,33], mean pay-offs and stability [11], long-run objectives [7,10], total average discounted rewards under PCTL [49], and stochastic shortest path objectives [46]. This has been extended to MDPs with unknown cost function [37], MDPs under a restricted class of policies [23], infinite-state MDPs [20] arising from two-player timed games in a stochastic environment, and stochastic two-player games [18]. To the best of our knowledge, this is the first work on multi-objective MDPs extended with *random timing*.

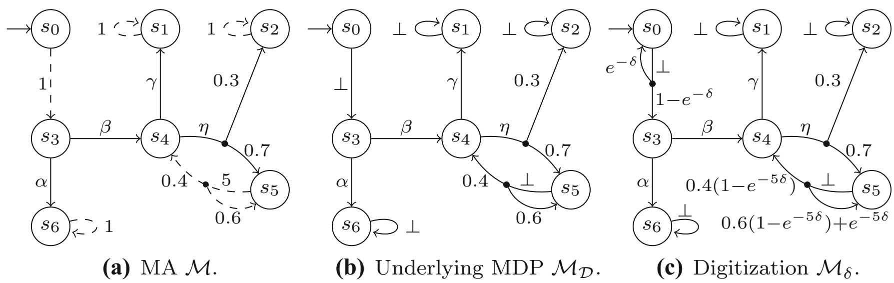


Fig. 2 MA $\mathcal{M}$ with underlying MDP $\mathcal{M}_\mathcal{D}$ and digitization $\mathcal{M}_\delta$

2 Preliminaries

Notations The set of real numbers is denoted by $\mathbb{R}$, and we write $\mathbb{R}_{>0} = \{x \in \mathbb{R} \mid x > 0\}$ and $\mathbb{R}_{\geq 0} = \mathbb{R}_{>0} \cup \{0\}$. For a finite set S, $Dist(S)$ denotes the set of probability distributions over S. $\mu \in Dist(S)$ is *Dirac* if $\mu(s) = 1$ for some $s \in S$.

2.1 Models

Markov automata generalize both Markov decision processes (MDPs) and continuous time Markov chains (CTMCs). They are extended with rewards (or, equivalently, costs) to allow modelling, e.g., energy consumption.

Definition 1 (Markov automaton) A *Markov automaton (MA)* is a tuple $\mathcal{M} = (S, Act, \rightarrow, s_0, (\rho_1, \ldots, \rho_\ell))$ where S is a finite set of *states* with *initial state* $s_0 \in S$, Act is a finite set of *actions* with $\bot \in Act$ and $Act \cap \mathbb{R}_{\geq 0} = \emptyset$,

- $\rightarrow \subseteq S \times (Act \cup \mathbb{R}_{>0}) \times Dist(S)$ is a set of *transitions* such that for all $s \in S$ there are no two transitions $(s, \lambda, \mu), (s, \lambda', \mu') \in \rightarrow$ with $\lambda, \lambda' \in \mathbb{R}_{>0}$, and
- $\rho_1, \ldots, \rho_\ell$ with $\ell \geq 0$ are *reward functions* $\rho_i : S \cup (S \times Act) \rightarrow \mathbb{R}_{\geq 0}$.

In the remainder of the article, let $\mathcal{M} = (S, Act, \rightarrow, s_0, (\rho_1, \ldots, \rho_\ell))$ denote an MA. Below, we introduce some notation and introduce some standard assumptions on MA. A transition $(s, \gamma, \mu) \in \rightarrow$, denoted by $s \xrightarrow{\gamma} \mu$, is called *probabilistic* if $\gamma \in Act$ and *Markovian* if $\gamma \in \mathbb{R}_{>0}$. In the latter case, γ is the rate of an exponential distribution, modeling a time-delayed transition. Probabilistic transitions fire instantaneously. The successor state is determined by μ, i.e., we move to s' with probability $\mu(s')$. A reward function ρ_i defines *state rewards* and *action rewards*. When sojourning in a state s for t time units , the state reward $\rho_i(s) \cdot t$ is obtained. Upon taking a transition $s \xrightarrow{\gamma} \mu$, we collect action reward $\rho_i(s, \gamma)$ (if $\gamma \in Act$) or $\rho(s, \bot)$ (if $\gamma \in \mathbb{R}_{>0}$).

Example 1 Figure 2a shows an MA $\mathcal{M}$ with seven states and no rewards. We draw direct edges for Dirac distributions, and Markovian transitions are illustrated by dashed arrows. For example: State s_0 has an outgoing Markovian transition with rate 1 to a distribution μ, where $\mu(s_3) = 1$. From state s_4 there are two probabilistic transitions labelled with two different actions γ and η. For $s_4 \xrightarrow{\eta} \mu$, μ is given by $\mu(s_5) = 0.7$ and $\mu(s_2) = 0.3$.

Probabilistic (Markovian) states PS (MS) have an outgoing probabilistic (Markovian) transition, respectively:

$$\text{PS} = \{s \in S \mid s \xrightarrow{\alpha} \mu, \alpha \in Act\},$$

$$\text{MS} = \{s \in S \mid s \xrightarrow{\lambda} \mu, \lambda \in \mathbb{R}_{>0}\}.$$

We exclude terminal states $s \notin \text{PS} \cup \text{MS}$ by adding a (Markovian) self-loop to all these states. As standard for MAs [24,26], we impose the *maximal progress assumption*, i.e., probabilistic transitions take precedence over Markovian ones as the probability to fire a Markovian transition instantly is zero. Thus, we assume that states are either probabilistic *or* Markovian: $S = \text{PS} \cup \text{MS}$. For Markovian states, we define the *exit rate* $\text{E}(s)$ of $s \in \text{MS}$ by $s \xrightarrow{\text{E}(s)} \mu$.

Example 2 Reconsider $\mathcal{M}$ in Fig. 2a. The states s_0, s_1, s_2, s_5 and s_6 are Markovian. All other states are probabilistic. The exit rate of s_0 is 1, the exit rate of s_5 is 5.

We assume *action-deterministic* MAs: $|\{\mu \in \mathit{Dist}(S) \mid s \xrightarrow{\alpha} \mu\}| \leq 1$ holds for all $s \in S$ and $\alpha \in \mathit{Act}$. For probabilistic states, this can be achieved by renaming. For Markovian states, this follows from the definition of Markov automata[1]. The *transition probabilities* of $\mathcal{M}$ are given by the function $\mathbf{P}\colon S \times \mathit{Act} \times S \to [0, 1]$ satisfying

$$\mathbf{P}(s, \alpha, s') = \begin{cases} \mu(s') & \text{if either } s \xrightarrow{\alpha} \mu \text{ or } \big(\alpha = \bot \text{ and } s \xrightarrow{\mathrm{E}(s)} \mu\big) \\ 0 & \text{otherwise.} \end{cases}$$

The value $\mathbf{P}(s, \alpha, s')$ corresponds to the probability to move from s with action α to s'. The *enabled actions* at state s are given by $\mathit{Act}(s) = \{\alpha \in \mathit{Act} \mid \exists s' \in S\colon \mathbf{P}(s, \alpha, s') > 0\}$.

Example 3 Reconsider $\mathcal{M}$ in Fig. 2a. The enabled actions in s_4 are γ and η. The probability $\mathbf{P}(s_4, \eta, s_2) = 0.3$, while $\mathbf{P}(s_5, \bot, s_4) = 0.4$.

Remark 1 In Sect. 4.3 we make one additional (but standard) assumption regarding the presence of Zeno-loops. We defer the discussion to that section.

Markov automata extend Markov decision processes by adding random timing. For concise notation, we define MDPs separately below.

Definition 2 (Markov decision process [43]) A *Markov decision process* (MDP) is a tuple $\mathcal{D} = (S, \mathit{Act}, \mathbf{P}, s_0, (\rho_1, \ldots, \rho_\ell))$, with $S, s_0, \mathit{Act}, \ell$ as in Def. 1, $\rho_1, \ldots, \rho_\ell$ are *action reward functions* $\rho_i : S \times \mathit{Act} \to \mathbb{R}_{\geq 0}$, and $\mathbf{P}\colon S \times \mathit{Act} \times S \to [0, 1]$ are the *transition probabilities* satisfying $\sum_{s' \in S} \mathbf{P}(s, \alpha, s') \in \{0, 1\}$ for all $s \in S$ and $\alpha \in \mathit{Act}$.

MDPs are MAs without Markovian states, i.e., $\mathrm{MS} = \emptyset$. Thus, MDPs exhibit probabilistic branching and non-determinism, but no random timing.

The reward $\rho(s, \alpha)$ is collected when taking action α at state s. We do not consider state rewards for MDPs. The *underlying MDP* of an MA abstracts away from its timing:

Definition 3 (Underlying MDP) For MA $\mathcal{M} = (S, \mathit{Act}, \to, s_0, (\rho_1, \ldots, \rho_\ell))$ with transition probabilities $\mathbf{P}$ the *underlying MDP of* $\mathcal{M}$ is the MDP $\mathcal{M}_{\mathcal{D}} = (S, \mathit{Act}, \mathbf{P}, s_0, (\rho_1^{\mathcal{D}}, \ldots, \rho_\ell^{\mathcal{D}}))$, where for each $i \in \{1, \ldots, \ell\}$:

$$\rho_i^{\mathcal{D}}(s, \alpha) = \begin{cases} \rho_i(s, \alpha) & \text{if } s \in \mathrm{PS}, \\ \rho_i(s, \bot) + \frac{1}{\mathrm{E}(s)} \cdot \rho_i(s) & \text{if } s \in \mathrm{MS} \text{ and } \alpha = \bot, \\ 0 & \text{otherwise.} \end{cases}$$

The reward functions $\rho_1^{\mathcal{D}}, \ldots, \rho_\ell^{\mathcal{D}}$ incorporate the action and state rewards of $\mathcal{M}$ where the state rewards are multiplied with the expected sojourn times $\frac{1}{\mathrm{E}(s)}$ of states $s \in \mathrm{MS}$.

Example 4 Figure 2 shows an MA $\mathcal{M}$ with its underlying MDP $\mathcal{M}_{\mathcal{D}}$.

1 Multiple outgoing Markovian transition could be reduced to a single Markovian transition by taking a weighted sum.

Paths and schedulers Paths represent runs of $\mathcal{M}$ starting in the initial state. Let $t(\kappa) = 0$ and $\alpha(\kappa) = \kappa$, if $\kappa \in Act$, and $t(\kappa) = \kappa$ and $\alpha(\kappa) = \bot$, if $\kappa \in \mathbb{R}_{\geq 0}$.

Definition 4 (Infinite path) An *infinite path* of MA $\mathcal{M}$ with transition probabilities $\mathbf{P}$ is an infinite sequence $\pi = s_0 \xrightarrow{\kappa_0} s_1 \xrightarrow{\kappa_1} \dots$ of states $s_0, s_1, \dots \in S$ and *stamps* $\kappa_0, \kappa_1, \dots \in Act \cup \mathbb{R}_{\geq 0}$ such that (1) $\sum_{i=0}^{\infty} t(\kappa_i) = \infty$, and for any $i \geq 0$ it holds that (2) $\mathbf{P}(s_i, \alpha(\kappa_i), s_{i+1}) > 0$, (3) $s_i \in \text{PS}$ implies $\kappa_i \in Act$, and (4) $s_i \in \text{MS}$ implies $\kappa_i \in \mathbb{R}_{\geq 0}$.

An infix $s_i \xrightarrow{\kappa_i} s_{i+1}$ of a path π represents that we stay at s_i for $t(\kappa_i)$ time units and then perform action $\alpha(\kappa_i)$ and move to state s_{i+1}. Condition (1) excludes Zeno paths, condition (2) ensures positive transition probabilities, and conditions (3) and (4) assert that stamps κ_i match the transition type at s_i.

A *finite path* is a finite prefix $\pi' = s_0 \xrightarrow{\kappa_0} \dots \xrightarrow{\kappa_{n-1}} s_n$ of an infinite path. The *length* of π' is $|\pi'| = n$, its *last state* is $last(\pi') = s_n$, and the *time duration* is $T(\pi') = \sum_{0 \leq i < |\pi'|} t(\kappa_i)$. We denote the sets of finite and infinite paths of $\mathcal{M}$ by $FPaths^{\mathcal{M}}$ and $IPaths^{\mathcal{M}}$, respectively. The superscript $\mathcal{M}$ is omitted if the model is clear from the context. For a finite or infinite path $\pi = s_0 \xrightarrow{\kappa_0} s_1 \xrightarrow{\kappa_1} \dots$ the *prefix* of π of length n is denoted by $pref(\pi, n)$. The ith state visited by π is given by $\pi[i] = s_i$. The *time-abstraction* $\text{ta}(\pi)$ of π removes all sojourn times and is a path of the underlying MDP $\mathcal{M}_{\mathcal{D}}$: $\text{ta}(\pi) = s_0 \xrightarrow{\alpha(\kappa_0)} s_1 \xrightarrow{\alpha(\kappa_1)} \dots$. Paths of $\mathcal{M}_{\mathcal{D}}$ are also referred to as the *time-abstract paths of* $\mathcal{M}$.

Definition 5 (Generic scheduler) A *generic scheduler* for $\mathcal{M}$ is a measurable function $\sigma\colon FPaths \times Act \to [0, 1]$ such that $\sigma(\pi, \cdot) \in Dist(Act(last(\pi)))$ for each $\pi \in FPaths$.

A scheduler σ for $\mathcal{M}$ resolves the non-determinism of $\mathcal{M}$: $\sigma(\pi, \alpha)$ is the probability to take transition $last(\pi) \xrightarrow{\alpha} \mu$ after observing the run π. The set of such schedulers is denoted by $\text{GM}^{\mathcal{M}}$ (GM if $\mathcal{M}$ is clear from the context). $\sigma \in \text{GM}$ is *deterministic* if the distribution $\sigma(\pi, \cdot)$ is Dirac for any π. *Time-abstract schedulers* behave independently of the time-stamps of the given path, i.e., $\sigma(\pi, \alpha) = \sigma(\pi', \alpha)$ for all actions α and paths π, π' with $\text{ta}(\pi) = \text{ta}(\pi')$. We write $\text{TA}^{\mathcal{M}}$ to denote the set of time-abstract schedulers of $\mathcal{M}$. GM is the most general scheduler class for MAs, and TA is the most general for MDPs.

2.2 Objectives

An objective $\mathbb{O}_i$ is a representation of a *quantitative* property like the probability to reach an error state, or the expected energy consumption. To express *Boolean* properties (e.g., the probability to reach an error state is below p_i), $\mathbb{O}_i$ is combined with a *threshold* $\rhd_i\ p_i$ where $\rhd_i \in \{<, \leq, >, \geq\}$ is a *threshold relation* and $p_i \in \mathbb{R}$ is a *threshold value*. Let $\mathcal{M}, \sigma \models \mathbb{O}_i \rhd_i p_i$ denote that the MA $\mathcal{M}$ under scheduler $\sigma \in \text{GM}$ satisfies the property $\mathbb{O}_i \rhd_i p_i$. We consider reachability objectives and expected reward objectives. In the remainder of this section, we formalize these standard notions.

2.2.1 Probability measure

Given a scheduler $\sigma \in \text{GM}$, a probability measure $\Pr_{\sigma}^{\mathcal{M}}$ on measurable sets of infinite paths is defined, which generalizes both the standard probability measure on MDPs and on CTMCs. We briefly sketch the definition of $\Pr_{\sigma}^{\mathcal{M}}$. More information can be found in, e.g., [35,40].

We first consider the probability measure $\mathrm{Pr}^{Steps}_{\sigma,\pi}$ for transition steps. We let $\mathfrak{B}(\mathbb{R}_{\geq 0})$ denote the Borel $\boldsymbol{\sigma}$-algebra on $\mathbb{R}_{\geq 0}$ and define the $\boldsymbol{\sigma}$-algebra for transition steps $\mathfrak{F}^{Steps} = \boldsymbol{\sigma}(\mathfrak{B}(\mathbb{R}_{\geq 0}) \times 2^{Act} \times 2^{S})$ using the Cartesian product of $\boldsymbol{\sigma}$-algebra [1]. For a finite path $\pi \in FPaths$ with $s = last(\pi)$ and a measurable set of transition steps $T \in \mathfrak{F}^{Steps}$, the probability of $T \subseteq \mathbb{R}_{\geq 0} \times Act \times S$ is

$$\mathrm{Pr}^{Steps}_{\sigma,\pi}(T) = \begin{cases} \sum_{(0,\alpha,s')\in T} \sigma(\pi,\alpha) \cdot \mathbf{P}(s,\alpha,s') & \text{if } s \in \mathrm{PS} \\ \int_{\{t \mid (t,\perp,s')\in T\}} \mathrm{E}(s) \cdot e^{-\mathrm{E}(s)t} \cdot \sum_{(t,\perp,s')\in T} \mathbf{P}(s,\perp,s')\,\mathrm{d}t & \text{if } s \in \mathrm{MS}. \end{cases}$$

Next, we define the probability measure Pr^{n}_{σ} for $n \in \mathbb{N}$ transition steps over the $\boldsymbol{\sigma}$-algebra $\mathfrak{F}^{n} = \boldsymbol{\sigma}\big(2^{S} \times (\times_{i=1}^{n} \mathfrak{F}^{Steps})\big)$. For $S' \subseteq S$ we set $\mathrm{Pr}^{0}_{\sigma}(S') = 1$ if $s_0 \in S'$ and $\mathrm{Pr}^{0}_{\sigma}(S') = 0$ otherwise. Moreover, given $n > 0$, $\Pi \in \mathfrak{F}^{n-1}$ and $T \in \mathfrak{F}^{Steps}$ we set $\Pi \circ T = \{(\pi,\ (t,\alpha,s')) \mid \pi \in \Pi, (t,\alpha,s') \in T\}$ and

$$\mathrm{Pr}^{n}_{\sigma}(\Pi \circ T) = \int_{\pi\in\Pi\cap FPaths} \mathrm{Pr}^{Steps}_{\sigma,\pi}(T)\,\mathrm{d}\mathrm{Pr}^{n-1}_{\sigma}(\{\pi\})\ .$$

Following standard constructions (e.g., [1,40]), we obtain $\mathrm{Pr}^{\mathcal{M}}_{\sigma}$ by lifting Pr^{n}_{σ} first to measurable sets of *finite* paths given by the $\boldsymbol{\sigma}$-algebra $\mathfrak{F}^{*} = \bigcup_{n=0}^{\infty} \mathfrak{F}^{n}$ and finally to measurable sets of infinite paths given by the $\boldsymbol{\sigma}$-algebra $\mathfrak{F}^{\omega}$. The latter is the smallest $\boldsymbol{\sigma}$-algebra containing the *cylinder set* $Cyl(\Pi)$ of all measurable $\Pi \in \mathfrak{F}^{*}$, where

$$Cyl(\Pi) = \{\pi \xrightarrow{\kappa_n} s_{n+1} \xrightarrow{\kappa_{n+1}} \cdots \in IPaths^{\mathcal{M}} \mid \pi \in \Pi\}.$$

Let $\Pi \in \mathfrak{F}^{*}$ and $\pi \in IPaths$. For simplicity, we may write $\mathrm{Pr}^{\mathcal{M}}_{\sigma}(\Pi)$ instead of $\mathrm{Pr}^{\mathcal{M}}_{\sigma}(Cyl(\Pi))$ and $\mathrm{Pr}^{\mathcal{M}}_{\sigma}(\pi)$ instead of $\mathrm{Pr}^{\mathcal{M}}_{\sigma}(\{\pi\})$. Note that $\mathrm{Pr}^{\mathcal{M}}_{\sigma}(\{\pi\}) = 0$ if π contains one or more Markovian transitions.

For some $\Lambda \in \mathfrak{F}^{\omega}$ with $\mathrm{Pr}^{\mathcal{M}}_{\sigma}(\Lambda) > 0$ we consider the *conditional probability measure* $\mathrm{Pr}^{\mathcal{M}}_{\sigma}(\cdot \mid \Lambda)\colon \mathfrak{F}^{\omega} \to [0,1]$, where for $\Pi \in \mathfrak{F}^{\omega}$:

$$\mathrm{Pr}^{\mathcal{M}}_{\sigma}(\Pi \mid \Lambda) = \frac{\mathrm{Pr}^{\mathcal{M}}_{\sigma}(\Pi \cap \Lambda)}{\mathrm{Pr}^{\mathcal{M}}_{\sigma}(\Lambda)}.$$

2.2.2 Reachability objectives

$I \subseteq \mathbb{R}$ is a *time interval* if it is of the form $I = [a,b]$ or $I = [a,\infty)$, where $0 \leq a < b$. The set of paths reaching a set of goal states $G \subseteq S$ in time I is defined as

$$\begin{aligned} \Diamond^{I} G =& \{\pi = s_0 \xrightarrow{\kappa_0} s_1 \xrightarrow{\kappa_1} \cdots \in IPaths \mid \exists n \geq 0\colon \pi[n] \in G \text{ and} \\ & I \cap [t, t + t(\kappa_n)] \neq \emptyset \text{ for } t = T(pref(\pi,n))\}. \end{aligned}$$

$\Diamond^{I}$ is measurable as it can be expressed via countable unions of measurable cylinder sets. We write $\Diamond G$ instead of $\Diamond^{[0,\infty)} G$.

We formulate reachability objectives as follows:

Definition 6 (Reachability objective) A *reachability objective* has the form $\mathbb{P}(\Diamond^{I} G)$ for time interval I and goal states G. The objective is *timed* if $I \neq [0,\infty)$ and *untimed* otherwise. For MA $\mathcal{M}$ and scheduler $\sigma \in \mathrm{GM}$, let

$$\mathcal{M}, \sigma \models \mathbb{P}(\Diamond^{I} G) \rhd_i p_i \quad \text{iff} \quad \mathrm{Pr}^{\mathcal{M}}_{\sigma}(\Diamond^{I} G) \rhd_i p_i\,.$$

2.2.3 Expected reward objectives

Expected rewards define the expected amount of reward collected (w.r.t. some ρ_j) until a goal state in $G \subseteq S$ is reached, generalizing the notion on CTMCs and MDPs. More precisely, we fix a reward function ρ of the MA $\mathcal{M}$. The reward of a finite path $\pi' = s_0 \xrightarrow{\kappa_0} \ldots \xrightarrow{\kappa_{n-1}} s_n \in$ *FPaths* is given by

$$rew^{\mathcal{M}}(\rho, \pi') = \sum_{i=0}^{|\pi'|-1} \rho(s_i) \cdot t(\kappa_i) + \rho(s_i, \alpha(\kappa_i)).$$

Intuitively, $rew^{\mathcal{M}}(\rho, \pi')$ is the sum over the rewards obtained in every step $s_i \xrightarrow{\kappa_i}$ in the path π'. The reward obtained in step i is composed of the state reward of s_i multiplied with the sojourn time $t(\kappa_i)$ plus the action reward given by s_i and $\alpha(\kappa_i)$. State rewards assigned to probabilistic states do not affect the reward of a path as the sojourn time in such states is zero.

For an infinite path $\pi = s_0 \xrightarrow{\kappa_0} s_1 \xrightarrow{\kappa_1} \cdots \in$ *IPaths*, the reward of π up to a set of goal states $G \subseteq S$ is given by

$$rew^{\mathcal{M}}(\rho, \pi, G) = \begin{cases} rew^{\mathcal{M}}(\rho, pref(\pi, n)) & \text{if } n = \min\{i \geq 0 \mid s_i \in G\} \\ \lim_{n\to\infty} rew^{\mathcal{M}}(\rho, pref(\pi, n)) & \text{if } s_i \notin G \text{ for all } i \geq 0\ . \end{cases}$$

Intuitively, we stop collecting reward as soon as π reaches a state in G. If no state in G is reached, reward is accumulated along the infinite path, which potentially yields an infinite reward. The function $rew^{\mathcal{M}}(\rho, \cdot, G)\colon IPaths^{\mathcal{M}} \to \mathbb{R}_{\geq 0}$ is measurable. Its expected value with respect to scheduler $\sigma \in \mathrm{GM}$ is called the *expected reward* $\mathrm{ER}_{\sigma}^{\mathcal{M}}(\rho, G)$, i.e.,

$$\mathrm{ER}_{\sigma}^{\mathcal{M}}(\rho, G) = \int_{\pi \in IPaths^{\mathcal{M}}} rew^{\mathcal{M}}(\rho, \pi, G)\, \mathrm{d}\mathrm{Pr}_{\sigma}^{\mathcal{M}}(\pi).$$

With this definition, we formulate expected reward objectives as follows:

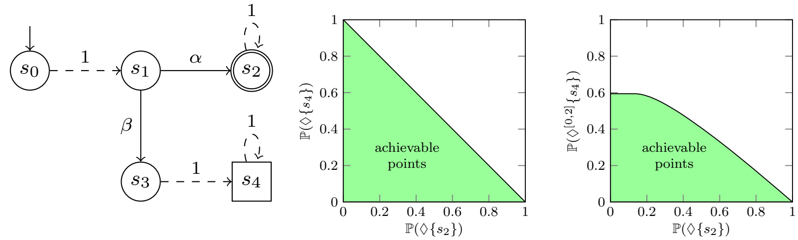
Definition 7 (Expected reward objective) An *expected reward objective* has the form $\mathbb{E}(\#j, G)$ where j is the index of reward function ρ_j and $G \subseteq S$. For MA $\mathcal{M}$ and scheduler $\sigma \in \mathrm{GM}$, let

$$\mathcal{M}, \sigma \models \mathbb{E}(\#j, G) \rhd_i p_i \quad \text{iff} \quad \mathrm{ER}_{\sigma}^{\mathcal{M}}(\rho_j, G) \rhd_i p_i.$$

Expected *time* objectives $\mathbb{E}(T, G)$ are expected reward objectives that consider the reward function ρ_T with $\rho_T(s) = 1$ if $s \in \mathrm{MS}$ and all other rewards are zero.

3 Multi-objective model checking

Standard model checking considers objectives individually. This approach is not feasible when we are interested in multiple objectives that should be fulfilled by the same scheduler, e.g., a scheduler that maximizes the expected profit might violate certain safety constraints. *Multi-objective* model checking aims to analyze multiple objectives at once and reveals potential trade-offs.



(a) MA $\mathcal{M}$. **(b)** Untimed objectives. **(c)** Timed objectives.

Fig. 3 Markov automaton and achievable points

Definition 8 (Satisfaction of multiple objectives) Let $\mathcal{M}$ be an MA and $\sigma \in GM$. For objectives $\mathbb{O} = (\mathbb{O}_1, \dots, \mathbb{O}_d)$ with threshold relations $\rhd = (\rhd_1, \dots, \rhd_d) \in \{<, \leq, >, \geq\}^d$ and threshold values $\mathbf{p} = (p_1, \dots, p_d) \in \mathbb{R}^d$ let

$$\mathcal{M}, \sigma \models \mathbb{O} \rhd \mathbf{p} \iff \mathcal{M}, \sigma \models \mathbb{O}_i \rhd_i p_i \text{ for all } 1 \leq i \leq d.$$

Furthermore, let $\textit{achieve}^{\mathcal{M}}(\mathbb{O} \rhd \mathbf{p}) \iff \exists \sigma \in GM$ such that $\mathcal{M}, \sigma \models \mathbb{O} \rhd \mathbf{p}$.

If $\mathcal{M}, \sigma \models \mathbb{O} \rhd \mathbf{p}$, the point $\mathbf{p} \in \mathbb{R}^d$ is *achievable* in $\mathcal{M}$ with scheduler σ. The *set of achievable points* of $\mathcal{M}$ w.r.t. $\mathbb{O}$ and $\rhd$ is $\{\mathbf{p} \in \mathbb{R}^d \mid \textit{achieve}^{\mathcal{M}}(\mathbb{O} \rhd \mathbf{p})\}$. We straightforwardly also apply these definitions to MDPs. This is compatible with the notions on multi-objective MDPs as given in [27,28].

Example 5 Figure 3b depicts the set of achievable points of the MA $\mathcal{M}$ from Fig. 3a w.r.t. relations $\rhd = (\geq, \geq)$ and objectives $(\mathbb{P}(\lozenge\{s_2\}), \mathbb{P}(\lozenge\{s_4\}))$. Likewise, Fig. 3c depicts the achievable points for $\mathcal{M}, \rhd$, and $(\mathbb{P}(\lozenge\{s_2\}), \mathbb{P}(\lozenge^{[0,2]}\{s_4\}))$. Using the set of achievable points, we can answer Pareto, numerical, and achievability queries as considered in [28], e.g., the Pareto front lies on the border of the set. We detail this computation in Sect. 5.

Schedulers For single-objective model checking on MAs, it suffices to consider deterministic schedulers [41]. For untimed reachability and expected rewards even *time-abstract* deterministic schedulers suffice [41]. Multi-objective model checking on MDPs requires history-dependent, randomized schedulers [27]. The following theorem and its proof show that schedulers on MA may also employ *timing* information to make optimal choices, even if only *untimed* objectives are considered. However, we show in Sect. 4 that for untimed objectives, such schedulers can always be converted to time-abstract schedulers (potentially considering the history and randomization). Put differently, timing information *can* be employed to achieve untimed objectives but it is not *necessary* to do so.

Theorem 1 *For some MA $\mathcal{M}$ with $\textit{achieve}^{\mathcal{M}}(\mathbb{O} \rhd \boldsymbol{p})$, no deterministic time-abstract scheduler σ satisfies $\mathcal{M}, \sigma \models \mathbb{O} \rhd \boldsymbol{p}$.*

Proof Consider the MA $\mathcal{M}$ in Fig. 3a with relations $\rhd = (\geq, \geq)$, objectives $\mathbb{O} = (\mathbb{P}(\lozenge\{s_2\}), \mathbb{P}(\lozenge\{s_4\}))$, and point $\mathbf{p} = (0.5, 0.5)$. We have $\textit{achieve}^{\mathcal{M}}(\mathbb{O} \rhd \mathbf{p})$: A simple graph argument yields that both properties are only satisfied if action α is taken with probability exactly a half. As the probability mass to stay in s_0 for at most $\ln(2)$ is exactly 0.5, a timed scheduler σ with $\sigma(s_0 \xrightarrow{t} s_1, \alpha) = 1$ if $t \leq \ln(2)$ and 0 otherwise does satisfy both objectives.

There are only two deterministic time abstract schedulers for $\mathcal{M}$:

$$\sigma_\alpha: \text{ always choose } \alpha \qquad \text{and} \qquad \sigma_\beta: \text{ always choose } \beta$$

and it holds that $\mathcal{M}, \sigma_\alpha \not\models \mathbb{P}(\Diamond\{s_4\}) \geq 0.5$ and $\mathcal{M}, \sigma_\beta \not\models \mathbb{P}(\Diamond\{s_2\}) \geq 0.5$. □

The geometric shape of the achievable points Like for MDPs [27], the set of achievable points of any combination of aforementioned objectives is convex.

Proposition 1 *The set* $\{\boldsymbol{p} \in \mathbb{R}^d \mid achieve^{\mathcal{M}}(\mathbb{O} \rhd \boldsymbol{p})\}$ *is convex.*

Proof Let $\mathcal{M}$ be an MA and let $\mathbb{O} = (\mathbb{O}_1, \ldots, \mathbb{O}_d)$ be objectives with relations $\rhd = (\rhd_1, \ldots, \rhd_d)$ and points $\mathbf{p}_1, \mathbf{p}_2 \in \mathbb{R}^d$ such that $achieve^{\mathcal{M}}(\mathbb{O} \rhd \mathbf{p}_1)$ and $achieve^{\mathcal{M}}(\mathbb{O} \rhd \mathbf{p}_2)$ holds. To show that the set of achievable points is convex, we need to show that for $w \in [0, 1]$ the point $\mathbf{p} = w \cdot \mathbf{p}_1 + (1 - w) \cdot \mathbf{p}_2$ is achievable as well, i.e., $achieve^{\mathcal{M}}(\mathbb{O} \rhd \mathbf{p})$. For $i \in 1, 2$, let $\sigma_i \in \text{GM}$ be a scheduler satisfying $\mathcal{M}, \sigma_i \models \mathbb{O} \rhd \mathbf{p}_i$. The point $\mathbf{p}$ is achievable with the scheduler σ that intuitively makes an initial one-off random choice:

- with probability w mimic σ_1 and
- with probability $1 - w$ mimic σ_2.

Specifying such a scheduler as a function $\sigma^w : FPaths \times Act \to [0, 1]$ as in Def. 5 is technically involved because σ^w can not memorize the outcome of the initial one-off random choice. For a path $\pi = s_0 \xrightarrow{\kappa_0} \ldots \xrightarrow{\kappa_{n-1}} s_n \in FPaths$ and $\alpha \in Act$, the value $\sigma^w(\pi, \alpha)$ needs to depend on the previously chosen actions $\alpha(\kappa_j)$ in π and the probability that these choices adhere to σ_1 and σ_2, respectively. Let $w_1 = w$ and $w_2 = 1 - w$. We set [2]

$$\sigma^w(\pi, \alpha) = \frac{\sum_{i=1}^{2} \Big(w_i \cdot \sigma_i(\pi, \alpha) \cdot \prod_{j=0}^{n-1} \sigma_i(\mathit{pref}(\pi, j), \alpha(\kappa_j)) \Big)}{\sum_{i=1}^{2} \Big(w_i \cdot \prod_{j=0}^{n-1} \sigma_i(\mathit{pref}(\pi, j), \alpha(\kappa_j)) \Big)}$$

In App. A we show that for any measurable $\Pi \subseteq IPaths$ we have $\Pr^{\mathcal{M}}_{\sigma^w}(\Pi) = w_1 \cdot \Pr^{\mathcal{M}}_{\sigma_1}(\Pi) + w_2 \cdot \Pr^{\mathcal{M}}_{\sigma_2}(\Pi)$. Lifting this observation to expected rewards is straightforward. It follows that σ^w achieves the point $\mathbf{p} = w_1 \cdot \mathbf{p}_1 + w_2 \cdot \mathbf{p}_2$. □

For MDPs, the set of achievable points is a convex polytope where the vertices can be realized by deterministic schedulers that use memory bounded by 2^d, where d is the number of objectives. As there are finitely many such schedulers, the polytope has a finite $\mathcal{V}$-representation [27], i.e., it can be represented by a finite number of vertices. This result does not carry over to MAs.

Theorem 2 *For some MA* $\mathcal{M}$ *and objectives* $\mathbb{O}$, *the polytope* $\{\boldsymbol{p} \in \mathbb{R}^d \mid achieve^{\mathcal{M}}(\mathbb{O} \rhd \boldsymbol{p})\}$ *does not have a finite* $\mathcal{V}$*-representation.*

Proof We show that the claim holds for the MA $\mathcal{M}$ in Fig. 3a with objectives $\mathbb{O} = (\mathbb{P}(\Diamond\{s_2\}), \mathbb{P}(\Diamond^{[0,2]}\{s_4\}))$ and relations $\rhd = (\geq, \geq)$. The insight here is that for any sojourn

[2] Our construction is roughly inspired by a construction in [10, Sect. 6], where schedulers for MDPs with stochastic memory-updates are considered. Lifting the approach of [10] to Markov automata is not obvious.

time $t \leq 2$ in s_0, the timing information is relevant for optimal schedulers: The shorter the sojourn time in s_0, the higher the probability to reach s_4 within the time bound. For the sake of contradiction assume that the polytope $A = \{\mathbf{p} \in \mathbb{R}^2 \mid achieve^{\mathcal{M}}(\mathbb{O} \rhd \mathbf{p})\}$ has a finite $\mathcal{V}$-representation. Then, there must be two distinct vertices $\mathbf{p}_1, \mathbf{p}_2$ of A such that $\{w \cdot \mathbf{p}_1 + (1 - w) \cdot \mathbf{p}_2 \mid w \in [0, 1]\}$ is a face of A. In particular, this means that $\mathbf{p} = 0.5 \cdot \mathbf{p}_1 + 0.5 \cdot \mathbf{p}_2$ is achievable but $\mathbf{p}_\varepsilon = \mathbf{p} + (0, \varepsilon)$ is not achievable for all $\varepsilon > 0$. We show that there is in fact an ε for which $\mathbf{p}_\varepsilon$ is achievable, contradicting our assumption that A has a finite $\mathcal{V}$-representation.

For $i \in 1, 2$, let $\sigma_i \in$ GM be a scheduler satisfying $\mathcal{M}, \sigma_i \models \mathbb{O} \rhd \mathbf{p}_i$. $\sigma_1 \neq \sigma_2$ has to hold as the schedulers achieve different vertices of A. The point $\mathbf{p}$ is achievable with the *randomized* scheduler σ that mimics σ_1 with probability 0.5 and mimics σ_2 otherwise. Consider $t = -\log(\Pr^{\mathcal{M}}_{\sigma}(\Diamond\{s_2\}))$ and the *deterministic* scheduler σ' given by

$$\sigma'(s_0 \xrightarrow{t_0} s_1, \alpha) = \begin{cases} 1 & \text{if } t_0 > t \\ 0 & \text{otherwise.} \end{cases}$$

σ' satisfies $\Pr^{\mathcal{M}}_{\sigma'}(\Diamond\{s_2\}) = e^{-t} = \Pr^{\mathcal{M}}_{\sigma}(\Diamond\{s_2\})$. Moreover, we have

$$\Pr^{\mathcal{M}}_{\sigma'}(\Diamond^{[0,t]}\{s_3\}) = \Pr^{\mathcal{M}}_{\sigma'}(\Diamond\{s_3\}) = \Pr^{\mathcal{M}}_{\sigma}(\Diamond\{s_3\}) > \Pr^{\mathcal{M}}_{\sigma}(\Diamond^{[0,t]}\{s_3\}),$$

where the last inequality is due to $\sigma \neq \sigma'$. While the probability to reach s_3 is equal under both schedulers, s_3 is reached earlier when σ' is considered. This increases the probability to reach s_4 in time, i.e., $\Pr^{\mathcal{M}}_{\sigma'}(\Diamond^{[0,2]}\{s_4\}) > \Pr^{\mathcal{M}}_{\sigma}(\Diamond^{[0,2]}\{s_4\})$. It follows that $\mathcal{M}, \sigma' \models \mathbb{O} \rhd \mathbf{p}_\varepsilon$ for some $\varepsilon > 0$. □

Since convex polytopes without a finite $\mathcal{V}$-representation cannot be represented by a finite number of vertices, any method extending the approach of [28]—which computes these vertices—can only approximate the set of achievable points.

Problem Statement: For an MA and objectives with threshold relations, construct arbitrarily tight over- and under-approximations of the achievable points.

4 Analysis of Markov automata with multiple objectives

The state-of-the-art in single-objective model checking of MA is to reduce the MA to an MDP, cf. [30,31,35], for which efficient algorithms exist. We aim to lift this approach to multi-objective model checking. Assume MA $\mathcal{M}$ and objectives $\mathbb{O}$ with threshold relations $\rhd$. We discuss how the set of achievable points of $\mathcal{M}$ relates to the set of achievable points of an MDP. The key challenge is to deal with timing information—even for *un*timed objectives—and to consider schedulers beyond those optimizing single objectives. We obtain:

- For untimed reachability and expected reward objectives, the achievable points of $\mathcal{M}$ *equal* those of its *underlying* MDP, cf. Theorems 3 and 4.
- For timed reachability objectives, the set of achievable points of a *digitized* MDP $\mathcal{M}_\delta$ provides a *sound approximation* of the achievable points of $\mathcal{M}$, cf. Theorem 5. Corollary 1 gives the precision of the approximation.

4.1 Untimed reachability objectives

Although timing information is essential for *deterministic* schedulers, cf. Theorem 1, timing information does not strengthen randomized schedulers:

Theorem 3 *For MA $\mathcal{M}$ and untimed reachability objectives $\mathbb{O}$ it holds that $achieve^{\mathcal{M}}(\mathbb{O} \rhd \boldsymbol{p}) \iff achieve^{\mathcal{M}_{\mathcal{D}}}(\mathbb{O} \rhd \boldsymbol{p})$.*

The main idea for proving Theorem 3 is to construct for scheduler $\sigma \in \mathrm{GM}^{\mathcal{M}}$ a time-abstract scheduler $\mathrm{ta}(\sigma) \in \mathrm{TA}^{\mathcal{M}_{\mathcal{D}}}$ such that they both induce the same untimed reachability probabilities. To this end, we discuss the connection between probabilities of paths of MA $\mathcal{M}$ and paths of MDP $\mathcal{M}_{\mathcal{D}}$.

Definition 9 (Induced paths of a time-abstract path) The set of *induced paths on MA $\mathcal{M}$* of a path $\hat{\pi}$ of $\mathcal{M}_{\mathcal{D}}$ is given by

$$\langle\hat{\pi}\rangle = \mathrm{ta}^{-1}(\hat{\pi}) = \{\pi \in \mathit{FPaths}^{\mathcal{M}} \cup \mathit{IPaths}^{\mathcal{M}} \mid \mathrm{ta}(\pi) = \hat{\pi}\}.$$

The set $\langle\hat{\pi}\rangle$ contains all paths of $\mathcal{M}$ where replacing sojourn times by $\perp$ yields $\hat{\pi}$.

For $\sigma \in \mathrm{GM}$, the probability distribution $\sigma(\pi, \cdot) \in \mathit{Dist}(\mathit{Act})$ might depend on the sojourn times of the path π. The time-abstract scheduler $\mathrm{ta}(\sigma)$ weights the distribution $\sigma(\pi, \cdot)$ with the probability masses of the paths $\pi \in \langle\hat{\pi}\rangle$.

Definition 10 (Time-abstraction of a scheduler) The time-abstraction of $\sigma \in \mathrm{GM}^{\mathcal{M}}$ is defined as $\mathrm{ta}(\sigma) \in \mathrm{TA}^{\mathcal{M}_{\mathcal{D}}}$ such that for any $\hat{\pi} \in \mathit{FPaths}^{\mathcal{M}_{\mathcal{D}}}$

$$\mathrm{ta}(\sigma)(\hat{\pi}, \alpha) = \begin{cases} \displaystyle\int_{\pi \in \langle\hat{\pi}\rangle} \sigma(\pi, \alpha)\, \mathrm{dPr}^{\mathcal{M}}_{\sigma}(\pi \mid \langle\hat{\pi}\rangle) & \text{if } \mathrm{Pr}^{\mathcal{M}}_{\sigma}(\langle\hat{\pi}\rangle) > 0 \\ 1 \,/\, |\mathit{Act}(\mathit{last}(\hat{\pi}))| & \text{otherwise.} \end{cases}$$

Intuitively, the term $\mathrm{Pr}^{\mathcal{M}}_{\sigma}(\pi \mid \langle\hat{\pi}\rangle)$ represents the probability for a path in $\langle\hat{\pi}\rangle$ to have sojourn times as given by π. The value $\mathrm{ta}(\sigma)(\hat{\pi}, \alpha)$ coincides with the probability that σ picks action α, given that the time-abstract path $\hat{\pi}$ was observed. If $\mathrm{Pr}^{\mathcal{M}}_{\sigma}(\langle\hat{\pi}\rangle) = 0$, the value for $\mathrm{ta}(\sigma)(\hat{\pi}, \alpha)$ is arbitrary. For simplicity, we picked a uniform choice.

Example 6 Consider the MA $\mathcal{M}$ in Fig. 2a and the scheduler σ choosing α at state s_3 iff the sojourn time at s_0 is at most one. Then $\mathrm{ta}(\sigma)(s_0 \xrightarrow{\perp} s_3, \alpha) = 1 - e^{-\mathrm{E}(s_0)}$, the probability that s_0 is left within one time unit. For $\bar{\pi} = s_0 \xrightarrow{\perp} s_3 \xrightarrow{\alpha} s_6$ we have

$$\mathrm{Pr}^{\mathcal{M}}_{\sigma}(\Diamond\{s_6\}) = \mathrm{Pr}^{\mathcal{M}}_{\sigma}(\langle\bar{\pi}\rangle) = 1 - e^{-\mathrm{E}(s_0)} = \mathrm{Pr}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\bar{\pi}) = \mathrm{Pr}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\Diamond\{s_6\}).$$

In the example, the considered scheduler and its time-abstraction induce the same untimed reachability probabilities. We generalize this observation.

Lemma 1 *For any $\hat{\pi} \in \mathit{FPaths}^{\mathcal{M}_{\mathcal{D}}}$ we have $Pr^{\mathcal{M}}_{\sigma}(\langle\hat{\pi}\rangle) = Pr^{\mathcal{M}_{\mathcal{D}}}_{ta(\sigma)}(\hat{\pi})$.*

Proof The proof is by induction over the length of the considered path $|\hat{\pi}| = n$. Let $\mathcal{M} = (S, \mathit{Act}, \rightarrow, s_0, (\rho_1, \ldots, \rho_\ell))$ and $\mathcal{M}_{\mathcal{D}} = (S, \mathit{Act}, \mathbf{P}, s_0, (\rho^{\mathcal{D}}_1, \ldots, \rho^{\mathcal{D}}_\ell))$. If $n = 0$, then $\{\hat{\pi}\} = \langle\hat{\pi}\rangle = \{s_0\}$. Hence, $\mathrm{Pr}^{\mathcal{M}}_{\sigma}(\langle\hat{\pi}\rangle) = 1 = \mathrm{Pr}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\hat{\pi})$. In the induction step, we assume that the lemma holds for a fixed path $\hat{\pi} \in \mathit{FPaths}^{\mathcal{M}_{\mathcal{D}}}$ with length $|\hat{\pi}| = n$ and $\mathit{last}(\hat{\pi}) = s$. Consider the path $\hat{\pi} \xrightarrow{\alpha} s' \in \mathit{FPaths}^{\mathcal{M}_{\mathcal{D}}}$. If $\mathrm{Pr}^{\mathcal{M}}_{\sigma}(\langle\hat{\pi}\rangle) = \mathrm{Pr}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\hat{\pi}) = 0$, then $\mathrm{Pr}^{\mathcal{M}}_{\sigma}(\langle\hat{\pi} \xrightarrow{\alpha} s'\rangle) = \mathrm{Pr}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\hat{\pi} \xrightarrow{\alpha} s') = 0$ because $\mathit{Cyl}(\langle\hat{\pi} \xrightarrow{\alpha} s'\rangle) \subseteq \mathit{Cyl}(\langle\hat{\pi}\rangle)$ and $\mathit{Cyl}(\{\hat{\pi} \xrightarrow{\alpha} s'\}) \subseteq \mathit{Cyl}(\{\hat{\pi}\})$.

Now assume $\Pr_{\sigma}^{\mathcal{M}}(\langle\hat{\pi}\rangle) > 0$.

<u>$Case\ s \in \mathrm{PS}$:</u> It follows that

$$\begin{aligned}
\Pr_{\sigma}^{\mathcal{M}}(\langle\hat{\pi} \xrightarrow{\alpha} s'\rangle) &= \int_{\pi\in\langle\hat{\pi}\rangle} \sigma(\pi,\alpha)\cdot\mathbf{P}(s,\alpha,s')\,\mathrm{d}\Pr_{\sigma}^{\mathcal{M}}(\pi)\\
&= \mathbf{P}(s,\alpha,s')\cdot\int_{\pi\in\langle\hat{\pi}\rangle} \sigma(\pi,\alpha)\,\mathrm{d}\Pr_{\sigma}^{\mathcal{M}}(\{\pi\}\cap\langle\hat{\pi}\rangle)\\
&= \mathbf{P}(s,\alpha,s')\cdot\int_{\pi\in\langle\hat{\pi}\rangle} \sigma(\pi,\alpha)\,\mathrm{d}\big[\Pr_{\sigma}^{\mathcal{M}}(\pi\mid\langle\hat{\pi}\rangle)\cdot\Pr_{\sigma}^{\mathcal{M}}(\langle\hat{\pi}\rangle)\big]\\
&= \Pr_{\sigma}^{\mathcal{M}}(\langle\hat{\pi}\rangle)\cdot\mathbf{P}(s,\alpha,s')\cdot\int_{\pi\in\langle\hat{\pi}\rangle} \sigma(\pi,\alpha)\,\mathrm{d}\Pr_{\sigma}^{\mathcal{M}}(\pi\mid\langle\hat{\pi}\rangle)\\
&= \Pr_{\sigma}^{\mathcal{M}}(\langle\hat{\pi}\rangle)\cdot\mathbf{P}(s,\alpha,s')\cdot\mathrm{ta}(\sigma)(\hat{\pi},\alpha)\\
&\overset{IH}{=} \Pr_{\mathrm{ta}(\sigma)}^{\mathcal{M}_{\mathcal{D}}}(\hat{\pi})\cdot\mathbf{P}(s,\alpha,s')\cdot\mathrm{ta}(\sigma)(\hat{\pi},\alpha)\\
&= \Pr_{\mathrm{ta}(\sigma)}^{\mathcal{M}_{\mathcal{D}}}(\hat{\pi} \xrightarrow{\alpha} s').
\end{aligned}$$

<u>$Case\ s \in \mathrm{MS}$:</u> As $s \in \mathrm{MS}$ we have $\alpha = \perp$ and it follows

$$\begin{aligned}
\Pr_{\sigma}^{\mathcal{M}}(\langle\hat{\pi} \xrightarrow{\perp} s'\rangle) &= \int_{\pi\in\langle\hat{\pi}\rangle}\int_0^{\infty} \mathrm{E}(s)\cdot e^{-\mathrm{E}(s)t}\cdot\mathbf{P}(s,\perp,s')\,\mathrm{d}t\,\mathrm{d}\Pr_{\sigma}^{\mathcal{M}}(\pi)\\
&= \mathbf{P}(s,\perp,s')\cdot\int_{\pi\in\langle\hat{\pi}\rangle}\int_0^{\infty} \mathrm{E}(s)\cdot e^{-\mathrm{E}(s)t}\,\mathrm{d}t\,\mathrm{d}\Pr_{\sigma}^{\mathcal{M}}(\pi)\\
&= \mathbf{P}(s,\perp,s')\cdot\Pr_{\sigma}^{\mathcal{M}}(\langle\hat{\pi}\rangle)\\
&\overset{IH}{=} \mathbf{P}(s,\perp,s')\cdot\Pr_{\mathrm{ta}(\sigma)}^{\mathcal{M}_{\mathcal{D}}}(\hat{\pi})\\
&= \Pr_{\mathrm{ta}(\sigma)}^{\mathcal{M}_{\mathcal{D}}}(\hat{\pi} \xrightarrow{\perp} s').
\end{aligned}$$

□

The result is lifted to untimed reachability probabilities.

Proposition 2 *For any $G \subseteq S$ it holds that $Pr_{\sigma}^{\mathcal{M}}(\Diamond G) = Pr_{ta(\sigma)}^{\mathcal{M}_{\mathcal{D}}}(\Diamond G)$.*

Proof Let Π be the set of finite time-abstract paths of $\mathcal{M}_{\mathcal{D}}$ that end at the first visit of a state in G, i.e.,

$$\Pi = \{s_0 \xrightarrow{\alpha_0} \ldots \xrightarrow{\alpha_{n-1}} s_n \in \mathit{FPaths}^{\mathcal{M}_{\mathcal{D}}} \mid s_n \in G \text{ and } \forall i < n\colon s_i \notin G\}.$$

Every path $\pi \in \Diamond G \subseteq \mathit{IPaths}^{\mathcal{M}}$ has a unique prefix π' with $\mathrm{ta}(\pi') \in \Pi$. We have

$$\Diamond G = \biguplus_{\hat{\pi}\in\Pi} \mathit{Cyl}(\langle\hat{\pi}\rangle).$$

The claim follows with Lemma 1 since

$$\Pr_{\sigma}^{\mathcal{M}}(\Diamond G) = \sum_{\hat{\pi}\in\Pi}\Pr_{\sigma}^{\mathcal{M}}(\langle\hat{\pi}\rangle) \overset{\text{Lem. 1}}{=} \sum_{\hat{\pi}\in\Pi}\Pr_{\mathrm{ta}(\sigma)}^{\mathcal{M}_{\mathcal{D}}}(\hat{\pi}) = \Pr_{\mathrm{ta}(\sigma)}^{\mathcal{M}_{\mathcal{D}}}(\Diamond G).$$

□

As the definition of $\mathrm{ta}(\sigma)$ is independent of the considered set of goal states $G \subseteq S$, Proposition 2 can be lifted to multiple untimed reachability objectives.

Proof (*Theorem* 3) Let $\mathbb{O} = (\mathbb{P}(\lozenge G_1), \ldots, \mathbb{P}(\lozenge G_d))$ be the considered list of objectives with threshold relations $\rhd = (\rhd_1, \ldots, \rhd_d)$. The following equivalences hold for any $\sigma \in \mathrm{GM}^{\mathcal{M}}$ and $\mathbf{p} \in \mathbb{R}^d$.

$$\begin{aligned}
\mathcal{M}, \sigma \models \mathbb{O} \rhd \mathbf{p} &\iff \forall i : \mathcal{M}, \sigma \models \mathbb{P}(\lozenge G_i) \rhd_i p_i \\
&\iff \forall i : \mathrm{Pr}^{\mathcal{M}}_{\sigma}(\lozenge G_i) \rhd_i p_i \\
&\overset{\text{Prop. 2}}{\iff} \forall i : \mathrm{Pr}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\lozenge G_i) \rhd_i p_i \\
&\iff \forall i : \mathcal{M}_{\mathcal{D}}, \mathrm{ta}(\sigma) \models \mathbb{P}(\lozenge G_i) \rhd_i p_i \\
&\iff \mathcal{M}_{\mathcal{D}}, \mathrm{ta}(\sigma) \models \mathbb{O} \rhd \mathbf{p}\,.
\end{aligned}$$

Assume that $\mathit{achieve}^{\mathcal{M}}(\mathbb{O} \rhd \mathbf{p})$ holds, i.e., there is a $\sigma \in \mathrm{GM}^{\mathcal{M}}$ such that $\mathcal{M}, \sigma \models \mathbb{O} \rhd \mathbf{p}$. It follows that $\mathcal{M}_{\mathcal{D}}, \mathrm{ta}(\sigma) \models \mathbb{O} \rhd \mathbf{p}$ which means that $\mathit{achieve}^{\mathcal{M}_{\mathcal{D}}}(\mathbb{O} \rhd \mathbf{p})$ holds as well. For the other direction assume $\mathit{achieve}^{\mathcal{M}_{\mathcal{D}}}(\mathbb{O} \rhd \mathbf{p})$, i.e., $\mathcal{M}_{\mathcal{D}}, \sigma \models \mathbb{O} \rhd \mathbf{p}$ for some time-abstract scheduler $\sigma \in \mathrm{TA}$. We have $\mathrm{ta}(\sigma) = \sigma$. It follows that $\mathcal{M}_{\mathcal{D}}, \mathrm{ta}(\sigma) \models \mathbb{O} \rhd \mathbf{p}$. Applying the equivalences above yields $\mathcal{M}, \sigma \models \mathbb{O} \rhd \mathbf{p}$ and thus $\mathit{achieve}^{\mathcal{M}}(\mathbb{O} \rhd \mathbf{p})$. □

4.2 Expected reward objectives

The results for expected reward objectives are similar to untimed reachability objectives: An analysis of the underlying MDP suffices. We show the following extension of Theorem 3 to expected reward objectives.

Theorem 4 *For MA $\mathcal{M}$ and untimed reachability and expected reward objectives $\mathbb{O}$: $\mathit{achieve}^{\mathcal{M}}(\mathbb{O} \rhd \boldsymbol{p}) \iff \mathit{achieve}^{\mathcal{M}_{\mathcal{D}}}(\mathbb{O} \rhd \boldsymbol{p})$.*

To prove this, we show that a scheduler $\sigma \in \mathrm{GM}^{\mathcal{M}}$ and its time-abstraction $\mathrm{ta}(\sigma) \in \mathrm{TA}$ induce the same expected rewards on $\mathcal{M}$ and $\mathcal{M}_{\mathcal{D}}$, respectively. Theorem 4 follows then analogously to Theorem 3.

Proposition 3 *Let ρ be some reward function of $\mathcal{M}$ and let $\rho^{\mathcal{D}}$ be its counterpart for $\mathcal{M}_{\mathcal{D}}$. For $G \subseteq S$ we have $\mathrm{ER}^{\mathcal{M}}_{\sigma}(\rho, G) = \mathrm{ER}^{\mathcal{M}_{\mathcal{D}}}_{ta(\sigma)}(\rho^{\mathcal{D}}, G)$.*

Notice that $\rho^{\mathcal{D}}$ encodes the *expected* reward of $\mathcal{M}$ obtained in a state s by assuming the sojourn time to be the expected sojourn time $\frac{1}{E(s)}$. Although the claim is similar to Proposition 2, its proof cannot be adapted straightforwardly. In particular, the analogon to Lemma 1 does not hold: The expected reward collected along a time-abstract path $\hat{\pi} \in \mathit{FPaths}^{\mathcal{M}_{\mathcal{D}}}$ does not coincide in general for $\mathcal{M}$ and $\mathcal{M}_{\mathcal{D}}$.

Example 7 Let $\mathcal{M}$ be the MA with underlying MDP $\mathcal{M}_{\mathcal{D}}$ as shown in Fig. 2. Let $\rho(s_0) = 1$ and zero otherwise. Reconsider the scheduler σ from Example 6. Let $\hat{\pi}_\alpha = s_0 \xrightarrow{\perp} s_3 \xrightarrow{\alpha} s_6$. The probability $\mathrm{Pr}^{\mathcal{M}}_{\sigma}(\{s_0 \xrightarrow{t} s_3 \xrightarrow{\alpha} s_6 \in \langle \hat{\pi}_\alpha \rangle \mid t > 1\})$ is zero since σ chooses β on such paths. For the remaining paths in $\langle \hat{\pi}_\alpha \rangle$, action α is chosen with probability one. The expected reward in $\mathcal{M}$ along $\hat{\pi}_\alpha$ is:

$$\int_{\pi \in \langle \hat{\pi}_\alpha \rangle} rew^{\mathcal{M}}(\rho, \pi)\, \mathrm{dPr}^{\mathcal{M}}_{\sigma}(\pi) = \int_0^1 \rho(s_0) \cdot t \cdot \mathrm{E}(s_0) \cdot e^{-\mathrm{E}(s_0)t}\, \mathrm{d}t = 1 - 2e^{-1}.$$

The expected reward in $\mathcal{M}_{\mathcal{D}}$ along $\hat{\pi}_\alpha$ differs as

$$rew^{\mathcal{M}_{\mathcal{D}}}(\rho^{\mathcal{D}}, \hat{\pi}_\alpha) \cdot \mathrm{Pr}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\hat{\pi}_\alpha) = \rho^{\mathcal{D}}(s_0, \perp) \cdot \mathrm{ta}(\sigma)(s_0 \xrightarrow{\perp} s_3, \alpha) = 1 - e^{-1}.$$

The intuition is as follows: If path $s_0 \xrightarrow{t} s_3 \xrightarrow{\alpha} s_6$ of $\mathcal{M}$ under σ occurs, we have $t \leq 1$ since σ chose α. Hence, the reward collected from paths in $\langle \hat{\pi}_\alpha \rangle$ is at most $1 \cdot \rho(s_0) = 1$. There is thus a dependency between the choice of the scheduler at s_3 and the collected reward at s_0. This dependency is absent in $\mathcal{M}_{\mathcal{D}}$ as the reward at a state is independent of the subsequent performed actions.

Let $\hat{\pi}_\beta = s_0 \xrightarrow{\perp} s_3 \xrightarrow{\beta} s_4$. The expected reward along $\hat{\pi}_\beta$ is $2e^{-1}$ for $\mathcal{M}$ and e^{-1} for $\mathcal{M}_{\mathcal{D}}$. As the rewards for $\hat{\pi}_\alpha$ and $\hat{\pi}_\beta$ sum up to one in both $\mathcal{M}$ and $\mathcal{M}_{\mathcal{D}}$, the expected reward along all paths of length two coincides for $\mathcal{M}$ and $\mathcal{M}_{\mathcal{D}}$.

This observation can be generalized to arbitrary MA and paths of arbitrary length. We first formalize the step-bounded expected reward.

Let $n \geq 0$ and $G \subseteq S$. The set of time-abstract paths that end after n steps or at the first visit of a state in G is denoted by

$$\begin{aligned} \Pi^n_G = &\{s_0 \xrightarrow{\alpha_0} \dots \xrightarrow{\alpha_{m-1}} s_m \in FPaths^{\mathcal{M}_{\mathcal{D}}} \mid (m = n \text{ or } s_m \in G) \text{ and} \\ &s_i \notin G \text{ for all } 0 \leq i < m\}. \end{aligned}$$

For $\mathcal{M}$ under $\sigma \in \mathrm{GM}^{\mathcal{M}}$ and $\mathcal{M}_{\mathcal{D}}$ under $\mathrm{ta}(\sigma) \in \mathrm{TA}$, we define the expected reward collected along the paths of Π^n_G as

$$\begin{aligned} \mathrm{ER}^{\mathcal{M}}_{\sigma}(\rho, \Pi^n_G) &= \sum_{\hat{\pi} \in \Pi^n_G} \int_{\pi \in \langle \hat{\pi} \rangle} rew^{\mathcal{M}}(\rho, \pi)\, \mathrm{dPr}^{\mathcal{M}}_{\sigma}(\pi) \text{ and} \\ \mathrm{ER}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\rho^{\mathcal{D}}, \Pi^n_G) &= \sum_{\hat{\pi} \in \Pi^n_G} rew^{\mathcal{M}_{\mathcal{D}}}(\rho^{\mathcal{D}}, \hat{\pi}) \cdot \mathrm{Pr}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\hat{\pi}), \end{aligned}$$

respectively. Intuitively, $\mathrm{ER}^{\mathcal{M}}_{\sigma}(\rho, \Pi^n_G)$ corresponds to $\mathrm{ER}^{\mathcal{M}}_{\sigma}(\rho, G)$ assuming that no more reward is collected after the n-th transition. It follows that the value $\mathrm{ER}^{\mathcal{M}}_{\sigma}(\rho, \Pi^n_G)$ approaches $\mathrm{ER}^{\mathcal{M}}_{\sigma}(\rho, G)$ for large n. Similarly, $\mathrm{ER}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\rho^{\mathcal{D}}, \Pi^n_G)$ approaches $\mathrm{ER}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\rho^{\mathcal{D}}, G)$ for large n. This observation is formalized by the following lemma.

Lemma 2 *For MA $\mathcal{M} = (S, Act, \rightarrow, s_0, (\rho_1, \dots, \rho_\ell))$ with $G \subseteq S$, $\sigma \in GM$, and reward function ρ it holds that*

$$\lim_{n \to \infty} \mathrm{ER}^{\mathcal{M}}_{\sigma}(\rho, \Pi^n_G) = \mathrm{ER}^{\mathcal{M}}_{\sigma}(\rho, G).$$

Furthermore, any reward function $\rho^{\mathcal{D}}$ for $\mathcal{M}_{\mathcal{D}}$ satisfies

$$\lim_{n \to \infty} \mathrm{ER}^{\mathcal{M}_{\mathcal{D}}}_{ta(\sigma)}(\rho^{\mathcal{D}}, \Pi^n_G) = \mathrm{ER}^{\mathcal{M}_{\mathcal{D}}}_{ta(\sigma)}(\rho^{\mathcal{D}}, G).$$

Proof *(Sketch)* Essentially, for each $n \geq 0$, consider the function $f_n \colon IPaths^{\mathcal{M}} \to \mathbb{R}_{\geq 0}$ given by

$$f_n(\pi) = \begin{cases} rew^{\mathcal{M}}(\rho, pref(\pi, m)) & \text{if } m = \min\left\{i \in \{0, \dots, n\} \mid s_i \in G\right\} \\ rew^{\mathcal{M}}(\rho, pref(\pi, n)) & \text{if } s_i \notin G \text{ for all } i \leq n \end{cases}$$

for every path $\pi = s_0 \xrightarrow{\kappa_0} s_1 \xrightarrow{\kappa_1} \cdots \in \mathit{IPaths}^{\mathcal{M}}$. Intuitively, $f_n(\pi)$ is the reward collected on π within the first n steps and only up to the first visit of G. This allows us to express the expected reward collected along the paths of Π_G^n, and note that the sequence of functions $f_0, f_1, \ldots$ is non-decreasing, we may apply the *monotone convergence theorem* [1] and conclude the proof. A full proof may be found in the appendix. □

The next step is to show that the expected reward collected along the paths of Π_G^n coincides for $\mathcal{M}$ under σ and $\mathcal{M}_{\mathcal{D}}$ under $\mathrm{ta}(\sigma)$.

Lemma 3 *Let ρ be some reward function of $\mathcal{M}$ and let $\rho^{\mathcal{D}}$ be its counterpart for $\mathcal{M}_{\mathcal{D}}$. Let $\mathcal{M} = (S, Act, \rightarrow, s_0, (\rho_1, \ldots, \rho_\ell))$ be an MA with $G \subseteq S$ and $\sigma \in GM$. For all $G \subseteq S$ and $n \geq 0$ it holds that*

$$\mathrm{ER}_{\sigma}^{\mathcal{M}}(\rho, \Pi_G^n) = \mathrm{ER}_{ta(\sigma)}^{\mathcal{M}_{\mathcal{D}}}(\rho^{\mathcal{D}}, \Pi_G^n).$$

Proof The proof is by induction over the path length n. To simplify the notation, we often omit the reward functions ρ and $\rho^{\mathcal{D}}$ and write, e.g., $\mathit{rew}^{\mathcal{M}_{\mathcal{D}}}(\pi)$ instead of $\mathit{rew}^{\mathcal{M}_{\mathcal{D}}}(\rho^{\mathcal{D}}, \pi)$ or $\mathrm{ER}_{\sigma}^{\mathcal{M}}(\Pi_G^n)$ instead of $\mathrm{ER}_{\sigma}^{\mathcal{M}}(\rho, \Pi_G^n)$.

If $n = 0$, then $\Pi_G^n = \{s_0\}$. The claim holds: $\mathit{rew}^{\mathcal{M}}(s_0) = \mathit{rew}^{\mathcal{M}_{\mathcal{D}}}(s_0) = 0$.

In the induction step, we assume that the lemma is true for some fixed $n \geq 0$. We define $\mathit{pref}(\pi, n)$ for paths with $|\pi| \leq n$ such that $\mathit{pref}(\pi, n) = \pi$. We split the term $\mathrm{ER}_{\sigma}^{\mathcal{M}}(\Pi_G^{n+1})$ into the reward that is obtained by paths which reach G within n steps and the reward obtained by paths of length $n+1$. In a second step, we consider the sum of the reward collected within the first n steps and the reward obtained in the $(n+1)$-th step.

$$\begin{aligned}
\mathrm{ER}_{\sigma}^{\mathcal{M}}(\Pi_G^{n+1}) &= \sum_{\substack{\hat{\pi} \in \Pi_G^{n+1} \\ |\hat{\pi}| \leq n}} \int_{\pi \in \langle \hat{\pi} \rangle} \mathit{rew}^{\mathcal{M}}(\pi) \, \mathrm{dPr}_{\sigma}^{\mathcal{M}}(\pi) \\
&\quad + \sum_{\substack{\hat{\pi} \in \Pi_G^{n+1} \\ |\hat{\pi}| = n+1}} \int_{\substack{\pi = \pi' \xrightarrow{\kappa} s' \in \langle \hat{\pi} \rangle \\ \mathit{last}(\pi') = s}} \mathit{rew}^{\mathcal{M}}(\pi') + \rho(s) \cdot t(\kappa) + \rho(s, \alpha(\kappa)) \, \mathrm{dPr}_{\sigma}^{\mathcal{M}}(\pi) \\
&= \sum_{\hat{\pi} \in \Pi_G^{n+1}} \int_{\pi \in \langle \hat{\pi} \rangle} \mathit{rew}^{\mathcal{M}}(\mathit{pref}(\pi, n)) \, \mathrm{dPr}_{\sigma}^{\mathcal{M}}(\pi) \qquad (1) \\
&\quad + \sum_{\substack{\hat{\pi} \in \Pi_G^{n+1} \\ |\hat{\pi}| = n+1}} \int_{\substack{\pi = \pi' \xrightarrow{\kappa} s' \in \langle \hat{\pi} \rangle \\ \mathit{last}(\pi') = s}} \rho(s) \cdot t(\kappa) + \rho(s, \alpha(\kappa)) \, \mathrm{dPr}_{\sigma}^{\mathcal{M}}(\pi), \\
&= \sum_{\hat{\pi} \in \Pi_G^{n+1}} \mathit{rew}^{\mathcal{M}_{\mathcal{D}}}(\mathit{pref}(\hat{\pi}, n)) \cdot \mathrm{Pr}_{\mathrm{ta}(\sigma)}^{\mathcal{M}_{\mathcal{D}}}(\hat{\pi}) \\
&\quad + \sum_{\substack{\hat{\pi} = \hat{\pi}' \xrightarrow{\alpha} s' \in \Pi_G^{n+1} \\ |\hat{\pi}| = n+1}} \rho^{\mathcal{D}}(\mathit{last}(\hat{\pi}'), \alpha) \cdot \mathrm{Pr}_{\mathrm{ta}(\sigma)}^{\mathcal{M}_{\mathcal{D}}}(\hat{\pi}) \\
&= \sum_{\hat{\pi} \in \Pi_G^{n+1}} \mathit{rew}^{\mathcal{M}_{\mathcal{D}}}(\hat{\pi}) \cdot \mathrm{Pr}_{\mathrm{ta}(\sigma)}^{\mathcal{M}_{\mathcal{D}}}(\hat{\pi}) = \mathrm{ER}_{\mathrm{ta}(\sigma)}^{\mathcal{M}_{\mathcal{D}}}(\Pi_G^{n+1}). \qquad (2)
\end{aligned}$$

Detailed reformulations for the terms (1) and (2) are treated separately and discussed in "Appendix B.2". □

We now first show Proposition 3 and then Theorem 4.

Proof (*Proposition* 3) The proposition is a direct consequence of Lemma 2 and Lemma 3 as

$$\begin{aligned}\mathrm{ER}^{\mathcal{M}}_{\sigma}(\rho, G) &= \lim_{n\to\infty} \mathrm{ER}^{\mathcal{M}}_{\sigma}(\rho, \Pi^n_G)\\ &= \lim_{n\to\infty} \mathrm{ER}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\rho^{\mathcal{D}}, \Pi^n_G) = \mathrm{ER}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\rho^{\mathcal{D}}, G).\end{aligned}$$

□

Proof (*Theorem* 4) Let $\mathbb{O} = (\mathbb{O}_1, \dots, \mathbb{O}_d)$ be the considered list of untimed reachability and expected reward objectives with threshold relations $\rhd = (\rhd_1, \dots, \rhd_d)$. The following equivalences hold for any $\sigma \in \mathrm{GM}^{\mathcal{M}}$ and $\mathbf{p} \in \mathbb{R}^d$.

$$\begin{aligned}\mathcal{M}, \sigma \models \mathbb{O} \rhd \mathbf{p} &\iff \forall i\colon \mathcal{M}, \sigma \models \mathbb{O}_i \rhd_i p_i\\ &\overset{*}{\iff} \forall i\colon \mathcal{M}_{\mathcal{D}}, \mathrm{ta}(\sigma) \models \mathbb{O}_i \rhd_i p_i \iff \mathcal{M}_{\mathcal{D}}, \mathrm{ta}(\sigma) \models \mathbb{O} \rhd \mathbf{p}\,,\end{aligned}$$

where for the equivalence marked with $*$ we consider two cases: If $\mathbb{O}_i$ is of the form $\mathbb{P}(\Diamond G)$, Proposition 2 yields

$$\begin{aligned}\mathcal{M}, \sigma \models \mathbb{O}_i \rhd_i p_i &\iff \mathrm{Pr}^{\mathcal{M}}_{\sigma}(\Diamond G) \rhd_i p_i\\ &\iff \mathrm{Pr}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\Diamond G) \rhd_i p_i \iff \mathcal{M}_{\mathcal{D}}, \mathrm{ta}(\sigma) \models \mathbb{O}_i \rhd_i p_i\,.\end{aligned}$$

Otherwise, $\mathbb{O}_i$ is of the form $\mathbb{E}(\#j, G)$ and with Proposition 3 it follows that

$$\begin{aligned}\mathcal{M}, \sigma \models \mathbb{O}_i \rhd_i p_i &\iff \mathrm{ER}^{\mathcal{M}}_{\sigma}(\rho_j, G) \rhd_i p_i\\ &\iff \mathrm{ER}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\rho^{\mathcal{D}}_j, G) \rhd_i p_i \iff \mathcal{M}_{\mathcal{D}}, \mathrm{ta}(\sigma) \models \mathbb{O}_i \rhd_i p_i\,.\end{aligned}$$

The remaining steps of the proof are completely analogous to the proof of Theorem 3 conducted on page 15. □

Thus, queries on MA with mixtures of untimed reachability and expected reward objectives can be analyzed on the underlying MDP $\mathcal{M}_{\mathcal{D}}$.

4.3 Timed reachability objectives

Timed reachability objectives cannot be analyzed on $\mathcal{M}_{\mathcal{D}}$ as it abstracts away from sojourn times. We lift the digitization approach for single-objective timed reachability [30,35] to multiple objectives. Instead of abstracting timing information, it is *digitized*. In this section, we make an additional but standard assumption on MAs (e.g., [14,30,35]). MAs with *Zeno behavior*, where infinitely many actions can be taken within finite time with non-zero probability, are unrealistic and considered a modeling error. MAs that exhibit Zeno behavior can be easily detected using the notion of end-components (that we discuss in the next section).

4.3.1 Digitized reachability

The *digitization* $\mathcal{M}_\delta$ of $\mathcal{M}$ w.r.t. some digitization constant $\delta \in \mathbb{R}_{>0}$ is an MDP which digitizes the time [30,35]. The main difference between $\mathcal{M}_{\mathcal{D}}$ and $\mathcal{M}_\delta$ is that the latter also introduces *self-loops* which describe the probability to stay in a Markovian state for δ time units. More precisely, the outgoing transitions of states $s \in \mathrm{MS}$ in $\mathcal{M}_\delta$ represent that either (1) a Markovian transition in $\mathcal{M}$ was taken within δ time units, or (2) no transition is taken within δ time units—which is captured by taking the self-loop in $\mathcal{M}_\delta$. Counting the taken self-loops at $s \in \mathrm{MS}$ allows to approximate the sojourn time in s.

Definition 11 (Digitization of an MA) For MA $\mathcal{M} = (S, Act, \rightarrow, s_0, \{\rho_1, \dots, \rho_\ell\})$ with transition probabilities $\mathbf{P}$ and *digitization constant* $\delta \in \mathbb{R}_{>0}$, the *digitization of* $\mathcal{M}$ w.r.t. δ is the MDP $\mathcal{M}_\delta = (S, Act, \mathbf{P}_\delta, s_0, (\rho_1^\delta, \dots, \rho_\ell^\delta))$, where

$$\mathbf{P}_\delta(s, \alpha, s') = \begin{cases} \mathbf{P}(s, \bot, s') \cdot (1 - e^{-\mathrm{E}(s)\delta}) & \text{if } s \in \mathrm{MS}, \alpha = \bot, s \neq s' \\ \mathbf{P}(s, \bot, s') \cdot (1 - e^{-\mathrm{E}(s)\delta}) + e^{-\mathrm{E}(s)\delta} & \text{if } s \in \mathrm{MS}, \alpha = \bot, s = s' \\ \mathbf{P}(s, \alpha, s') & \text{otherwise.} \end{cases}$$

and for each $i \in \{1, \dots, \ell\}$:

$$\rho_i^\delta(s, \alpha) = \begin{cases} \rho_i(s, \alpha) & \text{if } s \in \mathrm{PS} \\ \big(\rho_i(s, \bot) + \frac{1}{\mathrm{E}(s)} \cdot \rho_i(s)\big) \cdot \big(1 - e^{-\mathrm{E}(s)\delta}\big) & \text{if } s \in \mathrm{MS} \text{ and } \alpha = \bot \\ 0 & \text{otherwise.} \end{cases}$$

Example 8 Figure 2 on page 4 shows an MA $\mathcal{M}$ with its underlying MDP $\mathcal{M}_\mathcal{D}$ and a digitization $\mathcal{M}_\delta$ for unspecified $\delta \in \mathbb{R}_{>0}$.

A time interval $I \subseteq \mathbb{R}_{\geq 0}$ of the form $[a, \infty)$ or $[a, b]$ with $\mathrm{di}_a := \frac{a}{\delta} \in \mathbb{N}$ and $\mathrm{di}_b := \frac{b}{\delta} \in \mathbb{N}$ is called *well-formed*. For the remainder, we only consider well-formed intervals. If the interval boundaries a and b are rationals with $a = \frac{a_1}{a_2}, b = \frac{b_1}{b_2}$ for integers $a_1, a_2, b_1, b_2 \in \mathbb{Z}$, well-formedness can always be ensured by setting the digitization constant δ to $\frac{1}{k}$ for some common multiple k of a_2 and b_2.

An interval for time-bounds I is transformed to digitization step bounds $\mathrm{di}(I) \subseteq \mathbb{N}$. Let $a = \min I$, we set $\mathrm{di}(I) = \{m \in \mathbb{N} \mid m \cdot \delta \in I\}$.

We first relate paths in $\mathcal{M}$ to paths in its digitization.

Definition 12 (Digitization of a path) The *digitization* $\mathrm{di}(\pi)$ of path $\pi = s_0 \xrightarrow{\kappa_0} s_1 \xrightarrow{\kappa_1} \dots$ in $\mathcal{M}$ is the path in $\mathcal{M}_\delta$ given by

$$\mathrm{di}(\pi) = \big(s_0 \xrightarrow{\alpha(\kappa_0)}\big)^{m_0} s_0 \xrightarrow{\alpha(\kappa_0)} \big(s_1 \xrightarrow{\alpha(\kappa_1)}\big)^{m_1} s_1 \xrightarrow{\alpha(\kappa_1)} \dots$$

where $m_i = \max\{m \in \mathbb{N} \mid m \cdot \delta \leq t(\kappa_i)\}$ for each $i \geq 0$.

Example 9 For the path $\pi = s_0 \xrightarrow{1.1} s_3 \xrightarrow{\beta} s_4 \xrightarrow{\eta} s_5 \xrightarrow{0.3} s_4$ of the MA $\mathcal{M}$ in Fig. 2a and $\delta = 0.4$, we get $\mathrm{di}(\pi) = s_0 \xrightarrow{\bot} s_0 \xrightarrow{\bot} s_0 \xrightarrow{\bot} s_3 \xrightarrow{\beta} s_4 \xrightarrow{\eta} s_5 \xrightarrow{\bot} s_4$.

The m_i in the definition above represent a digitization of the sojourn times $t(\kappa_i)$ such that $m_i\delta \leq t(\kappa_i) < (m_i{+}1)\delta$. These digitized times are incorporated into the digitization of a path by taking the self-loop at state $s_i \in \mathrm{MS}$ m_i times. We also refer to the paths of $\mathcal{M}_\delta$ as *digital paths (of* $\mathcal{M}$*)*. The number $|\bar{\pi}|_{\mathrm{ds}}$ of *digitization steps* of a digital path $\bar{\pi}$ is the number of transitions emerging from Markovian states, i.e., $|\bar{\pi}|_{\mathrm{ds}} = |\{i < |\bar{\pi}| \mid \bar{\pi}[i] \in \mathrm{MS}\}|$. One digitization step represents the elapse of at most δ time units—either by staying at some $s \in \mathrm{MS}$ for δ time or by leaving s within δ time. The number $|\mathrm{di}(\pi)|_{\mathrm{ds}}$ multiplied with δ yields an estimate for the duration $T(\pi)$. A digital path $\bar{\pi}$ can be interpreted as representation of the set of paths of $\mathcal{M}$ whose digitization is $\bar{\pi}$.

Definition 13 (Induced paths of a digital path) The set of *induced paths* of a (finite or infinite) digital path $\bar{\pi}$ of $\mathcal{M}_\delta$ is

$$[\bar{\pi}] = \mathrm{di}^{-1}(\bar{\pi}) = \{\pi \in \mathit{FPaths}^{\mathcal{M}} \cup \mathit{IPaths}^{\mathcal{M}} \mid \mathrm{di}(\pi) = \bar{\pi}\}.$$

For sets of digital paths Π we define the *induced paths* $[\Pi] = \bigcup_{\bar{\pi} \in \Pi} [\bar{\pi}]$. To relate timed reachability probabilities for $\mathcal{M}$ under scheduler $\sigma \in \mathrm{GM}^{\mathcal{M}}$ with digitization step-bounded reachability probabilities for $\mathcal{M}_\delta$, relating σ to a scheduler for $\mathcal{M}_\delta$ is necessary.

Definition 14 (Digitization of a scheduler) The *digitization* of $\sigma \in \mathrm{GM}^{\mathcal{M}}$ is given by $\mathrm{di}(\sigma) \in \mathrm{TA}^{\mathcal{M}_\delta}$ such that for any $\bar{\pi} \in \mathit{FPaths}^{\mathcal{M}_\delta}$ with $\mathit{last}(\bar{\pi}) \in \mathrm{PS}$

$$\mathrm{di}(\sigma)(\bar{\pi}, \alpha) = \begin{cases} \displaystyle\int_{\pi \in [\bar{\pi}]} \sigma(\pi, \alpha)\, \mathrm{dPr}_{\sigma}^{\mathcal{M}}(\pi \mid [\bar{\pi}]) & \text{if } \mathrm{Pr}_{\sigma}^{\mathcal{M}}([\bar{\pi}]) > 0 \\ 1 \,/\, |\mathit{Act}(\mathit{last}(\bar{\pi}))| & \text{otherwise.} \end{cases}$$

The digitization $\mathrm{di}(\sigma)$ is similar to the time-abstraction $\mathrm{ta}(\sigma)$ as both schedulers get a path with restricted timing information as input and mimic the choice of σ. However, while $\mathrm{ta}(\sigma)$ receives no information regarding sojourn times, $\mathrm{di}(\sigma)$ receives the digital estimate. Intuitively, $\mathrm{di}(\sigma)(\bar{\pi}, \alpha)$ considers $\sigma(\pi, \alpha)$ for each $\pi \in [\bar{\pi}]$, weighted with the probability that the sojourn times of a path in $[\bar{\pi}]$ are as given by π. The restriction $\mathit{last}(\bar{\pi}) \in \mathrm{PS}$ asserts that $\bar{\pi}$ does not end with a self-loop on a Markovian state, implying $[\bar{\pi}] \neq \emptyset$.

Example 10 Let MA $\mathcal{M}$ in Fig. 2a and $\delta = 0.4$. Again, $\sigma \in \mathrm{GM}^{\mathcal{M}}$ chooses α at state s_3 iff the sojourn time at s_0 is at most one. Consider the digital paths $\bar{\pi}_m = (s_0 \xrightarrow{\perp})^m s_0 \xrightarrow{\perp} s_3$. For $\pi \in [\bar{\pi}_1] = \{s_0 \xrightarrow{t} s_3 \mid 0.4 \leq t < 0.8\}$ we have $\sigma(\pi, \alpha) = 1$. It follows $\mathrm{di}(\sigma)(\pi_1, \alpha) = 1$. For $\pi \in [\bar{\pi}_2] = \{s_0 \xrightarrow{t} s_3 \mid 0.8 \leq t < 1.2\}$ it is unclear whether σ chooses α or β. Hence, $\mathrm{di}(\sigma)$ randomly guesses:

$$\mathrm{di}(\sigma)(\bar{\pi}_2, \alpha) = \int_{\pi \in [\bar{\pi}_2]} \sigma(\pi, \alpha)\, \mathrm{dPr}_{\sigma}^{\mathcal{M}}(\pi \mid [\bar{\pi}_2]) = \frac{\int_{0.8}^{1.0} \mathrm{E}(s_0) e^{-\mathrm{E}(s_0)t}\, \mathrm{d}t}{\int_{0.8}^{1.2} \mathrm{E}(s_0) e^{-\mathrm{E}(s_0)t}\, \mathrm{d}t} \approx 0.55\,.$$

On $\mathcal{M}_\delta$ we consider ds-bounded reachability instead of timed reachability.

Definition 15 (ds-bounded reachability) The set of infinite digital paths that reach $G \subseteq S$ within the interval $J \subseteq \mathbb{N}$ of consecutive natural numbers is

$$\Diamond_{\mathrm{ds}}^{J} G = \{\bar{\pi} \in \mathit{IPaths}^{\mathcal{M}_\delta} \mid \exists n \geq 0 \colon \bar{\pi}[n] \in G \text{ and } |\mathit{pref}(\bar{\pi}, n)|_{\mathrm{ds}} \in J\}.$$

The timed reachability probabilities for $\mathcal{M}$ are estimated by ds-bounded reachability probabilities for $\mathcal{M}_\delta$. The induced ds-bounded reachability probability for $\mathcal{M}$ (under σ) coincides with ds-bounded reachability probability on $\mathcal{M}_\delta$ (under $\mathrm{di}(\sigma)$).

Proposition 4 *Let $\mathcal{M}$ be an MA with $G \subseteq S$, $\sigma \in GM$, and digitization $\mathcal{M}_\delta$. Further, let $J \subseteq \mathbb{N}$ be a set of consecutive natural numbers. It holds that*

$$Pr_{\sigma}^{\mathcal{M}}([\Diamond_{ds}^{J} G]) = Pr_{di(\sigma)}^{\mathcal{M}_\delta}(\Diamond_{ds}^{J} G).$$

Before we prove this proposition, we need to aggregate paths which differ a bit in their timing. Let $\mathcal{M} = (S, \mathit{Act}, \rightarrow, s_0, (\rho_1, \ldots, \rho_\ell))$ be an MA and let $\mathcal{M}_\delta$ be the digitization of $\mathcal{M}$ with respect to some $\delta \in \mathbb{R}_{>0}$. We consider the *infinite* paths of $\mathcal{M}$ that are represented by a *finite* digital path.

Definition 16 (Induced cylinder of a digital path) Given a digital path $\bar{\pi} \in FPaths^{\mathcal{M}_\delta}$ of MA $\mathcal{M}$, the *induced cylinder of* $\bar{\pi}$ is given by

$$[\bar{\pi}]_{cyl} = \{\pi \in IPaths^{\mathcal{M}} \mid \bar{\pi} \text{ is a prefix of } \mathrm{di}(\pi)\}.$$

Recall the definition of the cylinder of a set of finite paths (cf. Sect. 2.2.1). If $\bar{\pi} \in FPaths^{\mathcal{M}_\delta}$ does not end with a self-loop at a Markovian state, then $[\bar{\pi}]_{cyl} = Cyl([\bar{\pi}])$ holds.

Example 11 Let $\mathcal{M}$ and $\mathcal{M}_\delta$ be as in Fig. 2. We consider the path $\bar{\pi}_1 = s_0 \xrightarrow{\perp} s_0 \xrightarrow{\perp} s_0 \xrightarrow{\perp} s_3 \xrightarrow{\beta} s_4$ and digitization constant $\delta = 0.4$. The set $[\bar{\pi}_1]_{cyl}$ contains each infinite path whose digitization has the prefix $\bar{\pi}_1$, i.e.,

$$[\bar{\pi}_1]_{cyl} = \{s_0 \xrightarrow{t} s_3 \xrightarrow{\beta} s_4 \xrightarrow{\kappa} \cdots \in IPaths^{\mathcal{M}} \mid 0.8 \le t < 1.2\}.$$

We observe that these are exactly the paths that have a prefix in $[\bar{\pi}_1]$. Put differently, we have $[\bar{\pi}_1]_{cyl} = Cyl([\bar{\pi}_1])$.

Next, consider the digital path $\bar{\pi}_2 = s_0 \xrightarrow{\perp} s_0 \xrightarrow{\perp} s_0$. Note that there is no path $\pi \in FPaths^{\mathcal{M}}$ with $\mathrm{di}(\pi) = \bar{\pi}_2$, implying $[\bar{\pi}_2] = \emptyset$. Intuitively, $\bar{\pi}_2$ depicts a sojourn time at $last(\bar{\pi}_2)$ but finite paths of MAs do not depict sojourn times at their last state. On the other hand, the induced cylinder of $\bar{\pi}_2$ contains all paths that sojourn at least 2δ time units at s_0, i.e.,

$$[\bar{\pi}_2]_{cyl} = \{s_0 \xrightarrow{t} s_1 \xrightarrow{\kappa} \cdots \in IPaths^{\mathcal{M}} \mid t \ge 0.8\}.$$

The schedulers σ and $\mathrm{di}(\sigma)$ induce the same probabilities for a given digital path. This is formalized by the following lemma. Note that a similar statement for $\mathrm{ta}(\sigma)$ and time-abstract paths was shown in Lemma 1.

Lemma 4 *Let $\mathcal{M}$ be an MA with scheduler $\sigma \in GM$, digitization $\mathcal{M}_\delta$, and digital path $\bar{\pi} \in FPaths^{\mathcal{M}_\delta}$. It holds that*

$$Pr_\sigma^{\mathcal{M}}([\bar{\pi}]_{cyl}) = Pr_{di(\sigma)}^{\mathcal{M}_\delta}(\bar{\pi}).$$

The proof is included in App. C.1. To show Proposition 4, we now apply Lemma 4. The idea of the proof is similar to the proof of Proposition 2 conducted on page 14. We include a formal proof in App. C.2.

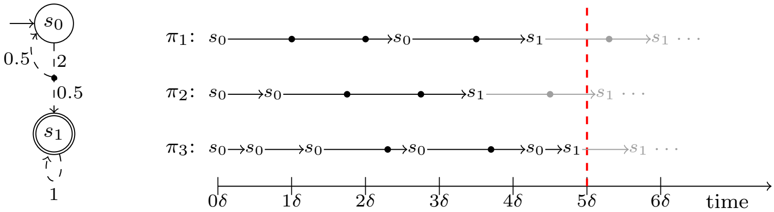
4.3.2 Timed reachability via digitization

Thus, induced ds-bounded reachability on MAs can be computed on their digitization. Next, we relate ds-bounded and timed reachability on MAs, i.e., we quantify the maximum difference between time-bounded and ds-bounded reachability probabilities.

The notation $|\bar{\pi}|_{\mathrm{ds}}$ for paths $\bar{\pi}$ of $\mathcal{M}_\delta$ is also applied to paths of $\mathcal{M}$, where $|\pi|_{\mathrm{ds}} = |\mathrm{di}(\pi)|_{\mathrm{ds}}$ for any $\pi \in FPaths^{\mathcal{M}}$. Intuitively, one digitization step represents the elapse of at most δ time units. Consequently, the duration of a path with $k \in \mathbb{N}$ digitization steps is at most $k\delta$.

Lemma 5 *For a path $\pi \in FPaths^{\mathcal{M}}$ and digitization constant δ it holds that*

$$T(\pi) \le |\pi|_{ds} \cdot \delta\ .$$



(a) MA $\mathcal{M}$. **(b)** Sample paths of $\mathcal{M}$.

Fig. 4 MA $\mathcal{M}$ and illustration of paths of $\mathcal{M}$ (cf. Example 12)

Proof Let $\pi = s_0 \xrightarrow{\kappa_0} \ldots \xrightarrow{\kappa_{n-1}} s_n$ and let $m_i = \max\{m \in \mathbb{N} \mid m\delta \leq t(\kappa_i)\}$ for each $i \in \{0, \ldots, n-1\}$ (as in Definition 12). The number $|\pi|_{\text{ds}}$ is given by $\sum_{0 \leq i < n,\, s_i \in \text{MS}} (m_i + 1)$. With $t(\kappa_i) \leq (m_i + 1)\delta$ it follows that

$$T(\pi) = \sum_{\substack{0 \leq i < n \\ s_i \in \text{MS}}} t(\kappa_i) \leq \sum_{\substack{0 \leq i < n \\ s_i \in \text{MS}}} (m_i + 1)\delta = |\pi|_{\text{ds}} \cdot \delta\ .$$

□

For a path π and $t \in \mathbb{R}_{\geq 0}$, the prefix of π up to time point t is given by $\mathit{pref}_T(\pi, t) = \mathit{pref}(\pi, \max\{n \mid T(\mathit{pref}(\pi, n)) \leq t\})$. Observe that the digitization approach yields an inaccurate estimate of the actual time. This inaccuracy is the probability that more than $k \in \mathbb{N}$ digitization steps have been performed within $k\delta$ time units.

Definition 17 (Digitization step bounded paths) Assume an MA $\mathcal{M}$ and a digitization constant $\delta \in \mathbb{R}_{>0}$. For some $t \in \mathbb{R}_{\geq 0}$, $k \in \mathbb{N}$, and $\rhd \in \{<, \leq, >, \geq\}$ the set of paths whose prefix up to time point t has $\rhd\, k$ digitization steps is defined as

$$\#[t]^{\rhd k} = \{\pi \in \mathit{IPaths}^{\mathcal{M}} \mid |\mathit{pref}_T(\pi, t)|_{\text{ds}} \rhd k\}.$$

Example 12 Let $\mathcal{M}$ be the MA given in Fig. 4a. We consider the set $\#[5\delta]^{\leq 5}$. The digitization constant δ remains unspecified in this example. Fig. 4b illustrates paths π_1, π_2, and π_3 of $\mathcal{M}$. We depict sojourn times by arrow length. For instance, the path π_1 corresponds to $s_0 \xrightarrow{2.5\delta} s_0 \xrightarrow{1.8\delta} s_1 \xrightarrow{1.7\delta} \cdots \in \mathit{IPaths}^{\mathcal{M}}$. Digitization steps that are "earned" by sojourning at some state for a multiple of δ time units are indicated by black dots. Transitions of π_i (where $i \in \{1, 2, 3\}$) that do not belong to $\mathit{pref}_T(\pi_i, 5\delta)$ are depicted in gray. We obtain

$$\begin{aligned}
|\mathit{pref}_T(\pi_1, 5\delta)|_{\text{ds}} = 5 \quad &\Longrightarrow \quad \pi_1 \in \#[5\delta]^{\leq 5} \\
|\mathit{pref}_T(\pi_2, 5\delta)|_{\text{ds}} = 4 \quad &\Longrightarrow \quad \pi_2 \in \#[5\delta]^{\leq 5} \\
|\mathit{pref}_T(\pi_3, 5\delta)|_{\text{ds}} = 7 \quad &\Longrightarrow \quad \pi_3 \notin \#[5\delta]^{\leq 5}\ .
\end{aligned}$$

Note that only the digitization steps of the prefix up to time point 5δ are considered. For example, the step of π_2 at time point 4.5δ is not considered since the corresponding transition is not part of $\mathit{pref}_T(\pi_2, 5\delta)$. However, we have $|\mathit{pref}_T(\pi_2, 5.5\delta)|_{\text{ds}} = 6$, implying $\pi_2 \notin \#[5.5\delta]^{\leq 5}$.

All considered paths reach $G = \{s_1\}$ within 5δ time units but $\pi_3 \in \#[5\delta]^{>5}$ requires more than 5 digitization steps. This yields

$$\pi_1, \pi_2 \in \Diamond^I G \cap [\Diamond^{\text{di}(I)}_{\text{ds}} G] \quad \text{and} \quad \pi_3 \in \Diamond^I G \setminus [\Diamond^{\text{di}(I)}_{\text{ds}} G].$$

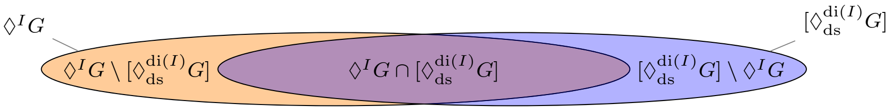


Fig. 5 Illustration of the sets $\Diamond^I G$ and $[\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G]$

The sets $\Diamond^I G$ and $[\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G]$ are illustrated in Fig. 5. We have

$$\mathrm{Pr}_\sigma(\Diamond^I G) = \mathrm{Pr}_\sigma([\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G]) + \mathrm{Pr}_\sigma(\Diamond^I G \setminus [\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G]) - \mathrm{Pr}_\sigma([\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G] \setminus \Diamond^I G).$$

One then shows

$$\mathrm{Pr}^{\mathcal{M}}_\sigma(\Diamond^I G \setminus [\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G]) \leq \varepsilon^\uparrow(I) \quad \text{and} \quad \mathrm{Pr}^{\mathcal{M}}_\sigma([\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G] \setminus \Diamond^I G) \leq \varepsilon^\downarrow(I).$$

for adequate error bounds $\varepsilon^\uparrow(I), \varepsilon^\downarrow(I)$. In particular, the following lemma gives an upper bound for the probability $\mathrm{Pr}^{\mathcal{M}}_\sigma(\#[k\delta]^{>k})$, i.e., the probability that more than $k \in \mathbb{N}$ digitization steps have been performed within $k\delta$ time units. Then, this probability can be related to the probability of paths in $\Diamond^I G \setminus [\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G]$ and $[\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G] \setminus \Diamond^I G$, respectively.

Lemma 6 *Let $\mathcal{M}$ be an MA with $\sigma \in GM$ and maximum rate $\lambda = \max\{E(s) \mid s \in MS\}$. Further, let $\delta \in \mathbb{R}_{>0}$ and $k \in \mathbb{N}$. It holds that*

$$Pr^{\mathcal{M}}_\sigma(\#[k\delta]^{>k}) \leq 1 - (1 + \lambda\delta)^k \cdot e^{-\lambda\delta k}$$

For the proof of Lemma 6 we employ the following auxiliary lemma.

Lemma 7 *Let $\mathcal{M}$ be an MA with $\sigma \in GM$ and maximum rate $\lambda = \max\{E(s) \mid s \in MS\}$. For each $\delta \in \mathbb{R}_{>0}$, $k \in \mathbb{N}$, and $t \in \mathbb{R}_{\geq 0}$ it holds that*

$$Pr^{\mathcal{M}}_\sigma(\#[k\delta + t]^{\leq k}) \geq Pr^{\mathcal{M}}_\sigma(\#[k\delta]^{\leq k}) \cdot e^{-\lambda t} \ .$$

Proofs for both lemmas are given in App. C.3.

We have now obtained all necessary insights to bound $\mathrm{Pr}^{\mathcal{M}}_\sigma(\Diamond^I G)$ from above and below, using the digitized steps and the following error bounds. Let $\lambda = \max\{\mathrm{E}(s) \mid s \in \mathrm{MS}\}$ be the maximum exit rate of $\mathcal{M}$. For $a \neq 0$ define

$$\varepsilon^\downarrow([a,b]) = \varepsilon^\downarrow([a,\infty)) = 1 - (1+\lambda\delta)^{\mathrm{di}_a} \cdot e^{-\lambda a} \ , \quad \varepsilon^\downarrow([0,b)) = \varepsilon^\downarrow([0,\infty]) = 0,$$
$$\varepsilon^\uparrow([a,b]) = \underbrace{1 - (1+\lambda\delta)^{\mathrm{di}_b} \cdot e^{-\lambda b}}_{=\varepsilon^\uparrow([0,b])} + \underbrace{1 - e^{-\lambda\delta}}_{=\varepsilon^\uparrow([a,\infty))} \ , \text{ and } \ \varepsilon^\uparrow([0,\infty)) = 0.$$

$\varepsilon^\downarrow(I)$ and $\varepsilon^\uparrow(I)$ approach 0 for small digitization constants $\delta \in \mathbb{R}_{>0}$.

Proposition 5 *For MA $\mathcal{M}$, scheduler $\sigma \in GM$, goal states $G \subseteq S$, digitization constant $\delta \in \mathbb{R}_{>0}$ and time interval I*

$$Pr^{\mathcal{M}}_\sigma(\Diamond^I G) \in Pr^{\mathcal{M}}_\sigma([\Diamond^I_{ds} G]) + \Big[-\varepsilon^\downarrow(I), \ \varepsilon^\uparrow(I)\Big].$$

Proof We already discussed that

$$\Pr_\sigma(\Diamond^I G) = \Pr_\sigma([\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G]) + \Pr_\sigma(\Diamond^I G \backslash [\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G]) - \Pr_\sigma([\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G] \backslash \Diamond^I G).$$

The main part of the proof is to show that

$$\Pr^{\mathcal{M}}_\sigma([\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G] \backslash \Diamond^I G) \leq \varepsilon^{\downarrow}(I) \text{ and } \Pr^{\mathcal{M}}_\sigma(\Diamond^I G \backslash [\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G]) \leq \varepsilon^{\uparrow}(I). \quad (3)$$

Then, the proposition follows directly. We show Equation 3 for the different forms of the time interval I. Here, we consider intervals of the form $I = [0, b]$. Recall that by assumption $b = \mathrm{di}_b \cdot \delta$. For the other forms, we refer to App. C.4.

We have $\mathrm{di}(I) = \{0, 1, \ldots, \mathrm{di}_b\}$.

– We show that $[\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G] \subseteq \Diamond^I G$ which implies

$$\Pr^{\mathcal{M}}_\sigma([\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G] \backslash \Diamond^I G) = \Pr^{\mathcal{M}}_\sigma(\emptyset) = 0 = \varepsilon^{\downarrow}(I).$$

Let $\pi \in [\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G]$ and let π' be the smallest prefix of π with $last(\pi') \in G$. It follows that $\mathrm{di}(\pi')$ is also the smallest prefix of $\mathrm{di}(\pi)$ with $last(\mathrm{di}(\pi')) \in G$. Since $\mathrm{di}(\pi) \in \Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G$, it follows that $|\pi'|_{\mathrm{ds}} = |\mathrm{di}(\pi')|_{\mathrm{ds}} \leq \mathrm{di}_b$. From Lemma 5 we obtain

$$T(\pi') \leq |\pi'|_{\mathrm{ds}} \cdot \delta = |\mathrm{di}(\pi')|_{\mathrm{ds}} \cdot \delta \leq \mathrm{di}_b \delta = b\ .$$

Hence, the prefix π' reaches G within b time units, implying $\pi \in \Diamond^I G$.

– Next, we show $\Diamond^I G \backslash [\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G] \subseteq \#[b]^{>\mathrm{di}_b}$. With Lemma 6 we obtain

$$\Pr^{\mathcal{M}}_\sigma(\Diamond^I G \backslash [\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G]) \leq \Pr^{\mathcal{M}}_\sigma(\#[b]^{>\mathrm{di}_b}) \leq 1 - (1 + \lambda\delta)^{\mathrm{di}_b} \cdot e^{-\lambda b} = \varepsilon^{\uparrow}(I)$$

Consider a path $\pi \in \Diamond^I G \backslash [\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G]$. Note that π reaches G within b time units but with more than di_b digitization steps. Hence, the prefix of π up to time point b certainly has more than di_b digitization steps, i.e., π satisfies $|pref_T(\pi, b)|_{\mathrm{ds}} > \mathrm{di}_b$ which means $\pi \in \#[b]^{>\mathrm{di}_b}$.

□

From Prop. 4 and Prop. 5, we immediately have Cor. 1, which ensures that the value $\Pr^{\mathcal{M}}_\sigma(\Diamond^I G)$ can be approximated with arbitrary precision by computing $\Pr^{\mathcal{M}_\delta}_{\mathrm{di}(\sigma)}(\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G)$ for a sufficiently small δ.

Corollary 1 *For MA $\mathcal{M}$, scheduler $\sigma \in GM$, goal states $G \subseteq S$, digitization constant $\delta \in \mathbb{R}_{>0}$ and time interval I*

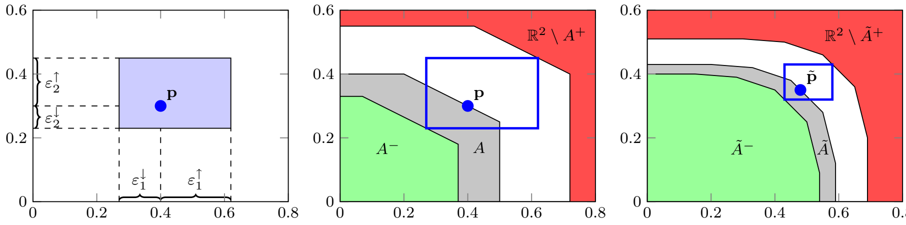
$$Pr^{\mathcal{M}}_\sigma(\Diamond^I G) \in Pr^{\mathcal{M}_\delta}_{di(\sigma)}(\Diamond^{di(I)}_{ds} G) + \left[-\varepsilon^{\downarrow}(I),\ \varepsilon^{\uparrow}(I)\right].$$

Corollary 1 generalizes existing results from single-objective timed reachability analysis: For MA $\mathcal{M}$, goal states G, time bound $b \in \mathbb{R}_{>0}$, and digitization constant $\delta \in \mathbb{R}_{>0}$ with $\frac{b}{\delta} = \mathrm{di}_b \in \mathbb{N}$, [30, Theorem 5.3] states that

$$\sup_{\sigma \in \mathrm{GM}^{\mathcal{M}}} \Pr^{\mathcal{M}}_\sigma(\Diamond^{[0,b]} G) \in \sup_{\sigma \in \mathrm{TA}^{\mathcal{M}_\delta}} \Pr^{\mathcal{M}_\delta}_\sigma(\Diamond^{\{0,\ldots,\mathrm{di}_b\}}_{\mathrm{ds}} G) + \left[-\varepsilon^{\downarrow}([0,b]),\ \varepsilon^{\uparrow}([0,b])\right].$$

Corollary 1 generalizes this result by explicitly referring to the schedulers $\sigma \in \mathrm{GM}^{\mathcal{M}}$ and $\mathrm{di}(\sigma) \in \mathrm{TA}^{\mathcal{M}_\delta}$ under which the claim holds. This extension is necessary as a multi-objective analysis can not be restricted to schedulers that only optimize a single objective. More details are given in App. D.

Next, we lift Cor. 1 to multiple objectives $\mathbb{O} = (\mathbb{O}_1, \ldots, \mathbb{O}_d)$. We define the satisfaction of a *timed* reachability objective $\mathbb{P}(\Diamond^I G)$ for the digitization $\mathcal{M}_\delta$ as $\mathcal{M}_\delta, \sigma \models \mathbb{P}(\Diamond^I G) \rhd_i$



(a) The set $\varepsilon(\mathbb{O}, \mathbf{p})$. **(b)** Coarse approximation. **(c)** Refined approximation.

Fig. 6 Approximation of achievable points

p_i iff $\Pr_{\sigma}^{\mathcal{M}_\delta}(\Diamond_{\text{ds}}^{\text{di}(I)} G) \rhd_i p_i$. This allows us to consider notations like $achieve^{\mathcal{M}_\delta}(\mathbb{O} \rhd \mathbf{p})$, where $\mathbb{O}$ contains one or more timed reachability objectives. For a point $\mathbf{p} = (p_1, \dots, p_d) \in \mathbb{R}^d$ we consider the hyperrectangle

$$\varepsilon(\mathbb{O}, \mathbf{p}) = \bigtimes_{i=1}^{d} \left[p_i - \varepsilon_i^{\downarrow},\ p_i + \varepsilon_i^{\uparrow}\right] \subseteq \mathbb{R}^d \text{ , where } \varepsilon_i^{\uparrow} = \begin{cases} \varepsilon^{\uparrow}(I) & \text{if } \mathbb{O}_i = \mathbb{P}(\Diamond^I G) \\ 0 & \text{if } \mathbb{O}_i = \mathbb{E}(\#j, G). \end{cases}$$

and $\varepsilon_i^{\downarrow}$ is defined similarly. The next example shows how the set of achievable points of $\mathcal{M}$ can be approximated using achievable points of $\mathcal{M}_\delta$.

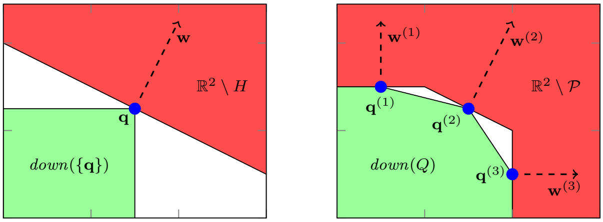
Example 13 Let $\mathbb{O} = (\mathbb{P}(\Diamond^{I_1} G_1), \mathbb{P}(\Diamond^{I_2} G_2))$ be two timed reachability objectives for an MA $\mathcal{M}$ with digitization $\mathcal{M}_\delta$ such that $\varepsilon_1^{\downarrow} = 0.13$, $\varepsilon_1^{\uparrow} = 0.22$, $\varepsilon_2^{\downarrow} = 0.07$, and $\varepsilon_2^{\uparrow} = 0.15$. The blue rectangle in Fig. 6a illustrates the set $\varepsilon(\mathbb{O}, \mathbf{p})$ for the point $\mathbf{p} = (0.4, 0.3)$. Assume $achieve^{\mathcal{M}_\delta}(\mathbb{O} \rhd \mathbf{p})$ holds for threshold relations $\rhd = \{\geq, \geq\}$, i.e., $\mathbf{p}$ is achievable for the digitization $\mathcal{M}_\delta$. From Cor. 1, we infer that $\varepsilon(\mathbb{O}, \mathbf{p})$ contains at least one point $\mathbf{p}'$ that is achievable for $\mathcal{M}$. Hence, the bottom left corner point of the rectangle is achievable for $\mathcal{M}$. This holds for any rectangle $\varepsilon(\mathbb{O}, \mathbf{q})$ with $\mathbf{q} \in A$, where A is the set of achievable points of $\mathcal{M}_\delta$ denoted by the gray area[3] in Fig. 6b. It follows that any point in A^- (depicted by the green area) is achievable for $\mathcal{M}$. On the other hand, an achievable point of $\mathcal{M}$ has to be contained in a set $\varepsilon(\mathbb{O}, \mathbf{q})$ for at least one $\mathbf{q} \in A$. The red area depicts the points $\mathbb{R}^d \setminus A^+$ for which this is not the case, i.e., points that are not achievable for $\mathcal{M}$. The digitization constant δ controls the accuracy of the resulting approximation. Figure 6c depicts a possible result when a smaller digitization constant $\tilde{\delta} < \delta$ is considered.

The observations from the example above are formalized in the following theorem. The theorem also covers unbounded reachability objectives by considering the time interval $I = [0, \infty)$. For expected reward objectives of the form $\mathbb{E}(\#j, G)$ it can be shown that $\mathrm{ER}_{\sigma}^{\mathcal{M}}(\rho_j, G) = \mathrm{ER}_{\mathrm{di}(\sigma)}^{\mathcal{M}_\delta}(\rho_j^\delta, G)$. This claim is similar to Proposition 3 and can be shown analogously. This enables multi-objective model checking of MAs with timed reachability objectives.

Theorem 5 *Let $\mathcal{M}$ be an MA with digitization $\mathcal{M}_\delta$. Furthermore, let $\mathbb{O}$ be (un)timed reachability or expected reward objectives with threshold relations $\rhd$ and $|\mathbb{O}| = d$. It holds that $A^- \subseteq \{\boldsymbol{p} \in \mathbb{R}^d \mid achieve^{\mathcal{M}}(\mathbb{O} \rhd \boldsymbol{p})\} \subseteq A^+$ with:*

$$A^- = \{\boldsymbol{p}' \in \mathbb{R}^d \mid \forall \boldsymbol{p} \in \mathbb{R}^d : \boldsymbol{p}' \in \varepsilon(\mathbb{O}, \boldsymbol{p}) \text{ implies } achieve^{\mathcal{M}_\delta}(\mathbb{O} \rhd \boldsymbol{p})\} \text{ and}$$
$$A^+ = \{\boldsymbol{p}' \in \mathbb{R}^d \mid \exists \boldsymbol{p} \in \mathbb{R}^d : \boldsymbol{p}' \in \varepsilon(\mathbb{O}, \boldsymbol{p}) \text{ and } achieve^{\mathcal{M}_\delta}(\mathbb{O} \rhd \boldsymbol{p})\}.$$

[3] In the figure, A^- partly overlaps A, i.e., the green area also belongs to A.



(a) Single weight vector. **(b)** Multiple weight vectors.

Fig. 7 Illustration of the Pareto curve approximation algorithm (cf. Example 14)

Proof For simplicity, we assume that only the threshold relation $\geq$ is considered, i.e., $\rhd = (\geq, \dots, \geq)$. Furthermore, we restrict ourself to (un)timed reachability objectives. The remaining cases are treated analogously.

First assume a point $\mathbf{p}' = (p'_1, \dots, p'_d) \in A^-$. Consider the point $\mathbf{p} = (p_1, \dots, p_d)$ satisfying $p'_i = p_i - \varepsilon_i^{\downarrow}$ for each index i. It follows that $\mathbf{p}' \in \varepsilon(\mathbb{O}, \mathbf{p})$ and thus $\mathcal{M}_\delta, \bar{\sigma} \models \mathbb{O} \rhd \mathbf{p}$ for some scheduler $\bar{\sigma} \in \mathrm{TA}^{\mathcal{M}_\delta}$. Consider the scheduler $\sigma \in \mathrm{GM}^{\mathcal{M}}$ given by $\sigma(\pi, \alpha) = \bar{\sigma}(\mathrm{di}(\pi), \alpha)$ for each path $\pi \in \mathit{FPaths}^{\mathcal{M}}$ and action $\alpha \in \mathit{Act}$. Notice that $\bar{\sigma} = \mathrm{di}(\sigma)$. For an index i let $\mathbb{O}_i$ be the objective $\mathbb{P}(\Diamond^I G)$. It follows that

$$\mathcal{M}_\delta, \bar{\sigma} \models \mathbb{O}_i \geq p_i \iff \mathcal{M}_\delta, \mathrm{di}(\sigma) \models \mathbb{O}_i \geq p_i \iff \mathrm{Pr}^{\mathcal{M}_\delta}_{\mathrm{di}(\sigma)}(\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G) \geq p_i \ ,$$

With Corollary 1 it follows that

$$p'_i = p_i - \varepsilon_i^{\downarrow} \leq \mathrm{Pr}^{\mathcal{M}_\delta}_{\mathrm{di}(\sigma)}(\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G) - \varepsilon_i^{\downarrow} \overset{Cor.\,1}{\leq} \mathrm{Pr}^{\mathcal{M}}_{\sigma}(\Diamond^I G).$$

As this observation holds for all objectives in $\mathbb{O}$, it follows that $\mathcal{M}, \sigma \models \mathbb{O} \rhd \mathbf{p}'$, implying $\mathit{achieve}^{\mathcal{M}}(\mathbb{O} \rhd \mathbf{p}')$.

The proof of the second inclusion is similar. Assume that $\mathcal{M}, \sigma \models \mathbb{O} \rhd \mathbf{p}'$ holds for a point $\mathbf{p}' = (p'_1, \dots, p'_d) \in \mathbb{R}^d$ and a scheduler $\sigma \in \mathrm{GM}^{\mathcal{M}}$. For some index i, consider $\mathbb{O}_i = \mathbb{P}(\Diamond^I G)$. It follows that $\mathrm{Pr}^{\mathcal{M}}_{\sigma}(\Diamond^I G) \geq p'_i$. With Corollary 1 we obtain

$$p'_i - \varepsilon_i^{\uparrow} \leq \mathrm{Pr}^{\mathcal{M}}_{\sigma}(\Diamond^I G) - \varepsilon_i^{\uparrow} \overset{Cor.\,1}{\leq} \mathrm{Pr}^{\mathcal{M}_\delta}_{\mathrm{di}(\sigma)}(\Diamond^{\mathrm{di}(I)}_{\mathrm{ds}} G).$$

Applying this for all objectives in $\mathbb{O}$ yields $\mathcal{M}_\delta, \mathrm{di}(\sigma) \models \mathbb{O} \rhd \mathbf{p}$, where the point $\mathbf{p} = (p_1, \dots, p_d) \in \mathbb{R}^d$ satisfies $p_i = p'_i - \varepsilon_i^{\uparrow}$ or, equivalently, $p'_i = p_i + \varepsilon_i^{\uparrow}$ for each index i. Note that $\mathbf{p}' \in \varepsilon(\mathbb{O}, \mathbf{p})$ which implies $\mathbf{p}' \in A^+$. □

5 Computation of Pareto curves

We have seen that we can reduce the analysis of multiple objectives on MA to multi-objective MDPs to compute the achievable points of the underlying MDP $\mathcal{M}_{\mathcal{D}}$ or a digitization $\mathcal{M}_\delta$ of MA $\mathcal{M}$. To analyze the MDPs, we adapt the approach of [28]. In this section, we briefly recap that approach and report on the necessary changes. The approach repeatedly checks weighted combinations of the objectives (by means of *value iteration* [43]—a standard technique in single-objective MDP model checking) to refine an approximation of the set of achievable points. The following example demonstrates this idea for two objectives.

Example 14 Consider an MDP $\mathcal{D} = (S, Act, \mathbf{P}, s_0, (\rho_1, \ldots, \rho_\ell))$ with two objectives $\mathbb{O} = (\mathbb{O}_1, \mathbb{O}_2)$ and relations $\rhd = \{\geq, \geq\}$. The approach of [28] iteratively considers *weight vectors* $\mathbf{w} = (w_1, w_2) \in (\mathbb{R}_{\geq 0})^2$ with $\mathbf{w} \neq (0, 0)$ that assign a weight $w_i \geq 0$ to each objective $\mathbb{O}_i$. Optimizing a combination of the weighted objectives (see Def. 18) yields a point $\mathbf{q} \in \mathbb{R}^2$ such that

- $\mathbf{q}$ is achievable and
- all achievable points of $\mathcal{D}$ are contained in the *half-space* $H = \{\mathbf{p} \in \mathbb{R}^2 \mid \mathbf{p} \cdot \mathbf{w} \leq \mathbf{q} \cdot \mathbf{w}\}$.

Since $\mathbf{q}$ is achievable, any point in the set $down(\{\mathbf{q}\}) = \{\mathbf{p} \in \mathbb{R}^2 \mid \mathbf{p} \leq \mathbf{q}\}$ depicted by the green area of Fig. 7a is achievable as well. On the contrary, there is no achievable point in $\mathbb{R}^2 \setminus H$, illustrated by the red area. For the points in the white area, it is still unknown whether they are achievable or not. The set of achievable points of $\mathcal{D}$ is explored by combining the results for multiple weight-vectors as indicated in Fig. 7b.

The sketched approach theoretically converges to an exact representation of the set of achievable points, but the number of required calls to value iteration can be exponential in the size of the MDP and the number of objectives [28]. However, experiments in [28] and in Sect. 6 of this work indicate that, in practice, only a small number of weight vectors need to be considered in order to obtain "good" approximations.

We extend [28] towards

- the simultaneous analysis of minimizing and maximizing expected reward objectives, and
- the analysis of ds-bounded reachability objectives.

These extensions only concern the computation of *optimal points*. The remaining aspects of the approach are as in [28]. We first restrict our attention to maximizing *expected total reward objectives*, i.e., expected reward objectives of the form $\mathbb{E}(\#j, G)$ with $G = \emptyset$.

Definition 18 (Optimal Scheduler, Optimal Point) Let $\mathcal{D}$ be an MDP and let $\mathbb{O} = (\mathbb{E}(\#j_1, \emptyset), \ldots, \mathbb{E}(\#j_d, \emptyset))$ be expected total reward objectives with threshold relations $\rhd = (\rhd_1, \ldots, \rhd_d)$ and $\rhd_i \in \{\geq, >\}$. A scheduler $\sigma \in \mathrm{TA}$ is called *optimal* for a weight vector $\mathbf{w} = (w_1, \ldots, w_d) \in \mathbb{R}^d_{\geq 0} \setminus \{\mathbf{0}\}$ iff

$$\sigma \in \arg\max_{\sigma' \in \mathrm{TA}} \Big(\sum_{i=0}^{d} w_i \cdot \mathrm{ER}^{\mathcal{D}}_{\sigma'}(\rho_{j_i}, \emptyset) \Big).$$

A point $\mathbf{p} = (p_1, \ldots, p_d) \in \mathbb{R}^d$ is *optimal* for $\mathbf{w}$ iff $p_i = \mathrm{ER}^{\mathcal{D}}_{\sigma}(\rho_{j_i}, \emptyset)$ for all i and optimal σ.

The computation of optimal points for a weighted sum of expected total reward objectives with thresholds $\rhd_i \in \{\geq, >\}$ is detailed in [28]. The idea is to use *value iteration* [43] to compute an optimal scheduler σ_{opt} for the maximal expected reward $\max_{\sigma \in \mathrm{TA}} \mathrm{ER}^{\mathcal{D}}_{\sigma}(\rho_{\mathbf{w}}, \emptyset)$ with the weighted reward function $\rho_{\mathbf{w}}$ given by $\rho_{\mathbf{w}}(s, \alpha) = \sum_{i=1}^{d} w_i \cdot \rho_{j_i}(s, \alpha)$ for each $s \in S$ and $\alpha \in Act$. Evaluating the individual objectives with respect to σ_{opt} yields the entries of an optimal point $\mathbf{p} = (p_1, \ldots, p_d)$, i.e., $p_i = \mathrm{ER}^{\mathcal{D}}_{\sigma_{\mathrm{opt}}}(\rho_{j_i}, \emptyset)$ for all i. In general, we might have $p_i = \infty$ and thus $\mathbf{p} \notin \mathbb{R}^d$. In this case, the sketched approach is not applicable. We therefore impose some assumptions.

Assumptions on reward finiteness For maximizing objectives $\mathbb{E}(\#j_i, \emptyset)$ with $\rhd_i \in \{\geq, >\}$ we assume $\max_{\sigma \in \text{TA}} \text{ER}^{\mathcal{D}}_{\sigma}(\rho_{j_i}, \emptyset) < \infty$—following the suggestions of [28,29]. Note that if a scheduler σ_∞ with $\text{ER}^{\mathcal{D}}_{\sigma_\infty}(\rho_{j_i}, \emptyset) = \infty$ exists, we can also construct schedulers σ with $\text{ER}^{\mathcal{D}}_{\sigma}(\rho_{j_i}, \emptyset) = \infty$ that mimic σ_∞ with an arbitrarily small (but non-zero) probability and otherwise focus on the remaining objectives. We further assume that there is at least one scheduler σ inducing $\text{ER}^{\mathcal{D}}_{\sigma}(\rho_{j_i}, \emptyset) < \infty$ for *all* minimizing objectives $\mathbb{E}(\#j_i, \emptyset)$ with $\rhd_i \in \{\leq, <\}$. If this is not the case, there is no achievable point $\mathbf{p} \in \mathbb{R}^d$ at all. These assumptions can be checked algorithmically via a graph analysis.

Remark 2 (Reachability and (general) expected reward objectives) A transformation of untimed reachability objectives $\mathbb{P}(\Diamond G)$ to expected total reward objectives is given in [28,29]. Roughly, the state space of the MDP is unfolded, yielding two copies of each state. Transitions leading to a state in G are redirected to the second copy, allowing us to store whether a state in G has already been visited. A reward of 1 is collected whenever a goal state is visited for the first time, i.e., when we move from the first copy to the second one. A similar unfolding technique can be applied to transform expected reward objectives $\mathbb{E}(\#j, G)$ with $G \neq \emptyset$ to total reward objectives. This approach increases the number of considered states by a factor of up to 2^d, where d is the number of objectives.

5.1 Treatment of minimizing objectives

We now consider expected total reward objectives with arbitrary threshold relations. Let $\mathcal{D} = (S, Act, \mathbf{P}, s_0, (\rho_1, \ldots, \rho_\ell))$ be an MDP and let $\mathbb{O} = (\mathbb{E}(\#j_1, \emptyset), \ldots, \mathbb{E}(\#j_d, \emptyset))$ be expected total reward objectives with threshold relations $\rhd = (\rhd_1, \ldots, \rhd_d)$. Without loss of generality, let each objective consider a different reward function, i.e., the indices $j_1, \ldots, j_d$ are pairwise distinct. We further simplify the notations by assuming $j_i = i$ for $i \in \{1, \ldots, d\}$, i.e. the i-th objective considers reward function ρ_i. We proceed in three steps:

1. Convert all *minimizing* objectives $\mathbb{E}(\#j_i, \emptyset)$ with $\rhd_i \in \{\leq, <\}$ to *maximizing* objectives, potentially introducing negative rewards.
2. Compute an optimal scheduler for the maximal expected reward for a weighted reward function that considers positive- and negative rewards.
3. Lift further reward finiteness assumptions imposed in Step 2.

5.1.1 From minimizing to maximizing objectives

We convert minimizing objectives $\mathbb{E}(\#i, \emptyset)$ with $\rhd_i \in \{\leq, <\}$ to maximizing objectives by negating the considered rewards, and thereby deviate from our definition of MDPs from Sect. 2 by allowing negative rewards.

More precisely, we consider the reward functions $\overline{\rho}_1, \ldots, \overline{\rho}_d \colon S \times Act \to \mathbb{R}$ and relations $\overline{\rhd} = (\overline{\rhd}_1, \ldots, \overline{\rhd}_d)$, where for $i \in \{1, \ldots, d\}$, $s \in S$, and $\alpha \in Act$:

$$\overline{\rho}_i(s, \alpha) = \begin{cases} -\rho_i(s, \alpha) & \text{if } \rhd_i \in \{\leq, <\} \\ \rho_i(s, \alpha) & \text{otherwise} \end{cases} \quad \text{and} \quad \overline{\rhd}_i = \begin{cases} \geq & \text{if } \rhd_i = \leq \\ > & \text{if } \rhd_i = < \\ \rhd_i & \text{otherwise.} \end{cases}$$

Let $\overline{\mathcal{D}} = (S, Act, \mathbf{P}, s_0, (\overline{\rho}_1, \ldots, \overline{\rho}_d))$. For each scheduler σ for $\mathcal{D}$ (and $\overline{\mathcal{D}}$), $i \in \{1, \ldots, d\}$, and $p_i \in \mathbb{R}$ we have

$$\text{ER}^{\mathcal{D}}_{\sigma}(\rho_i, \emptyset) \leq p_i \iff -\text{ER}^{\mathcal{D}}_{\sigma}(\rho_i, \emptyset) \geq -p_i \iff \text{ER}^{\mathcal{D}}_{\sigma}(\overline{\rho}_i, \emptyset) \geq -p_i \; .$$

Lifting to multiple objectives yields for all $\mathbf{p} = (p_1, \dots, p_d) \in \mathbb{R}^d$:

$$achieve^{\mathcal{D}}(\mathbb{O} \rhd \mathbf{p}) \iff achieve^{\overline{\mathcal{D}}}(\mathbb{O} \,\overline{\rhd}\, \overline{\mathbf{p}}),$$

where $\overline{\mathbf{p}} = (\overline{p}_1, \dots, \overline{p}_d)$ with $\overline{p}_i = -p_i$ if $\rhd_i \in \{\leq, <\}$ and otherwise $\overline{p}_i = p_i$.

We can thus compute (or approximate) the set of achievable points for $\overline{\mathcal{D}}$ and $\overline{\rhd}$ instead of $\mathcal{D}$ and $\rhd$. More concretely, we employ the approach of [28] to approximate the set $\{\mathbf{p} \in \mathbb{R}^d \mid achieve^{\overline{\mathcal{D}}}(\mathbb{O} \,\overline{\rhd}\, \mathbf{p})\}$ where only maximizing objectives are considered. The result for $\mathcal{D}$ and $\rhd$ can then be obtained by a simple transformation, essentially multiplying entries of minimizing objectives by -1.

5.1.2 Mixtures of positive and negative rewards

As mentioned above, the approach of [28] requires to repeatedly compute an optimal scheduler for weighted reward functions. Considering the MDP $\overline{\mathcal{D}} = (S, Act, \mathbf{P}, s_0, (\overline{\rho}_1, \dots, \overline{\rho}_d))$ from above results in some technical complications due to the presence of positive and negative rewards.

For simplicity, we further strengthen our assumptions on reward finiteness by assuming that *all* induced expected rewards are finite, i.e., $\mathrm{ER}^{\overline{\mathcal{D}}}_{\sigma}(\overline{\rho}_i, \emptyset) \notin \{-\infty, \infty\}$ for all $\sigma \in \mathrm{TA}$ and $i \in \{1, \dots, d\}$. The next section discusses how this assumption can be lifted again.

For a weight vector $\mathbf{w} = (w_1, \dots, w_d) \in \mathbb{R}^d_{\geq 0} \setminus \{\mathbf{0}\}$ we consider the weighted reward function $\rho_{\mathbf{w}}$ given by $\rho_{\mathbf{w}}(s, \alpha) = \sum_{i=1}^{d} w_i \cdot \overline{\rho}_i(s, \alpha)$ for $s \in S$ and $\alpha \in Act$ and let $\mathcal{D}^{\mathbf{w}} = (S, Act, \mathbf{P}, s_0, (\rho_{\mathbf{w}}))$ denote the MDP which arises from $\overline{\mathcal{D}}$ by replacing the reward functions $(\overline{\rho}_1, \dots, \overline{\rho}_d)$ by $\rho_{\mathbf{w}}$. Our goal in this section is to compute an optimal scheduler for $\mathbf{w}$ (cf. Def. 18), i.e., a scheduler $\sigma_{\mathrm{opt}} \in \arg\max_{\sigma \in \mathrm{TA}} \mathrm{ER}^{\mathcal{D}^{\mathbf{w}}}_{\sigma}(\rho_{\mathbf{w}}, \emptyset)$ inducing the maximal expected total reward for $\rho_{\mathbf{w}}$. Since $\rho_{\mathbf{w}}$ considers both, positive- and negative rewards, conventional value iteration as considered in [28] yields incorrect results.

Example 15 Consider the MDP $\mathcal{D}^{\mathbf{w}}$ with the weighted reward function $\rho_{\mathbf{w}}$ depicted in Fig. 8a. Action rewards are depicted next to the action label, e.g., $\rho_{\mathbf{w}}(s_0, \alpha) = 2$. The maximal expected total reward is obtained for a scheduler σ_{opt} that always chooses action α, yielding $\mathrm{ER}_{\sigma_{\mathrm{opt}}}(\rho_{\mathbf{w}}, \emptyset) = 2 - 1 = 1$.

On the other hand, value iteration as suggested in [28, Alg. 2] yields a value of 2 and a suboptimal scheduler σ with $\sigma(s_0) = \beta$. Roughly speaking, this is because value iteration computes the expected reward accumulated within n steps for an increasing value of n. However, for the example MDP, this step-bounded expected reward does not converge to the *unbounded* expected total reward: For a given step-bound n, an optimal scheduler can avoid to collect the reward $\rho_{\mathbf{w}}(s_1, \alpha) = -1$ by taking the β-self-loop at s_0 $n - 1$ times and then taking the α-transition to s_1 only in the very last step.

In the example, the problem is introduced by the action β at state s_0. This action allows a scheduler to stay at s_0 arbitrarily long without collecting any reward. We refer to such model components as 0-EC.

Definition 19 (End Component, 0-EC) A non-empty set of state-action pairs $\mathcal{E} \subseteq S \times Act$ is an *end component (EC)* of $\mathcal{D}^{\mathbf{w}}$ if

1. $\forall (s, \alpha) \in \mathcal{E}$: $\mathbf{P}(s, \alpha, s') > 0$ implies $s' \in states(\mathcal{E}) = \{\tilde{s} \mid \exists (\tilde{s}, \tilde{\alpha}) \in \mathcal{E}\}$ and
2. the graph $\big(states(\mathcal{E}), \ \{(s, s') \mid \exists (s, \alpha) \in \mathcal{E}\colon \mathbf{P}(s, \alpha, s') > 0\}\big)$ is strongly connected.

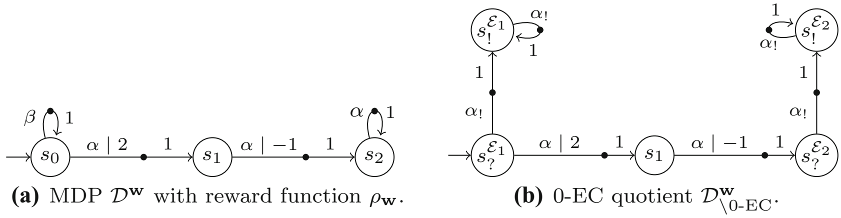


Fig. 8 MDP where value iteration yields wrong results and the corresponding 0-EC quotient (cf. Examples 15 to 17)

An EC $\mathcal{E}$ is a 0-EC w.r.t. reward function $\rho_{\mathbf{w}}$ if $\rho_{\mathbf{w}}(s, \alpha) = 0$ for all $(s, \alpha) \in \mathcal{E}$. A (0-)EC $\mathcal{E}$ is *maximal*, if there is no other (0-)EC $\mathcal{E}'$ with $\mathcal{E} \subsetneq \mathcal{E}'$.

Example 16 The MDP $\mathcal{D}^{\mathbf{w}}$ in Fig. 8a has two maximal 0-EC w.r.t. the depicted reward function $\rho_{\mathbf{w}}$: $\mathcal{E}_1 = \{(s_0, \beta)\}$ and $\mathcal{E}_2 = \{(s_2, \alpha)\}$.

The set of all maximal ECs of an MDP can be computed efficiently [17]. The set of maximal 0-ECs can be obtained by computing the maximal ECs of a modified MDP in which all transitions incurring non-zero reward are erased. However, in this particular case we know that all ECs (reachable from s_0) actually are 0-ECs: If an EC is not a 0-EC, we could construct a scheduler inducing infinite reward for at least one objective which violates our assumption that all induced expected rewards are finite.

For an EC $\mathcal{E}$ we define $\mathit{states}(\mathcal{E}) := \{s \in S \mid \exists \alpha \in \mathit{Act}(s) : (s, \alpha) \in \mathcal{E}\}$ and $\mathit{exits}(\mathcal{E}) := \{(s, \alpha) \in S \times \mathit{Act} \mid s \in \mathit{states}(\mathcal{E}), \alpha \in \mathit{Act}(s), \text{ and } (s, \alpha) \notin \mathcal{E}\}$.

Intuitively, once a maximal 0-EC $\mathcal{E}$ is reached, a scheduler can choose to stay within the states of $\mathcal{E}$ for arbitrary many steps by only picking actions α at states $s \in \mathit{states}(\mathcal{E})$ with $(s, \alpha) \in \mathcal{E}$. Since ECs are strongly connected, it is possible to reach any state $s \in \mathit{states}(\mathcal{E})$ almost surely. At some point, the scheduler may decide to leave the EC by choosing an exiting state-action pair $(s, \alpha) \in \mathit{exits}(\mathcal{E})$. Although this decision *could* be delayed for arbitrary many steps, such a delay has no effect on the induced expected total reward since no reward is accumulated in 0-ECs.

Our approach is to replace each maximal 0-EC $\mathcal{E}$ of $\mathcal{D}^{\mathbf{w}}$ by a state $s_?^{\mathcal{E}}$ in which a scheduler immediately has to choose either an exiting state-action pair $(s, \alpha) \in \mathit{exits}(\mathcal{E})$ or an action $\alpha_!$, indicating that the EC should never be left. This procedure coincides with the computation of EC quotients in, e.g., [3,4,21,32].

Definition 20 (0-EC Quotient) The 0-EC quotient of $\mathcal{D}^{\mathbf{w}}$ w.r.t. reward function $\rho_{\mathbf{w}}$ is the MDP $\mathcal{D}^{\mathbf{w}}_{\setminus\text{0-EC}} = (S', \mathit{Act}', \mathbf{P}', s_0', (\rho'_{\mathbf{w}}))$ obtained by applying the following steps on $\mathcal{D}^{\mathbf{w}}$ for every maximal 0-EC $\mathcal{E}$:

1. Add two new states $s_?^{\mathcal{E}}$ and $s_!^{\mathcal{E}}$ and remove all states from $\mathit{states}(\mathcal{E})$.
2. Add a new action α_s for each $(s, \alpha) \in \mathit{exits}(\mathcal{E})$.
3. Redirect the target of transitions that *enter* the states in $\mathit{states}(\mathcal{E})$ to $s_?^{\mathcal{E}}$, i.e., $\mathbf{P}'(s, \alpha, s_?^{\mathcal{E}}) = \sum_{s' \in \mathit{states}(\mathcal{E})} \mathbf{P}(s, \alpha, s')$ for every $(s, \alpha) \in (S \times \mathit{Act}) \setminus \mathcal{E}$.
4. Redirect the origin of transitions that *exit* the states in $\mathit{states}(\mathcal{E})$ to $s_?^{\mathcal{E}}$, i.e., $\mathbf{P}'(s_?^{\mathcal{E}}, \alpha_s, s') = \mathbf{P}(s, \alpha, s')$ for every $(s, \alpha) \in \mathit{exits}(\mathcal{E})$.
5. Add a new action $\alpha_!$ and set $\mathbf{P}(s, \alpha_!, s_!^{\mathcal{E}}) = 1$ for $s \in \{s_?^{\mathcal{E}}, s_!^{\mathcal{E}}\}$.
6. Restrict the reward function $\rho_{\mathbf{w}}$ to the remaining state-action pairs and set $\rho'_{\mathbf{w}}(s_?^{\mathcal{E}}, \alpha_!) = \rho'_{\mathbf{w}}(s_!^{\mathcal{E}}, \alpha_!) = 0$ and $\rho'_{\mathbf{w}}(s_?^{\mathcal{E}}, \alpha_s) = \rho_{\mathbf{w}}(s, \alpha)$ for all $(s, \alpha) \in \mathit{exits}(\mathcal{E})$.

Example 17 Fig. 8b depicts the 0-EC quotient of the MDP in Fig. 8a w.r.t. the depicted reward function $\rho_{\mathbf{w}}$.

Considering the 0-EC quotient $\mathcal{D}^{\mathbf{w}}_{\backslash 0\text{-EC}} = (S', Act', \mathbf{P}', s_0', (\rho'_{\mathbf{w}}))$ instead of $\mathcal{D}^{\mathbf{w}}$ preserves maximal expected total rewards, i.e.,

$$\max_{\sigma \in \mathrm{TA}^{\mathcal{D}^{\mathbf{w}}}} \mathrm{ER}^{\mathcal{D}^{\mathbf{w}}}_{\sigma}(\rho_{\mathbf{w}}, \emptyset) = \max_{\sigma' \in \mathrm{TA}^{\mathcal{D}^{\mathbf{w}}_{\backslash 0\text{-EC}}}} \mathrm{ER}^{\mathcal{D}^{\mathbf{w}}_{\backslash 0\text{-EC}}}_{\sigma'}(\rho'_{\mathbf{w}}, \emptyset).$$

Moreover, we can transform any scheduler σ' for $\mathcal{D}^{\mathbf{w}}_{\backslash 0\text{-EC}}$ to a scheduler σ for $\mathcal{D}^{\mathbf{w}}$ with the same expected reward, i.e., $\mathrm{ER}^{\mathcal{D}^{\mathbf{w}}}_{\sigma}(\rho_{\mathbf{w}}, \emptyset) = \mathrm{ER}^{\mathcal{D}^{\mathbf{w}}_{\backslash 0\text{-EC}}}_{\sigma'}(\rho'_{\mathbf{w}}, \emptyset)$. In particular, if σ' chooses an action α_s at a state $s^{\mathcal{E}}_{?}$, σ can mimic this by choosing α at state $s \in \mathit{states}(\mathcal{E})$ and enforcing that s is reached almost surely from any other state $s' \in \mathit{states}(\mathcal{E}) \backslash \{s\}$ of the 0-EC $\mathcal{E}$. Similarly, if σ' chooses $\alpha_!$ at $s^{\mathcal{E}}_{?}$, σ can mimic this by only picking actions α at states $s \in \mathit{states}(\mathcal{E})$ with $(s, \alpha) \in \mathcal{E}$. We refer to [4,21] for more details on the correctness of this construction.

Since there are no more ECs in $\mathcal{D}^{\mathbf{w}}_{\backslash 0\text{-EC}}$ (other than the 0-ECs of the form $\{(s^{\mathcal{E}}_{!}, \alpha_!)\}$ which can not be left), it can be shown that the value iteration algorithm on $\mathcal{D}^{\mathbf{w}}_{\backslash 0\text{-EC}}$ approaches the correct expected total rewards [4,6].

5.1.3 Negative infinite rewards

We now lift our strengthened reward finiteness assumptions from the previous section (but still impose the assumptions mentioned on page 29). More precisely, we now allow that $\mathrm{ER}^{\overline{\mathcal{D}}}_{\sigma_\infty}(\overline{\rho}_i, \emptyset) = -\infty$ for some $i \in \{1, \dots, d\}$ with $\triangleright_i \in \{\leq, <\}$ and some scheduler σ_∞. Observe that such a scheduler does not achieve any point $\mathbf{p} \in \mathbb{R}^d$ and can thus be excluded from the analysis. Thus, for the computation of $\{\mathbf{p} \in \mathbb{R}^d \mid \mathit{achieve}^{\overline{\mathcal{D}}}(\mathbb{O} \triangleright \mathbf{p})\}$ it suffices to restrict the analysis to schedulers in $\{\sigma \in \mathrm{TA} \mid \forall i \in \{1, \dots, d\} \colon \mathrm{ER}^{\overline{\mathcal{D}}}_{\sigma}(\overline{\rho}_i, \emptyset) \neq -\infty\}$. Our assumptions for reward finiteness imply that this set is not empty.

If the considered weight vector $\mathbf{w} = (w_1, \dots, w_d)$ satisfies $w_i > 0$ for all $i \in \{0, \dots, d\}$, our approach from the previous section can be applied without further modifications: If $\mathrm{ER}^{\overline{\mathcal{D}}}_{\sigma_\infty}(\overline{\rho}_i, \emptyset) = -\infty$ for some $i \in \{1, \dots, d\}$ and some scheduler σ_∞, we also have $\mathrm{ER}^{\mathcal{D}^{\mathbf{w}}}_{\sigma_\infty}(\rho_{\mathbf{w}}, \emptyset) = -\infty < \max_{\sigma \in \mathrm{TA}} \mathrm{ER}^{\mathcal{D}^{\mathbf{w}}}_{\sigma}(\rho_{\mathbf{w}}, \emptyset)$.

However, if $w_i = 0$ for some $i \in \{0, \dots, d\}$, there might be a 0-EC w.r.t. $\rho_{\mathbf{w}}$ that is not a 0-EC w.r.t. $\overline{\rho}_i$. This is because the rewards of ρ_i are not considered in the weighted reward function $\rho_{\mathbf{w}}$. As a consequence, a scheduler for $\mathcal{D}^{\mathbf{w}}_{\backslash 0\text{-EC}}$ as defined above might choose action $\alpha_!$ at a state $s^{\mathcal{E}}_{?}$ for an EC $\mathcal{E}$ that is not a 0-EC w.r.t. $\overline{\rho}_i$. Transforming such a scheduler back to $\overline{\mathcal{D}}$ can then induce an expected total reward of $-\infty$ w.r.t. $\overline{\rho}_i$. We exclude such schedulers during the computations by tweaking the elimination of 0-ECs in Def. 20: The action $\alpha_!$ is only inserted (Step 5) if there is a scheduler for $\overline{\mathcal{D}}$ that stays in the given EC forever and that yields finite expected reward for all individual objectives.

5.2 ds-bounded reachability

We generalize reachability objectives from Def. 6 to ds-bounded reachability objectives of the form $\mathbb{P}(\lozenge^{J}_{\mathrm{ds}} G)$ which concern ds-bounded reachability probabilities (cf. Def. 15). As explained in Sect. 4.3, multi-objective queries for MA considering timed reachability objectives can be approximated by checking ds-bounded reachability objectives on a digitization of the MA. This section extends the approach of [28] to process such objectives.

Assume the digitization $\mathcal{M}_\delta = (S, Act, \mathbf{P}_\delta, s_0, (\rho_1^\delta, \ldots, \rho_\ell^\delta))$ of an MA $\mathcal{M}$ w.r.t. $\delta > 0$ and ds-bounded reachability objectives $\mathbb{P}(\Diamond_{\text{ds}}^{\leq k_1} G_1), \ldots, \mathbb{P}(\Diamond_{\text{ds}}^{\leq k_d} G_d)$. Other digitization step bounds (e.g. lower step bounds) are treated similarly. We further denote by $\text{MS} \subseteq S$ the set of states of $\mathcal{M}_\delta$ that represent the Markovian states of the initially considered MA. We present a transformation from ds-bounded reachability objectives to untimed reachability objectives. The idea is to store the number of visited Markovian states in the state space. Upon reaching a state in G_i, this information allows us to distinguish the cases where the step bound k_i has been exceeded or not. The procedure is based on the idea of [35, Algorithm 1] for single-objective MA.

For the largest occurring step bound $k_{\max} = \max_i k_i$ let $\mathcal{D}$ be the MDP consisting of $k_{\max} + 2$ copies of $\mathcal{M}_\delta$ such that transitions to states $s' \in \text{MS}$ are redirected to the next copy. Formally, $\mathcal{D} = (S \times \{0, \ldots, k_{\max} + 1\}, Act, \mathbf{P}, (s_0, k_{\text{init}}), \{\rho_1, \ldots, \rho_\ell\})$ where

- $k_{\text{init}} = 1$ if $s_0 \in \text{MS}$, and $k_{\text{init}} = 0$ otherwise,
- $\mathbf{P}((s, k),\ \alpha,\ (s', k')) = \mathbf{P}_\delta(s, \alpha, s')$ if $s' \in \text{MS}$ and $k' = \min(k{+}1, k_{\max}{+}1)$,
- $\mathbf{P}((s, k),\ \alpha,\ (s', k)) = \mathbf{P}_\delta(s, \alpha, s')$ if $s' \notin \text{MS}$,
- $\mathbf{P}((s, k),\ \alpha,\ (s', k')) = 0$ in all other cases, and
- every reward function ρ_i satisfies $\rho_i((s, k),\ \alpha) = \rho_i^\delta(s, \alpha)$.

The second component of a state (s, k) of $\mathcal{D}$ reflects the number of Markovian states visited so far. Hence, eventually reaching G_i in $\mathcal{M}_\delta$ with at most k_i digitization steps is equivalent to reaching $G_i' = \{(s, k) \mid s \in G_i \text{ and } k \leq k_i\}$ in $\mathcal{D}$. In particular, there is a mapping from schedulers σ_δ for $\mathcal{M}_\delta$ to schedulers σ for $\mathcal{D}$ such that $\Pr_{\sigma_\delta}^{\mathcal{M}_\delta}(\Diamond_{\text{ds}}^{\leq k_i} G_i) = \Pr_\sigma^{\mathcal{D}}(\Diamond G_i')$.

Example 18 We illustrate the construction for the digitization $\mathcal{M}_\delta$ depicted in Fig. 9a and the objective $\mathbb{P}(\Diamond_{\text{ds}}^{\leq 2}\{s_2\})$. States that correspond to Markovian states of the original MA are depicted with rectangles.

Figure 9b shows the reachable states of the MDP $\mathcal{D}$ as given by the construction above.

With the construction above, a digitization $\mathcal{M}_\delta$ with objectives $\mathbb{O}$ can be transformed to an MDP $\mathcal{D}$ and untimed reachability or expected reward objectives $\mathbb{O}'$ such that

$$achieve^{\mathcal{M}_\delta}(\mathbb{O}) = achieve^{\mathcal{D}}(\mathbb{O}').$$

The latter can be computed using the approach of [28]. In practice, however, this approach is not feasible as $k_{\max}$ may take high values, leading to huge model sizes. To avoid the problematic, the different copies of $\mathcal{M}_\delta$ are analyzed individually. For $k \in \{0, \ldots, k_{\max}{+}1\}$ let S_k denote the states (s, k) of $\mathcal{D}$ with second entry k. A transition emerging from a state in S_k can only point to a state in $S_k \cup S_{k+1}$. As a consequence, for the computation of optimal values for the states in S_k, only the results of the states in S_{k+1} are relevant. We thus analyze the sub-models of $\mathcal{D}$ induced by the state sets $S_{k_{\max}+1}, S_{k_{\max}}, \ldots, S_1$ and S_0 sequentially in the given order. Details on the sequential approach can be found in [35] for single-objective MA and in [33] for multi-objective MDPs.

Dealing with digitization errors Recall that the analysis of the digitization $\mathcal{M}_\delta$ only approximates the timed reachability probabilities of the MA. More precisely, given a weight vector $\mathbf{w}$, the approach outlined above computes an optimal point $\mathbf{p}$ for $\mathbf{w}$ on $\mathcal{M}_\delta$ and the *digitized* query. Applying our results from Sect. 4.3 (in particular Theorem 5), we can lift the point $\mathbf{p}$ to the original time-bounded reachability query by considering the hyperrectangle $\varepsilon(\mathbb{O}, \mathbf{p})$. Fig. 10a illustrates this, where $\varepsilon(\mathbb{O}, \mathbf{p})$ is depicted by the blue rectangle. The point $\mathbf{q}$ at the

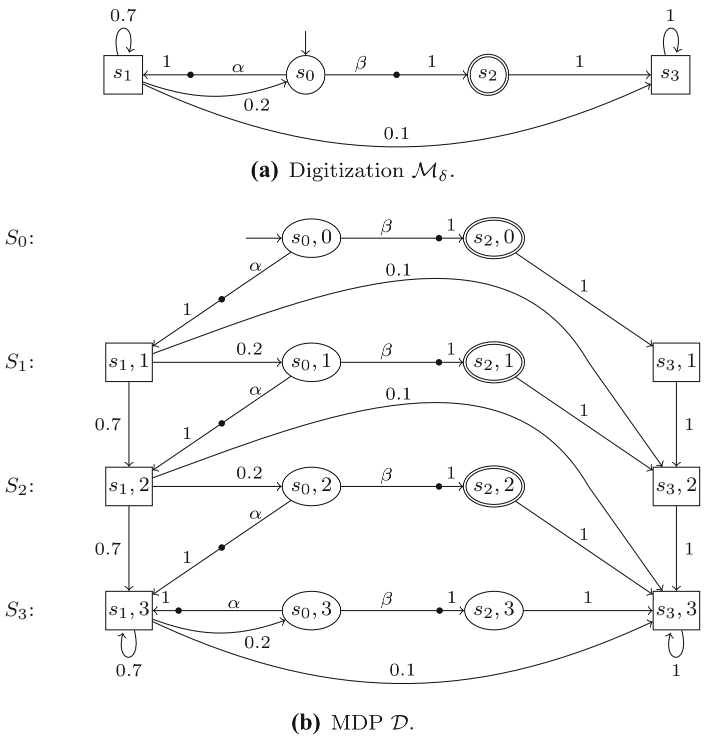


Fig. 9 Transformation for ds-bounded reachability objectives (cf. Example 18)

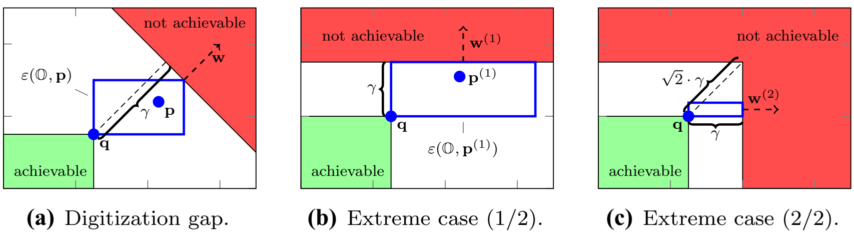


Fig. 10 Illustration of digitization errors

bottom left of the rectangle is known to be achievable, implying that all points in the green area are achievable, whereas the points in the red area are known to be unachievable. This introduces a gap between the achievable and unachievable area. We measure the size γ of this gap by considering the smallest distance between the achievable and the unachievable points, i.e., the distance between the point $\mathbf{q}$ and the red area in Fig. 10a. The value of γ depends on the size of the rectangle $\varepsilon(\mathbb{O}, \mathbf{p})$ and thus on the selection of the digitization constant δ. However, γ also depends on the considered weight vector $\mathbf{w}$ which motivates a dynamic selection of the δ, depending on the currently considered weight vector.

Example 19 Assume that an optimal point $\mathbf{p}$ has been computed for the weight vector $\mathbf{w}^{(1)} = (0, 1)$, resulting in the (un-)achievable points and gap γ depicted in Fig. 10b. Note that γ is not affected by the comparably large approximation error for the first objective (x-axis).

We continue the example in Fig. 10c. To achieve the same gap γ also for the weight vector $\mathbf{w}^{(2)} = (1, 0)$, a smaller digitization constant has been chosen so that the rectangle $\varepsilon(\mathbb{O}, \mathbf{p}^{(2)})$ becomes smaller. We assume that—by coincidence—the analysis for $\mathbf{w}^{(2)} = (1, 0)$ produces the same point $\mathbf{q}$ as in the previous step from Fig. 10b. For example, this is possible when the nondeterminism of the model is spurious. Observe that the largest distance between a point in the unknown (white) area and the green area is $\sqrt{\gamma^2 + \gamma^2} = \sqrt{2} \cdot \gamma$. Put differently, the obtained approximation in Fig. 10c is a $(\sqrt{2} \cdot \gamma)$-approximation[4] of the set of achievable points for the considered MA.

In general, we can show the following. If all gaps γ are below $\eta/\sqrt{d}$ for some $\eta > 0$ and the number of objectives d, the approach of [28] converges to an η-approximation of the set of achievable points.

In order to obtain an η-approximation, our implementation that we discuss in Sect. 6 therefore implements the following heuristic for the selection of the digitization constants. Given a weight vector $\mathbf{w}$, we choose the largest possible digitization constant δ for which the resulting gap γ w.r.t. $\mathbf{w}$ is at most $(\eta/\sqrt{d}) \cdot 0.9$. Here, the factor 0.9 ensures that we do not have to explore the full set of achievable points of $\mathcal{M}_\delta$, which otherwise could result in analyzing an unnecessarily large amount of weight vectors.

6 Experimental evaluation

Implementation We implemented multi-objective model checking of MAs into `Storm` [22, 36]. The input model is given in the `PRISM` language[5] and translated into a sparse representation. For MA $\mathcal{M}$, the implementation performs a multi-objective analysis on the underlying MDP $\mathcal{M}_\mathcal{D}$ or a digitization $\mathcal{M}_\delta$ and infers (an approximation of) the achievable points of $\mathcal{M}$ by exploiting the results from Sect. 4. For computing the achievable points of $\mathcal{M}_\mathcal{D}$ and $\mathcal{M}_\delta$, we apply the approach of [28] with the extensions explained in Sect. 5.

All material to replicate the experiments is available at [45]. The implementation is part of `Storm` since release 1.6.3 available at http://stormchecker.org.

Setup Our implementation uses a single core of an Intel Xeon Platinum 8160 CPU with memory limited to 20GB RAM. The timeout (TO) is two hours. For a model, a set of objectives, and a precision $\eta \in \mathbb{R}_{>0}$, we measure the time to compute an η-approximation of the set of achievable points. This set-up coincides with Pareto queries as considered in [28]. The digitization constant δ is chosen heuristically as discussed in Sect. 5.2. We set the precision for value-iteration to $\varepsilon = 10^{-6}$. Similar to [28], we use the conventional, unsound variant of value iteration. The use of improved algorithms providing sound precision guarantees [32] is left for future work.

Results for MAs We consider four case studies: (i) a *job scheduler* [12], see Sect. 1; (ii) a *polling system* [48,50] containing a server processing jobs that arrive at two stations; (iii) a *video streaming client* buffering received packages and deciding when to start playback; and (iv) a randomized *mutual exclusion algorithm* [50], a variant of [42] with a process-dependent

[4] An η-approximation of $A \subseteq \mathbb{R}^d$ is given by $A^-, A^+ \subseteq \mathbb{R}^d$ with $A^- \subseteq A \subseteq A^+$ and for all $\mathbf{p} \in A^+$ exists a $\mathbf{q} \in A^-$ such that the distance between $\mathbf{p}$ and $\mathbf{q}$ is at most η.

[5] We slightly extend the `PRISM` language in order to describe MAs.

Table 1 Experimental results for multi-objective MAs

Benchmark			$(\Diamond, \text{ER}, \Diamond^I)$		$(\Diamond, \text{ER}, \Diamond^I)$		$(\Diamond, \text{ER}, \Diamond^I)$		$(\Diamond, \text{ER}, \Diamond^I)$	
N (-K)	# states	$\log_{10}(\eta)$	Pts	Time	Pts	Time	Pts	Time	Pts	Time
Job scheduling			(0, 3, 0)		(0, 1, 1)		(1, 3, 0)		(1, 1, 2)	
10-2	12554	− 2	10	2.19	9	13.4	17	120	18	695
		− 3	48	22.1	21	277	TO		TO	
12-3	116814	− 2	11	7.3	10	273	23	855	TO	
		− 3	58	65.7	23	5955	TO		TO	
17-2	4587537	− 2	14	480	TO		22	1238	TO	
		− 3	62	995	TO		TO		TO	
Polling			(0, 2, 0)		(0, 4, 0)		(0, 0, 2)		(0, 2, 2)	
3-2	990	− 2	2	0.18	4	0.25	3	20.4	10	97.4
		− 3	2	0.13	4	0.47	7	449	32	6906
3-3	9522	− 2	4	0.36	8	4.63	6	359	33	6040
		− 3	6	0.49	20	405	TO		TO	
4-4	813287	− 2	8	158	20	936	TO		TO	
		− 3	10	287	TO		TO		TO	
Stream			(0, 2, 0)		(0, 1, 1)		(0, 0, 2)		(0, 2, 1)	
30	1426	− 2	20	0.23	16	8.47	16	6.49	28	81
		− 3	51	2.01	44	241	38	160	99	3269
250	94376	− 2	31	7.63	16	542	16	409	23	5291
		− 3	90	19.3	TO		TO		TO	
1000	1502501	− 2	41	288	TO		16	7054	TO	
		− 3	118	488	TO		TO		TO	
Mutex			(0, 0, 3)		(0, 0, 3)					
2	10560	− 2	18	221	16	869				
		− 3	16	2232	TO					
3	31733	− 2	15	3227	TO					
		− 3	TO		TO					

random delay in the critical section. Details on the benchmarks and the objectives are given in App. E.1.

Tab. 1 lists results. For each instance we give the defining constants, the number of states of the MA and the used η-approximation. A multi-objective query is given by the triple (l, m, n) indicating l untimed, m expected reward, and n timed objectives. For each MA and query we depict the total run-time of our implementation (time) and the number of vertices of the obtained under-approximation (*pts*).

Queries analyzed on the underlying MDP are solved efficiently on large models with up to millions of states. For timed objectives the run-times increase drastically due to the costly analysis of digitized reachability objectives on the digitization, cf. [30]. Queries with up to four objectives can be dealt with within the time limit. Furthermore, for an approximation one order of magnitude better, the number of vertices of the result increases approximately by a factor three. In addition, a lower digitization constant has then to be considered which often leads to timeouts in experiments with timed objectives.

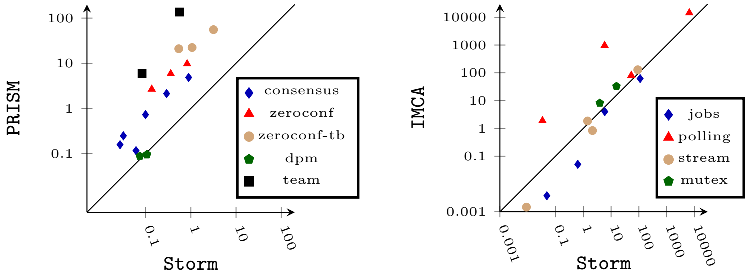


Fig. 11 Verification times (in seconds) of our implementation and other tools

Comparison with PRISM [39] *and* IMCA [30]. We compared the performance of our implementation with both PRISM[6] and IMCA[7].

For the comparison with PRISM (no MAs), we considered the multi-objective MDP benchmarks from [28,29]. We conducted our experiments on PRISM with both variants of the value iteration-based implementation (standard and Gauss-Seidel) and chose the faster variant for each benchmark instance. For all experiments the approximation precision $\eta = 0.001$ was considered.

For the comparison with IMCA (no multi-objective queries) we used the benchmarks from Tab. 1, with just a single objective. The experiments on IMCA have been conducted with and without enabling value-iteration and we chose the faster variant for each benchmark instance. For timed reachability objectives, the precision $\eta = 0.001$ was considered in all experiments.

Verification times are summarized in Fig. 11: On points above the diagonal, our implementation is faster. Both implementations are based on [28]. We observe that our implementation is competitive with these tools. Further details are given in App. E.2 and App. E.3.

7 Conclusion

We considered multi-objective verification of Markov automata, including in particular timed reachability objectives. The next step is to apply our algorithms to the manifold applications of MA, such as generalized stochastic Petri nets to enrich the analysis possibilities of such nets, and to investigate whether recent advances in the single-objective timed reachability analysis, in particular the methods of [14,15] may be lifted to the multi-objective case.

A Proofs about sets of achievable points

Let $\mathcal{M} = (S, Act, \rightarrow, s_0, (\rho_1, \ldots, \rho_\ell))$ be an MA and $\sigma_1, \sigma_2 \in \text{GM}$ be two schedulers for $\mathcal{M}$. Further let $w_1 \in [0, 1]$ and $w_2 = 1 - w_1 \in [0, 1]$. The proof of Proposition 1 considers the scheduler $\sigma^w \in \text{GM}$, where for a path $\pi = s_0 \xrightarrow{\kappa_0} \ldots \xrightarrow{\kappa_{n-1}} s_n \in \mathit{FPaths}$ and action

[6] We considered PRISM 4.6 obtained from www.prismmodelchecker.org.

[7] We considered IMCA 1.6 obtained from https://github.com/buschko/imca.

$\alpha \in Act$ we have

$$\sigma^w(\pi,\alpha) = \frac{\sum_{i=1}^{2}\left(w_i \cdot \sigma_i(\pi,\alpha)\cdot \prod_{j=0}^{n-1}\sigma_i(\mathit{pref}(\pi,j),\alpha(\kappa_j))\right)}{\sum_{i=1}^{2}\left(w_i \cdot \prod_{j=0}^{n-1}\sigma_i(\mathit{pref}(\pi,j),\alpha(\kappa_j))\right)}$$

We now show the following lemma.

Lemma 8 *For $\mathcal{M}$, $\sigma_1, \sigma_2, w_1, w_2$, and σ as above and arbitrary measurable set $\Pi \subseteq IPaths$ we have*

$$Pr_{\sigma^w}^{\mathcal{M}}(\Pi) = w_1 \cdot Pr_{\sigma_1}^{\mathcal{M}}(\Pi) + w_2 \cdot Pr_{\sigma_2}^{\mathcal{M}}(\Pi).$$

To show the Lemma, we fix a time-abstract path $\hat{\pi} = s_0 \xrightarrow{\alpha_0} \dots \xrightarrow{\alpha_{n-1}} s_n$ of $\mathcal{M}$ and show that the claim holds for the cylinder set $Cyl(\Pi)$ of some measurable $\Pi \subseteq \{\pi \in FPaths \mid \mathrm{ta}(\pi) = \hat{\pi}\}$. The lemma also follows for arbitrary measurable sets as these can be described via unions and complements of such cylinder sets.

We define the scheduler $\sigma_{\hat{\pi}}$ where for $\pi \in FPaths$ and $\alpha \in Act$ we set

$$\sigma_{\hat{\pi}}(\pi,\alpha) = \begin{cases} 1 & \text{if } \exists\, j < n\colon \mathrm{ta}(\pi) = \mathit{pref}(\hat{\pi}, j) \text{ and } \alpha = \alpha_j \\ 0 & \text{if } \exists\, j < n\colon \mathrm{ta}(\pi) = \mathit{pref}(\hat{\pi}, j) \text{ and } \alpha \neq \alpha_j \\ \frac{1}{|Act(last(\pi))|} & \text{otherwise.} \end{cases}$$

On a path whose time abstraction is a proper prefix of $\hat{\pi}$, $\sigma_{\hat{\pi}}$ will choose exactly the action given in $\hat{\pi}$. In other cases, the choice is arbitrary (for simplicity, we picked a uniform distribution over available actions). We first show two auxiliary lemmas.

Lemma 9 *For a scheduler $\sigma \in GM$ and $\hat{\pi}$, $\Pi \subseteq \{\pi \in FPaths \mid ta(\pi) = \hat{\pi}\}$, and $\sigma_{\hat{\pi}}$ as above we have*

$$Pr_{\sigma}^{\mathcal{M}}(\Pi) = \int_{\pi\in\Pi} \left(\prod_{j=0}^{n-1} \sigma(\mathit{pref}(\pi,j),\alpha_j)\right) dPr_{\sigma_{\hat{\pi}}}^{\mathcal{M}}(\pi).$$

Proof The proof is by induction over the length n of $\hat{\pi} = s_0 \xrightarrow{\alpha_0} \dots \xrightarrow{\alpha_{n-1}} s_n$. If $n = 0$ we have either $\Pi = \{s_0\}$ or $\Pi = \emptyset$ and thus either $Cyl(\Pi) = IPaths$ or $Cyl(\Pi) = \emptyset$. The lemma follows immediately in both cases. Now assume that Lemma 9 holds for $\Pi' = \{\mathit{pref}(\pi, n-1) \mid \pi \in \Pi\}$, i.e., for paths of length $n-1$. Notice that for $\pi' \in \Pi'$ we have $last(\pi') = s_{n-1}$.

<u>*Case* $s_{n-1} \in \mathrm{PS}$:</u>

$$\begin{aligned}
\mathrm{Pr}_{\sigma}^{\mathcal{M}}(\Pi) &= \int_{\pi'\in\Pi'} \sigma(\pi',\alpha_{n-1}) \cdot \mathbf{P}(s,\alpha_{n-1},s')\, \mathrm{dPr}_{\sigma}^{\mathcal{M}}(\pi') \\
&= \int_{\pi'\in\Pi'} \sigma(\pi',\alpha_{n-1}) \cdot \mathbf{P}(s,\alpha_{n-1},s') \cdot \left(\prod_{j=0}^{n-2} \sigma(\mathit{pref}(\pi',j),\alpha_j)\right) \mathrm{dPr}_{\sigma_{\hat{\pi}}}^{\mathcal{M}}(\pi') \\
&= \int_{\pi'\in\Pi'} \left(\prod_{j=0}^{n-1} \sigma(\mathit{pref}(\pi',j),\alpha_j)\right) \cdot \mathbf{P}(s,\alpha_{n-1},s')\, \mathrm{dPr}_{\sigma_{\hat{\pi}}}^{\mathcal{M}}(\pi') \\
&= \int_{\pi'\in\Pi'} \left(\prod_{j=0}^{n-1} \sigma(\mathit{pref}(\pi',j),\alpha_j)\right) \cdot \underbrace{\sigma_{\hat{\pi}}(\pi',\alpha_{n-1})}_{=1} \cdot \mathbf{P}(s,\alpha_{n-1},s')\, \mathrm{dPr}_{\sigma_{\hat{\pi}}}^{\mathcal{M}}(\pi')
\end{aligned}$$

$$= \int_{\pi \in \Pi} \Big(\prod_{j=0}^{n-1} \sigma(\mathit{pref}(\pi, j), \alpha_j) \Big) \,\mathrm{dPr}^{\mathcal{M}}_{\sigma_{\hat{\pi}}}(\pi).$$

In the last step we use that for $\pi' \in \Pi'$ we have $\pi = \pi' \xrightarrow{\alpha_{n-1}} s_n \in \Pi$.

<u>$Case\, s_{n-1} \in \mathrm{MS}$:</u> For $\pi' \in \Pi'$ let $T_{\pi'} = \{(t, \alpha_{n-1}, s_n) \mid \pi' \xrightarrow{t} s_n \in \Pi\}$. We note that $\alpha_{n-1} = \bot$, $\sigma'(\pi', \bot) = 1$, and that the probability of the transition step $\mathrm{Pr}^{\mathit{Steps}}_{\sigma,\pi}(T_{\pi'})$ does not depend on σ since $s_{n-1} \in \mathrm{MS}$. We get:

$$\begin{aligned}
\mathrm{Pr}^{\mathcal{M}}_{\sigma}(\Pi) &= \int_{\pi' \in \Pi'} \mathrm{Pr}^{\mathit{Steps}}_{\sigma,\pi'}(T_{\pi'}) \,\mathrm{dPr}^{\mathcal{M}}_{\sigma}(\pi') \\
&= \int_{\pi' \in \Pi'} \mathrm{Pr}^{\mathit{Steps}}_{\sigma,\pi'}(T_{\pi'}) \cdot \Big(\prod_{j=0}^{n-2} \sigma(\mathit{pref}(\pi', j), \alpha_j) \Big) \,\mathrm{dPr}^{\mathcal{M}}_{\sigma_{\hat{\pi}}}(\pi') \\
&= \int_{\pi' \in \Pi'} \Big(\prod_{j=0}^{n-1} \sigma(\mathit{pref}(\pi', j), \alpha_j) \Big) \cdot \mathrm{Pr}^{\mathit{Steps}}_{\sigma_{\hat{\pi}},\pi}(T_{\pi'}) \,\mathrm{dPr}^{\mathcal{M}}_{\sigma_{\hat{\pi}}}(\pi') \\
&= \int_{\pi \in \Pi} \Big(\prod_{j=0}^{n-1} \sigma(\mathit{pref}(\pi, j), \alpha_j) \Big) \,\mathrm{dPr}^{\mathcal{M}}_{\sigma_{\hat{\pi}}}(\pi).
\end{aligned}$$

□

Lemma 10 *For $\sigma_1, \sigma_2, \sigma^w$ as above and $\pi = s_0 \xrightarrow{\kappa_0} \ldots \xrightarrow{\kappa_{n-1}} s_n \in \mathit{FPaths}$ we have*

$$\prod_{j=0}^{n-1} \sigma^w(\mathit{pref}(\pi, j), \alpha(\kappa_j)) = \sum_{i=1}^{2} \left(w_i \cdot \prod_{j=0}^{n-1} \sigma_i(\mathit{pref}(\pi, j), \alpha(\kappa_j)) \right).$$

Proof

$$\begin{aligned}
\prod_{j=0}^{n-1} \sigma^w(\mathit{pref}(\pi, j), \alpha(\kappa_j)) &= \prod_{j=0}^{n-1} \frac{\sum_{i=1}^{2} \Big(w_i \cdot \prod_{k=0}^{j} \sigma_i(\mathit{pref}(\pi, k), \alpha(\kappa_k)) \Big)}{\sum_{i=1}^{2} \Big(w_i \cdot \prod_{k=0}^{j-1} \sigma_i(\mathit{pref}(\pi, j), \alpha(\kappa_j)) \Big)} \\
&= \frac{\sum_{i=1}^{2} \Big(w_i \cdot \prod_{k=0}^{n-1} \sigma_i(\mathit{pref}(\pi, k), \alpha(\kappa_k)) \Big)}{\sum_{i=1}^{2} \Big(w_i \cdot \prod_{k=0}^{0-1} \sigma_i(\mathit{pref}(\pi, j), \alpha(\kappa_j)) \Big)} \\
&= \sum_{i=1}^{2} \left(w_i \cdot \prod_{j=0}^{n-1} \sigma_i(\mathit{pref}(\pi, j), \alpha(\kappa_j)) \right)
\end{aligned}$$

Proof (*of Lemma* 8) Using the auxiliary Lemmas 9 and 10, we can prove Lemma 8 as follows:

$$\begin{aligned}
\mathrm{Pr}^{\mathcal{M}}_{\sigma^w}(\Pi) &\overset{\text{Lem. 9}}{=} \int_{\pi \in \Pi} \Big(\prod_{j=0}^{n-1} \sigma^w(\mathit{pref}(\pi, j), \alpha_j) \Big) \,\mathrm{dPr}^{\mathcal{M}}_{\sigma^w_{\hat{\pi}}}(\pi) \\
&\overset{\text{Lem. 10}}{=} \int_{\pi \in \Pi} \sum_{i=1}^{2} \left(w_i \cdot \prod_{j=0}^{n-1} \sigma_i(\mathit{pref}(\pi, j), \alpha_j) \right) \mathrm{dPr}^{\mathcal{M}}_{\sigma^w_{\hat{\pi}}}(\pi) \\
&= \sum_{i=1}^{2} \left(w_i \cdot \int_{\pi \in \Pi} \cdot \prod_{j=0}^{n-1} \sigma_i(\mathit{pref}(\pi, j), \alpha_j) \,\mathrm{dPr}^{\mathcal{M}}_{\sigma^w_{\hat{\pi}}}(\pi) \right)
\end{aligned}$$

$$\overset{\text{Lem. 9}}{=} \sum_{i=1}^{2} \left(w_i \cdot \Pr^{\mathcal{M}}_{\sigma_i}(\Pi) \right).$$

□

B Proofs for expected reward

B.1 Proof of Lemma 2

Lemma 2 *For MA $\mathcal{M} = (S, Act, \rightarrow, s_0, (\rho_1, \ldots, \rho_\ell))$ with $G \subseteq S$, $\sigma \in GM$, and reward function ρ it holds that*

$$\lim_{n\to\infty} \mathrm{ER}^{\mathcal{M}}_{\sigma}(\rho, \Pi^n_G) = \mathrm{ER}^{\mathcal{M}}_{\sigma}(\rho, G).$$

Furthermore, any reward function $\rho^{\mathcal{D}}$ for $\mathcal{M}_{\mathcal{D}}$ satisfies

$$\lim_{n\to\infty} \mathrm{ER}^{\mathcal{M}_{\mathcal{D}}}_{ta(\sigma)}(\rho^{\mathcal{D}}, \Pi^n_G) = \mathrm{ER}^{\mathcal{M}_{\mathcal{D}}}_{ta(\sigma)}(\rho^{\mathcal{D}}, G).$$

Proof We show the first claim. The second claim follows analogously. For each $n \geq 0$, consider the function $f_n \colon IPaths^{\mathcal{M}} \rightarrow \mathbb{R}_{\geq 0}$ given by

$$f_n(\pi) = \begin{cases} rew^{\mathcal{M}}(\rho, pref(\pi, m)) & \text{if } m = \min \left\{ i \in \{0, \ldots, n\} \mid s_i \in G \right\} \\ rew^{\mathcal{M}}(\rho, pref(\pi, n)) & \text{if } s_i \notin G \text{ for all } i \leq n \end{cases}$$

for every path $\pi = s_0 \xrightarrow{\kappa_0} s_1 \xrightarrow{\kappa_1} \cdots \in IPaths^{\mathcal{M}}$. Intuitively, $f_n(\pi)$ is the reward collected on π within the first n steps and only up to the first visit of G. This allows us to express the expected reward collected along the paths of Π^n_G as

$$\mathrm{ER}^{\mathcal{M}}_{\sigma}(\Pi^n_G) = \sum_{\hat{\pi} \in \Pi^n_G} \int_{\pi \in \langle \hat{\pi} \rangle} rew^{\mathcal{M}}(\rho, \pi)\, \mathrm{d}\Pr^{\mathcal{M}}_{\sigma}(\pi) = \int_{\pi \in IPaths^{\mathcal{M}}} f_n(\pi)\, \mathrm{d}\Pr^{\mathcal{M}}_{\sigma}(\pi).$$

It holds that $\lim_{n\to\infty} f_n(\pi) = rew^{\mathcal{M}}(\rho, \pi, G)$ which is a direct consequence from the definition of the reward of π up to G (cf. Sect. 2.2.3). Furthermore, note that the sequence of functions $f_0, f_1, \ldots$ is non-decreasing, i.e., we have $f_n(\pi) \leq f_{n+1}(\pi)$ for all $n \geq 0$ and $\pi \in IPaths^{\mathcal{M}}$. By applying the *monotone convergence theorem* [1] we obtain

$$\begin{aligned} \lim_{n\to\infty} \mathrm{ER}^{\mathcal{M}}_{\sigma}(\Pi^n_G) &= \lim_{n\to\infty} \int_{\pi \in IPaths^{\mathcal{M}}} f_n(\pi)\, \mathrm{d}\Pr^{\mathcal{M}}_{\sigma}(\pi) \\ &= \int_{\pi \in IPaths^{\mathcal{M}}} \lim_{n\to\infty} f_n(\pi)\, \mathrm{d}\Pr^{\mathcal{M}}_{\sigma}(\pi) \\ &= \int_{\pi \in IPaths^{\mathcal{M}}} rew^{\mathcal{M}}(\rho, \pi, G)\, \mathrm{d}\Pr^{\mathcal{M}}_{\sigma}(\pi) = \mathrm{ER}^{\mathcal{M}}_{\sigma}(\rho, G). \end{aligned}$$

□

The next step is to show that the expected reward collected along the paths of Π^n_G coincides for $\mathcal{M}$ under σ and $\mathcal{M}_{\mathcal{D}}$ under $\mathrm{ta}(\sigma)$.

B.2 Proof of Lemma 3

Lemma 3 *Let ρ be some reward function of $\mathcal{M}$ and let $\rho^{\mathcal{D}}$ be its counterpart for $\mathcal{M}_{\mathcal{D}}$. Let $\mathcal{M} = (S, Act, \rightarrow, s_0, (\rho_1, \ldots, \rho_\ell))$ be an MA with $G \subseteq S$ and $\sigma \in GM$. For all $G \subseteq S$ and $n \geq 0$ it holds that*

$$\mathrm{ER}_{\sigma}^{\mathcal{M}}(\rho, \Pi_G^n) = \mathrm{ER}_{ta(\sigma)}^{\mathcal{M}_{\mathcal{D}}}(\rho^{\mathcal{D}}, \Pi_G^n).$$

We detail the terms (1) and (2) from the proof of Lemma 3 separately.

Term (1): Let $\Lambda_G^{\leq n} = \{\hat{\pi} \in \Pi_G^{n+1} \mid |\hat{\pi}| \leq n\}$ be the paths in Π_G^{n+1} of length at most n. We have $\Lambda_G^{\leq n} \subseteq \Pi_G^n$ and every path in $\Lambda_G^{\leq n}$ visits a state in G. Correspondingly, $\Lambda_{\neg G}^{=n} = \Pi_G^n \setminus \Lambda_G^{\leq n}$ is the set of time-abstract paths of length n that do not visit a state in G. Hence, the paths in Π_G^{n+1} with length $n+1$ have a prefix in $\Lambda_{\neg G}^{=n}$.

The set Π_G^{n+1} is partitioned such that

$$\begin{aligned}\Pi_G^{n+1} &= \Lambda_G^{\leq n} \cup \left\{\hat{\pi} \in \Pi_G^{n+1} \mid |\hat{\pi}| = n+1\right\} \\ &= \Lambda_G^{\leq n} \cup \{\hat{\pi} = \hat{\pi}' \xrightarrow{\alpha} s' \in FPaths^{\mathcal{M}_{\mathcal{D}}} \mid \hat{\pi}' \in \Lambda_{\neg G}^{=n}\}.\end{aligned}$$

The reward obtained within the first n steps is independent of the $(n+1)$-th transition. To show this formally, we fix a path $\hat{\pi}' \in \Lambda_{\neg G}^{=n}$ with $last(\hat{\pi}') = s$ and derive

$$\begin{aligned}&\sum_{\hat{\pi}' \xrightarrow{\alpha} s' \in FPaths^{\mathcal{M}_{\mathcal{D}}}} \int_{\pi \in \langle \hat{\pi}' \xrightarrow{\alpha} s' \rangle} rew^{\mathcal{M}}(pref(\pi, n))\, \mathrm{dPr}_{\sigma}^{\mathcal{M}}(\pi) \\ &= \begin{cases} \displaystyle\int_{\pi' \in \langle \hat{\pi}' \rangle} rew^{\mathcal{M}}(\pi') \cdot \sum_{(\alpha, s') \in Act \times S} \sigma(\pi', \alpha) \cdot \mathbf{P}(s, \alpha, s')\, \mathrm{dPr}_{\sigma}^{\mathcal{M}}(\pi') & \text{if } s \in \mathrm{PS} \\ \displaystyle\int_{\pi' \in \langle \hat{\pi}' \rangle} rew^{\mathcal{M}}(\pi') \cdot \sum_{s' \in S} \mathbf{P}(s, \bot, s')\, \mathrm{dPr}_{\sigma}^{\mathcal{M}}(\pi') & \text{if } s \in \mathrm{MS} \end{cases} \\ &= \int_{\pi' \in \langle \hat{\pi}' \rangle} rew^{\mathcal{M}}(\pi')\, \mathrm{dPr}_{\sigma}^{\mathcal{M}}(\pi'). \end{aligned} \tag{4}$$

With the above-mentioned partition of the set Π_G^{n+1}, it follows that the expected reward obtained within the first n steps is given by

$$\begin{aligned}&\sum_{\hat{\pi} \in \Pi_G^{n+1}} \int_{\pi \in \langle \hat{\pi} \rangle} rew^{\mathcal{M}}(pref(\pi, n))\, \mathrm{dPr}_{\sigma}^{\mathcal{M}}(\pi) \\ &= \sum_{\hat{\pi} \in \Lambda_G^{\leq n}} \int_{\pi \in \langle \hat{\pi} \rangle} rew^{\mathcal{M}}(\pi)\, \mathrm{dPr}_{\sigma}^{\mathcal{M}}(\pi) \\ &\quad + \sum_{\hat{\pi}' \in \Lambda_{\neg G}^{=n}} \sum_{\hat{\pi}' \xrightarrow{\alpha} s' \in FPaths^{\mathcal{M}_{\mathcal{D}}}} \int_{\pi \in \langle \hat{\pi}' \xrightarrow{\alpha} s' \rangle} rew^{\mathcal{M}}(pref(\pi, n))\, \mathrm{dPr}_{\sigma}^{\mathcal{M}}(\pi) \\ &\overset{(4)}{=} \sum_{\hat{\pi} \in \Lambda_G^{\leq n}} \int_{\pi \in \langle \hat{\pi} \rangle} rew^{\mathcal{M}}(\pi)\, \mathrm{dPr}_{\sigma}^{\mathcal{M}}(\pi) + \sum_{\hat{\pi} \in \Lambda_{\neg G}^{=n}} \int_{\pi \in \langle \hat{\pi} \rangle} rew^{\mathcal{M}}(\pi)\, \mathrm{dPr}_{\sigma}^{\mathcal{M}}(\pi) \\ &= \mathrm{ER}_{\sigma}^{\mathcal{M}}(\Pi_G^n)\end{aligned}$$

$$
\begin{aligned}
&\overset{IH}{=} \mathrm{ER}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\Pi^n_G)\\
&= \sum_{\hat{\pi}\in\Lambda^{\leq n}_G} rew^{\mathcal{M}_{\mathcal{D}}}(\hat{\pi})\cdot \mathrm{Pr}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\hat{\pi}) + \sum_{\hat{\pi}\in\Lambda^{=n}_{\neg G}} rew^{\mathcal{M}_{\mathcal{D}}}(\hat{\pi})\cdot \mathrm{Pr}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\hat{\pi})\\
&= \sum_{\hat{\pi}\in\Lambda^{\leq n}_G} rew^{\mathcal{M}_{\mathcal{D}}}(\hat{\pi})\cdot \mathrm{Pr}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\hat{\pi})\\
&\quad + \sum_{\hat{\pi}'\in\Lambda^{=n}_{\neg G}} \sum_{\substack{\hat{\pi}\in FPaths^{\mathcal{M}_{\mathcal{D}}}\\ \hat{\pi}=\hat{\pi}'\xrightarrow{\alpha}s'}} rew^{\mathcal{M}_{\mathcal{D}}}(pref(\hat{\pi},n))\cdot \mathrm{Pr}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\hat{\pi})\\
&= \sum_{\hat{\pi}\in\Pi^{n+1}_G} rew^{\mathcal{M}_{\mathcal{D}}}(pref(\hat{\pi},n))\cdot \mathrm{Pr}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\hat{\pi}). \qquad (5)
\end{aligned}
$$

Term (2): For the expected reward obtained in step $n+1$, consider a path $\hat{\pi} = \hat{\pi}' \xrightarrow{\alpha} s' \in \Pi^{n+1}_G$ such that $|\hat{\pi}'| = n$ and $last(\hat{\pi}') = s$.

– If $s \in \mathrm{MS}$, we have $\hat{\pi} = \hat{\pi}' \xrightarrow{\perp} s'$. It follows that

$$
\begin{aligned}
&\int_{\pi=\pi'\xrightarrow{t}s'\in\langle\hat{\pi}\rangle} \rho(s)\cdot t + \rho(s,\perp)\,\mathrm{dPr}^{\mathcal{M}}_{\sigma}(\pi)\\
&= \int_{\pi=\pi'\xrightarrow{t}s'\in\langle\hat{\pi}\rangle} \rho(s)\cdot t\,\mathrm{dPr}^{\mathcal{M}}_{\sigma}(\pi) + \int_{\pi\in\langle\hat{\pi}\rangle} \rho(s,\perp)\,\mathrm{dPr}^{\mathcal{M}}_{\sigma}(\pi)\\
&= \rho(s)\cdot \int_{\pi'\in\langle\hat{\pi}'\rangle}\int_0^\infty t\cdot \mathrm{E}(s)\cdot e^{-\mathrm{E}(s)t}\cdot \mathbf{P}(s,\perp,s')\,\mathrm{d}t\,\mathrm{dPr}^{\mathcal{M}}_{\sigma}(\pi')\\
&\quad + \rho(s,\perp)\cdot \mathrm{Pr}^{\mathcal{M}}_{\sigma}(\langle\hat{\pi}\rangle)\\
&= \frac{\rho(s)}{\mathrm{E}(s)}\cdot \mathrm{Pr}^{\mathcal{M}}_{\sigma}(\langle\hat{\pi}\rangle) + \rho(s,\perp)\cdot \mathrm{Pr}^{\mathcal{M}}_{\sigma}(\langle\hat{\pi}\rangle)\\
&= \rho^{\mathcal{D}}(s,\perp)\cdot \mathrm{Pr}^{\mathcal{M}}_{\sigma}(\langle\hat{\pi}\rangle) \overset{Lem. 1}{=} \rho^{\mathcal{D}}(s,\perp)\cdot \mathrm{Pr}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\hat{\pi}). \qquad (6)
\end{aligned}
$$

– If $s \in \mathrm{PS}$, then $\int_{\pi=\pi'\xrightarrow{\alpha}s'\in\langle\hat{\pi}\rangle} \rho(s,\alpha)\,\mathrm{dPr}^{\mathcal{M}}_{\sigma}(\pi) = \rho^{\mathcal{D}}(s,\alpha)\cdot \mathrm{Pr}^{\mathcal{M}_{\mathcal{D}}}_{\mathrm{ta}(\sigma)}(\hat{\pi})$ follows similarly.

C Proofs for timed reachability

C.1 Proof of Lemma 4

Lemma 4 *Let $\mathcal{M}$ be an MA with scheduler $\sigma \in GM$, digitization $\mathcal{M}_\delta$, and digital path $\bar{\pi} \in FPaths^{\mathcal{M}_\delta}$. It holds that*

$$
Pr^{\mathcal{M}}_{\sigma}([\bar{\pi}]_{cyl}) = Pr^{\mathcal{M}_\delta}_{di(\sigma)}(\bar{\pi}).
$$

Proof The proof is by induction over the length n of $\bar{\pi}$. Let $\mathcal{M} = (S, Act, \rightarrow, s_0, (\rho_1,\dots,\rho_\ell))$ and $\mathcal{M}_\delta = (S, Act, \mathbf{P}_\delta, s_0, (\rho^\delta_1,\dots,\rho^\delta_\ell))$. If $n=0$, then $\bar{\pi} = s_0$ and $[\bar{\pi}]_{cyl} = IPaths^{\mathcal{M}}$. Hence, $\mathrm{Pr}^{\mathcal{M}}_{\sigma}([s_0]_{cyl}) = 1 = \mathrm{Pr}^{\mathcal{M}_\delta}_{\mathrm{di}(\sigma)}(s_0)$. In the induction step it is assumed that the lemma holds for a fixed path $\bar{\pi} \in FPaths^{\mathcal{M}_\delta}$ with $|\bar{\pi}| = n$ and $last(\bar{\pi}) = s$. Consider a path $\bar{\pi} \xrightarrow{\alpha} s' \in FPaths^{\mathcal{M}_\delta}$.

If $\Pr_\sigma^{\mathcal{M}}([\bar{\pi}]_{cyl}) = \Pr_{\mathrm{di}(\sigma)}^{\mathcal{M}_\delta}(\bar{\pi}) = 0$, then $\Pr_\sigma^{\mathcal{M}}([\bar{\pi} \xrightarrow{\alpha} s']_{cyl}) = \Pr_{\mathrm{di}(\sigma)}^{\mathcal{M}_\delta}(\bar{\pi} \xrightarrow{\alpha} s') = 0$ because $[\bar{\pi} \xrightarrow{\alpha} s']_{cyl} \subseteq [\bar{\pi}]_{cyl}$ and $Cyl(\{\bar{\pi} \xrightarrow{\alpha} s'\}) \subseteq Cyl(\{\bar{\pi}\})$.

Now assume $\Pr_\sigma^{\mathcal{M}}([\bar{\pi}]_{cyl}) > 0$. We distinguish the following cases.

<u>*Cases* $\in$ PS :</u> It follows that $[\bar{\pi} \xrightarrow{\alpha} s']_{cyl} = Cyl([\bar{\pi} \xrightarrow{\alpha} s'])$ since $\bar{\pi} \xrightarrow{\alpha} s'$ ends with a probabilistic transition. Hence,

$$\begin{aligned}
\Pr_\sigma^{\mathcal{M}}([\bar{\pi} \xrightarrow{\alpha} s']_{cyl}) &= \Pr_\sigma^{\mathcal{M}}([\bar{\pi} \xrightarrow{\alpha} s']) \\
&= \int_{\pi \in [\bar{\pi}]} \sigma(\pi, \alpha) \cdot \mathbf{P}(s, \alpha, s') \, \mathrm{d}\Pr_\sigma^{\mathcal{M}}(\pi) \\
&= \int_{\pi \in [\bar{\pi}]} \sigma(\pi, \alpha) \cdot \mathbf{P}(s, \alpha, s') \, \mathrm{d}\Pr_\sigma^{\mathcal{M}}(\{\pi\} \cap [\bar{\pi}]) \\
&= \int_{\pi \in [\bar{\pi}]} \sigma(\pi, \alpha) \cdot \mathbf{P}(s, \alpha, s') \, \mathrm{d}\big[\Pr_\sigma^{\mathcal{M}}(\pi \mid [\bar{\pi}]) \cdot \Pr_\sigma^{\mathcal{M}}([\bar{\pi}])\big] \\
&= \Pr_\sigma^{\mathcal{M}}([\bar{\pi}]) \cdot \mathbf{P}(s, \alpha, s') \cdot \int_{\pi \in [\bar{\pi}]} \sigma(\pi, \alpha) \, \mathrm{d}\Pr_\sigma^{\mathcal{M}}(\pi \mid [\bar{\pi}]) \\
&= \Pr_\sigma^{\mathcal{M}}([\bar{\pi}]) \cdot \mathbf{P}(s, \alpha, s') \cdot \mathrm{di}(\sigma)(\bar{\pi}, \alpha) \\
&\overset{IH}{=} \Pr_{\mathrm{di}(\sigma)}^{\mathcal{M}_\delta}(\bar{\pi}) \cdot \mathbf{P}(s, \alpha, s') \cdot \mathrm{di}(\sigma)(\bar{\pi}, \alpha) \\
&= \Pr_{\mathrm{di}(\sigma)}^{\mathcal{M}_\delta}(\bar{\pi} \xrightarrow{\alpha} s').
\end{aligned}$$

<u>*Cases* $\in$ MS :</u> As $s \in \mathrm{MS}$ we have $\alpha = \bot$ and it follows

$$\begin{aligned}
\Pr_\sigma^{\mathcal{M}}([\bar{\pi} \xrightarrow{\bot} s']_{cyl}) &= \Pr_\sigma^{\mathcal{M}}([\bar{\pi}]_{cyl} \cap [\bar{\pi} \xrightarrow{\bot} s']_{cyl}) \\
&= \Pr_\sigma^{\mathcal{M}}([\bar{\pi}]_{cyl}) \cdot \Pr_\sigma^{\mathcal{M}}([\bar{\pi} \xrightarrow{\bot} s']_{cyl} \mid [\bar{\pi}]_{cyl}). \qquad (7)
\end{aligned}$$

Assume that a path $\pi \in [\bar{\pi}]_{cyl}$ has been observed, i.e., $pref(\mathrm{di}(\pi), m) = \bar{\pi}$ holds for some $m \geq 0$. The term $\Pr_\sigma^{\mathcal{M}}([\bar{\pi} \xrightarrow{\bot} s']_{cyl} \mid [\bar{\pi}]_{cyl})$ coincides with the probability that also $pref(\mathrm{di}(\pi), m+1) = \bar{\pi} \xrightarrow{\bot} s'$ holds.

We have either

- $s \neq s'$ which means that the transition from s to s' has to be taken during a period of δ time units or
- $s = s'$ where we additionally have to consider the case that no transition is taken at s for δ time units.

It follows that

$$\begin{aligned}
\Pr_\sigma^{\mathcal{M}}([\bar{\pi} \xrightarrow{\bot} s']_{cyl} \mid [\bar{\pi}]_{cyl}) &= \begin{cases} \mathbf{P}(s, \bot, s')(1 - e^{-\mathrm{E}(s)\delta}) & \text{if } s \neq s' \\ \mathbf{P}(s, \bot, s')(1 - e^{-\mathrm{E}(s)\delta}) + e^{-\mathrm{E}(s)\delta} & \text{if } s = s' \end{cases} \\
&= \mathbf{P}_\delta(s, \bot, s'). \qquad (8)
\end{aligned}$$

We conclude that

$$\begin{aligned}
\Pr_\sigma^{\mathcal{M}}([\bar{\pi} \xrightarrow{\bot} s']_{cyl}) &\overset{(7),(8)}{=} \Pr_\sigma^{\mathcal{M}}([\bar{\pi}]_{cyl}) \cdot \mathbf{P}_\delta(s, \bot, s') \\
&\overset{IH}{=} \Pr_{\mathrm{di}(\sigma)}^{\mathcal{M}_\delta}(\bar{\pi}) \cdot \mathbf{P}_\delta(s, \bot, s') = \Pr_{\mathrm{di}(\sigma)}^{\mathcal{M}_\delta}(\bar{\pi} \xrightarrow{\bot} s').
\end{aligned}$$

□

C.2 Proof of Proposition 4

Proposition 4 *Let $\mathcal{M}$ be an MA with $G \subseteq S$, $\sigma \in GM$, and digitization $\mathcal{M}_\delta$. Further, let $J \subseteq \mathbb{N}$ be a set of consecutive natural numbers. It holds that*

$$Pr_\sigma^{\mathcal{M}}([\Diamond_{ds}^J G]) = Pr_{di(\sigma)}^{\mathcal{M}_\delta}(\Diamond_{ds}^J G).$$

Proof Consider the set $\Pi_G^J \subseteq FPaths^{\mathcal{M}_\delta}$ of paths that (i) visit G within J digitization steps and (ii) do not have a proper prefix that satisfies (i). Every path in $\Diamond_{\text{ds}}^J G$ has a unique prefix in Π_G^J, yielding

$$\Diamond_{\text{ds}}^J G = \biguplus_{\bar{\pi} \in \Pi_G^J} Cyl(\{\bar{\pi}\})$$

For the corresponding paths of $\mathcal{M}$ we obtain

$$\begin{aligned}[\Diamond_{\text{ds}}^J G] &= \{\pi \in IPaths^{\mathcal{M}} \mid \text{di}(\pi) \in \Diamond_{\text{ds}}^J G\} \\ &= \{\pi \in IPaths^{\mathcal{M}} \mid \text{di}(\pi) \text{ has a unique prefix in } \Pi_G^J\} \\ &= \biguplus_{\bar{\pi} \in \Pi_G^J} [\bar{\pi}]_{cyl} .\end{aligned}$$

The proposition follows with Lemma 4 since

$$\text{Pr}_{\text{di}(\sigma)}^{\mathcal{M}_\delta}(\Diamond_{\text{ds}}^J G) = \sum_{\bar{\pi} \in \Pi_G^J} \text{Pr}_{\text{di}(\sigma)}^{\mathcal{M}_\delta}(\bar{\pi}) \overset{Lem. 4}{=} \sum_{\bar{\pi} \in \Pi_G^J} \text{Pr}_\sigma^{\mathcal{M}}([\bar{\pi}]_{cyl}) = \text{Pr}_\sigma^{\mathcal{M}}([\Diamond_{\text{ds}}^J G]).$$

□

C.3 Proofs of Lemmas 6 and 7

Lemma 7 *Let $\mathcal{M}$ be an MA with $\sigma \in GM$ and maximum rate $\lambda = \max\{E(s) \mid s \in MS\}$. For each $\delta \in \mathbb{R}_{>0}$, $k \in \mathbb{N}$, and $t \in \mathbb{R}_{\geq 0}$ it holds that*

$$Pr_\sigma^{\mathcal{M}}(\#[k\delta + t]^{\leq k}) \geq Pr_\sigma^{\mathcal{M}}(\#[k\delta]^{\leq k}) \cdot e^{-\lambda t} .$$

Proof First, we show that the set $\#[k\delta + t]^{\leq k}$ corresponds to the paths of $\#[k\delta]^{\leq k}$ with the additional requirement that no transition is taken between the time points $k\delta$ and $k\delta + t$, i.e.,

$$\#[k\delta + t]^{\leq k} = \{\pi \in \#[k\delta]^{\leq k} \mid \text{there is no prefix } \pi' \text{ of } \pi \text{ with } k\delta < T(\pi') \leq k\delta + t\}.$$

"$\subseteq$": If $\pi \in \#[k\delta + t]^{\leq k}$, then $\pi \in \#[k\delta]^{\leq k}$ follows immediately. Furthermore, assume towards a contradiction that there is a prefix π' of π with $k\delta < T(\pi') \leq k\delta + t$. Then, $k < \frac{T(\pi')}{\delta} \leq |\pi'|_{\text{ds}}$ (cf. Lemma 5). As $T(\pi') \leq k\delta + t$, this means that $|pref_T(\pi, k\delta + t)|_{\text{ds}} \geq |\pi'|_{\text{ds}} > k$ which contradicts $\pi \in \#[k\delta + t]^{\leq k}$.
"$\supseteq$": For $\pi \in \#[k\delta]^{\leq k}$ with no prefix π' such that $k\delta < T(\pi') \leq k\delta + t$, it holds that $pref_T(\pi, k\delta + t) = pref_T(\pi, k\delta)$. Hence, $|pref_T(\pi, k\delta + t)|_{\text{ds}} = |pref_T(\pi, k\delta)|_{\text{ds}} \leq k$ and it follows that $\pi \in \#[k\delta + t]^{\leq k}$.

The probability for no transition to be taken between $k\delta$ and $k\delta + t$ only depends on the current state at time point $k\delta$. More precisely, for some state $s \in \text{MS}$ assume the set of paths $\{\pi \in \#[k\delta]^{\leq k} \mid last(pref_T(\pi, k\delta)) = s\}$. The probability that a path in this set takes no transition between time points $k\delta$ and $k\delta + t$ is given by $e^{-\mathrm{E}(s)t}$. With $\lambda \geq \mathrm{E}(s)$ for all $s \in \mathrm{MS}$ it follows that

$$
\begin{aligned}
&\Pr_{\sigma}^{\mathcal{M}}(\#[k\delta + t]^{\leq k}) \\
&= \Pr_{\sigma}^{\mathcal{M}}(\{\pi \in \#[k\delta]^{\leq k} \mid \text{there is no prefix } \pi' \text{ of } \pi \text{ with } k\delta < T(\pi') \leq k\delta + t\}) \\
&= \sum_{s \in \mathrm{MS}} \Pr_{\sigma}^{\mathcal{M}}(\{\pi \in \#[k\delta]^{\leq k} \mid \mathit{last}(\mathit{pref}_T(\pi, k\delta)) = s\}) \cdot e^{-\mathrm{E}(s)t} \\
&\geq \sum_{s \in \mathrm{MS}} \Pr_{\sigma}^{\mathcal{M}}(\{\pi \in \#[k\delta]^{\leq k} \mid \mathit{last}(\mathit{pref}_T(\pi, k\delta)) = s\}) \cdot e^{-\lambda t} \\
&= \Pr_{\sigma}^{\mathcal{M}}(\#[k\delta]^{\leq k}) \cdot e^{-\lambda t}\ .
\end{aligned}
$$

□

Lemma 6 *Let $\mathcal{M}$ be an MA with $\sigma \in GM$ and maximum rate $\lambda = \max\{E(s) \mid s \in MS\}$. Further, let $\delta \in \mathbb{R}_{>0}$ and $k \in \mathbb{N}$. It holds that*

$$
Pr_{\sigma}^{\mathcal{M}}(\#[k\delta]^{>k}) \leq 1 - (1 + \lambda\delta)^k \cdot e^{-\lambda\delta k}
$$

Proof Let $\mathcal{M} = (S, Act, \rightarrow, s_0, (\rho_1, \ldots, \rho_\ell))$. By induction over k we show that

$$
\Pr_{\sigma}^{\mathcal{M}}(\#[k\delta]^{\leq k}) \geq (1 + \lambda\delta)^k \cdot e^{-\lambda\delta k}.
$$

The claim follows as $\#[k\delta]^{>k} = \mathit{IPaths}^{\mathcal{M}} \setminus \#[k\delta]^{\leq k}$.

For $k = 0$, we have $\pi \in \#[0 \cdot \delta]^{\leq 0}$ iff π takes no Markovian transition at time point zero. As this happens with probability one, it follows that

$$
\Pr_{\sigma}^{\mathcal{M}}(\#[0 \cdot \delta]^{\leq 0}) = 1 = (1 + \lambda\delta)^0 \cdot e^{-\lambda\delta \cdot 0}\ .
$$

We assume in the induction step that the proposition holds for some fixed k. We distinguish between two cases for the initial state s_0 of $\mathcal{M}$.

Case $s_0 \in \mathrm{MS}$: We partition the set $\#[k\delta + \delta]^{\leq k+1} = \Lambda^{\geq\delta} \cup \Lambda^{<\delta}$ with

$$
\begin{aligned}
\Lambda^{\geq\delta} &= \{s_0 \xrightarrow{t} s_1 \xrightarrow{\kappa_1} \cdots \in \#[k\delta + \delta]^{\leq k+1} \mid t \geq \delta\} \text{ and} \\
\Lambda^{<\delta} &= \{s_0 \xrightarrow{t} s_1 \xrightarrow{\kappa_1} \cdots \in \#[k\delta + \delta]^{\leq k+1} \mid t < \delta\}.
\end{aligned}
$$

Hence, $\Lambda^{\geq\delta}$ contains the paths where we wait at least δ time units at s_0 and $\Lambda^{<\delta}$ contains the paths where the first transition is taken within $t < \delta$ time units. It follows that $\Pr_{\sigma}^{\mathcal{M}}(\#[k\delta + \delta]^{\leq k+1}) = \Pr_{\sigma}^{\mathcal{M}}(\Lambda^{\geq\delta}) + \Pr_{\sigma}^{\mathcal{M}}(\Lambda^{<\delta})$. We consider the probabilities for $\Lambda^{\geq\delta}$ and $\Lambda^{<\delta}$ separately.

– $\Pr_{\sigma}^{\mathcal{M}}(\Lambda^{\geq\delta})$: For a path $s_0 \xrightarrow{t+\delta} s_1 \xrightarrow{\kappa_1} \cdots \in \Lambda^{\geq\delta}$, after the first δ time units there are at most k digitization steps within the next $k\delta$ time units, i.e.,

$$
s_0 \xrightarrow{t+\delta} s_1 \xrightarrow{\kappa_1} \cdots \in \Lambda^{\geq\delta} \iff s_0 \xrightarrow{t} s_1 \xrightarrow{\kappa_1} \cdots \in \#[k\delta]^{\leq k}.
$$

The probability for $\Lambda^{\geq\delta}$ can therefore be derived from the probability to wait at s_0 for at least δ time units and the probability for $\#[k\delta]^{\leq k}$. In order to apply this, we need to modify the considered scheduler as it might depend on the sojourn time in s_0. Let σ_δ be the scheduler for $\mathcal{M}$ that mimics σ on paths where the first transition is delayed by δ, i.e., σ_δ satisfies

$$
\sigma_\delta(s_0 \xrightarrow{t} \ldots \xrightarrow{\kappa_{n-1}} s_n, \alpha) = \sigma(s_0 \xrightarrow{t+\delta} \ldots \xrightarrow{\kappa_{n-1}} s_n, \alpha).
$$

for all $s_0 \xrightarrow{t} \dots \xrightarrow{\kappa_{n-1}} s_n \in \mathit{FPaths}^{\mathcal{M}}$ and $\alpha \in \mathit{Act}$. It holds that

$$\begin{aligned}\Pr_\sigma^{\mathcal{M}}(\Lambda^{\geq\delta}) &= e^{-\mathrm{E}(s_0)\delta} \cdot \Pr_{\sigma_\delta}^{\mathcal{M}}(\#[k\delta]^{\leq k}) \\ &\overset{IH}{\geq} e^{-\mathrm{E}(s_0)\delta} \cdot (1+\lambda\delta)^k \cdot e^{-\lambda\delta k} \\ &= e^{-\mathrm{E}(s_0)\delta} \cdot (1+\lambda\delta)^k \cdot e^{-\lambda\delta k} \cdot e^{-\lambda\delta} \cdot e^{\lambda\delta} \\ &= (1+\lambda\delta)^k \cdot e^{-\lambda\delta(k+1)} \cdot e^{(\lambda-\mathrm{E}(s_0))\delta}\,. \end{aligned} \tag{9}$$

– $\Pr_\sigma^{\mathcal{M}}(\Lambda^{<\delta})$: For a path $s_0 \xrightarrow{t} s_1 \xrightarrow{\kappa_1} \dots \in \Lambda^{<\delta}$, the first digitization step happens at less than δ time units, i.e., $0 \leq t < \delta$. It follows that there are at most k digitization steps in the remaining $k\delta + \delta - t$ time units, i.e.,

$$s_0 \xrightarrow{t} s_1 \xrightarrow{\kappa_1} s_2 \xrightarrow{\kappa_2} \dots \in \Lambda^{<\delta} \iff s_1 \xrightarrow{\kappa_1} s_2 \xrightarrow{\kappa_2} \dots \in \#^{s_1}[k\delta+\delta-t]^{\leq k}\,,$$

where $\#^{s_1}[k\delta+\delta-t]^{\leq k}$ refers to the paths $\#[k\delta+\delta-t]^{\leq k}$ of $\mathcal{M}^{s_1} = (S, \mathit{Act}, \rightarrow, s_1, (\rho_1, \dots, \rho_\ell))$, the MA obtained from $\mathcal{M}$ by changing the initial state to s_1. Hence, the probability for $\Lambda^{<\delta}$ can be derived from the probability to take a transition from s_0 to some state s within $t < \delta$ time units and the probability for $\#^{s}[k\delta+\delta-t]^{\leq k}$. Again, we need to adapt the considered scheduler. Let $\pi \in \mathit{FPaths}^{\mathcal{M}}$ with $\mathit{last}(\pi) = s$. The scheduler $\sigma[\pi]$ for $\mathcal{M}^s$ mimics the scheduler σ for $\mathcal{M}$, where π is prepended to the given path, i.e., we set

$$\sigma[\pi](s \xrightarrow{\kappa_j} \dots \xrightarrow{\kappa_{n-1}} s_n, \alpha) = \sigma(\pi \xrightarrow{\kappa_j} \dots \xrightarrow{\kappa_{n-1}} s_n, \alpha)$$

for all $s \xrightarrow{\kappa_j} \dots \xrightarrow{\kappa_{n-1}} s_n \in \mathit{FPaths}^{\mathcal{M}^s}$ and $\alpha \in \mathit{Act}$. With Lemma 7 it follows that

$$\begin{aligned}&\Pr_\sigma^{\mathcal{M}}(\Lambda^{<\delta}) \\ &= \int_0^\delta \mathrm{E}(s_0) \cdot e^{-\mathrm{E}(s_0)t} \cdot \left(\sum_{s\in S} \mathbf{P}(s_0, \bot, s) \cdot \Pr_{\sigma[\pi]}^{\mathcal{M}^s}(\#^s[k\delta+\delta-t]^{\leq k})\right) \mathrm{d}t \\ &\geq \int_0^\delta \mathrm{E}(s_0) \cdot e^{-\mathrm{E}(s_0)t} \cdot \left(\sum_{s\in S} \mathbf{P}(s_0, \bot, s) \cdot \Pr_{\sigma[\pi]}^{\mathcal{M}^s}(\#^s[k\delta]^{\leq k}) \cdot e^{-\lambda(\delta-t)}\right) \mathrm{d}t \\ &\overset{IH}{\geq} \int_0^\delta \mathrm{E}(s_0) \cdot e^{-\mathrm{E}(s_0)t} \cdot \left(\sum_{s\in S} \mathbf{P}(s_0, \bot, s) \cdot (1+\lambda\delta)^k \cdot e^{-\lambda\delta k} \cdot e^{-\lambda(\delta-t)}\right) \mathrm{d}t \\ &= (1+\lambda\delta)^k \cdot e^{-\lambda\delta k} \cdot \mathrm{E}(s_0) \cdot \int_0^\delta e^{-\mathrm{E}(s_0)t} \cdot e^{-\lambda(\delta-t)} \cdot \left(\sum_{s\in S} \mathbf{P}(s_0, \bot, s)\right) \mathrm{d}t \\ &= (1+\lambda\delta)^k \cdot e^{-\lambda\delta k} \cdot \mathrm{E}(s_0) \cdot \int_0^\delta e^{-\mathrm{E}(s_0)t} \cdot e^{-\lambda\delta} \cdot e^{\lambda t}\, \mathrm{d}t \\ &= (1+\lambda\delta)^k \cdot e^{-\lambda\delta(k+1)} \cdot \mathrm{E}(s_0) \cdot \int_0^\delta e^{(\lambda-\mathrm{E}(s_0))t}\, \mathrm{d}t\,. \end{aligned} \tag{10}$$

Combining the results for $\Lambda^{\geq\delta}$ and $\Lambda^{<\delta}$ (i.e., Equations 9 and 10), we obtain

$$\begin{aligned}&\Pr_\sigma^{\mathcal{M}}(\#[k\delta+\delta]^{\leq k+1}) \\ &\quad= \Pr_\sigma^{\mathcal{M}}(\Lambda^{\geq\delta}) + \Pr_\sigma^{\mathcal{M}}(\Lambda^{<\delta}) \\ &\quad\geq (1+\lambda\delta)^k \cdot e^{-\lambda\delta(k+1)} \cdot \left(e^{(\lambda-\mathrm{E}(s_0))\delta} + \mathrm{E}(s_0) \cdot \int_0^\delta e^{(\lambda-\mathrm{E}(s_0))t}\, \mathrm{d}t\right)\end{aligned}$$

$$\stackrel{*}{\geq} (1+\lambda\delta)^k \cdot e^{-\lambda\delta(k+1)} \cdot (1+\lambda\delta) = (1+\lambda\delta)^{k+1} \cdot e^{-\lambda\delta(k+1)} ,$$

where the inequality marked with $*$ is due to

$$\begin{aligned}
&e^{(\lambda-\mathrm{E}(s_0))\delta} + \mathrm{E}(s_0) \cdot \int_0^\delta e^{(\lambda-\mathrm{E}(s_0))t}\,\mathrm{d}t \\
&= e^{(\lambda-\mathrm{E}(s_0))\delta} + (\mathrm{E}(s_0) - \lambda + \lambda) \cdot \int_0^\delta e^{(\lambda-\mathrm{E}(s_0))t}\,\mathrm{d}t \\
&= e^{(\lambda-\mathrm{E}(s_0))\delta} - \big(\lambda - \mathrm{E}(s_0)\big) \cdot \int_0^\delta e^{(\lambda-\mathrm{E}(s_0))t}\,\mathrm{d}t + \lambda \cdot \int_0^\delta e^{(\lambda-\mathrm{E}(s_0))t}\,\mathrm{d}t \\
&= \begin{cases} 1 - 0 + \lambda \cdot \int_0^\delta e^{(\lambda-\mathrm{E}(s_0))t}\,\mathrm{d}t & \text{if } \mathrm{E}(s_0) = \lambda \\ e^{(\lambda-\mathrm{E}(s_0))\delta} - \big(e^{(\lambda-\mathrm{E}(s_0))\delta} - 1\big) + \lambda \cdot \int_0^\delta e^{(\lambda-\mathrm{E}(s_0))t}\,\mathrm{d}t & \text{if } \mathrm{E}(s_0) < \lambda \end{cases} \\
&= 1 + \lambda \cdot \int_0^\delta e^{(\lambda-\mathrm{E}(s_0))t}\,\mathrm{d}t \geq 1 + \lambda \cdot \int_0^\delta 1\,\mathrm{d}t = 1 + \lambda\delta .
\end{aligned}$$

Case $s_0 \in \mathrm{PS}$: Since $\mathcal{M}$ is non-zeno, a state $s \in \mathrm{MS}$ is reached from s_0 within zero time almost surely (i.e., with probability one). From the previous case, it already follows that the Proposition holds for $\mathcal{M}^s$ with $s \in \mathrm{MS}$ and the set $\#^s[k\delta + \delta]^{\leq k+1}$. With $\Pi_{\mathrm{MS}} = \{s_0 \xrightarrow{\kappa_0} \dots \xrightarrow{\kappa_{n-1}} s_n \in \mathit{FPaths}^{\mathcal{M}} \mid s_n \in \mathrm{MS} \text{ and } \forall i < n\colon s_i \in \mathrm{PS}\}$ we obtain

$$\begin{aligned}
\mathrm{Pr}_\sigma^{\mathcal{M}}(\#[k\delta+\delta]^{\leq k+1}) &= \int_{\substack{\pi \in \Pi_{\mathrm{MS}} \\ \mathit{last}(\pi)=s}} \mathrm{Pr}_{\sigma[\pi]}^{\mathcal{M}^s}(\#^s[k\delta+\delta]^{\leq k+1})\,\mathrm{d}\mathrm{Pr}_\sigma^{\mathcal{M}}(\pi) \\
&\geq \int_{\substack{\pi \in \Pi_{\mathrm{MS}} \\ \mathit{last}(\pi)=s}} (1+\lambda\delta)^{k+1} \cdot e^{-\lambda\delta(k+1)}\,\mathrm{d}\mathrm{Pr}_\sigma^{\mathcal{M}}(\pi) \\
&= (1+\lambda\delta)^{k+1} \cdot e^{-\lambda\delta(k+1)} \cdot \mathrm{Pr}_\sigma^{\mathcal{M}}(\Pi_{\mathrm{MS}}) \\
&= (1+\lambda\delta)^{k+1} \cdot e^{-\lambda\delta(k+1)} .
\end{aligned}$$

□

C.4 Proof of Proposition 5

Proposition 5 *For MA $\mathcal{M}$, scheduler $\sigma \in GM$, goal states $G \subseteq S$, digitization constant $\delta \in \mathbb{R}_{>0}$ and time interval I*

$$Pr_\sigma^{\mathcal{M}}(\Diamond^I G) \in Pr_\sigma^{\mathcal{M}}([\Diamond_{ds}^I G]) + \Big[-\varepsilon^{\downarrow}(I),\ \varepsilon^{\uparrow}(I)\Big].$$

We show Eq. 3, that is,

$$\mathrm{Pr}_\sigma^{\mathcal{M}}([\Diamond_{\mathrm{ds}}^{\mathrm{di}(I)} G] \setminus \Diamond^I G) \leq \varepsilon^{\downarrow}(I) \text{ and } \mathrm{Pr}_\sigma^{\mathcal{M}}(\Diamond^I G \setminus [\Diamond_{\mathrm{ds}}^{\mathrm{di}(I)} G]) \leq \varepsilon^{\uparrow}(I)$$

for the remaining forms of the time interval I.

Case $I = [0, \infty)$: In this case we have $\mathrm{di}(I) = \mathbb{N}$. It follows that

$$[\Diamond_{\mathrm{ds}}^{\mathrm{di}(I)} G] = \Diamond^I G = \{s_0 \xrightarrow{\kappa_0} s_1 \xrightarrow{\kappa_1} \cdots \in \mathit{IPaths}^{\mathcal{M}} \mid s_i \in G \text{ for some } i \geq 0\}.$$

Hence,

$$\Pr_\sigma^{\mathcal{M}}([\Diamond_{\mathrm{ds}}^{\mathrm{di}(I)} G] \setminus \Diamond^I G) = \Pr_\sigma^{\mathcal{M}}(\Diamond^I G \setminus [\Diamond_{\mathrm{ds}}^{\mathrm{di}(I)} G]) = \Pr_\sigma^{\mathcal{M}}(\emptyset) = 0 = \varepsilon^{\downarrow}(I) = \varepsilon^{\uparrow}(I).$$

<u>*Case* $I = [a, \infty)$ *for* $a = \mathrm{di}_a \delta$:</u> We have $\mathrm{di}(I) = \{\mathrm{di}_a + 1, \mathrm{di}_a + 2, \dots\}$.

– We show that $[\Diamond_{\mathrm{ds}}^{\mathrm{di}(I)} G] \setminus \Diamond^I G \subseteq \#[a]^{>\mathrm{di}_a}$. With Lemma 6 we obtain

$$\Pr_\sigma^{\mathcal{M}}([\Diamond_{\mathrm{ds}}^{\mathrm{di}(I)} G] \setminus \Diamond^I G) \le \Pr_\sigma^{\mathcal{M}}(\#[a]^{>\mathrm{di}_a}) \le 1 - (1 + \lambda\delta)^{\mathrm{di}_a} \cdot e^{-\lambda a} = \varepsilon^{\downarrow}(I).$$

Consider a path $\pi \in [\Diamond_{\mathrm{ds}}^{\mathrm{di}(I)} G] \setminus \Diamond^I G$. As $\pi \notin \Diamond^I G$, it follows that π has to reach (and leave) G within less than a time units. Let $\bar{\pi}$ be the largest prefix of $\mathrm{di}(\pi)$ that satisfies $\mathit{last}(\bar{\pi}) \in G$. Our observations yield that π leaves $\mathit{last}(\bar{\pi})$ before time point a. Hence, $\bar{\pi}$ is a prefix of $\mathrm{di}(\mathit{pref}_T(\pi, a))$. Moreover, $|\bar{\pi}|_{\mathrm{ds}} \in \mathrm{di}(I)$ as $\mathrm{di}(\pi) \in \Diamond_{\mathrm{ds}}^{\mathrm{di}(I)} G$. It follows that $|\mathit{pref}_T(\pi, a)|_{\mathrm{ds}} \ge |\bar{\pi}|_{\mathrm{ds}} > \mathrm{di}_a$ which implies $\pi \in \#[a]^{>\mathrm{di}_a}$.

– Now consider a path $\pi \in \Diamond^I G \setminus [\Diamond_{\mathrm{ds}}^{\mathrm{di}(I)} G]$. π visits G at least once since $\pi \in \Diamond^I G$. Moreover, $\mathrm{di}(\pi)$ does not visit G after di_a digitization steps due to $\pi \notin [\Diamond_{\mathrm{ds}}^{\mathrm{di}(I)} G]$. This means π visits G only finitely often. Let $\pi' = s_0 \xrightarrow{\kappa_0} \dots \xrightarrow{\kappa_{n-1}} s_n$ be the largest prefix of π such that $s_n \in G$. Notice that $|\pi'|_{\mathrm{ds}} \le \mathrm{di}_a$ holds. Let $\pi' \xrightarrow{\kappa} s$ be the prefix of π of length $|\pi'| + 1$. We show by contradiction that $a \le T(\pi' \xrightarrow{\kappa} s) < a + \delta$ holds:

  – If $T(\pi' \xrightarrow{\kappa} s) < a$, then $\mathit{last}(\pi') \in G$ is left before time point a which contradicts $\pi \in \Diamond^I G$.
  – Further, assume that $T(\pi' \xrightarrow{\kappa} s) \ge a + \delta$. With Lemma 5 we obtain

$$\begin{aligned} t(\kappa) &\ge a + \delta - T(\pi') \\ &\ge a + \delta - |\pi'|_{\mathrm{ds}} \cdot \delta \\ &\ge (\mathrm{di}_a + 1 - \underbrace{|\pi'|_{\mathrm{ds}}}_{\le \mathrm{di}_a}) \cdot \delta > 0\,. \end{aligned}$$

  Hence, π stays at $\mathit{last}(\pi')$ for at least $(j + 1 - |\pi'|_{\mathrm{ds}}) \cdot \delta$ time units which means that $\mathrm{di}(\pi')\big(\xrightarrow{\perp} \mathit{last}(\pi')\big)^{j+1-|\pi'|_{\mathrm{ds}}} = \bar{\pi}$ is a prefix of $\mathrm{di}(\pi)$. Since $|\bar{\pi}|_{\mathrm{ds}} = j + 1$, this contradicts $\pi \notin [\Diamond_{\mathrm{ds}}^{\mathrm{di}(I)} G]$.

We infer that π takes at least one transition in the time interval $[a, a + \delta)$. The probability for this can be upper bounded by $1 - e^{-\lambda\delta}$, i.e.,

$$\begin{aligned} &\Pr_\sigma^{\mathcal{M}}(\Diamond^I G \setminus [\Diamond_{\mathrm{ds}}^{\mathrm{di}(I)} G]) \\ &\quad \le \Pr_\sigma^{\mathcal{M}}(\{\pi \in \mathit{IPaths}^{\mathcal{M}} \mid \pi \text{ takes a transition in time interval } [a, a + \delta)\}) \\ &\quad \le 1 - e^{-\lambda\delta} = \varepsilon^{\uparrow}(I). \end{aligned}$$

<u>*Case* $I = [a, b]$ *for* $a = \mathrm{di}_a \delta$ *and* $b = \mathrm{di}_b \delta$:</u> We have $\mathrm{di}(I) = \{\mathrm{di}_a + 1, \mathrm{di}_a + 2, \dots, \mathrm{di}_b\}$.

– As in the case "$I = [a, \infty)$", we show that $[\Diamond_{\mathrm{ds}}^{\mathrm{di}(I)} G] \setminus \Diamond^I G \subseteq \#[a]^{>\mathrm{di}_a}$. With Lemma 6 we obtain

$$\Pr_\sigma^{\mathcal{M}}([\Diamond_{\mathrm{ds}}^{\mathrm{di}(I)} G] \setminus \Diamond^I G) \le \Pr_\sigma^{\mathcal{M}}(\#[a]^{>\mathrm{di}_a}) \le 1 - (1 + \lambda\delta)^{\mathrm{di}_a} \cdot e^{-\lambda a} = \varepsilon^{\downarrow}(I).$$

Let $\pi \in [\Diamond_{\mathrm{ds}}^{\mathrm{di}(I)} G] \setminus \Diamond^I G$ and let $\bar{\pi}$ be the largest prefix of $\mathrm{di}(\pi)$ with $\mathit{last}(\bar{\pi}) \in G$ and $|\bar{\pi}|_{\mathrm{ds}} \in \mathrm{di}(I)$. Such a prefix exists due to $\pi \in [\Diamond_{\mathrm{ds}}^{\mathrm{di}(I)} G]$. π reaches $\mathit{last}(\bar{\pi})$ with at most di_b digitization steps and therefore within at most b time units (cf. Lemma 5). As $\pi \notin \Diamond^I G$, we conclude that π has to reach (and leave) $last(\bar{\pi})$ within less than a time units. It follows that $|pref_T(\pi, a)|_{\mathrm{ds}} \geq |\bar{\pi}|_{\mathrm{ds}} > \mathrm{di}_a$ which implies $\pi \in \#[a]^{>\mathrm{di}_a}$.

– Next, let $\pi \in \Diamond^I G \backslash [\Diamond_{\mathrm{ds}}^{\mathrm{di}(I)} G]$ and let $\pi' = s_0 \xrightarrow{\kappa_0} \ldots \xrightarrow{\kappa_{n-1}} s_n$ be the largest prefix of π such that $s_n \in G$ and $T(\pi') \leq b$. Such a prefix exists due to $\pi \in \Diamond^I G$. We distinguish two cases.

  – If $|\pi'|_{\mathrm{ds}} > \mathrm{di}_b$, then $\pi \in \#[b]^{>\mathrm{di}_b}$ since $|pref_T(\pi, b)|_{\mathrm{ds}} \geq |\pi'|_{\mathrm{ds}} > \mathrm{di}_b$.
  – If $|\pi'|_{\mathrm{ds}} \leq \mathrm{di}_b$, then $|\pi'|_{\mathrm{ds}} \leq \mathrm{di}_a$ holds due to $\pi \notin [\Diamond_{\mathrm{ds}}^{\mathrm{di}(I)} G]$. Similar to the case "$I = [a, \infty)$" we can show that π takes at least one transition in time interval $[a, a+\delta)$.

It follows that

$$\Diamond^I G \backslash [\Diamond_{\mathrm{ds}}^{\mathrm{di}(I)} G] \subseteq \#[b]^{>\mathrm{di}_b} \cup \{\pi \in IPaths^{\mathcal{M}} \mid \pi \text{ takes a transition in time interval } [a, a+\delta)\}$$

Hence,

$$\Pr_\sigma^{\mathcal{M}}(\Diamond^I G \backslash [\Diamond_{\mathrm{ds}}^{\mathrm{di}(I)} G]) \leq 1 - (1+\lambda\delta)^{\mathrm{di}_b} \cdot e^{-\lambda b} + 1 - e^{-\lambda\delta} = \varepsilon^{\uparrow}(I).$$

D Comparison to single-objective analysis

We remark that the proof in [30, Theorem 5.3] can not be adapted to show our result. The main reason is that the proof relies on an auxiliary lemma which claims that[8]

$$\Pr_\sigma^{\mathcal{M}}(\Diamond^{[0,b]} G \mid \#[\delta]^{<2}) \leq \Pr_\sigma^{\mathcal{M}}(\Diamond^{[0,b]} G) \tag{11}$$

holds for *all* schedulers $\sigma \in \mathrm{GM}^{\mathcal{M}}$. We show that this claim does *not* hold. The intuition is as follows. Assume we observe that at most one Markovian transition is taken in $\mathcal{M}$ within the first δ time units (i.e., we observe a path in $\#[\delta]^{<2}$). The lemma claims that under this observation the probability to reach G within b time units does not increase. We give a counterexample to illustrate that there are schedulers for which this is not true. Consider the MA $\mathcal{M}$ from Fig. 12 and let σ be the scheduler for $\mathcal{M}$ satisfying

$$\sigma(s_0 \xrightarrow{t_1} s_1 \xrightarrow{t_2} s_2, \alpha) = \begin{cases} 1 & \text{if } t_1 + t_2 > \delta \\ 0 & \text{otherwise.} \end{cases}$$

Hence, σ chooses α iff there are less than two digitization steps within the first δ time units. It follows that the probability to reach $G = \{s_3\}$ on a path in $\#[\delta]^{\geq 2}$ is zero. We conclude that

$$\begin{aligned}\Pr_\sigma^{\mathcal{M}}(\Diamond^{[0,b]}\{s_3\}) &= \Pr_\sigma^{\mathcal{M}}(\Diamond^{[0,b]}\{s_3\} \cap \#[\delta]^{<2}) + \underbrace{\Pr_\sigma^{\mathcal{M}}(\Diamond^{[0,b]}\{s_3\} \cap \#[\delta]^{\geq 2})}_{=0} \\ &= \Pr_\sigma^{\mathcal{M}}(\Diamond^{[0,b]}\{s_3\} \mid \#[\delta]^{<2}) \cdot \underbrace{\Pr_\sigma^{\mathcal{M}}(\#[\delta]^{<2})}_{<1} \\ &< \Pr_\sigma^{\mathcal{M}}(\Diamond^{[0,b]}\{s_3\} \mid \#[\delta]^{<2})\end{aligned}$$

which contradicts Eq. 11.

[8] We adapt [30, Lemma G.2] to our notations from "Appendix C.4".

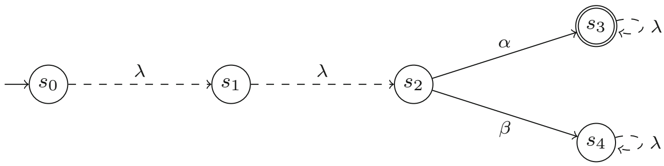


Fig. 12 MA $\mathcal{M}$ (cf. "Appendix D")

Table 2 Additional model details

N(-K)	#states	#choices	#transitions	#MS	λ_{max}
Job scheduling					
10-2	1×10^4	2×10^4	3×10^4	1×10^4	5.7
12-3	1×10^5	2×10^5	5×10^5	1×10^5	8.5
17-2	5×10^6	9×10^6	1×10^7	4×10^6	5.9
Polling					
3-2	990	1762	2387	508	14
3-3	9522	2×10^4	2×10^4	4801	14
4-4	8×10^5	2×10^6	2×10^6	5×10^5	16
Stream					
30	1426	1861	2731	931	8
250	9×10^4	1×10^5	2×10^5	6×10^4	8
1000	2×10^6	2×10^6	3×10^6	1×10^6	8
Mutex					
2	1×10^4	2×10^4	3×10^4	216	2
3	3×10^4	7×10^4	8×10^4	729	3

E Further details for the experiments

E.1 Benchmark details

We provide additional information regarding our experiments on multi-objective MAs. Table 2 provides details of the considered MA. We further describe the considered case studies and objectives.

Job scheduling The job scheduling case study originates from [12] and was already discussed in Sect. 1. We consider N jobs that are executed on K identical processors. Each of the N jobs gets a different rate between 1 and 3. We consider the following objectives.

$\mathbb{E}_1$: Minimize the expected time until all jobs are completed.
$\mathbb{E}_2$: Minimize the expected time until $\lceil \frac{N}{2} \rceil$ jobs are completed.
$\mathbb{E}_3$: Minimize the expected waiting time of the jobs.
$\mathbb{P}$: Minimize the probability that the job with the lowest rate is completed before the job with the highest rate.
$\mathbb{P}_1^{\leq}$: Maximize the probability that all jobs are completed within $\frac{N}{2K}$ time units.
$\mathbb{P}_2^{\leq}$: Maximize the probability that $\lceil \frac{N}{2} \rceil$ jobs are completed within $\frac{N}{4K}$ time units.

The objectives have been combined as follows: ($\mathbb{O}^i$ refers to the objectives considered in Column i of Table 1):

$$\mathbb{O}^1 = (\mathbb{E}_1, \mathbb{E}_2, \mathbb{E}_3) \quad \mathbb{O}^2 = (\mathbb{E}_1, \mathbb{P}_2^{\leq}) \quad \mathbb{O}^3 = (\mathbb{P}, \mathbb{E}_1, \mathbb{E}_2, \mathbb{E}_3) \quad \mathbb{O}^4 = (\mathbb{P}, \mathbb{E}_3, \mathbb{P}_1^{\leq}, \mathbb{P}_2^{\leq})$$

Polling The polling system is based on [48,50]. It considers two stations, each having a separate queue storing up to K jobs of N different types. The jobs arrive at Station i (for $i \in \{1, 2\}$) with some rate λ_i as long as the queue of the station is not full. A server polls the two stations and processes the jobs by (nondeterministically) taking a job from a non-empty queue. The time for processing a job is given by a rate which depends on the type of the job. Erasing a job from a queue is unreliable, i.e., there is a 10 % chance that an already processed job stays in the queue. For $i \in \{1, 2\}$ we assume the following objectives:

$\mathbb{E}_i$: Maximize the expected number of processed jobs of Station i until its queue is full.
$\mathbb{E}_{2+i}$: Minimize the expected sum of all waiting times of the jobs arriving at Station i until the queue of Station i is full.
$\mathbb{P}_i^{\leq}$: Minimize the probability that the queue of Station i is full within two time units.

The objectives have been combined as follows: ($\mathbb{O}^i$ refers to the objectives considered in Column i of Table 1):

$$\mathbb{O}^1 = (\mathbb{E}_1, \mathbb{E}_2) \quad \mathbb{O}^2 = (\mathbb{E}_1, \mathbb{E}_2, \mathbb{E}_3, \mathbb{E}_4) \quad \mathbb{O}^3 = (\mathbb{P}_1^{\leq}, \mathbb{P}_2^{\leq}) \quad \mathbb{O}^4 = (\mathbb{E}_1, \mathbb{E}_2, \mathbb{P}_1^{\leq}, \mathbb{P}_2^{\leq})$$

Stream This case study considers a client of a video streaming platform. The client consecutively receives N data packages and stores them into a buffer. The buffered packages are processed during the playback of the video. The time it takes to receive (or to process) a single package is modeled by an exponentially distributed delay. Whenever a package is received and the video is not playing, the client nondeterministically chooses whether it starts the playback or whether it keeps on buffering. The latter choice is not reliable, i.e., there is a 1 % chance that the playback is started anyway. In case of a buffer underrun[9], the playback is paused and the client waits for new packages to arrive. We analyzed the following objectives:

$\mathbb{E}_1$: Minimize the expected buffering time until the playback is finished.
$\mathbb{E}_2$: Minimize the expected number of buffer underruns during the playback.
$\mathbb{E}_3$: Minimize the expected time to start the playback.
$\mathbb{P}_1^{\leq}$: Minimize the probability for a buffer underrun within 2 time units.
$\mathbb{P}_2^{\leq}$: Maximize the probability that the playback starts within 0.5 time units.

The objectives have been combined as follows: ($\mathbb{O}^i$ refers to the objectives considered in Column i of Table 1):

$$\mathbb{O}^1 = (\mathbb{E}_1, \mathbb{E}_2) \quad \mathbb{O}^2 = (\mathbb{E}_3, \mathbb{P}_1^{\leq}) \quad \mathbb{O}^3 = (\mathbb{P}_1^{\leq}, \mathbb{P}_2^{\leq}) \quad \mathbb{O}^4 = (\mathbb{E}_1, \mathbb{E}_3, \mathbb{P}_1^{\leq})$$

Mutex This case study regards a randomized mutual exclusion protocol based on [42,50]. Three processes nondeterministically choose a job for which they need to enter the critical section. The amount of time a process spends in its critical section is given by a rate which depends on the chosen job. There are N different types of jobs. For each $i \in \{1, 2, 3\}$ the following objective are considered:

[9] A buffer underrun occurs when the next package needs to be processed while the buffer is empty.

Table 3 Results for our implementation (Storm) and PRISM on the multi-objective MDP benchmarks from [28]. All run-times are in seconds

Benchmark			PRISM				Storm			
Instance	#states	($\Diamond$, ER, $\leq$)	Pts	Iter	Verif	Total	Pts	Iter	Verif	Total
Consensus										
2-3-2	691	(2,0,0)	4	< 0.01	0.12	1.25	2	0.03	0.06	1.74
2-4-2	1517	(2,0,0)	4	0.01	0.16	1.26	2	0.02	0.03	0.14
2-5-2	3169	(2,0,0)	4	0.02	0.25	1.25	2	0.03	0.03	0.15
3-3-2	2×10^4	(2,0,0)	4	0.10	0.73	1.28	2	0.06	0.10	0.31
3-4-2	6×10^4	(2,0,0)	4	0.35	2.13	2.87	2	0.16	0.29	0.87
3-5-2	2×10^5	(2,0,0)	4	1.08	4.84	5.8	2	0.48	0.89	2.56
dpm										
100	636	(0,0,2)	12	0.07	0.09	1.29	6	0.07	0.07	0.18
200	636	(0,0,2)	8	0.08	0.09	1.23	4	0.10	0.10	0.20
300	636	(0,0,2)	6	0.08	0.10	1.23	3	0.11	0.11	0.21
Scheduler										
05	3×10^4	(0,2,0)	No convergence				∞ reward detected			
25	6×10^5	(0,2,0)	No convergence				∞ reward detected			
50	2×10^6	(0,2,0)	No convergence				∞ reward detected			
Team										
3	1×10^4	(1,1,0)	6	0.10	5.92	6.61	3	0.07	0.08	0.29
4	1×10^5	(1,1,0)	6	1.13	136	138	3	0.47	0.56	1.22
5	9×10^5	(1,1,0)	6	5.9	3814	3818	3	9.32	10.6	17.1
3	1×10^4	(2,1,0)	>2 objectives not supported				6	0.17	0.18	0.35
4	1×10^5	(2,1,0)	>2 objectives not supported				6	1.19	1.32	1.97
5	9×10^5	(2,1,0)	>2 objectives not supported				6	20.6	22.4	32.2
Zeroconf										
4	5449	(2,0,0)	4	0.05	2.66	3.22	2	0.12	0.14	0.28
6	1×10^4	(2,0,0)	4	0.16	5.87	6.55	2	0.33	0.36	0.76
8	2×10^4	(2,0,0)	4	0.14	9.71	10.5	2	0.78	0.83	1.09
Zeroconf-tb										
2-14	3×10^4	(2,0,0)	2	0.08	21	21.6	2	0.47	0.54	0.84
4-10	2×10^4	(2,0,0)	2	0.12	22.3	23	2	1.02	1.06	1.29
4-14	4×10^4	(2,0,0)	2	0.18	55.5	56.5	1	3.07	3.17	3.54

$\mathbb{P}_i^{\leq}$: Maximize the probability that Process i enters its critical section within 0.5 time units.
$\mathbb{P}_{3+i}^{\leq}$: Maximize the probability that Process i enters its critical section within 1 time unit.

The objectives have been combined as follows: ($\mathbb{O}^i$ refers to the objectives considered in Column i of Table 1):

$$\mathbb{O}^1 = (\mathbb{P}_1^{\leq}, \mathbb{P}_2^{\leq}, \mathbb{P}_3^{\leq}) \quad \mathbb{O}^2 = (\mathbb{P}_4^{\leq}, \mathbb{P}_5^{\leq}, \mathbb{P}_6^{\leq})$$

E.2 Comparison with PRISM

The detailed results of our experiments with PRISM are given in Table 3. We depict the different benchmark instances with the number of states of the MDP (Column *#states*) and the considered combination of objectives ($\Diamond$ represents an (untimed) probabilistic objective, ER an expected reward objective, and $\leq$ a step-bounded reward objective). We list the number of vertices of the obtained under-approximation (Column *pts*). Column *iter* lists the time required for the iterative exploration of the set of achievable points as described in [28]. In Column *verif* we depict the verification time—including the time for the iterations as well as the conducted preprocessing steps. Column *total* indicates the total runtime of the tool which includes model building time and verification time.

During our experiments we observed that PRISM does not detect that both objectives considered for the **scheduler**-instances yield infinite rewards under every possible resolution of non-determinism. As a result, the value iteration-based procedure does not converge and PRISM reports that the maximal number of iterations are exceeded. Storm detects this issue and shows a proper warning to the user.

We further note that PRISM can not compute Pareto curves for more than two objectives. However, it can answer achievability- and numerical queries as introduced in [28] with three or more objectives.

Table 4 Results for our implementation (Storm) and IMCA for single-objective MAs. All run-times are in seconds

Benchmark			IMCA	Storm (multi)	Storm (single)
N(-K)	# states	$\mathbb{O}$	Verif. time	Verif. time	Verif. time
Job scheduling					
10-2	1×10^4	$\mathbb{E}_1$	< 0.01	0.05	0.02
10-2	1×10^4	$\mathbb{P}_2^{\leq}$	4.02	5.94	0.02
12-3	1×10^5	$\mathbb{E}_1$	0.05	0.64	0.24
12-3	1×10^5	$\mathbb{P}_2^{\leq}$	62.1	111	0.29
Polling					
3-3	9522	$\mathbb{E}_1$	1.9	0.03	0.01
3-3	9522	$\mathbb{P}_1^{\leq}$	81.1	53.3	0.05
4-4	8×10^5	$\mathbb{E}_1$	966	5.85	3.2
4-4	8×10^5	$\mathbb{P}_1^{\leq}$	TO	6513	4.66
Stream					
30	1426	$\mathbb{E}_1$	< 0.01	< 0.01	< 0.01
30	1426	$\mathbb{P}_1^{\leq}$	1.85	1.42	< 0.01
250	9×10^4	$\mathbb{E}_1$	0.84	2.14	0.14
250	9×10^4	$\mathbb{P}_1^{\leq}$	129	91.2	0.13
Mutex					
2	1×10^4	$\mathbb{P}_1^{\leq}$	8.19	3.92	0.51
2	1×10^4	$\mathbb{P}_4^{\leq}$	32.8	15.6	0.71

E.3 Comparison with IMCA

The resulting verification times are given in Table 4. We depict the different benchmark instances with the number of states of the MA (Column *#states*) and the considered objective (as discussed in App. E.1). Besides the run-times of IMCA, we depict the run-times of our implementation (effectively performing multi-objective model checking with only one objective) in Column Storm (multi). Column Storm (single) shows the run-times obtained when Storm is invoked with standard (single-objective) model checking methods. The latter uses the more recent *Unif*+ algorithm [15].

Acknowledgements We thank the anonymous reviewers for their detailed feedback on an earlier version of this draft. This work has been supported by the DFG RTG 2236 "UnRAVeL". S. Junges was supported in part by NSF grants 1545126 (VeHICaL) and 1646208, by the DARPA Assured Autonomy program and the DARPA SDCPS Program (Contract FA8750-20-C-0156), by Berkeley Deep Drive, and by Toyota under the iCyPhy center.

References

1. Ash RB, Doléans-Dade C (2000) Probability and measure theory. Academic Press, New York
2. Ashok P, Butkova Y, Hermanns H, Kretínský J (2018) Continuous-time Markov decisions based on partial exploration. In: ATVA, LNCS, vol 11138. Springer, pp 317–334
3. Ashok P, Chatterjee K, Daca P, Kretínský J, Meggendorfer T (2017) Value iteration for long-run average reward in Markov decision processes. In: CAV (1), LNCS, vol 10426. Springer, pp 201–221
4. Baier C, Bertrand N, Dubslaff C, Gburek D, Sankur O (2018) Stochastic shortest paths and weight-bounded properties in Markov decision processes. In: LICS. ACM, pp 86–94
5. Baier C, Dubslaff C, Klüppelholz S (2014) Trade-off analysis meets probabilistic model checking. In: CSL-LICS. ACM, pp 1:1–1:10
6. Baier C, Klein J, Leuschner L, Parker D, Wunderlich S (2017) Ensuring the reliability of your model checker: interval iteration for Markov decision processes. In: CAV (1), LNCS, vol 10426. Springer, pp 160–180
7. Basset N, Kwiatkowska MZ, Topcu U, Wiltsche C (2015) Strategy synthesis for stochastic games with multiple long-run objectives. In: Proceedings of of TACAS, LNCS, vol 9035. Springer, pp 256–271
8. Boudali H, Crouzen P, Stoelinga M (2010) A rigorous, compositional, and extensible framework for dynamic fault tree analysis. IEEE Trans Dependable Secur Comput 7(2):128–143
9. Bozzano M, Cimatti A, Katoen JP, Nguyen VY, Noll T, Roveri M (2011) Safety, dependability and performance analysis of extended AADL models. Comput J 54(5):754–775
10. Brázdil T, Brozek V, Chatterjee K, Forejt V, Kucera A (2014) Markov decision processes with multiple long-run average objectives. LMCS 10(1):1156
11. Brázdil T, Chatterjee K, Forejt V, Kucera A (2017) Trading performance for stability in Markov decision processes. J Comput Syst Sci 84:144–170
12. Bruno JL, Downey PJ, Frederickson GN (1981) Sequencing tasks with exponential service times to minimize the expected flow time or makespan. J ACM 28(1):100–113
13. Bruyère V, Filiot E, Randour M, Raskin JF (2014) Meet your expectations with guarantees: beyond worst-case synthesis in quantitative games. In: Proceedings of STACS, LIPIcs, vol 25. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, pp 199–213
14. Butkova Y, Fox G (2019) Optimal time-bounded reachability analysis for concurrent systems. In: TACAS (2), LNCS, vol 11428. Springer, pp 191–208
15. Butkova Y, Hatefi H, Hermanns H, Krcál J (2015) Optimal continuous time Markov decisions. In: ATVA, LNCS, vol 9364. Springer, pp 166–182
16. Butkova Y, Wimmer R, Hermanns H (2017) Long-run rewards for Markov automata. In: Proceedings of TACAS, LNCS. Springer https://link.springer.com/chapter/10.1007%2F978-3-662-54580-5_11
17. Chatterjee K, Henzinger M (2011) Faster and dynamic algorithms for maximal end-component decomposition and related graph problems in probabilistic verification. In: Proceedings of SODA, pp 1318–1336
18. Chen T, Forejt V, Kwiatkowska MZ, Simaitis A, Wiltsche C (2013) On stochastic games with multiple objectives. In: Proceedings of MFCS, LNCS, vol 8087. Springer, pp 266–277

19. Coste N, Hermanns H, Lantreibecq E, Serwe W (2009) Towards performance prediction of compositional models in industrial GALS designs. In: Proceedings of CAV, LNCS, vol 5643. Springer, pp 204–218
20. David A, Jensen PG, Larsen KG, Legay A, Lime D, Sørensen MG, Taankvist JH (2014) On time with minimal expected cost! In: Proceedings of ATVA, LNCS, vol 8837. Springer, pp 129–145
21. de Alfaro L (1997) Formal verification of probabilistic systems. Ph.D. thesis, Stanford University
22. Dehnert C, Junges S, Katoen JP, Volk M (2017) A storm is coming: a modern probabilistic model checker. In: Proceedings of CAV
23. Delgrange F, Katoen J, Quatmann T, Randour M (2020) Simple strategies in multi-objective MDPs. In: TACAS (1), LNCS, vol 12078. Springer, pp 346–364
24. Deng Y, Hennessy M (2013) On the semantics of Markov automata. Inf Comput 222:139–168
25. Eisentraut C, Hermanns H, Katoen JP, Zhang L (2013) A semantics for every GSPN. In: Petri Nets, LNCS, vol 7927. Springer, pp 90–109
26. Eisentraut C, Hermanns H, Zhang L (2010) On probabilistic automata in continuous time. In: Proceedings of LICS. IEEE CS, pp 342–351
27. Etessami K, Kwiatkowska MZ, Vardi MY, Yannakakis M (2008) Multi-objective model checking of Markov decision processes. LMCS 4(4):1–21
28. Forejt V, Kwiatkowska M, Parker D (2012) Pareto curves for probabilistic model checking. In: Proceedings of ATVA, LNCS, vol 7561. Springer, pp 317–332
29. Forějt V, Kwiatkowska MZ, Norman G, Parker D, Qu H ()20119 Quantitative multi-objective verification for probabilistic systems. In: Proceedings of TACAS, LNCS, vol 6605. Springer, pp 112–127
30. Guck D, Hatefi H, Hermanns H, Katoen JP, Timmer M (2014) Analysis of timed and long-run objectives for Markov automata. LMCS 10(3):943
31. Guck D, Timmer M, Hatefi H, Ruijters E, Stoelinga M (2014) Modelling and analysis of Markov reward automata. In: Proceedings of ATVA, LNCS, vol 8837. Springer, pp 168–184
32. Haddad S, Monmege B (2014) Reachability in MDPs: refining convergence of value iteration. In: RP, LNCS, vol 8762. Springer, pp 125–137
33. Hartmanns A, Junges S, Katoen J, Quatmann T (2020) Multi-cost bounded tradeoff analysis in MDP. J Autom Reason 64(7):1483–1522
34. Hatefi H, Braitling B, Wimmer R, Fioriti LMF, Hermanns H, Becker B (2015) Cost vs. time in stochastic games and Markov automata. In: Proceedings of SETTA, LNCS, vol 9409. Springer, pp 19–34
35. Hatefi H, Hermanns H (2012) Model checking algorithms for Markov automata. ECEASST, vol 53. http://journal.ub.tu-berlin.de/eceasst/article/view/783
36. Hensel C, Junges S, Katoen J, Quatmann T, Volk M (2020) The probabilistic model checker Storm. CoRR abs/2002.07080
37. Junges S, Jansen N, Dehnert C, Topcu U, Katoen JP (2016) Safety-constrained reinforcement learning for MDPs. In: Proceedings of TACAS, LNCS, vol 9636. Springer, pp 130–146
38. Katoen JP, Wu H (2016) Probabilistic model checking for uncertain scenario-aware data flow. ACM Trans Embedded Comput Syst 22(1):15:1–15:27
39. Kwiatkowska M, Norman G, Parker D (2011) PRISM 4.0: verification of probabilistic real-time systems. In: Proceedings of CAV, LNCS, vol 6806. Springer, pp 585–591
40. Neuhäußer MR (2010) Model checking nondeterministic and randomly timed systems. Ph.D. thesis, RWTH Aachen University
41. Neuhäußer MR, Stoelinga M, Katoen JP (2009) Delayed nondeterminism in continuous-time Markov decision processes. In: Proceedings of FOSSACS, LNCS, vol 5504. Springer, pp 364–379
42. Pnueli A, Zuck L (1986) Verification of multiprocess probabilistic protocols. Distrib. Comput. 1(1):53–72
43. Puterman ML (1994) Markov decision processes: discrete stochastic dynamic programming. Wiley, New York
44. Quatmann T, Junges S, Katoen J (2017) Markov automata with multiple objectives. In: CAV (1), LNCS, vol 10426. Springer, pp 140–159
45. Quatmann T, Junges S, Katoen J (2020) Markov automata with multiple objectives: supplemental material. Zenodo. https://doi.org/10.5281/zenodo.4298642
46. Randour M, Raskin JF, Sankur O (2015) Variations on the stochastic shortest path problem. In: Proceedings of VMCAI, LNCS, vol 8931. Springer, pp 1–18
47. Roijers DM, Vamplew P, Whiteson S, Dazeley R (2013) A survey of multi-objective sequential decision-making. J Artif Intell Res 48:67–113
48. Srinivasan MM (1991) Nondeterministic polling systems. Manag Sci 37(6):667–681
49. Teichteil-Königsbuch F (2012) Path-constrained Markov decision processes: bridging the gap between probabilistic model-checking and decision-theoretic planning. In: Proceedings of ECAI, frontiers in AI and applications, vol 242. IOS Press, pp 744–749

50. Timmer M, Katoen JP, van de Pol J, Stoelinga M (2012) Efficient modelling and generation of Markov automata. In: Proceedings of CONCUR, LNCS, vol 7454. Springer, pp 364–379
51. Volk M, Junges S, Katoen J (2018) Fast dynamic fault tree analysis by model checking techniques. IEEE Trans Ind Inform 14(1):370–379