Taming Infinity one Chunk at a Time: Concisely Represented Strategies in One-Counter MDPs

Michal Ajdarów
Masaryk University, Czech Republic

James C. A. Main
F.R.S.-FNRS and UMONS – Université de Mons, Belgium

Petr Novotný
Masaryk University, Czech Republic

Mickael Randour
F.R.S.-FNRS and UMONS – Université de Mons, Belgium

Abstract

Markov decision processes (MDPs) are a canonical model to reason about decision making within a stochastic environment. We study a fundamental class of *infinite* MDPs: one-counter MDPs (OC-MDPs). They extend finite MDPs via an associated counter taking natural values, thus inducing an infinite MDP over the set of configurations (current state and counter value). We consider two characteristic objectives: reaching a target state (state-reachability), and reaching a target state with counter value zero (selective termination). The synthesis problem for the latter is not known to be decidable and connected to major open problems in number theory. Furthermore, even seemingly simple strategies (e.g., memoryless ones) in OC-MDPs might be impossible to build in practice (due to the underlying infinite configuration space): we need finite, and preferably small, representations.

To overcome these obstacles, we introduce two natural classes of *concisely represented* strategies based on a (possibly infinite) partition of counter values in intervals. For both classes, and both objectives, we study the verification problem (does a given strategy ensure a high enough probability for the objective?), and two synthesis problems (does there exist such a strategy?): one where the interval partition is fixed as input, and one where it is only parameterized. We develop a generic approach based on a compression of the induced infinite MDP that yields decidability in all cases, with all complexities within PSPACE.

2012 ACM Subject Classification Theory of computation → Probabilistic computation; Theory of computation → Logic and verification

Keywords and phrases one-counter Markov decision processes, randomised strategies, termination, reachability

Funding This work has been supported by the Fonds de la Recherche Scientifique - FNRS under Grant n° T.0188.23 (PDR ControlleRS).
James C. A. Main: Research Fellow of the Fonds de la Recherche Scientifique – FNRS and member of the TRAIL institute.
Mickael Randour: Research Associate of the Fonds de la Recherche Scientifique – FNRS and member of the TRAIL institute.

Acknowledgements We thank Sougata Bose for pointing out the reference [5].

1 Introduction

A *Markov decision process* (MDP) models the continuous interaction of a controllable system with a stochastic environment. MDPs are used notably in the fields of formal methods (e.g., [3, 29]) and reinforcement learning (e.g., [46]). In each step of an MDP, the system selects an action that is available in the current state, and the state is updated randomly following a distribution depending only on the current state and chosen action. This interaction goes on

forever and yields a *play* (i.e., an execution) of the MDP. The resolution of non-determinism in an MDP is done via a *strategy*: strategies prescribe choices depending on the current history of the play and may use randomisation. Strategies are the formal counterpart of controllers for the modelled system [42, 9]. A strategy is *pure* if it does not use randomisation.

A classical problem in MDPs is *synthesis*, which asks to construct a good strategy with respect to a specification. Specifications can be defined, e.g., via *payoff functions* that quantify the quality of plays, and via *objectives*, which are sets of well-behaved plays. The goal is thus to synthesise a strategy that, in the former case, has a high expected payoff, and, in the latter case, enforces the objective with high probability. If possible, it is desirable to construct an *optimal strategy*, i.e., a strategy whose performance is not (strictly) exceeded by another. The decision variant of the synthesis problem, which asks whether a well-performing strategy exists, is called *realisability*. We focus on variants of *reachability objectives*, asking that a target set of states be visited. These canonical objectives are central in synthesis [21].

Strategies and their representations. Traditionally, strategies are represented by *Mealy machines*, i.e., finite automata with outputs (see, e.g., [12, 37]). Strategies that admit such a representation are called *finite-memory strategies*. A special subclass is that of *memoryless strategies*, i.e., strategies whose decisions depend only on the current state.

In finite MDPs, pure memoryless strategies often suffice to play optimally with a single payoff or objective, e.g., for reachability [3] and parity objectives [23], and mean [7] and discounted-sum payoffs [45]. In countable MDPs, for reachability objectives, pure memoryless strategies are also as powerful as general ones, though optimal strategies need not exist in general [39, 34]. In the multi-objective setting, memory and randomisation are necessary in general (e.g., [43, 21]). Nonetheless, finite memory often suffices in finite MDPs, e.g., when considering several ω-regular objectives [27].

The picture is more complicated in infinite-state MDPs, where even pure memoryless strategies need not admit finite representations. Our contribution focuses on small (and particularly, finite) counter-based representations of memoryless strategies in a fundamental class of infinite-state MDPs: *one-counter MDPs*.

One-counter Markov decision processes. *One-counter MDPs* (OC-MDPs, [16]) are finite MDPs augmented with a counter that can be incremented (by one), decremented (by one) or left unchanged on each transition.[1] Such a finite OC-MDP induces an infinite MDP over a set of *configurations* given by states of the underlying MDP and counter values. In this induced MDP, any play that reaches counter value zero is interrupted; this event is called *termination*. We consider two variants of the model: *unbounded OC-MDPs*, where counter values can grow arbitrarily large, and *bounded OC-MDPs*, in which plays are interrupted when a fixed counter upper bound is reached.

The counter in OC-MDPs can, e.g., model resource consumption along plays [16], or serve as an abstraction of unbounded data types and structures [20]. It can also model a passage of time: indeed, OC-MDPs generalise *finite-horizon MDPs*, in which a bound is imposed on the number of steps (see, e.g., [4]). OC-MDPs can be also seen as an extension of *one-counter Markov chains* with non-determinism. One-counter Markov chains are equivalent to (discrete-time) quasi-birth-death processes [28], a model studied in queuing theory.

[1] Considering such counter updates is not restrictive for modelling: any integer counter update can be obtained with several transitions. However, this impacts the complexity of decision problems.

Termination is the canonical objective in OC-MDPs [16]. Also relevant is the more general *selective termination* objective, which requires terminating in a target set of states. In this work, we study both the selective termination objective and the *state-reachability* objective, which requires visiting a target set of states regardless of the counter value.

Optimal strategies need not exist in unbounded OC-MDPs for these objectives [15]. The general synthesis problem in unbounded OC-MDPs for selective termination is not known to be decidable, and it is at least as hard as the positivity problem for linear recurrence sequences [41], whose decidability would yield a major breakthrough in number theory [40]. Optimal strategies exist in bounded OC-MDPs: the induced MDP is finite and we consider reachability objectives. However, constructing optimal strategies is already EXPTIME-hard for reachability in finite-horizon MDPs [4].

In this work, we propose to tame the inherent complexity of analysing OC-MDPs by restricting our analysis to a class of succinctly representable (yet natural and expressive) strategies called *interval strategies*.

Interval strategies. The monotonic structure of OC-MDPs makes them amenable to analysis through strategies of special structure. For instance, in their study of *solvency games* (stateless variant of OC-MDPs with binary counter updates), Berger et al. [6] identify a natural class of *rich man's strategies,* which, whenever the counter value is above some threshold, always select the same action in the same state. They argue that the question of existence of rich man's optimal strategies "*gets at a real phenomenon which . . . should be reflected in any good model of risk-averse investing.*" While [6] shows that optimal rich man's strategies do not always exist, if they do, their existence substantially simplifies the model analysis.

In this work, we make such finitely-represented strategies the primary object of study. We consider so-called *interval* strategies: each such strategy is based on some (finite or infinite but finitely-representable) partitioning of $\mathbb{N}$ into intervals, and the strategy's decision depends on the current state and on the partition containing the current counter value.

More precisely, we focus on two classes of these strategies: On the one hand, in bounded and unbounded OC-MDPs, we consider *open-ended interval strategies* (OEISs): here, the underlying partitioning is finite, and thus contains an open-ended interval $[n, \infty)$ on which the strategy behaves memorylessly, akin to rich man's strategies.

On the other hand, in unbounded OC-MDPs, we also consider *cyclic interval strategies* (CISs): strategies for which there exists a (positive integer) period such that, for any two counter values that differ by the period, we take the same decisions. A CIS can be represented similarly to an OEIS using a partition of the set of counter values up to the period.

We collectively refer to OEISs and CISs as *interval strategies*. While interval strategies are not sufficient to play optimally in unbounded OC-MDPs [6], they can be used to approximate the supremum probability for the objectives we consider [15].

Our contributions. We formally introduce the class of interval strategies. We show that OEISs in bounded OC-MDPs and CISs in unbounded OC-MDPs can be exponentially more concise than equivalent Mealy machines (Section 3), and unbounded OEISs can even represent infinite-memory strategies.

For selective termination and state reachability, we consider the interval strategy verification problem and two realisability problems for structure-constrained interval strategies. On the one hand, the *fixed-interval realisability* problem asks, given an interval partition, whether there is an interval strategy built on this partition that ensures the objective with a probability greater than a given threshold. Intuitively, in this case, the system designer

<table>
<tr><td>Semantics</td><td colspan="2">Bounded</td><td colspan="4">Unbounded</td></tr>
<tr><td>Strategy type</td><td colspan="2">Open-ended</td><td colspan="2">Open-ended</td><td colspan="2">Cyclic</td></tr>
<tr><td rowspan="2">Verification</td><td colspan="2">$\mathsf{P}^{\mathsf{PosSLP}}$
Thm. 20</td><td colspan="2">co-ETR
Thm. 24</td><td colspan="2">co-ETR
Thm. 28</td></tr>
<tr><td colspan="2">sqrt-sum-hard
Thm. 46</td><td colspan="4">sqrt-sum-hard [28]</td></tr>
<tr><td rowspan="2">Realisability (both fixed-interval and parameterised)</td><td>Pure
$\mathsf{NP}^{\mathsf{PosSLP}}$
Thm. 30</td><td>Random
$\mathsf{NP}^{\mathsf{ETR}}$
Thm. 32</td><td>Pure
$\mathsf{NP}^{\mathsf{ETR}}$
Thm. 33</td><td>Random
PSPACE
Thm. 35</td><td>Pure
$\mathsf{NP}^{\mathsf{ETR}}$
Thm. 36</td><td>Random
PSPACE
Thm. 38</td></tr>
<tr><td colspan="6">NP-hard (termination, Thm. 47) and sqrt-sum-hard ([28], Thm. 46)</td></tr>
</table>

Table 1 Complexity bound summary for our problems. All bounds are below PSPACE. Square-root-sum-hardness results are derived from instances of the form $\sum \sqrt{x_i} \geq y$.

specifies the desired structure of the controller. On the other hand, the *parameterised realisability* problem (for interval strategies), asks whether there exists a well-performing strategy built on a partition of size no more than a parameter d such that no finite interval is larger than a second parameter n. We consider two variants of the realisability problem: one for checking the existence of a suitable pure strategy and another for randomised strategies. Randomisation allows for better performance when imposing structural constraints on strategies (Example 6), but pure strategies are however often preferred for synthesis [25].

Our complexity results are summarised in Table 1. Analysing the performance of a memoryless strategy amounts to studying the Markov chain it induces on the MDP. Our results rely on the analysis of a *compressed Markov chain* derived from the (potentially infinite) Markov chain induced by an interval strategy. We remove certain configurations and aggregate several transitions into one (Section 4). This compressed Markov chain preserves the probability of selective termination and of hitting counter upper bounds (Theorem 11). However, its transition probabilities may require exponential-size representations or even be irrational. To represent these probabilities, we characterise them as the least solutions of quadratic systems of equations (Theorems 13 and 16), similarly to the result of [35] for termination probabilities in probabilistic pushdown automata. Compressed Markov chains are finite for OEISs and are induced by a one-counter Markov chain for CISs (Section 4.5).

The crux of our algorithmic results is the aforementioned compression. For verification, we reduce the problem to checking the validity of a universal formula in the theory of the reals, by exploiting our characterisation of transition probabilities in compressed Markov chains. This induces a PSPACE upper bound. For bounded OC-MDPs, we can do better: verification can be solved in polynomial time in the unit-cost arithmetic RAM model of computation of Blum, Shub and Smale [10], by computing transition and reachability probabilities of the compressed Markov chain. This yields a $\mathsf{P}^{\mathsf{PosSLP}}$ complexity in the Turing model (see [1]).

Both realisability variants exploit the verification approach through the theory of the reals. For fixed-interval realisability for pure strategies, we exploit non-determinism to select good strategies and then verify them with the above. In the randomised case, in essence, we build on the verification formulae and existentially quantify over the probabilities of actions under the sought strategy. Finally, for parameterised realisability, we build on our algorithms for the fixed-interval case by first non-deterministically building an appropriate partition.

We also provide complexity lower bounds. We show that all of our considered problems

are hard for the square-root-sum problem, a problem that is not known to be solvable in polynomial time but that is solvable in polynomial time in the Blum-Shub-Smale model [47]. We also prove NP-hardness for our realisability problems for selective termination, already when checking the existence of good single-interval strategies.

Impact. Our results provide a natural class of strategies for which realisability is decidable (whereas the general case remains open and known to be difficult [41]), and with arguably low complexity (for synthesis). Furthermore, the class of interval strategies is of practical interest thanks to their concise representation and their inherently understandable structure (in contrast to the corresponding Mealy machine representation).

Related work. In addition to the main references cited previously, we mention some (non-exhaustive) related work. The closest is [8]: interval strategies are similar to counter selector strategies, studied in *consumption* MDPs. These differ from OC-MDPs in key aspects: all transitions consume resources (i.e., bear negative weights), and recharge can only be done in special reload states, where it is considered as an atomic action up to a given capacity. Consumption and counter-based (or energy) models have different behaviors (e.g., [19]). The authors of [8] also study incomparable objectives: almost-sure (i.e., probability one) Büchi objectives.

Restricting oneself to subclasses of strategies that prove to be of practical interest is a common endeavor in synthesis: e.g., strategies represented by decision trees [17, 18, 2, 33], pure strategies [31, 25, 13], finite-memory strategies [24, 14, 13], or strategies using limited forms of randomness [37, 38].

Outline. In Section 2, we introduce all prerequisite notions. We define interval strategies and formalise our decision problems in Section 3. The compressed Markov chain construction, which is central in our verification and realisability algorithms, is presented in Section 4. Section 5 focuses on the verification problem via compressed Markov chains. We study the fixed-interval and parameterised realisability problems in Section 6. Finally, we present our complexity lower bounds in Section 7. Appendix A presents additional details used in the correctness proof of our reduction from the square-root-sum-problem to the verification problem in bounded OC-MDPs.

2 Preliminaries

Set and probability theory. We write $\mathbb{N}$, $\mathbb{Q}$ and $\mathbb{R}$ for the sets of non-negative integers, rational numbers and real numbers respectively. We let $\mathbb{N}_{>0}$ denote the set of positive integers. We let $\bar{\mathbb{N}} = \mathbb{N} \cup \{\infty\}$ and $\bar{\mathbb{N}}_{>0} = \mathbb{N}_{>0} \cup \{\infty\}$. Given $n, n' \in \bar{\mathbb{N}}$, we let $[\![n, n']\!]$ denote the set $\{k \in \mathbb{N} \mid n \leq k \leq n'\}$, and if $n = 0$, we shorten the notation to $[\![n']\!]$.

Let A be a finite or countable set. We write $\mathcal{D}(A)$ for the set of distributions over A, i.e., of functions $\mu \colon A \to [0, 1]$ such that $\sum_{a \in A} \mu(a) = 1$. Given a distribution $\mu \in \mathcal{D}(A)$, we let $\mathsf{supp}(\mu) = \{a \in A \mid \mu(a) > 0\}$ denote the support of μ.

Markov decision processes. Formally, a *Markov decision process* (MDP) is a tuple $\mathcal{M} = (S, A, \delta)$ where S is a countable set of states, A is a finite set of actions and $\delta \colon S \times A \to \mathcal{D}(S)$ is a (partial) probabilistic transition function. Given $s \in S$, we write $A(s)$ for the set of actions $a \in A$ such that $\delta(s, a)$ is defined. We say that an action a is enabled in s if $a \in A(s)$.

We require, without loss of generality, that there are no deadlocks in MDPs, i.e., that for all states s, $A(s)$ is non-empty. We say that $\mathcal{M}$ is *finite* if its state space is finite.

A *Markov chain* can be seen as an MDP with a single enabled action per state. We omit actions in Markov chains and denote them as tuples $\mathcal{C} = (S, \delta)$ with $\delta \colon S \to \mathcal{D}(S)$.

Let $\mathcal{M} = (S, A, \delta)$. A *play* of $\mathcal{M}$ is an infinite sequence $\pi = s_0 a_0 s_1 a_1 \ldots \in (SA)^\omega$ such that for all $\ell \in \mathbb{N}$, $s_{\ell+1} \in \mathsf{supp}(\delta(s_\ell, a_\ell))$. A *history* is a finite prefix of a play ending in a state. We let $\mathsf{Plays}(\mathcal{M})$ and $\mathsf{Hist}(\mathcal{M})$ respectively denote the set of plays and of histories of $\mathcal{M}$. Given a play $\pi = s_0 a_0 s_1 a_1 \ldots$ and $n \in \mathbb{N}$, we let $\pi_{|n} = s_0 a_0 \ldots a_{n-1} s_n$ and extend this notation to histories. For any given history $h = s_0 a_0 \ldots a_{n-1} s_n$, we let $\mathsf{first}(h) = s_0$ and $\mathsf{last}(h) = s_n$. Given two histories $h_1 = s_0 a_0 \ldots a_{n-1} s_\ell$ and $h_2 = s_\ell a_\ell \ldots a_{n-1} s_n$ such that $\mathsf{last}(h_1) = \mathsf{first}(h_2)$, we let $h_1 \cdot h_2 = s_0 a_0 \ldots a_{n-1} s_n$ denote their concatenation without repetition of state s_ℓ; we abusively call $h_1 \cdot h_2$ the *concatenation* of h_1 and h_2. The *cylinder* of a history $h = s_0 a_0 \ldots a_{n-1} s_n$ is the set $\mathsf{Cyl}(h) = \{\pi \in \mathsf{Plays}(\mathcal{M}) \mid \pi_{|n} = h\}$ of plays extending h.

Let $\mathcal{H}$ be a set of histories. We let $\mathsf{Cyl}(\mathcal{H}) = \bigcup_{h \in \mathcal{H}} \mathsf{Cyl}(h)$. We say that $\mathcal{H}$ is *prefix-free* if there are no distinct histories $h, h' \in \mathcal{H}$ such that h is a prefix of h'. Whenever $\mathcal{H}$ is prefix-free, the union $\mathsf{Cyl}(\mathcal{H}) = \bigcup_{h \in \mathcal{H}} \mathsf{Cyl}(h)$ is disjoint.

In a Markov chain, due to the lack of actions, plays and histories are simply sequences of states coherent with transitions. We extend the notation introduced above for plays and histories of MDPs to plays and histories of Markov chains.

Strategies. A strategy describes the action choices made throughout a play of an MDP. In general, strategies can use randomisation and all previous knowledge to make decisions. Formally, a *strategy* is a function $\sigma \colon \mathsf{Hist}(\mathcal{M}) \to \mathcal{D}(A)$ such that for any history $h \in \mathsf{Hist}(\mathcal{M})$, $\mathsf{supp}(\sigma(h)) \subseteq A(\mathsf{last}(h))$, i.e., a strategy can only prescribe actions that are available in the last state of a history. A play $\pi = s_0 a_0 s_1 \ldots$ is *consistent* with a strategy σ if for all $\ell \in \mathbb{N}$, $\sigma(\pi_{|\ell})(a_\ell) > 0$. Consistency of a history with respect to a strategy is defined similarly.

A strategy is *pure* if it does not use any randomisation, i.e., if it maps histories to Dirac distributions. We view pure strategies as functions $\sigma \colon \mathsf{Hist}(\mathcal{M}) \to A$. A strategy σ is *memoryless* if the distribution it suggests depends only on the last state of a history, i.e., if for all histories h and h', $\mathsf{last}(h) = \mathsf{last}(h')$ implies $\sigma(h) = \sigma(h')$. We view memoryless strategies and pure memoryless strategies as functions $S \to \mathcal{D}(A)$ and $S \to A$ respectively.

Fix a strategy σ and an initial state $s_{\mathsf{init}} \in S$. We obtain a purely stochastic process when following σ from s_{init}, i.e., a countable Markov chain over the set of histories with the initial state s_{init}. Rather than defining this Markov chain, we instead define a distribution, denoted by $\mathbb{P}^\sigma_{\mathcal{M}, s_{\mathsf{init}}}$, over the σ-algebra induced by cylinder sets. For all histories $h = s_0 a_0 s_1 \ldots s_n \in \mathsf{Hist}(\mathcal{M})$, if $s_0 = s_{\mathsf{init}}$, we let $\mathbb{P}^\sigma_{\mathcal{M}, s_{\mathsf{init}}}(\mathsf{Cyl}(h)) = \prod_{\ell=0}^{n-1} \sigma(h_{|\ell})(a_\ell) \cdot \delta(s_\ell, a_\ell)(s_{\ell+1})$, and otherwise, if $s_0 \neq s_{\mathsf{init}}$, we let $\mathbb{P}^\sigma_{\mathcal{M}, s_{\mathsf{init}}}(\mathsf{Cyl}(h)) = 0$. Carathéodory's extension theorem [26, Thm. A.1.3.] guarantees that this probability distribution extends in a unique fashion to the σ-algebra induced by cylinder sets. For a Markov chain $\mathcal{C}$, due to the absence of non-determinism, we drop the strategy from the notation and write $\mathbb{P}_{\mathcal{C}, s_{\mathsf{init}}}$. If the relevant MDP or Markov chain is clear from the context, we omit them from the notation above.

In the following, we focus on memoryless strategies. If σ is a memoryless strategy of $\mathcal{M}$, we consider the *Markov chain induced by σ on $\mathcal{M}$* as a Markov chain over S. Formally, it is defined as the Markov chain (S, δ') such that for all s, $s' \in S$, $\delta'(s)(s') = \sum_{a \in A(s)} \sigma(s)(a) \cdot \delta(s, a)(s')$.

Finite-memory strategies. In general, strategies can use unlimited knowledge of the past and thus may not admit a finite representation. The classical finite representation for

strategies is based on Mealy machines, i.e., finite automata with outputs. At each step of a play, an action is chosen according to the outputs of the Mealy machine and the state of the Mealy machine is updated according to the current memory state, the current MDP state and chosen action. A strategy is *finite-memory* if it can be represented by a Mealy machine.

Let $\mathcal{M} = (S, A, \delta)$ be an MDP. Formally, a *Mealy machine* is a tuple $\mathfrak{M} = (M, m_{\mathsf{init}}, \mathsf{up}, \mathsf{nxt})$ where M is a finite set of memory states, $m_{\mathsf{init}} \in M$ is an initial memory state, $\mathsf{up}\colon M \times S \times A \to M$ is a memory update function and $\mathsf{nxt}\colon M \times S \to \mathcal{D}(A)$ is a next-move function. We formalise the finite-memory strategy $\sigma^{\mathfrak{M}}$ induced by a Mealy machine $\mathfrak{M} = (M, m_{\mathsf{init}}, \mathsf{up}, \mathsf{nxt})$ as follows. First, we define the iterated update function $\widehat{\mathsf{up}}\colon (SA)^* \to M$ by induction. We let $\widehat{\mathsf{up}}(\varepsilon) = m_{\mathsf{init}}$ (where ε denotes the empty word), and let, for all $usa \in (SA)^*$, $\widehat{\mathsf{up}}(usa) = \mathsf{up}(\widehat{\mathsf{up}}(u), s, a)$. The strategy $\sigma^{\mathfrak{M}}$ is then defined, for all histories $h = us \in \mathsf{Hist}(\mathcal{M})$, by $\sigma^{\mathfrak{M}}(h) = \mathsf{nxt}(\widehat{\mathsf{up}}(u), s)$. Memoryless strategies are induced by one-state Mealy machines.

▶ Remark 1. There exist more general definitions of finite-memory strategies, e.g., where the considered Mealy machines have a randomised initialisation and a randomisation update function. However, these models are not equivalent in general (see, e.g., [37]. The model presented here is sufficient to compare interval strategies introduced in Section 3 to strategies that rely on memory instead of observing counter values.

One-counter Markov decision processes. One-counter MDPs (OC-MDPs) extend MDPs with a counter that can be incremented, decremented or left unchanged on each transition. Formally, a *one-counter MDP* is a tuple $\mathcal{Q} = (Q, A, \delta, w)$ where (Q, A, δ) is a finite MDP and $w\colon S \times A \to \{-1, 0, 1\}$ is a (partial) weight function that assigns an integer weight from $\{-1, 0, 1\}$ to state-action pairs. For all $q \in Q$ and $a \in A$, we require that $w(q, a)$ be defined whenever $a \in A(q)$. A *configuration* of $\mathcal{Q}$ is a pair (q, k) where $q \in Q$ and $k \in \mathbb{N}$. In the sequel, by plays, histories, strategies etc. of $\mathcal{Q}$, we refer to the corresponding notion with respect to the MDP (Q, A, δ) underlying $\mathcal{Q}$.

We remark that the weight function is not subject to randomisation, i.e., the weight of any transition is determined by the outgoing state and chosen action. In particular, counter values can be inferred from histories and be taken in account to make decisions in $\mathcal{Q}$.

Let $\mathcal{Q} = (Q, A, \delta, w)$ be an OC-MDP. The OC-MDP $\mathcal{Q}$ induces an MDP over the infinite countable space of configurations. Transitions in this induced MDP are defined using δ for the probability of updating the state and w for the deterministic change of counter value. We interrupt any play whenever a configuration with counter value 0 is reached. Intuitively, such configurations can be seen as situations in which we have run out of an energy resource. We may also impose an upper bound on the counter value, and interrupt any plays that reach this counter upper bound. We refer to OC-MDPs with a finite upper bound on counter values as *bounded OC-MDPs* and OC-MDPs with no upper bounds as *unbounded OC-MDP*.

We provide a unified definition for both semantics. Let $\mathcal{Q} = (Q, A, \delta, w)$ be an OC-MDP and let $r \in \bar{\mathbb{N}}_{>0}$ be an upper bound on the counter value. We define the MDP $\mathcal{M}^{\leq r}(\mathcal{Q}) = (Q \times [\![r]\!], A, \delta^{\leq r})$ where $\delta^{\leq r}$ is defined, for all configurations $s = (q, k) \in Q \times [\![r]\!]$, actions $a \in A(q)$ and states $p \in Q$, by $\delta^{\leq r}(s, a)(p, k + w(q, a)) = \delta(q, a)(p)$ if $k \notin \{0, r\}$, and $\delta^{\leq r}(s, a)(s) = 1$ otherwise. The state space of $\mathcal{M}^{\leq r}(\mathcal{Q})$ is finite if and only if $r \neq \infty$.

For technical reasons, we also introduce one-counter Markov chains. In one-counter Markov chains, we authorise stochastic counter updates, i.e., counter updates are integrated in the transition function. This contrasts with OC-MDPs where deterministic counter updates are used to allow strategies to observe counter updates. We only require the unbounded semantics in this case. Formally, a *one-counter Markov chain* is a tuple $\mathcal{P} = (Q, \delta)$ where Q is a finite set of states and $\delta\colon Q \to \mathcal{D}(Q \times \{-1, 0, 1\})$ is a probabilistic transition and counter update

function. The one-counter Markov chain $\mathcal{P}$ induces a Markov chain $\mathcal{C}^{\leq\infty}(\mathcal{P}) = (Q \times \mathbb{N}, \delta^{\leq\infty})$ such that for any configuration $s = (q, k) \in Q \times \mathbb{N}$, any $p \in Q$ and any $u \in \{-1, 0, 1\}$, we have $\delta^{\leq\infty}(s)((p, k + u)) = \delta(q)(p, u)$ if $k \neq 0$ and $\delta^{\leq\infty}(s)(s) = 1$ otherwise.

Objectives. Let $\mathcal{M} = (S, A, \delta)$ be an MDP. An *objective* is a measurable set of plays that represents a specification. We focus on variants of reachability. Given a set of target states $T \subseteq S$, the *reachability* objective for T is defined as $\mathsf{Reach}(T) = \{s_0 a_0 s_1 a_1 \ldots \in \mathsf{Plays}(\mathcal{M}) \mid \exists \ell \in \mathbb{N},\, s_\ell \in T\}$. If $T = \{s\}$, we write $\mathsf{Reach}(s)$ for $\mathsf{Reach}(\{s\})$.

Let $\mathcal{Q} = (Q, A, \delta, w)$ be an OC-MDP, $r \in \bar{\mathbb{N}}_{>0}$ be a counter upper bound and $T \subseteq Q$. We consider two variants of reachability on $\mathcal{M}^{\leq r}(\mathcal{Q})$. The first goal we consider requires visiting T regardless of the counter value and the second requires visiting T with a counter value of zero. The first objective is called *state-reachability* and we denote it by $\mathsf{Reach}(T)$ instead of $\mathsf{Reach}(T \times [\![r]\!])$. The second objective is called *selective termination*, which is formalised as $\mathsf{Term}(T) = \mathsf{Reach}(T \times \{0\})$. Like above, if $T = \{q\}$, we write $\mathsf{Reach}(q)$ and $\mathsf{Term}(q)$ instead of $\mathsf{Reach}(\{q\})$ and $\mathsf{Term}(\{q\})$.

Reachability probabilities in Markov chains. Let $\mathcal{C} = (S, \delta)$ be a finite Markov chain and let $T \subseteq S$ be a set of targets. The probabilities $(\mathbb{P}_s(\mathsf{Reach}(T)))_{s \in S}$ are the least solution of a system of linear equations (e.g., [3, Thm. 10.15]). We recall this system.

Let $\{S_{=0}, S_{=1}, S_?\}$ be a partition of S such that $S_{=0} \subseteq \{s \in S \mid \mathbb{P}_s(\mathsf{Reach}(T)) = 0\}$ and $T \subseteq S_{=1} \subseteq \{s \in S \mid \mathbb{P}_s(\mathsf{Reach}(T)) = 1\}$. We consider the system defined by $x_s = 0$ for all $s \in S_{=0}$, $x_s = 1$ for all $s \in S_{=1}$ and $x_s = \sum_{s' \in S} \delta(s)(s') \cdot x_{s'}$ for all $s \in S_?$. The least non-negative solution of this system is obtained by letting $x_s = \mathbb{P}_s(\mathsf{Reach}(T))$ for all $s \in S$. Furthermore, this system has a unique solution when $S_{=0} = \{s \in S \mid \mathbb{P}_s(\mathsf{Reach}(T)) = 0\}$. The set $\{s \in S \mid \mathbb{P}_s(\mathsf{Reach}(T)) = 0\}$ only depends on the topology of the Markov chains (i.e., which states are connected by a transition) and not the transition probabilities; a state is in this set if and only if there are no histories starting in this state and ending in T.

The Blum-Shub-Smale model of computation. Some of our complexity bounds rely on a model of computation introduced by Blum, Shub and Smale [10]. A Blum-Shub-Smale (BSS) machine, intuitively, is a random-access memory machine with registers storing real numbers. Arithmetic computations on the content of registers are in constant time in this model.

The class of decision problem that can be solved in polynomial time in the BSS model coincides with the class of decision problems that can be solved, in the Turing model, in polynomial time with a PosSLP oracle [1]. The PosSLP problem asks, given a division-free straight-line program (intuitively, an arithmetic circuit), whether its output is positive. The PosSLP problem lies in the counting hierarchy and can be solved in polynomial space [1].

Theory of the reals. The theory of the reals refers to the set of sentences (i.e., fully quantified first-order logical formulae) in the signature of ordered fields that hold in $\mathbb{R}$. The problem of deciding whether a sentence is in the theory of the reals is decidable; if the number of quantifier blocks is fixed, this can be done in PSPACE [5, Rmk. 13.10]. Furthermore, the problem of deciding the validity of an existential (resp. universal) formula is NP-hard (resp. co-NP-hard) [5, Rmk. 13.9]. Our complexity bounds refer to the complexity classes ETR (existential theory of the reals) and co-ETR, which contain the problems that can be reduced in polynomial time to checking the membership of an *existential sentence* and *universal sentence* in the theory of the reals respectively.

3 Interval strategies and decision problems

We now introduce interval strategies. Definitions are given in Section 3.1. We compare interval strategies to finite-memory strategies of the MDP underlying the OC-MDP in Section 3.2. We close this section by formulating our decision problems in Section 3.3.

We fix an OC-MDP $\mathcal{Q} = (Q, A, \delta, w)$ and a counter upper bound $r \in \bar{\mathbb{N}}_{>0}$ for this whole section.

3.1 Interval strategies

Interval strategies are a subclass of memoryless strategies of $\mathcal{M}^{\leq r}(\mathcal{Q})$ that admit finite compact representations based on families of intervals. Intuitively, a strategy is an interval strategy if there exists an *interval partition* (i.e., a partition containing only intervals) of the set of counter values such that decisions taken in a state for two counter values in the same interval coincide. We require that this partition has a finite representation to formulate verification and synthesis problems for interval strategies.

The set $[\![1, r-1]\!]$ contains all counter values for which decisions are relevant. Let $\mathcal{I}$ be an interval partition of $[\![1, r-1]\!]$. A memoryless strategy σ of $\mathcal{M}^{\leq r}(\mathcal{Q})$ is *based on the partition* $\mathcal{I}$ if for all $q \in Q$, all $I \in \mathcal{I}$ and all $k, k' \in I$, we have $\sigma(q, k) = \sigma(q, k')$. All memoryless strategies are based on the trivial partition of $[\![1, r-1]\!]$ into singleton sets. In practice, we are interested in strategies based on partitions with a small number of large intervals.

We define two classes of interval strategies: strategies that are based on finite partitions and, in unbounded OC-MDPs, strategies that are based on periodic partitions. An interval partition $\mathcal{I}$ of $\mathbb{N}_{>0}$ is *periodic* if there exists a period $\rho \in \mathbb{N}_{>0}$ such that for all $I \in \mathcal{I}$, $I + \rho = \{k + \rho \mid k \in I\} \in \mathcal{I}$. A periodic interval partition $\mathcal{I}$ with period ρ *induces* the interval partition $\mathcal{J} = \{I \in \mathcal{I} \mid I \subseteq [\![1, \rho]\!]\}$ of $[\![1, \rho]\!]$. Conversely, for any $\rho \in \mathbb{N}_{>0}$, an interval partition $\mathcal{J}$ of $[\![1, \rho]\!]$ *generates* the periodic partition $\mathcal{I} = \{I + k \cdot \rho \mid I \in \mathcal{J}, k \in \mathbb{N}\}$.

Let σ be a memoryless strategy of $\mathcal{M}^{\leq r}(\mathcal{Q})$. The strategy σ is an *open-ended interval strategy* (OEIS) if it is based on a finite interval partition of $[\![1, r-1]\!]$. The qualifier open-ended follows from there being an unbounded interval in any finite interval partition of $\mathbb{N}_{>0}$ (for the unbounded case). When $r = \infty$, σ is a *cyclic interval strategy* (CIS) if there exists a period $\rho \in \mathbb{N}_{>0}$ such that for all $q \in Q$ and all $k \in \mathbb{N}_{>0}$, we have $\sigma(q, k) = \sigma(q, k + \rho)$. A strategy is a CIS with period ρ if and only if it is based on a periodic interval partition of $\mathbb{N}_{>0}$ with period ρ.

▶ Remark 2. We do not consider strategies based on ultimately periodic interval partitions. However, our techniques can be adapted to analyse such strategies. The compressed Markov chain construction of Section 4 can be defined for any memoryless strategy of $\mathcal{M}^{\leq r}(\mathcal{Q})$. By combining our approaches for the analysis of OEISs and CISs via compressed Markov chains, we can analyse such strategies. Furthermore, we can show that our complexity bounds for the decision problems we study extend to these strategies. ◁

We represent interval strategies as follows. First, assume that σ is an OEIS and let $\mathcal{I}$ be the coarsest finite interval partition of $[\![1, r-1]\!]$ on which σ is based. We can represent σ by a table that lists, for each $I \in \mathcal{I}$, the bounds of I and a memoryless strategy of $\mathcal{Q}$ dictating the choices to be made when the current counter value lies in I. Next, assume that $r = \infty$ and that σ is a CIS with period ρ. Let $\mathcal{J}$ be an interval partition of $[\![1, \rho]\!]$ such that σ is based on the partition $\mathcal{I}$ generated by $\mathcal{J}$. We represent σ by ρ and an OEIS of $\mathcal{M}^{\leq \rho+1}(\mathcal{Q})$ based on $\mathcal{J}$ that specifies the behaviour of σ for counter values up to ρ.

▶ Remark 3. In practice, in the bounded setting, it is not necessary to encode the counter upper bound r in the representation of an OEIS; it is implicit from $\mathcal{M}^{\leq r}(\mathcal{Q})$. Nonetheless, we assume that r is part of the strategy representation for the sake of convenience: this allows us to treat all bounded intervals uniformly in complexity analyses, as though all interval bounds are in the encoding of considered OEIS. This has no impact on our complexity results, as r is part of the description of $\mathcal{M}^{\leq r}(\mathcal{Q})$. ◁

Interval strategies subsume *counter-oblivious strategies*, i.e., memoryless strategies that make choices based only on the state and disregard the current counter value. Counter-oblivious strategies can be viewed as memoryless strategies $\sigma\colon Q \to \mathcal{D}(A)$ of $\mathcal{Q}$.

3.2 Substituting counters with memory

Memoryless strategies of $\mathcal{M}^{\leq r}(\mathcal{Q})$ can be seen as strategies of $\mathcal{Q}$ when an initial counter value is fixed. The idea is to use memory to keep track of the counter instead of observing it. We can thus compare our representation of interval strategies to the classical Mealy machine representation of the corresponding strategies of $\mathcal{Q}$.

We first formalise how to derive a strategy of $\mathcal{Q}$ from a memoryless strategy of $\mathcal{M}^{\leq r}(\mathcal{Q})$ and an initial counter value. Let $k_{\mathsf{init}} \in [\![1, r-1]\!]$ be an initial counter value and let σ be a memoryless strategy of $\mathcal{M}^{\leq r}(\mathcal{Q})$. We build a (partially-defined) strategy $\tau^{\sigma}_{k_{\mathsf{init}}}$ of $\mathcal{Q}$ from σ. Intuitively, instead of having the counter value as an input of the strategy, we store the current counter value in memory. For any history $h_{\mathcal{Q}} = q_1 a_1 \dots a_{n-1} q_n \in \mathsf{Hist}(\mathcal{Q})$, we define $w(h_{\mathcal{Q}}) = \sum_{\ell=0}^{n-1} w(q_\ell, a_\ell)$. The strategy $\tau^{\sigma}_{k_{\mathsf{init}}}$ is defined for any history $h_{\mathcal{Q}} \in \mathsf{Hist}(\mathcal{Q})$ by $\tau^{\sigma}_{k_{\mathsf{init}}}(h_{\mathcal{Q}}) = \sigma((\mathsf{last}(h_{\mathcal{Q}}), k_{\mathsf{init}} + w(h_{\mathcal{Q}})))$ when $k_{\mathsf{init}} + w(h_{\mathcal{Q}}) \in [\![1, r-1]\!]$ and is left undefined otherwise.

Similarly, the state-reachability and selective termination objectives of $\mathcal{M}^{\leq r}(\mathcal{Q})$ can be translated into objectives of $\mathcal{Q}$ when an initial counter value k_{init} is specified. We can show that the probability in $\mathcal{M}^{\leq r}(\mathcal{Q})$ of such an objective under σ from a configuration (q, k_{init}) matches the probability of the counterpart objective in $\mathcal{Q}$ under the strategy $\tau^{\sigma}_{k_{\mathsf{init}}}$ from q.

If $r \in \mathbb{N}$, the counterpart in $\mathcal{Q}$ of any OEIS has finite memory: there are only finitely many counter values. Assume that $r \in \mathbb{N}$ and let σ be an OEIS. The strategy $\tau^{\sigma}_{k_{\mathsf{init}}}$ is induced by the following Mealy machine: we let $\mathfrak{M} = (M, m_{\mathsf{init}}, \mathsf{up}, \mathsf{nxt})$ where $M = [\![1, r-1]\!]$, $m_{\mathsf{init}} = k_{\mathsf{init}}$, and for all $q \in Q$, $k \in M$ and $a \in A$, we let $\mathsf{up}(k, q, a) = k + w(q, a)$ and $\mathsf{nxt}(k, q)(a) = \sigma((q, k))$. Intuitively, $\mathfrak{M}$ induces $\tau^{\sigma}_{k_{\mathsf{init}}}$ because it keeps track of the weight of the current history and this weight determines the choice prescribed by $\tau^{\sigma}_{k_{\mathsf{init}}}$.

This construction yields a Mealy machine whose size is exponential in the binary encoding size of r. The following example illustrates that such exponential-size Mealy machines may be required, even for OEISs based on the partition $\{\{1\}, [\![2, r-1]\!]\}$, i.e., OEISs that only have to distinguish 1 from other counter values. Additionally, this same example shows that the counterpart in $\mathcal{Q}$ of an OEIS may require infinite memory in the unbounded setting.

▶ **Example 4.** We consider the OC-MDP $\mathcal{Q}$ illustrated in Figure 1. Let $r \in \bar{\mathbb{N}}_{>0}$, $r \geq 3$. We consider the OEIS σ defined by $\sigma(q_1, k) = a$ for all $k \in [\![2, r-1]\!]$ and $\sigma(q_1, 1) = b$. The strategy σ maximises the probability of terminating in q_2 from the configuration $(q_0, 1)$.

We let τ_1^{σ} denote the counterpart of σ in $\mathcal{Q}$ for the initial counter value 1. We show that for all $2 \leq k < r$, a Mealy machine inducing τ_1^{σ} must have at least k states. In particular, if r is finite, it means that any such Mealy machine must have $r-1$ states, and if r is infinite, it means that there is no (finite) Mealy machine inducing τ_1^{σ}.

Let $2 \leq k < r$. We proceed by contradiction. Assume that there exists a Mealy machine $\mathfrak{M} = (M, m_{\mathsf{init}}, \mathsf{up}, \mathsf{nxt})$ inducing τ_1^{σ} such that $|M| \leq k-1$. For all $\ell \in [\![k-1]\!]$, we let

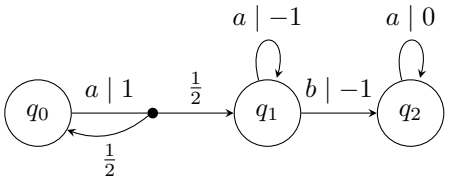


Figure 1 An OC-MDP. Edge splits following actions indicate probabilistic transitions. We indicate the weight of a state-action pair next to each action.

$m_\ell = \widehat{\mathsf{up}}((q_0a)^{k-1}(q_1a)^\ell)$ and let $h_\ell = (q_0a)^{k-1}(q_1a)^\ell q_1$. For all $\ell \in [\![k-1]\!]$, h_ℓ is a history of $\mathcal{Q}$ in the domain of τ_1^σ (because $k < r$ and the initial counter value is 1) that is consistent with τ_1^σ. Since $\mathfrak{M}$ induces τ_1^σ, it follows that for all $\ell \in [\![k-2]\!]$, we have $\tau_1^\sigma(h_\ell) = \mathsf{nxt}(m_\ell, q_1) = a$ and that $\tau_1^\sigma(h_{k-1}) = \mathsf{nxt}(m_{k-1}, q_1) = b$. However, m_{k-1} must occur more than once in the sequence $(m_\ell)_{\ell \in [\![k-1]\!]}$ because this sequence is obtained by repeating the update rule $\mathsf{up}(\cdot, q_1, a)$ from m_0 and there are no more than $k-1$ memory states in $\mathfrak{M}$. This is a contradiction. ◁

We now assume that $r = \infty$. The counterpart in $\mathcal{Q}$ of any CIS is a finite-memory strategy: it suffices to keep track of the remainder of division of the counter value by a period. Let σ be a CIS, ρ be a period of σ and k_{init} be an initial counter value. Formally, the strategy $\tau^\sigma_{k_{\mathsf{init}}}$ is induced by the Mealy machine $\mathfrak{M} = (M, m_{\mathsf{init}}, \mathsf{up}, \mathsf{nxt})$ defined as follows. We let $M = [\![\rho - 1]\!]$ and $m_{\mathsf{init}} = k_{\mathsf{init}} \bmod \rho$. Updates are defined, for all $q \in Q$, $k \in M$ and $a \in A$, by $\mathsf{up}(k, q, a) = (k + w(q, a)) \bmod \rho$. The next-move function is defined differently following whether the memory state is zero or not. We let, for all $q \in Q$, $k \in M$ and $a \in A$, $\mathsf{nxt}(k, q)(a) = \sigma((q, k))$ if $k \neq 0$ and $\mathsf{nxt}(0, q)(a) = \sigma((q, \rho))$. Intuitively, $\mathfrak{M}$ induces $\tau^\sigma_{k_{\mathsf{init}}}$ because the remainder of the current counter value for its division by the period is sufficient to mimic σ.

By adapting Example 4, we can show that a CIS representation may be exponentially more succinct than any Mealy machine for its counterpart in $\mathcal{Q}$.

3.3 Decision problems

We formulate the decision problems studied in following sections. The common inputs of these problems are an OC-MDP $\mathcal{Q}$ with rational transition probabilities, a counter bound $r \in \bar{\mathbb{N}}_{>0}$ (encoded in binary if it is finite), a target $T \subseteq Q$, an objective $\Omega \in \{\mathsf{Reach}(T), \mathsf{Term}(T)\}$, an initial configuration $s_{\mathsf{init}} = (q_{\mathsf{init}}, k_{\mathsf{init}})$ and a threshold $\theta \in [0, 1] \cap \mathbb{Q}$ against which we compare the probability of Ω. Problems that are related to CISs assume that $r = \infty$ as all strategies in the bounded case are OEISs. We lighten the probability notation below by omitting $\mathcal{M}^{\leq r}(\mathcal{Q})$ from it.

First, we are concerned with the verification of interval strategies. The *interval strategy verification problem* asks to decide, given an interval strategy σ, whether $\mathbb{P}^\sigma_{s_{\mathsf{init}}}(\Omega) \geq \theta$. We assume that the encoding of the input interval strategy matches the description of Section 3.1.

The other two problems relate to interval strategy synthesis. The corresponding decision problem is called *realisability*. We provide algorithms checking the existence of well-performing structurally-constrained interval strategies. We formulate two variants of this problem.

For the first variant, we fix an interval partition $\mathcal{I}$ of $[\![1, r-1]\!]$ beforehand and ask to check if there is a good strategy based on $\mathcal{I}$. Formally, the *fixed-interval OEIS realisability problem* asks whether, given a finite interval partition $\mathcal{I}$ of $[\![1, r-1]\!]$, there exists an OEIS σ based on $\mathcal{I}$ such that $\mathbb{P}^\sigma_{s_{\mathsf{init}}}(\Omega) \geq \theta$. The variant for CISs is defined similarly: the *fixed-interval*

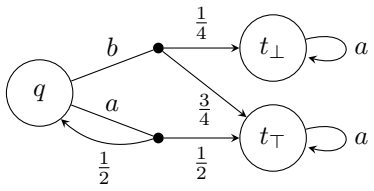


Figure 2 An OC-MDP where all weights are -1 and are omitted from the figure.

CIS realisability problem asks whether, given a period $\rho \in \mathbb{N}_{>0}$ and an interval partition $\mathcal{J}$ of $[\![1, \rho]\!]$, there exists a CIS σ based on the partition generated by $\mathcal{J}$ such that $\mathbb{P}^{\sigma}_{s_{\text{init}}}(\Omega) \geq \theta$.

For the second variant, we parameterise the number of intervals in the partition and the size of bounded intervals. The *parameterised OEISs realisability problem* asks, given a bound $d \in \mathbb{N}_{>0}$ on the number of intervals and a bound $n \in \mathbb{N}_{>0}$ on the size of bounded intervals, whether there exists an OEIS σ such that $\mathbb{P}^{\sigma}_{s_{\text{init}}}(\Omega) \geq \theta$ and σ is based on an interval partition $\mathcal{I}$ of $[\![1, r-1]\!]$ with $|\mathcal{I}| \leq d$ and, for all bounded $I \in \mathcal{I}$, $|I| \in [\![1, n]\!]$ (in particular, the parameter n does not constrain the required infinite interval in the unbounded setting $r = \infty$). The *parameterised CIS realisability problem* asks, given a bound $d \in \mathbb{N}_{>0}$ on the number of intervals and a bound $n \in \mathbb{N}_{>0}$ on the size of intervals, whether there exists a CIS σ such that $\mathbb{P}^{\sigma}_{s_{\text{init}}}(\Omega) \geq \theta$ and σ is based on an interval partition $\mathcal{I}$ of $\mathbb{N}_{>0}$ with period ρ such that $|I| \leq n$ for all $I \in \mathcal{I}$ and $\mathcal{I}$ induces a partition of $[\![1, \rho]\!]$ with at most d intervals. In both cases, we assume that the number d is given in unary. This ensures that witness strategies, when they exist, are based on interval partitions that have a representation of size polynomial in the size of the inputs.

▶ Remark 5. In bounded OC-MDPs, instances of the parameterised OEIS realisability problem such that no partitions are compatible with the input parameters d and n are trivially negative. If $r \in \mathbb{N}$, then there are no interval partitions of $[\![1, r-1]\!]$ compatible with d and n whenever $r - 1 > d \cdot n$. This issue does not arise for OEISs in unbounded OC-MDPs or for CISs, as counter-oblivious strategies are always possible witnesses no matter the parameters. ◁

For both realisability problems, we consider two variants, depending on whether we want the answer with respect to the set of *pure* or *randomised* interval strategies. For many objectives in MDPs (e.g., reachability, parity objectives [3, 13]), the maximum probability of satisfying the objective is the same for pure and randomised strategies. As highlighted by the following example, when we restrict the structure of the sought interval strategy, there may exist randomised strategies that perform better than all pure ones. Intuitively, randomisation can somewhat alleviate the loss in flexibility caused by structural restrictions [24].

▶ **Example 6.** The fixed-interval and parameterised realisability problems subsume the realisability problem for counter-oblivious strategies. The goal of this example is to provide an OC-MDP in which there exists a randomised counter-oblivious strategy that performs better than any pure counter-oblivious strategy from a given configuration.

We consider the OC-MDP $\mathcal{Q}$ depicted in Figure 2 in which all weights are -1, the objective $\mathsf{Reach}(t_\top) = \mathsf{Term}(t_\top)$, a counter bound $r \geq 3$ and the initial configuration $(q, 2)$. To maximise the probability of reaching $t_\top$ from $(q, 2)$ with an unrestricted strategy, we must select action a in $(q, 2)$ and then b in $(q, 1)$. Intuitively, a is used when two steps remain as it allows another attempt and we switch to b when one step is left because it is the action that maximises the probability of reaching $t_\top$ in a single step.

We now limit our attention to counter-oblivious strategies. For pure strategies, no matter whether action a or b is chosen in q, $t_\top$ is reached with probability $\frac{3}{4}$ from $(q,2)$. However, when playing both actions uniformly at random in q, the resulting reachability probability from $(q,2)$ is $\frac{25}{32} > \frac{3}{4}$. This shows that randomised counter-oblivious strategies can achieve better reachability (resp. selective termination) probabilities than pure strategies. ◁

4 Compressing induced Markov chains

This section introduces the construction underlying our algorithms for the decision problems presented in Section 3. To analyse the (possibly infinite) Markov chains induced by memoryless strategies over the space of configurations of an OC-MDP, we build a *compressed Markov chain*, i.e., a Markov chain with fewer configurations and adjusted transitions. A compressed Markov chain is defined with respect to a memoryless strategy and an interval partition on which the strategy is based. This construction is generic, in the sense that it can formally be defined not only for interval strategies, but for all memoryless strategies.

We introduce the compression construction in Section 4.1. Section 4.2 explains how to efficiently enforce a technical property on interval partitions such that compressed Markov chains are well-defined. In Section 4.3, we prove that termination probabilities and the probability of hitting a counter upper bound are preserved following compression. For algorithmic reasons, we show that the transition probabilities of the compressed Markov chain can be represented as solutions of systems of quadratic equations in Section 4.4. Although compressed Markov chains are finite for OEISs, they are not for CISs; we close the section by proving that compressed Markov chains for CISs are induced by one-counter Markov chains.

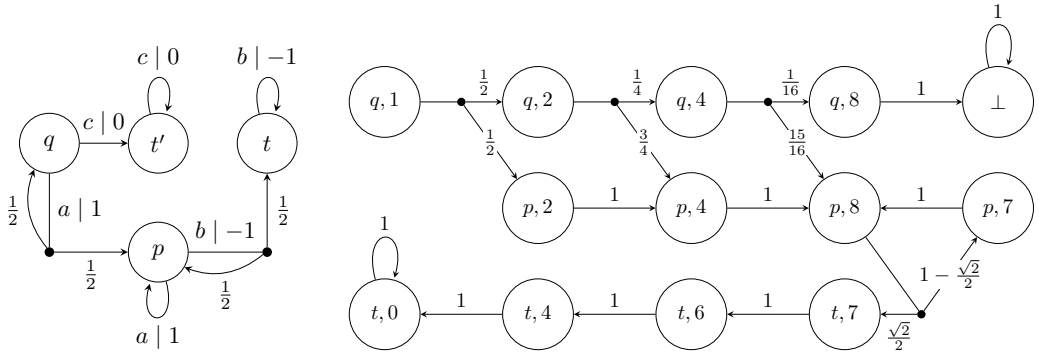
For the whole section, we fix an OC-MDP $\mathcal{Q} = (Q, A, \delta, w)$, a bound $r \in \bar{\mathbb{N}}_{>0}$ on counter values and a memoryless strategy σ of $\mathcal{M}^{\leq r}(\mathcal{Q})$ based on an interval partition $\mathcal{I}$ of $[\![1, r-1]\!]$.

4.1 Defining compressed Markov chains

We introduce the *compressed Markov chain* $\mathcal{C}^\sigma_\mathcal{I}(\mathcal{Q})$ derived from the Markov chain induced on $\mathcal{M}^{\leq r}(\mathcal{Q})$ by σ and the partition $\mathcal{I}$. We write $\mathcal{C}^\sigma_\mathcal{I}$ instead of $\mathcal{C}^\sigma_\mathcal{I}(\mathcal{Q})$ whenever $\mathcal{Q}$ is clear from the context. Intuitively, we keep some configurations in the state space of $\mathcal{C}^\sigma_\mathcal{I}$ and, to balance this, transitions of $\mathcal{C}^\sigma_\mathcal{I}$ aggregate several histories of the induced Markov chain. We also apply compression for bounded intervals: interval bounds are represented in binary and thus the size of an interval is exponential in its encoding size. We open with an example.

▶ **Example 7.** We consider the OC-MDP depicted in Figure 3a and counter upper bound $r = +\infty$. Let σ denote the randomised OEIS based on $\mathcal{I} = \{[\![1,7]\!], [\![8,\infty]\!]\}$ such that $\sigma(q,1)(a) = \sigma(p,1)(a) = \sigma(q,8)(c) = 1$ and $\sigma(p,8)(a) = \sigma(p,8)(b) = \frac{1}{2}$.

We define the compressed Markov chain $\mathcal{C}^\sigma_\mathcal{I}$ depicted in Figure 3b by processing each interval individually. First, we consider the bounded interval $I = [\![1,7]\!]$. When we enter I from its minimum or maximum, we only consider counter jumps by powers of 2, starting with $1 = 2^0$. If a counter value in I is reached by jumping by 2^β, we consider counter updates of $2^{\beta+1}$ from it; Figure 4 illustrates this counter update rule. This explains the counter progressions from $(q,1)$ to $(q,8)$ and from $(t,7)$ to $(t,0)$. The state space of $\mathcal{C}^\sigma_\mathcal{I}$ with respect to I contains the configurations whose counter values can be reached via this scheme. Transitions aggregate several histories of $\mathcal{M}^{\leq\infty}(\mathcal{Q})$, e.g., the probability from $s = (q,2)$ to $s' = (p,4)$ is the probability under σ of all histories of $\mathcal{M}^{\leq\infty}(\mathcal{Q})$ from s to s' along which counter values elsewhere than in s' remain between $\min I = 1$ and 3 (i.e., the counter



(a) An OC-MDP. Weights are written next to actions.

(b) Fragment of the compressed Markov chain of Ex. 7 reachable from $(q, 1)$. Configuration parentheses are omitted to lighten the figure.

Figure 3 An illustration of an OC-MDP and its compression for a specific strategy.

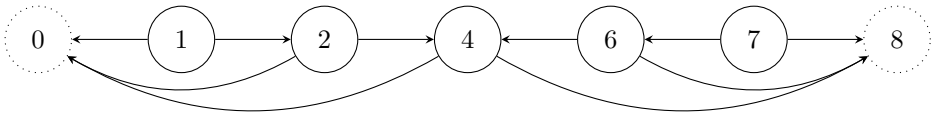


Figure 4 An illustration of the counter update scheme in a compressed Markov chain for the interval $[\![1, 7]\!]$.

value before the next step). The encoding of transition probabilities may be exponential in the number of retained configurations; this is highlighted by the progression of probability denominators between $(q, 1)$ and $(q, 8)$.

For the unbounded interval $I = [\![8, \infty]\!]$, we only consider configurations with counter value $\min I = 8$ and consider transitions to configurations with counter value $\min I - 1 = 7$. In this case, for instance, the transition probability from $(p, 8)$ to $(p, 7)$, corresponds the probability under σ in $\mathcal{M}^{\leq\infty}(\mathcal{Q})$ of hitting counter value 7 for the first time in p from $(p, 8)$. This example illustrates that this probability can be irrational. Here, the probability of moving from $(p, 8)$ to $(p, 7)$ is a solution of the quadratic equation $x = \frac{1}{4} + \frac{1}{2}x^2$ (see [35]): $\frac{1}{4}$ is the probability of directly moving from $(p, 8)$ to $(p, 7)$ and $\frac{1}{2}x^2$ is the probability of moving from $(p, 8)$ to $(p, 7)$ by first going through the intermediate configuration $(p, 9)$.

Finally, we comment on the absorbing state $\bot$. The rules making up transitions of $\mathcal{C}^\sigma_\mathcal{I}$ outlined above require a change in counter value. We redirect the probability of never seeing such a change to $\bot$. In this example, σ does not allow a counter decrease from $(q, 8)$. ◁

Example 7 outlines the main ideas to construct $\mathcal{C}^\sigma_\mathcal{I}$. To ensure that compressed Markov chains are well-defined, we impose the following assumption on $\mathcal{I}$ which guarantees that, in general, bounded intervals of $\mathcal{I}$ can only be entered by one of their bounds.

▶ **Assumption 8.** *For all bounded $I \in \mathcal{I}$, $\log_2(|I| + 1) \in \mathbb{N}$, i.e., $|I| = 2^{\beta_I} - 1$ for some $\beta_I \in \mathbb{N}$.*

Assumption 8 is not prohibitive: we prove in Section 4.2 that, given a bounded interval, we can partition it into sub-intervals satisfying the required size constraint in polynomial time. We assume that Assumption 8 is satisfied for $\mathcal{I}$ for the remainder of the section.

We now formalise $\mathcal{C}^\sigma_\mathcal{I} = (S_\mathcal{I}, \delta^\sigma_\mathcal{I})$. We start by defining its state space $S_\mathcal{I}$ which does not depend on σ. We first formalise the configurations that are retained for each interval.

Let $I \in \mathcal{I}$. First, let us assume that I is unbounded and let b_I^- denote its minimum. We set $S_I = Q \times \{b_I^-\}$, i.e., we only retain the configurations with minimal counter value in I.

Next, let us assume that I is bounded and of the form $[\![b_I^-, b_I^+]\!]$. Let $\beta_I = \log_2(|I|+1)$ (this is an integer by Assumption 8). We retain the counter values that can be reached via a generalisation of the scheme depicted in Figure 4. The set of retained configurations for I is given by

$$S_I = Q \times \left(\{b_I^- + 2^\alpha - 1 \mid \alpha \in [\![\beta_I - 1]\!]\} \cup \{b_I^+ - (2^\alpha - 1) \mid \alpha \in [\![\beta_I - 1]\!]\}\right).$$

Finally, we consider absorbing configurations and the new state $\perp$. We let $S_{\mathcal{I}}^\perp = \{\perp\} \cup (Q \times \{0, r\})$ if $r \in \mathbb{N}$ and $S_{\mathcal{I}}^\perp = \{\perp\} \cup (Q \times \{0\})$ otherwise. We define $S_{\mathcal{I}} = S_{\mathcal{I}}^\perp \cup \bigcup_{I \in \mathcal{I}} S_I$.

We now define the transition function $\delta_{\mathcal{I}}^\sigma$. For all $s \in S_{\mathcal{I}}^\perp$, we let $\delta_{\mathcal{I}}^\sigma(s)(s) = 1$. For configurations whose counter value lies in one of the intervals $I \in \mathcal{I}$, we provide a unified definition based on a notion of *successor counter values*, generalising the ideas of Example 7 and Figure 4.

Let $I \in \mathcal{I}$. If I is unbounded, we define the successor of $b_I^- = \min I$ to be $b_I^- - 1$. We now assume that $I = [\![b_I^-, b_I^+]\!]$ is bounded and let $\beta_I = \log_2(|I|+1)$ and $\alpha \in [\![\beta_I - 1]\!]$. The successors of $b_I^- + 2^\alpha - 1$ are $b_I^- - 1$ and $b_I^- + 2^{\alpha+1} - 1$. Symmetrically, for $b_I^+ - (2^\alpha - 1)$, its successors are $b_I^+ + 1$ and $b_I^+ - (2^{\alpha+1} - 1)$. Both cases entail a counter change by 2^α. Assumption 8 ensures that all successor counter values appear in $S_{\mathcal{I}}$.

Let $s = (q, k) \in S_{\mathcal{I}} \setminus S_{\mathcal{I}}^\perp$ and $s' = (q', k') \in S_{\mathcal{I}} \setminus \{\perp\}$. If k' is not a successor of k, we set $\delta_{\mathcal{I}}^\sigma(s)(s') = 0$. Assume now that k' is a successor of k. We let $\mathcal{H}_{\mathsf{succ}}(s, s') \subseteq \mathsf{Hist}(\mathcal{M}^{\leq r}(\mathcal{Q}))$ be the set of histories h such that $\mathsf{first}(h) = s$, $\mathsf{last}(h) = s'$ and for all configurations s'' along h besides s', the counter value of s'' is not a successor of k; outside of s' along h, the counter remains, in the bounded case, strictly between the two successors of k, and, in the unbounded case, strictly above the successor $k-1$ of k. We set $\delta_{\mathcal{I}}^\sigma(s)(s') = \mathbb{P}^\sigma_{\mathcal{M}^{\leq r}(\mathcal{Q}),s}(\mathsf{Cyl}(\mathcal{H}_{\mathsf{succ}}(s, s')))$. To ensure that $\delta_{\mathcal{I}}^\sigma(s)$ is a distribution we let $\delta_{\mathcal{I}}^\sigma(s)(\perp) = 1 - \sum_{s'' \neq \perp} \delta_{\mathcal{I}}^\sigma(s)(s'')$; this transition captures the probability of the counter never hitting a successor of k.

▶ Remark 9. Although we have formalised compressed Markov chains for OC-MDPs, the construction can be applied to one-counter Markov chains. In particular, the properties outlined below transfer to the compression of a one-counter Markov chain. ◁

In the following, we differentiate histories of $\mathcal{C}_{\mathcal{I}}^\sigma$ from histories of $\mathcal{M}^{\leq r}(\mathcal{Q})$ by denoting them with a bar, e.g., $\bar{h}$ indicates a history of $\mathcal{C}_{\mathcal{I}}^\sigma$.

4.2 Efficiently refining interval partitions

To define a compressed Markov chain with respect to an interval partition $\mathcal{J}$ of $[\![1, r-1]\!]$, we require that the size constraints of Assumption 8 hold, i.e., that for all bounded $I \in \mathcal{J}$, $\log_2(|I|+1) \in \mathbb{N}$. We present a refinement procedure for interval partitions that enforces this property while generating few intervals. At the end of this section, we provide an additional procedure that can be used to retain specific configurations in the state space of compressed Markov chains.

To refine an interval partition, we subdivide its bounded intervals one by one. Breaking up these intervals into singleton sets is not a valid approach for complexity reasons; any input interval partition is such that its interval bounds are represented in binary, i.e., the size of intervals is exponential in the size of their representation. We provide a polynomial-time refinement procedure that divides an interval into sub-intervals of the required size in Algorithm 1. To refine an interval, we determine a largest sub-interval of a suitable size and then continue by recursively partitioning its complement. This algorithm enables us, in the

Algorithm 1 Procedure Refine to split an interval into intervals of size in $\{2^\beta - 1 \mid \beta \in \mathbb{N}\}$.

Data: A bounded interval $I = [\![b^-, b^+]\!]$.
$\ell \leftarrow \lfloor \log_2(b^+ - b^- + 2) \rfloor (= \lfloor \log_2(|I| + 1) \rfloor)$;
if $|I| = 2^\ell - 1$ **then**
  return $\{I\}$;
else
  $I' \leftarrow [\![b^-, b^- + 2^\ell - 2]\!]$; $I'' \leftarrow [\![b^- + 2^\ell - 1, b^+]\!]$;
  return $\{I'\} \cup \mathsf{Refine}(I'')$;

context of verification, to modify the interval partition from the representation of an interval strategy into one suitable for compressed Markov chains.

We show that, for all bounded intervals I of $\mathbb{N}_{>0}$, the partition $\mathsf{Refine}(I)$ (from Algorithm 1) has a polynomial size (with respect to the binary encoding of the bounds of I) and all of its elements J satisfy $\log_2(|J| + 1) \in \mathbb{N}$.

▶ **Lemma 10.** *Let $I = [\![b^-, b^+]\!]$ be a bounded interval of $\mathbb{N}_{>0}$. The interval partition $\mathsf{Refine}(I)$ of I obtained via Algorithm 1 satisfies $|\mathsf{Refine}(I)| \leq \log_2(|I| + 1) + 1 \leq \log_2(b^+) + 1$ and, for all $J \in \mathsf{Refine}(I)$, we have $\log_2(|J| + 1) \in \mathbb{N}$.*

Proof. Let $\ell = \lfloor \log_2(|I| + 1) \rfloor$. We show both statements by induction.

For the size of the elements in $\mathsf{Refine}(I)$, we proceed by induction on $|I|$. If $|I| = 1$, then $\mathsf{Refine}(I) = \{I\}$ (since $1 = 2^1 - 1$) and the result follows. Now, assume that for all intervals smaller than I, the statement holds. If $I = 2^\ell - 1$, we have $\mathsf{Refine}(I) = \{I\}$ which satisfies the condition. Otherwise, we let $I' = [\![b^-, b^- + 2^\ell - 2]\!]$ and $I'' = [\![b^- + 2^\ell - 1, b^+]\!]$. In particular, we have $|I'| = b^- + 2^\ell - 2 - b^- + 1 = 2^\ell - 1$ and thus $\log_2(|I'| + 1) \in \mathbb{N}$. We conclude that all elements of $\mathsf{Refine}(I) = \{I'\} \cup \mathsf{Refine}(I'')$ satisfy the required constraint on their size.

We now show that $|\mathsf{Refine}(I)| \leq \ell + 1$. We proceed by induction on ℓ. If $\ell = 1$, then $|I| = 1$ and we have $|\mathsf{Refine}(I)| = 1$. This closes the base case.

We assume by induction that for all I' such that $\ell' = \lfloor \log_2(|I'| + 1) \rfloor < \ell$, we have $|\mathsf{Refine}(I')| \leq \ell' + 1$. If $|I| = 2^\ell - 1$, we have $|\mathsf{Refine}(I)| = 1 \leq \ell + 1$. We thus assume that $2^\ell - 1 < |I| < 2^{\ell+1} - 1$ (the upper bound follows from the definition of ℓ). We let $I' = [\![b^-, b^- + 2^\ell - 2]\!]$ and $I'' = [\![b^- + 2^\ell - 1, b^+]\!]$. It remains to show that $|\mathsf{Refine}(I'')| \leq \ell$ to conclude. It holds that $|I''| = |I| - (2^\ell - 1) < 2^{\ell+1} - 1 - (2^\ell - 1) = 2^\ell$. We distinguish two cases in light of this. First, we assume that $|I''| = 2^\ell - 1$. In this case, we have $|\mathsf{Refine}(I'')| = 1$, which implies that $|\mathsf{Refine}(I)| = 2 \leq \ell + 1$. Second, we assume that $|I''| < 2^\ell - 1$. By the induction hypothesis, we obtain that $|\mathsf{Refine}(I'')| \leq \ell$, ensuring that $|\mathsf{Refine}(I)| \leq \ell + 1$ and ending the argument. ◀

For the sake of conciseness, we extend the Refine operator to infinite intervals and interval partitions. For any infinite interval I of $\mathbb{N}_{>0}$, we let $\mathsf{Refine}(I) = \{I\}$. Let J be an interval of $\mathbb{N}_{>0}$ and let $\mathcal{J}$ be a partition of J. We let $\mathsf{Refine}(\mathcal{J}) = \bigcup_{I \in \mathcal{J}} \mathsf{Refine}(I)$. Lemma 10 implies that the constraints of Assumption 8 are satisfied by $\mathsf{Refine}(\mathcal{J})$. This result also yields bounds on the size of $\mathsf{Refine}(\mathcal{J})$ when $\mathcal{J}$ is finite.

We remark that if an interval partition $\mathcal{J}$ of $\mathbb{N}_{>0}$ has period ρ and is generated by an interval partition $\mathcal{J}'$ of $[\![1, \rho]\!]$, then $\mathsf{Refine}(\mathcal{J})$ is generated by $\mathsf{Refine}(\mathcal{J}')$.

We now introduce an operator ensuring that a specific counter value is retained in a compression by making it an interval bound. For any interval $I = [\![b^-, b^+]\!]$ of $\mathbb{N}_{>0}$ (not necessarily bounded) and $k \in \mathbb{N}$, we let $\mathsf{Isolate}(I, k)$ denote $\{I\}$ if $k \notin I$ and $\{[\![b^-, k]\!], [\![k +$

$1, b^+]\!]\} \setminus \{\emptyset\}$ if $k \in I$. We extend the $\mathsf{Isolate}$ operator to interval partitions as follows. For all intervals J of $\mathbb{N}_{>0}$, interval partitions $\mathcal{J}$ of J and $k \in \mathbb{N}$, we let $\mathsf{Isolate}(\mathcal{J}, k) = \bigcup_{I \in \mathcal{J}} \mathsf{Isolate}(I, k)$.

4.3 Validity of the compression approach

We establish that for all configurations $s \in S_{\mathcal{I}}$ and states $q \in Q$, the probability of terminating or reaching the counter upper bound r in q coincides in $\mathcal{C}^\sigma_{\mathcal{I}}$ and in the Markov chain induced on $\mathcal{M}^{\leq r}(\mathcal{Q})$ by σ. There is not such a direct correspondence for state-reachability probabilities. We prove that for all targets $T \subseteq Q$, there exists an OC-MDP $\mathcal{Q}'$ with state space Q derived by changing transitions of $\mathcal{Q}$ such that, for all $q \in T$, the probability of visiting T for the first time via a configuration with state q in $\mathcal{M}^{\leq r}(\mathcal{Q})$ under σ coincides with the probability of terminating or hitting the counter upper bound in q in $\mathcal{M}^{\leq r}(\mathcal{Q}')$ under σ.

For the first property, we rely on a relation between histories of $\mathcal{C}^\sigma_{\mathcal{I}}$ and of $\mathcal{M}^{\leq r}(\mathcal{Q})$. Let $h = s_0 a_0 \dots a_{n-1} s_n \in \mathsf{Hist}(\mathcal{M}^{\leq r}(\mathcal{Q}))$ such that $\mathsf{last}(h) \in Q \times \{0, r\}$ and $\mathsf{last}(h)$ occurs only once in h. By induction, we identify a sequence of configurations in $S_{\mathcal{I}}$ along h that is a well-formed history of $\mathcal{C}^\sigma_{\mathcal{I}}$. Let $\ell_0 = 0$. Assume that we have constructed an increasing sequence $\ell_0 < \ldots < \ell_\iota$ such that $s_{\ell_0} \dots s_{\ell_\iota} \in \mathsf{Hist}(\mathcal{C}^\sigma_{\mathcal{I}})$. If $\ell_\iota \neq n$, we let $\ell_{\iota+1}$ be the least index $\ell > \ell_\iota$ such that $s_\ell \in S_{\mathcal{I}}$ and $\delta^\sigma_{\mathcal{I}}(s_{\ell_\iota})(s_\ell) > 0$ and continue the induction. Such an index is guaranteed to exist. Since weights are in $\{-1, 0, 1\}$, we witness all counter values between that of s_{ℓ_ι} and s_n in the suffix $s_{\ell_\iota} \dots s_n$. This index necessarily exists: all counter values have a smaller successor, and those from a bounded interval have a greater successor. If $\ell_\iota = n$, the induction ends and we let $\bar{h} = s_0 s_{\ell_1} \dots s_{\ell_{n'-1}} s_{\ell_{n'}}$ be the resulting history. We say that $\bar{h}$ *abstracts* h, and it is the unique history of $\mathcal{C}^\sigma_{\mathcal{I}}$ that does so.

We now state the first theorem of this section. The crux of its proof is to establish that, for all histories $\bar{h}$ of $\mathcal{C}^\sigma_{\mathcal{I}}$ ending in $Q \times \{0, r\}$, the probability of its cylinder in $\mathcal{C}^\sigma_{\mathcal{I}}$ matches the probability that a history abstracted by $\bar{h}$ occurs in the Markov chain induced by σ on $\mathcal{M}^{\leq r}(\mathcal{Q})$. We conclude by writing reachability objectives as countable unions of cylinders. For the sake of clarity, in the following statement, we indicate the relevant MDP or Markov chain for each objective.

▶ **Theorem 11.** *Let $s \in S_{\mathcal{I}} \setminus \{\bot\}$. For all $q \in Q$, $\mathbb{P}^\sigma_{\mathcal{M}^{\leq r}(\mathcal{Q}),s}(\mathsf{Term}(q)) = \mathbb{P}_{\mathcal{C}^\sigma_{\mathcal{I}},s}(\mathsf{Reach}_{\mathcal{C}^\sigma_{\mathcal{I}}}(q, 0))$ and, if $r \in \mathbb{N}$, $\mathbb{P}^\sigma_{\mathcal{M}^{\leq r}(\mathcal{Q}),s}(\mathsf{Reach}_{\mathcal{M}^{\leq r}(\mathcal{Q})}(q, r)) = \mathbb{P}_{\mathcal{C}^\sigma_{\mathcal{I}},s}(\mathsf{Reach}_{\mathcal{C}^\sigma_{\mathcal{I}}}(q, r))$.*

Proof. Let $q \in Q$. We only prove the result for the target configuration $(q, 0)$. The argument is the same when the target is (q, r). To lighten notation, we let $\mathcal{M} = \mathcal{M}^{\leq r}(\mathcal{Q})$ for the remainder of the proof. For the sake of clarity, we indicate whether cylinder sets are with respect to $\mathcal{M}$ or $\mathcal{C}^\sigma_{\mathcal{I}}$ by indexing the notation with whichever is applicable.

Let $\bar{h} = s_0 s_1 \dots s_n \in \mathsf{Hist}(\mathcal{C}^\sigma_{\mathcal{I}})$ be such that $\mathsf{first}(\bar{h}) = s$ and $\mathsf{last}(\bar{h}) = (q, 0)$. We let $\mathcal{H}_{\mathsf{abs}}(\bar{h}) \subseteq \mathsf{Hist}(\mathcal{M})$ be the set of histories abstracted by $\bar{h}$. We show that $\mathbb{P}_{\mathcal{C}^\sigma_{\mathcal{I}},s}(\mathsf{Cyl}_{\mathcal{C}^\sigma_{\mathcal{I}}}(\bar{h})) = \mathbb{P}^\sigma_{\mathcal{M},s}(\mathsf{Cyl}_{\mathcal{M}}(\mathcal{H}_{\mathsf{abs}}(\bar{h})))$. By construction of the abstraction relation, all elements of $\mathcal{H}_{\mathsf{abs}}(\bar{h})$ are a uniquely defined concatenation of an element of $\mathcal{H}_{\mathsf{succ}}(s_0, s_1)$ with an element of $\mathcal{H}_{\mathsf{succ}}(s_1, s_2)$, …, with an element of $\mathcal{H}_{\mathsf{succ}}(s_{n-1}, s_n)$. Conversely, any such concatenation is an element of $\mathcal{H}_{\mathsf{abs}}(\bar{h})$. We obtain:

$$\begin{aligned}\mathbb{P}_{\mathcal{C}^\sigma_{\mathcal{I}},s}(\mathsf{Cyl}_{\mathcal{C}^\sigma_{\mathcal{I}}}(\bar{h})) &= \prod_{\ell=0}^{n-1} \mathbb{P}^\sigma_{\mathcal{M},s}(\mathsf{Cyl}_{\mathcal{M}}(\mathcal{H}_{\mathsf{succ}}(s_\ell, s_{\ell+1}))) \\ &= \prod_{\ell=0}^{n-1} \left(\sum_{h_\ell \in \mathcal{H}_{\mathsf{succ}}(s_\ell, s_{\ell+1})} \mathbb{P}^\sigma_{\mathcal{M},s}(\mathsf{Cyl}_{\mathcal{M}}(h_\ell)) \right)\end{aligned}$$

$$
\begin{aligned}
&= \sum_{h_0 \in \mathcal{H}_{\mathsf{succ}}(s_0,s_1)} \cdots \sum_{h_{n-1} \in \mathcal{H}_{\mathsf{succ}}(s_{n-1},s_n)} \left(\prod_{\ell=0}^{n-1} \mathbb{P}^{\sigma}_{\mathcal{M},s}(\mathsf{Cyl}_{\mathcal{M}}(h_\ell)) \right) \\
&= \sum_{h_0 \in \mathcal{H}_{\mathsf{succ}}(s_0,s_1)} \cdots \sum_{h_{n-1} \in \mathcal{H}_{\mathsf{succ}}(s_{n-1},s_n)} \left(\mathbb{P}^{\sigma}_{\mathcal{M},s}(\mathsf{Cyl}_{\mathcal{M}}(h_0 \cdot \ldots \cdot h_{n-1})) \right) \\
&= \mathbb{P}^{\sigma}_{\mathcal{M},s}(\mathsf{Cyl}_{\mathcal{M}}(\mathcal{H}_{\mathsf{abs}}(\bar{h}))).
\end{aligned}
$$

The first line follows by definition of $\delta^{\sigma}_{\mathcal{I}}$ and the definition of the probability distribution over plays of Markov chains. For the second line, we first observe that for all $\ell \in [\![n-1]\!]$, the set $\mathcal{H}_{\mathsf{succ}}(s_\ell, s_{\ell+1})$ is prefix-free and thus the cylinders of elements of $\mathcal{H}_{\mathsf{succ}}(s_\ell, s_{\ell+1})$ are pairwise disjoint. The third line is a rewriting of the second. The fourth line is obtained by definition of the probability distribution induced by a strategy in an MDP, using the fact that σ is a memoryless strategy. The last line is obtained because $\mathcal{H}_{\mathsf{abs}}(\bar{h})$ is the set of all concatenations occurring in the previous line.

We can now end the argument. Let $\mathcal{H}$ and $\bar{\mathcal{H}}$ denote the set of histories of $\mathcal{M}$ and $\mathcal{C}^{\sigma}_{\mathcal{I}}$ respectively that start in s and end in $(q,0)$ with only one occurrence of $(q,0)$. These two sets are prefix-free and we have $\mathcal{H} = \bigcup_{\bar{h} \in \bar{\mathcal{H}}} \mathcal{H}_{\mathsf{abs}}(\bar{h})$. Using the above, we conclude that:

$$
\mathbb{P}^{\sigma}_{\mathcal{M},s}(\mathsf{Term}(q)) = \sum_{h \in \mathcal{H}} \mathbb{P}^{\sigma}_{\mathcal{M},s}(\mathsf{Cyl}_{\mathcal{M}}(h)) = \sum_{\bar{h} \in \bar{\mathcal{H}}} \mathbb{P}_{\mathcal{C}^{\sigma}_{\mathcal{I}},s}(\mathsf{Cyl}_{\mathcal{C}^{\sigma}_{\mathcal{I}}}(\bar{h})) = \mathbb{P}_{\mathcal{C}^{\sigma}_{\mathcal{I}},s}(\mathsf{Reach}((q,0))).
$$

This is the claim of the theorem. ◀

We now discuss state-reachability probabilities. Let $T \subseteq Q$ be a target. Transitions of $\mathcal{C}^{\sigma}_{\mathcal{I}}$ group together (possibly infinitely many) transitions of $\mathcal{M}^{\leq r}(\mathcal{Q})$. In particular, this compression may result in some visits to T not being observed in $\mathcal{C}^{\sigma}_{\mathcal{I}}$ despite occurring in $\mathcal{M}^{\leq r}(\mathcal{Q})$. By slightly modifying $\mathcal{Q}$, we can guarantee that all of these visits are witnessed in the new compressed Markov chain.

The idea is to make all target states absorbing with self-loops of weight -1. Formally, we let $\mathcal{Q}' = (Q, A, \delta', w')$ be the OC-MDP defined by letting, for all $q \in Q$ and all $a \in A(q)$, $\delta'(q,a) = \delta(q,a)$ and $w'(q,a) = w(q,a)$ if $q \notin T$ and, otherwise, $\delta'(q,a)(q) = 1$ and $w'(q,a) = -1$. We remark that σ is a well-defined memoryless strategy of $\mathcal{M}^{\leq r}(\mathcal{Q}')$.

By design, any history of $\mathcal{M}^{\leq r}(\mathcal{Q})$ that ends in a configuration in $T \times [\![r]\!]$ and that does not visit this set before is also a history of $\mathcal{M}^{\leq r}(\mathcal{Q}')$. The cylinders of these histories in both MDPs have the same probability under σ, as transitions are the same in states outside of T. This implies that, under σ, the probability of terminating or hitting the counter upper bound in T in $\mathcal{M}^{\leq r}(\mathcal{Q}')$ is equal to the probability of reaching T in $\mathcal{M}^{\leq r}(\mathcal{Q})$. We conclude by Theorem 11 that the compressed Markov chain $\mathcal{C}^{\sigma}_{\mathcal{I}}(\mathcal{Q}')$ captures the state-reachability probabilities for the target T in $\mathcal{M}^{\leq r}(\mathcal{Q})$ under σ. We formalise this by the following theorem, in which, for the sake of clarity, we indicate the relevant MDP or Markov chain for each objective.

▶ **Theorem 12.** *Let $T \subseteq Q$. Let $\mathcal{Q}' = (Q, A, \delta', w')$ be defined as above. For all $s \in S_{\mathcal{I}}$, we have $\mathbb{P}^{\sigma}_{\mathcal{M}^{\leq r}(\mathcal{Q}),s}(\mathsf{Reach}_{\mathcal{M}^{\leq r}(\mathcal{Q})}(T)) = \mathbb{P}_{\mathcal{C}^{\sigma}_{\mathcal{I}}(\mathcal{Q}'),s}(\mathsf{Reach}_{\mathcal{C}^{\sigma}_{\mathcal{I}}(\mathcal{Q}')}(T \times \{0,r\}))$.*

Proof. To lighten notation, we let $\mathcal{M} = \mathcal{M}^{\leq r}(\mathcal{Q})$ and $\mathcal{M}' = \mathcal{M}^{\leq r}(\mathcal{Q}')$ for the remainder of the proof. We also indicate whether cylinder sets are with respect to $\mathcal{M}$ or $\mathcal{M}'$ by indexing the notation by the applicable MDP. By Theorem 11, it suffices to show that $\mathbb{P}^{\sigma}_{\mathcal{M},s}(\mathsf{Reach}_{\mathcal{M}}(T)) = \mathbb{P}^{\sigma}_{\mathcal{M}',s}(\mathsf{Reach}_{\mathcal{M}'}(T \times \{0,r\}))$.

Let $\mathcal{H} \subseteq \mathsf{Hist}(\mathcal{M})$ be the set of histories $h \in \mathsf{Hist}(\mathcal{M})$ such that $\mathsf{last}(h) \in T \times [\![r]\!]$ and no prior configuration is in $T \times [\![r]\!]$. The state-reachability objective $\mathsf{Reach}_{\mathcal{M}}(T)$ can be written

as $\mathsf{Cyl}_{\mathcal{M}}(\mathcal{H})$. Since $\mathcal{H}$ is prefix-free, we have $\mathbb{P}^{\sigma}_{\mathcal{M},s}(\mathsf{Reach}_{\mathcal{M}}(T)) = \sum_{h\in\mathcal{H}} \mathbb{P}^{\sigma}_{\mathcal{M},s}(\mathsf{Cyl}_{\mathcal{M}}(h))$. Furthermore, for all $h \in \mathcal{H}$, by definition of δ', since no configuration with a state in T occurs along h besides the last one, we have $h \in \mathsf{Hist}(\mathcal{M}')$ and $\mathbb{P}^{\sigma}_{\mathcal{M},s}(\mathsf{Cyl}_{\mathcal{M}}(h)) = \mathbb{P}^{\sigma}_{\mathcal{M}',s}(\mathsf{Cyl}_{\mathcal{M}'}(h))$.

To end the proof, it suffices to show that $\mathsf{Cyl}_{\mathcal{M}'}(\mathcal{H}) = \mathsf{Reach}_{\mathcal{M}'}(T \times \{0, r\})$. We show both inclusions. Let $\pi \in \mathsf{Cyl}_{\mathcal{M}'}(\mathcal{H})$. By definition of $\mathcal{H}$, there exists a configuration of π with a state in T. If the counter value of this configuration is r, then we have $\pi \in \mathsf{Reach}_{\mathcal{M}'}(T \times \{0, r\})$. If not, we are guaranteed to have a configuration in $T \times \{0\}$ along π because states of T are absorbing in $\mathcal{Q}'$ and their self-loops have weight -1. Conversely, let $\pi \in \mathsf{Reach}_{\mathcal{M}'}(T \times \{0, r\})$. By definition of the reachability objective, there must be a configuration with a state in T along π. The earliest occurrence of a state of T (regardless of the counter value) witnesses that $\pi \in \mathsf{Cyl}_{\mathcal{M}'}(\mathcal{H})$. This ends the proof. ◀

4.4 Characterising transition probabilities

Example 7 illustrates that the transition probabilities of a compressed Markov chain may require large representations or be irrational. This section presents characterisations of these transition probabilities via equation systems.

For the remainder of this section, we fix an interval $I \in \mathcal{I}$. We present a system characterising the outgoing transition probabilities from configurations of $\mathcal{C}^{\sigma}_{\mathcal{I}}$ with counter value in I. In Section 4.4.1, we assume that I is unbounded, and we handle the bounded case in Section 4.4.2. We also provide bounds on the size of the systems.

4.4.1 Unbounded intervals

We assume that I is an infinite interval and let $b^- = \min I$. This implies that $r = \infty$. We characterise the transition probabilities from the configurations of $S_{\mathcal{I}}$ with counter value b^- via existing results on termination probabilities in one-counter Markov chains.

For any q, $p \in Q$, the transition probability $\delta^{\sigma}_{\mathcal{I}}((q, b^-))((p, b^- - 1))$ can be seen as a termination probability in a one-counter Markov chain. Let τ denote the counter-oblivious strategy $\sigma(\cdot, b^-)$. We consider the one-counter Markov chain $\mathcal{P} = (Q, \delta^{\tau})$, where, for all q, $p \in Q$ and all $u \in \{-1, 0, 1\}$, we let $\delta^{\tau}(q)(p, u) = \sum_{a\in A(q), w(q,a)=u} \tau(q)(a) \cdot \delta(q, a)(p)$.

Let q, $p \in Q$ and let $s = (q, b^-)$. There is a bijection between $h \in \mathcal{H}_{\mathsf{succ}}(s, (p, b^- - 1))$ and the set of histories of $\mathcal{C}^{\leq\infty}(\mathcal{P})$ that start in $(q, 1)$ and end in $(p, 0)$: one omits all actions and subtracts $b^- - 1$ to all counter values in the history. By definition of σ and δ^{τ}, this bijection preserves the probability of cylinders. This implies that $\delta^{\sigma}_{\mathcal{I}}((q, b^-))((p, b^- - 1))$ is exactly the probability, in $\mathcal{C}^{\leq\infty}(\mathcal{P})$, of terminating in p from $(q, 1)$.

It follows that the characterisation of termination probabilities in one-counter Markov chains of [35] applies to our transition probabilities in this case. We obtain the following.

▶ **Theorem 13** ([35]). *For each $q, p \in Q$, we consider a variable $\langle q \searrow p\rangle$ and the system of equations formed by the equations, for all $q, p \in Q$,*

$$\langle q \searrow p\rangle = \delta^I(q)(p, -1) + \sum_{t\in Q} \delta^I(q)(t, 0) \cdot \langle t \searrow p\rangle + \sum_{t\in Q} \delta^I(q)(t, 1) \cdot \left(\sum_{t'\in Q} \langle t \searrow t'\rangle \cdot \langle t' \searrow p\rangle \right),$$

where $\delta^I(t)(t', u) = \sum_{a\in A(t), w(t,a)=u} \sigma(t, b^-)(a) \cdot \delta(t, a)(t')$ for all t, $t' \in Q$ and all $u \in \{-1, 0, 1\}$. The least non-negative solution of this system is obtained by substituting each variable $\langle q \searrow p\rangle$ by $\delta^{\sigma}_{\mathcal{I}}((q, b^-))((p, b^- - 1))$.

The equation system of Theorem 13 has one variable of the form $\langle q \searrow p \rangle$ for every two states q, $p \in Q$ and there is one equation per variable. Furthermore, the equations have length polynomial in the sizes of $|A|$ and $|Q|$. Indeed, if we distribute all products in the right-hand sides of the equations to rewrite them as a sum of products, there are at most $|A| \cdot |Q|^2$ products of at most four variables or constants. We obtain the following result.

▶ **Lemma 14.** *The equation system of Theorem 13 has $|Q|^2$ variables and equations. Its equations have length polynomial in $|Q|$ and $|A|$.*

4.4.2 Bounded intervals

We now assume that I is bounded. We write $I = [\![b^-, b^+]\!]$ and let $\beta = \log_2(|I| + 1) \in \mathbb{N}_{>0}$. To improve readability, we assume that $b^- = 1$ and $b^+ = 2^\beta - 1$. All results below can be recovered for the general case by adding $b^- - 1$ to the counter values in configurations.

Counter values of I that are kept in $S_{\mathcal{I}}$ can be partitioned in two sets: the set $\{2^\alpha \mid \alpha \in [\![\beta - 1]\!]\}$ of values reachable from b^- and the set $Q \times \{2^\beta - 2^\alpha \mid \alpha \in [\![\beta - 1]\!]\}$ of values reachable from b^+ (in the sense of Figure 4). By symmetry of the transition structure of the compressed Markov chain, the outgoing transitions from a configuration $(q, 2^\alpha)$ correspond to outgoing transitions from the configuration $(q, 2^\beta - 2^\alpha)$.

▶ **Lemma 15.** *Let $q, p \in Q$ and $\alpha \in [\![\beta - 1]\!]$. It holds that $\delta^\sigma_{\mathcal{I}}(q, 2^\alpha)(p, 2^{\alpha+1}) = \delta^\sigma_{\mathcal{I}}(q, 2^\beta - 2^\alpha)(p, 2^\beta)$ and $\delta^\sigma_{\mathcal{I}}(q, 2^\alpha)(p, 0) = \delta^\sigma_{\mathcal{I}}(q, 2^\beta - 2^\alpha)(p, 2^\beta - 2^{\alpha+1})$*

Proof. We only prove that $\delta^\sigma_{\mathcal{I}}(q, 2^\alpha)(p, 2^{\alpha+1}) = \delta^\sigma_{\mathcal{I}}(q, 2^\beta - 2^\alpha)(p, 2^\beta)$ as the other case is similar. We define a bijection $f \colon \mathcal{H}_{\mathsf{succ}}((q, 2^\alpha), (p, 2^{\alpha+1})) \to \mathcal{H}_{\mathsf{succ}}((q, 2^\beta - 2^\alpha), (p, 2^\beta))$ and prove that we have $\mathbb{P}^\sigma_{\mathcal{M}^{\leq r}(\mathcal{Q}),(q,2^\alpha)}(\mathsf{Cyl}(h)) = \mathbb{P}^\sigma_{\mathcal{M}^{\leq r}(\mathcal{Q}),(q,2^\beta - 2^\alpha)}(\mathsf{Cyl}(f(h)))$ for all $h \in \mathcal{H}_{\mathsf{succ}}((q, 2^\alpha), (p, 2^{\alpha+1}))$. This is sufficient to obtain our result.

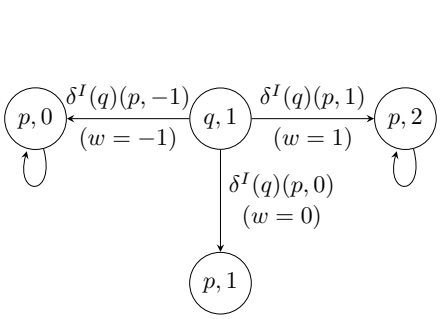
Let $h \in \mathcal{H}_{\mathsf{succ}}((q, 2^\alpha), (p, 2^{\alpha+1}))$. We let $f(h)$ be the history obtained by adding $2^\beta - 2^{\alpha+1}$ to all counter values along h. We must show that $f(h) \in \mathcal{H}_{\mathsf{succ}}((q, 2^\beta - 2^\alpha), (p, 2^\beta))$. For the first and last configurations, we observe that $2^\alpha + 2^\beta - 2^{\alpha+1} = 2^\beta - 2^\alpha$ and $2^{\alpha+1} + 2^\beta - 2^{\alpha+1} = 2^\beta$. For the other configurations, their counter values are in the interval $[\![1, 2^{\alpha+1} - 1]\!]$, thus their counterparts in $f(h)$ have a counter value in $[\![2^\beta - 2^{\alpha+1} + 1, 2^\beta - 1]\!]$.

We now establish that $\mathbb{P}^\sigma_{\mathcal{M}^{\leq r}(\mathcal{Q}),(q,2^\alpha)}(\mathsf{Cyl}(h)) = \mathbb{P}^\sigma_{\mathcal{M}^{\leq r}(\mathcal{Q}),(q,2^\beta - 2^\alpha)}(\mathsf{Cyl}(f(h)))$. Let $h = (q_0, k_0)a_0(q_1, k_1) \dots a_n(q_n, k_n)$. Because σ is memoryless, based on $\mathcal{I}$ and $I \in \mathcal{I}$, we obtain

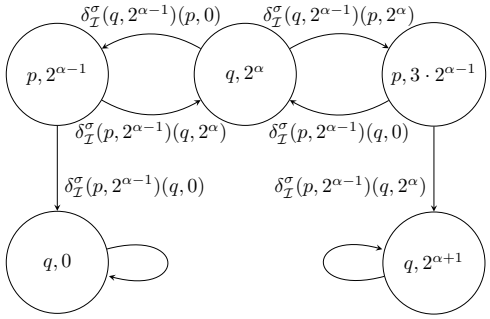
$$\begin{aligned}
\mathbb{P}^\sigma_{\mathcal{M}^{\leq r}(\mathcal{Q}),(q,2^\alpha)}(\mathsf{Cyl}(h)) &= \prod_{\ell=0}^{n-1} \delta(q_\ell, a_\ell)(q_{\ell+1}) \cdot \sigma(s_\ell, k_\ell)(a_\ell) \\
&= \prod_{\ell=0}^{n-1} \delta(q_\ell, a_\ell)(q_{\ell+1}) \cdot \sigma(s_\ell, k_\ell + 2^\beta - 2^{\alpha+1})(a_\ell) \\
&= \mathbb{P}^\sigma_{\mathcal{M}^{\leq r}(\mathcal{Q}),(q,2^\beta - 2^\alpha)}(\mathsf{Cyl}(f(h))).
\end{aligned}$$

To prove that f is bijective, we define its inverse. We let f^{-1} be the function over $\mathcal{H}_{\mathsf{succ}}((q, 2^\beta - 2^\alpha), (p, 2^\beta))$ that subtracts $2^\beta - 2^{\alpha+1}$ to all counter values along histories. It is easy to verify that f^{-1} is well-defined and that it is the inverse of f. ◀

Due to Lemma 15, it is sufficient for us to characterise the outgoing transition probabilities for the configurations in $Q \times \{2^\alpha \mid \alpha \in [\![\beta - 1]\!]\}$. We do so via a quadratic system of equations. We provide intuition on how to derive this system for our interval $I = [\![1, 2^\beta - 1]\!]$ by using Markov chains.



(a) Markov chain transition scheme for the characterisation of transitions from $Q \times \{1\}$, where $\delta^I(q)(p,u) = \sum_{\substack{a \in A(q) \\ w(q,a)=u}} \sigma((q,1))(a) \cdot \delta(q,a)(p)$.



(b) Markov chain transition scheme for the characterisation of transitions from $Q \times \{2^\alpha\}$ for $0 < \alpha < \beta$.

Figure 5 Fragments of Markov chain used to derive a characterisation for transition probabilities of $\mathcal{C}^\sigma_{\mathcal{I}}$ for a bounded interval of the form $[\![1, 2^\beta - 1]\!]$.

Let us first consider transitions from $Q \times \{1\}$. We illustrate the situation in Figure 5a: we consider a Markov chain over $Q \times \{0, 1, 2\}$ where states in $Q \times \{0, 2\}$ are absorbing and transitions from other states are inherited from the Markov chain induced by σ on $\mathcal{M}^{\leq r}(\mathcal{Q})$. We represent transitions in this Markov chain from a configuration $(q, 1) \in Q \times \{1\}$ to configurations with a state $p \in Q$. For any $q \in Q$, the probability of reaching a configuration $s' \in Q \times \{0, 2\}$ from $(q, 1)$ in this Markov chain is $\delta^\sigma_{\mathcal{I}}((q, 1))(s')$ by definition.

Next, we let $\alpha \in [\![1, \beta - 1]\!]$ and consider configurations in $Q \times \{2^\alpha\}$. The situation is depicted in Fig. 5b. We divide a counter change by 2^α into counter changes by $2^{\alpha-1}$ and, thus, rely on the transition probabilities from $Q \times \{2^{\alpha-1}\}$ in $\mathcal{C}^\sigma_{\mathcal{I}}$. In this case, we can see transition probabilities from $Q \times \{2^\alpha\}$ in $\mathcal{C}^\sigma_{\mathcal{I}}$ as reachability probabilities in a Markov chain over $Q \times \{0, 2^{\alpha-1}, 2^\alpha, 3 \cdot 2^{\alpha-1}, 2^{\alpha+1}\}$.

By putting together the reachability systems for $Q \times \{2^\alpha\}$ for all $\alpha \in [\![\beta - 1]\!]$, we obtain a quadratic system of equations. To formalise our system and prove its validity, we introduce some notation. Let $\alpha \in [\![\beta - 1]\!]$, $q, p \in Q$ and $k \in [\![1, 2^{\alpha+1} - 1]\!]$. We let $\mathcal{H}_\alpha((q,k) \nearrow (p, 2^{\alpha+1}))$ (resp. $\mathcal{H}_\alpha((q,k) \searrow (p,0))$) denote the set of histories h of $\mathcal{M}^{\leq r}(\mathcal{Q})$ such that $\mathsf{first}(h) = (q,k)$, $\mathsf{last}(h) = (p, 2^{\alpha+1})$ (resp. $(p,0)$) and no configuration along h besides its last one has a counter value in $\{0, 2^{\alpha+1}\}$. These sets are prefix-free. We let $[(q,k) \nearrow (p, 2^{\alpha+1})]_\alpha = \mathbb{P}^\sigma_{\mathcal{M}^{\leq r}(\mathcal{Q}),(q,k)}(\mathsf{Cyl}(\mathcal{H}_\alpha((q,k) \nearrow (p, 2^{\alpha+1}))))$ and $[(q,k) \searrow (p, 2^{\alpha+1})]_\alpha = \mathbb{P}^\sigma_{\mathcal{M}^{\leq r}(\mathcal{Q}),(q,k)}(\mathsf{Cyl}(\mathcal{H}_\alpha((q,k) \searrow (p, 2^{\alpha+1}))))$. The transition probabilities of $\mathcal{C}^\sigma_{\mathcal{I}}$ can be written with this notation. For all $q, p \in Q$ and $\alpha \in [\![\beta - 1]\!]$, we have $\delta^\sigma_{\mathcal{I}}((q, 2^\alpha))((p,0)) = [(q, 2^\alpha) \searrow (p,0)]_\alpha$ and $\delta^\sigma_{\mathcal{I}}((q, 2^\alpha))((p, 2^{\alpha+1})) = [(q, 2^\alpha) \nearrow (p, 2^{\alpha+1})]_\alpha$.

The following theorem formalises our characterisation of the transition probabilities of $\mathcal{C}^\sigma_{\mathcal{I}}$ for configurations in $S_{\mathcal{I}} \cap (Q \times I)$. The size of this system is polynomial in $|Q|$ and β. We provide a self-contained proof that does not refer to the Markov chains described in Figure 5. A corollary of this proof is that the Markov chains above yield an accurate characterisation of the transition probabilities.

▶ **Theorem 16.** *For each $q, p \in Q$, we consider variables $\langle (q,1) \nearrow (p,2) \rangle_0$ and $\langle (q,1) \searrow (p,0) \rangle_0$, and for all $\alpha \in [\![1, \beta - 1]\!]$ and $k \in \{2^{\alpha-1}, 2^\alpha, 3 \cdot 2^{\alpha-1}\}$, we consider variables $\langle (q,k) \nearrow (p, 2^{\alpha+1}) \rangle_\alpha$ and $\langle (q,k) \searrow (p,0) \rangle_\alpha$. For all q, $p \in Q$ and all $u \in \{-1, 0, 1\}$, let $\delta^I(q)(p,u) = \sum_{a \in A(q), w(q,a)=u} \sigma(q,1)(a) \cdot \delta(q,a)(p)$.*

Consider the system of equations given by, for all $q, p \in Q$:

$$\langle (q,1) \nearrow (p,2) \rangle_0 = \delta^I(q)(p,1) + \sum_{t \in Q} \delta^I(q)(t,0) \cdot \langle (t,1) \nearrow (p,2) \rangle_0 \tag{1}$$
$$\langle (q,1) \searrow (p,0) \rangle_0 = \delta^I(q)(p,-1) + \sum_{t \in Q} \delta^I(q)(t,0) \cdot \langle (t,1) \searrow (p,0) \rangle_0$$

and for all $\alpha \in [\![1, \beta - 1]\!]$,

$$\langle (q, 2^{\alpha-1}) \nearrow (p, 2^{\alpha+1}) \rangle_\alpha = \sum_{t \in Q} \langle (q, 2^{\alpha-1}) \nearrow (t, 2^\alpha) \rangle_{\alpha-1} \cdot \langle (t, 2^\alpha) \nearrow (p, 2^{\alpha+1}) \rangle_\alpha, \tag{2}$$
$$\begin{aligned} \langle (q, 2^{\alpha}) \nearrow (p, 2^{\alpha+1}) \rangle_\alpha = \sum_{t \in Q} \Big(& \langle (q, 2^{\alpha-1}) \nearrow (t, 2^\alpha) \rangle_{\alpha-1} \cdot \langle (t, 3 \cdot 2^{\alpha-1}) \nearrow (p, 2^{\alpha+1}) \rangle_\alpha \\ & + \langle (q, 2^{\alpha-1}) \searrow (t, 0) \rangle_{\alpha-1} \cdot \langle (t, 2^{\alpha-1}) \nearrow (p, 2^{\alpha+1}) \rangle_\alpha \Big), \end{aligned} \tag{3}$$
$$\begin{aligned} \langle (q, 3 \cdot 2^{\alpha-1}) \nearrow (p, 2^{\alpha+1}) \rangle_\alpha = \sum_{t \in Q} & \Big(\langle (q, 2^{\alpha-1}) \searrow (t, 0) \rangle_{\alpha-1} \cdot \langle (t, 2^\alpha) \nearrow (p, 2^{\alpha+1}) \rangle_\alpha \Big) \\ & + \langle (q, 2^{\alpha-1}) \nearrow (p, 2^\alpha) \rangle_{\alpha-1}, \end{aligned} \tag{4}$$

$$\langle (q, 3 \cdot 2^{\alpha-1}) \searrow (p, 0) \rangle_\alpha = \sum_{t \in Q} \langle (q, 2^{\alpha-1}) \searrow (t, 0) \rangle_{\alpha-1} \cdot \langle (t, 2^\alpha) \searrow (p, 0) \rangle_\alpha,$$
$$\begin{aligned} \langle (q, 2^{\alpha}) \searrow (p, 0) \rangle_\alpha = \sum_{t \in Q} \Big(& \langle (q, 2^{\alpha-1}) \searrow (t, 0) \rangle_{\alpha-1} \cdot \langle (t, 2^{\alpha-1}) \searrow (p, 0) \rangle_\alpha \\ & + \langle (q, 2^{\alpha-1}) \nearrow (t, 2^\alpha) \rangle_{\alpha-1} \cdot \langle (t, 3 \cdot 2^{\alpha-1}) \searrow (p, 0) \rangle_\alpha \Big), \end{aligned}$$
$$\begin{aligned} \langle (q, 2^{\alpha-1}) \searrow (p, 0) \rangle_\alpha = \sum_{t \in Q} & \Big(\langle (q, 2^{\alpha-1}) \nearrow (t, 2^\alpha) \rangle_{\alpha-1} \cdot \langle (t, 2^\alpha) \searrow (p, 0) \rangle_\alpha \Big) \\ & + \langle (q, 2^{\alpha-1}) \searrow (p, 0) \rangle_{\alpha-1}. \end{aligned}$$

The least non-negative solution of this system is obtained by substituting each variable $\langle (q,k) \nearrow (p, 2^{\alpha+1}) \rangle_\alpha$ by $[(q,k) \nearrow (p, 2^{\alpha+1})]_\alpha$ and $\langle (q,k) \searrow (p,0) \rangle_\alpha$ by $[(q,k) \searrow (p,0)]_\alpha$.

Proof. Occurrences of $\mathbb{P}$ in this proof refer to $\mathcal{M}^{\leq r}(\mathcal{Q})$, thus we omit it from the notation. We show a claim to shorten our arguments. Let s, $s' \in Q \times [\![r]\!]$. Let $\mathcal{H} \subseteq \mathsf{Hist}(\mathcal{M}^{\leq r}(\mathcal{Q}))$ be a prefix-free set of histories starting in s. Assume that there exist two prefix-free sets of histories $\mathcal{H}^{(1)}$ and $\mathcal{H}^{(2)}$ such that the last (resp. first) configuration of all elements of $\mathcal{H}^{(1)}$ (resp. $\mathcal{H}^{(2)}$) is s' and we have $\mathcal{H} = \{h_1 \cdot h_2 \mid h_1 \in \mathcal{H}^{(1)}, h_2 \in \mathcal{H}^{(2)}\}$. Then it holds that

$$\mathbb{P}^\sigma_s(\mathsf{Cyl}(\mathcal{H})) = \mathbb{P}^\sigma_s(\mathsf{Cyl}(\mathcal{H}^{(1)})) \cdot \mathbb{P}^\sigma_{s'}(\mathsf{Cyl}(\mathcal{H}^{(2)})). \tag{5}$$

Equation (5) can be proven as follows:

$$\begin{aligned}
\mathbb{P}^{\sigma}_{s}(\mathsf{Cyl}(\mathcal{H})) &= \sum_{h_1 \in \mathcal{H}^{(1)}} \sum_{h_2 \in \mathcal{H}^{(2)}} \mathbb{P}^{\sigma}_{s}(\mathsf{Cyl}(h_1 \cdot h_2)) \\
&= \sum_{h_1 \in \mathcal{H}^{(1)}} \sum_{h_2 \in \mathcal{H}^{(2)}} \mathbb{P}^{\sigma}_{s}(h_1) \cdot \mathbb{P}^{\sigma}_{s'}(h_2) \\
&= \left(\sum_{h_1 \in \mathcal{H}^{(1)}} \mathbb{P}^{\sigma}_{s}(\mathsf{Cyl}(h_1)) \right) \cdot \left(\sum_{h_2 \in \mathcal{H}^{(2)}} \mathbb{P}^{\sigma}_{s}(\mathsf{Cyl}(h_2)) \right) \\
&= \mathbb{P}^{\sigma}_{s}(\mathsf{Cyl}(\mathcal{H}^{(1)})) \cdot \mathbb{P}^{\sigma}_{s'}(\mathsf{Cyl}(\mathcal{H}^{(2)}))
\end{aligned}$$

The first line follows from $\mathcal{H}$ being prefix-free. The second line is obtained from the definition of $\mathbb{P}^{\sigma}_{s}$, using the fact that σ is memoryless. We obtain the third line by algebraic manipulations and the last one using the fact that $\mathcal{H}^{(1)}$ and $\mathcal{H}^{(2)}$ are prefix-free.

We now prove the theorem. We start by proving that the asserted solution is a solution of the system. We only verify Equations (1)–(4), i.e., the equations in which the left-hand side of the equation has a variable with the arrow $\nearrow$. Arguments for the others are analogous.

Let $q, p \in Q$. First, we consider the case $\alpha = 0$. We recall that $\mathcal{H}_0((q,1) \nearrow (p,2))$ is prefix-free. We partition $\mathcal{H}_0((q,1) \nearrow (p,2))$ into two sets $\mathcal{H}$ and $\mathcal{H}'$ such that $\mathcal{H}$ is the set of histories starting in $(q,1)$ whose second configuration is $(p,0)$ and $\mathcal{H}' = \mathcal{H}_0((q,1) \nearrow (p,2)) \setminus \mathcal{H}$. For all histories of $\mathcal{H}'$, their second configuration has counter value 1. We rewrite $\mathbb{P}^{\sigma}_{(q,1)}(\mathsf{Cyl}(\mathcal{H}))$ and $\mathbb{P}^{\sigma}_{(q,1)}(\mathsf{Cyl}(\mathcal{H}'))$ to prove the desired equality. On the one hand, we have

$$\mathbb{P}^{\sigma}_{(q,1)}(\mathsf{Cyl}(\mathcal{H})) = \sum_{\substack{a \in A(q) \\ w(q,a)=1}} \mathbb{P}^{\sigma}_{(q,1)}(\mathsf{Cyl}((q,1)a(p,2))) = \delta^{I}(q,1)(p).$$

For the other set, we partition $\mathcal{H}'$ according to the second configuration of the histories. We further partition the resulting sets following the first action and apply Equation (5) to obtain

$$\begin{aligned}
\mathbb{P}^{\sigma}_{(q,1)}(\mathsf{Cyl}(\mathcal{H}')) &= \sum_{t \in Q} \sum_{\substack{a \in A(q) \\ w(q,a)=0}} \sum_{h \in \mathcal{H}_0((t,1) \nearrow (p,2))} \mathbb{P}^{\sigma}_{(q,1)}(\mathsf{Cyl}((q,1)a(t,1) \cdot h)) \\
&= \sum_{t \in Q} \left(\sum_{\substack{a \in A(q) \\ w(q,a)=0}} \sigma(q,1)(a) \cdot \delta(q,a)(t) \right) \cdot \mathbb{P}^{\sigma}_{(t,1)}(\mathsf{Cyl}(\mathcal{H}_0((t,1) \nearrow (p,2)))) \\
&= \sum_{t \in Q} \delta^{I}(q,0)(t) \cdot [(t,1) \nearrow (p,2)]_0.
\end{aligned}$$

Equation (1) thus follows from the above and

$$[(q,1) \nearrow (p,2)]_0 = \mathbb{P}^{\sigma}_{(q,1)}(\mathsf{Cyl}(\mathcal{H})) + \mathbb{P}^{\sigma}_{(q,1)}(\mathsf{Cyl}(\mathcal{H}')).$$

This ends the case where $\alpha = 0$.

Let $\alpha \geq 1$. We start by considering Equation (2). All histories in $\mathcal{H}_\alpha((q, 2^{\alpha-1}) \nearrow (p, 2^{\alpha+1}))$ have a configuration with counter value 2^α. We let $(U_t)_{t \in Q}$ be a partition of $\mathcal{H}_\alpha((q, 2^{\alpha-1}) \nearrow (p, 2^{\alpha+1}))$ based on the state of the first configuration with counter value 2^α that is reached. For all $t \in Q$ and all $h \in U_t$, we let h_1 and h_2 such that $h = h_1 \cdot h_2$ where h_1 is the prefix of h up to the first occurrence of $(t, 2^\alpha)$, and let, for $i \in \{1,2\}$, $U_t^{(i)} = \{h_i \mid h_1 \cdot h_2 \in U_t\}$. We have $U_t^{(1)} = \mathcal{H}_{\alpha-1}((q, 2^{\alpha-1}) \nearrow (t, 2^\alpha))$ and $U_t^{(2)} = \mathcal{H}_\alpha((q, 2^\alpha) \nearrow (t, 2^{\alpha+1}))$

by construction. We conclude that Equation (2) is satisfied by the candidate solution via the following equations (the second line uses Equation (5)):

$$\begin{aligned}[(q,2^{\alpha-1}) \nearrow (p,2^{\alpha+1})]_\alpha &= \sum_{t\in Q} \mathbb{P}^\sigma_{(q,2^{\alpha-1})}(\mathsf{Cyl}(U_t)) \\ &= \sum_{t\in Q} \mathbb{P}^\sigma_{(q,2^{\alpha-1})}(\mathsf{Cyl}(U_t^{(1)})) \cdot \mathbb{P}^\sigma_{(t,2^\alpha)}(\mathsf{Cyl}(U_t^{(2)})) \\ &= \sum_{t\in Q} [(q,2^{\alpha-1}) \nearrow (t,2^\alpha)]_{\alpha-1} \cdot [(t,2^\alpha) \nearrow (p,2^{\alpha+1})]_\alpha.\end{aligned}$$

We now move on to Equation (3). We partition $\mathcal{H}_\alpha((q,2^\alpha) \nearrow (p,2^{\alpha+1}))$ as follows. Let $t \in Q$. We let U_t (resp. D_t) denote the subset of $\mathcal{H}_\alpha((q,2^\alpha) \nearrow (p,2^{\alpha+1}))$ containing the histories such that the first configuration with counter value $3\cdot 2^{\alpha-1}$ (resp. $2^{\alpha-1}$) that is visited has state t and no prior configuration has a counter value in $\{3\cdot 2^{\alpha-1}, 2^{\alpha-1}\}$. The sets D_t and U_t, $t \in Q$ partition, $\mathcal{H}_\alpha((q,2^\alpha) \nearrow (p,2^{\alpha+1}))$. Indeed, all of these sets are disjoint by definition and any history from $(q,2^\alpha)$ to $(p,2^{\alpha+1})$ must traverse a configuration with counter value $3\cdot 2^{\alpha-1}$. Similarly to above (for Equation (2)), for all $t \in Q$ and all $h \in U_t$ (resp. D_t), we let $h = h_1 \cdot h_2$ such that h_1 ends in the configuration witnessing that $h \in U_t$ (resp. D_t). For $i \in \{1,2\}$, we let $U_t^{(i)} = \{h_i \mid h_1 \cdot h_2 \in U_t\}$ and $D_t^{(i)} = \{h_i \mid h_1 \cdot h_2 \in D_t\}$. By applying Equation (5), we obtain:

$$\begin{aligned}[(q,2^\alpha) \nearrow (p,2^{\alpha+1})]_\alpha = &\sum_{t\in Q} \mathbb{P}^\sigma_{(q,2^\alpha)}(\mathsf{Cyl}(U_t^{(1)})) \cdot \mathbb{P}^\sigma_{(t,3\cdot 2^{\alpha-1})}(\mathsf{Cyl}(U_t^{(2)})) \\ &+ \sum_{t\in Q} \mathbb{P}^\sigma_{(q,2^\alpha)}(\mathsf{Cyl}(D_t^{(1)})) \cdot \mathbb{P}^\sigma_{(t,2^{\alpha-1})}(\mathsf{Cyl}(D_t^{(2)}))\end{aligned}$$

We now prove that the cylinder probabilities match the terms in Equation (3). Let $t \in Q$. We have $\mathbb{P}^\sigma_{(t,3\cdot 2^{\alpha-1})}(\mathsf{Cyl}(U_t^{(2)})) = [(t,3\cdot 2^{\alpha-1}) \nearrow (p,2^{\alpha+1})]_\alpha$ and $\mathbb{P}^\sigma_{(t,2^{\alpha-1})}(\mathsf{Cyl}(D_t^{(2)})) = [(t,2^{\alpha-1}) \nearrow (p,2^{\alpha+1})]_\alpha$ because $U_t^{(2)}$ and $D_t^{(2)}$ are respectively the sets $\mathcal{H}_\alpha((t,3\cdot 2^{\alpha-1}) \nearrow (p,2^{\alpha+1}))$ and $\mathcal{H}_\alpha((t,2^{\alpha-1}) \nearrow (p,2^{\alpha+1}))$.

The sets $U_t^{(1)}$ and $D_t^{(1)}$ do not directly match relevant sets of histories as above. However, there are bijections from $U_t^{(1)}$ to $\mathcal{H}_{\alpha-1}((q,2^{\alpha-1}) \nearrow (t,2^\alpha))$ and from $D_t^{(1)}$ to $\mathcal{H}_{\alpha-1}((q,2^{\alpha-1}) \searrow (t,0))$. Both bijections map a history to the history obtained by subtracting $2^{\alpha-1}$ to the counter values in all configurations along the history. All counter values in a history in $U_t^{(1)}$ or $D_t^{(1)}$ and its image lie in the interval I. Therefore, for all $h_1 \in U_t^{(1)} \cup D_t^{(1)}$ with s as its first configuration, if its image by the relevant bijection is h_1' and h_1' starts in s', then $\mathbb{P}^\sigma_s(\mathsf{Cyl}(h_1)) = \mathbb{P}^\sigma_{s'}(\mathsf{Cyl}(h_1'))$. We conclude that $\mathbb{P}^\sigma_{(q,2^\alpha)}(\mathsf{Cyl}(U_t^{(1)})) = [(q,2^{\alpha-1}) \nearrow (t,2^\alpha)]_{\alpha-1}$ and $\mathbb{P}^\sigma_{(q,2^\alpha)}(\mathsf{Cyl}(D_t^{(1)})) = [(q,2^{\alpha-1}) \searrow (t,0)]_{\alpha-1}$ (this argument is detailed further in the proof of Lemma 15). We have shown that the asserted solution satisfies Equation (3).

We now move on to Equation (4). We follow the same scheme as above, i.e., we partition $\mathcal{H}_\alpha((q,3\cdot 2^{\alpha-1}) \nearrow (p,2^{\alpha+1}))$. First, we let U_p be the subset with all histories that never hit counter value 2^α. For any $t \in Q$, we let D_t be the subset of $\mathcal{H}_\alpha((q,3\cdot 2^{\alpha-1}) \nearrow (p,2^{\alpha+1})) \setminus U_p$ consisting of histories that reach counter value 2^α for the first time in a configuration with state t. The sets D_t, $t \in Q$, and U_p partition $\mathcal{H}_\alpha((q,3\cdot 2^{\alpha-1}) \nearrow (p,2^{\alpha+1}))$. As above, for any $t \in Q$ and $h \in D_t$, we write $h = h_1 \cdot h_2$ such that h_1 is the prefix of h up to the first occurrence of $(t,2^\alpha)$. For $i \in \{1,2\}$, we let $D_t^{(i)} = \{h_i \mid h_1 \cdot h_2 \in D_t\}$. Like before, we obtain,

from Equation (5),

$$[(q,3\cdot 2^{\alpha-1}) \nearrow (p,2^{\alpha+1})]_\alpha = \sum_{t\in Q} \mathbb{P}^\sigma_{(q,3\cdot 2^{\alpha-1})}(\mathsf{Cyl}(D_t^{(1)})) \cdot \mathbb{P}^\sigma_{(t,2^\alpha)}(\mathsf{Cyl}(D_t^{(2)})) + \mathbb{P}^\sigma_{(q,3\cdot 2^\alpha)}(\mathsf{Cyl}(U_p)).$$

Let $t \in Q$. It follows from $D_t^{(2)} = \mathcal{H}_\alpha((t,2^\alpha) \nearrow (p,2^{\alpha+1}))$ that $\mathbb{P}^\sigma_{(t,2^\alpha)}(\mathsf{Cyl}(D_t^{(2)})) = [(t,2^\alpha) \nearrow (p,2^{\alpha+1})]_\alpha$. We can adapt the bijection-based argument used for Equation (3) to conclude that $\mathbb{P}^\sigma_{(q,3\cdot 2^{\alpha-1})}(\mathsf{Cyl}(D_t^{(1)})) = [(q,2^{\alpha-1}) \searrow (t,0)]_{\alpha-1}$ and $\mathbb{P}^\sigma_{(q,3\cdot 2^{\alpha-1})}(\mathsf{Cyl}(U_p)) = [(q,2^{\alpha-1}) \nearrow (p,2^\alpha)]_{\alpha-1}$. This shows that Equation (4) is verified by the asserted solution, and ends the argument that all equations hold.

It remains to show that the asserted solution is the least non-negative solution of the system. Once again, we only consider the case of variables with ascending arrows as the other case can be handled similarly. We fix an arbitrary non-negative solution of the system. We denote its component corresponding to a variable x by $x^\star$.

All probabilities in the asserted solution can be written as the probability of a cylinder of a set of histories. In particular, these probabilities can be approximated by only considering the histories with at most n actions (for $n \in \mathbb{N}$). It suffices therefore to show that each approximation is lesser or equal to the fixed arbitrary solution to end the proof.

For all $n \in \mathbb{N}$, $q,p \in Q$, $\alpha \in [\![\beta-1]\!]$ and $k \in \{2^{\alpha-1}, 2^\alpha, 3\cdot 2^{\alpha-1}\}$ if $\alpha \neq 0$ and $k = 1$ otherwise, we let $[(q,k) \nearrow (p,2^{\alpha+1})]_\alpha^{\leq n} = \mathbb{P}^\sigma_{(q,k)}(\mathsf{Cyl}(\mathcal{H}^{\leq n}))$ where $\mathcal{H}^{\leq n}$ is the subset of $\mathcal{H}_\alpha((q,k) \nearrow (p,2^{\alpha+1}))$ containing all histories with at most n actions We define $[(q,k) \searrow (p,0)]_\alpha^{\leq n}$ similarly.

Let q and $p \in Q$. We use nested induction arguments in the remainder of the proof: an outer induction on α and an inner induction on n.

First, we deal with the case $\alpha = 0$. Let $n \in \mathbb{N}$. For the base case $n = 0$, we have $[(q,1) \nearrow (p,2)]_0^{\leq 0} = 0 \leq \langle (q,1) \nearrow (p,2)\rangle_0^\star$ because we consider a non-negative solution. We now assume that $[(q,1) \nearrow (p,2)]_0^{\leq n-1} \leq \langle (q,1) \nearrow (p,2)\rangle_0^\star$ by induction. We can apply the reasoning used when considering Equation (1) in the first part of the proof (taking in account the length of histories) and then apply the induction hypothesis to obtain:

$$\begin{aligned}[(q,1) \nearrow (p,2)]_0^{\leq n} &= \delta^I(q)(p,1) + \sum_{t\in Q} \delta^I(q)(t,0) \cdot [(t,1) \nearrow (p,2)]_0^{\leq n-1} \\ &\leq \delta^I(q)(p,1) + \sum_{t\in Q} \delta^I(q)(t,0) \cdot \langle (t,1) \nearrow (p,2)\rangle_0^\star \\ &= \langle (q,1) \nearrow (p,2)\rangle_0^\star.\end{aligned}$$

This closes the proof for the case $\alpha = 0$.

Next, let $\alpha \geq 1$. We assume, by induction on α, that we have shown that for all $t,t' \in Q$, we have $[(t,2^{\alpha-1}) \nearrow (t',2^\alpha)]_{\alpha-1} \leq \langle (t,2^{\alpha-1}) \nearrow (t',0)\rangle_{\alpha-1}^\star$ and $[(t,2^{\alpha-1}) \searrow (t',2^\alpha)]_{\alpha-1} \leq \langle (t,2^{\alpha-1}) \searrow (t',0)\rangle_{\alpha-1}^\star$. The base case $n = 0$ of the inner induction is direct because for all $k \in \{2^{\alpha-1}, 2^\alpha, 3\cdot 2^{\alpha-1}\}$, we have $[(q,k) \nearrow (p,2^{\alpha+1})]_\alpha^{\leq 0} = 0$.

We assume by induction that $[(t,k) \nearrow (t',2^{\alpha+1})]_\alpha^{\leq n} \leq \langle (t,k) \nearrow (t',2^{\alpha+1})\rangle_\alpha^\star$ for all $t,t' \in Q$ and all $k \in \{2^{\alpha-1}, 2^\alpha, 3\cdot 2^{\alpha-1}\}$. All required inequalities are obtained by an adaptation of the argument used in the first part of the proof for Equations (2), (3) and (4), i.e., partitioning the set of histories while taking in account the length of histories and invoking Equation (5). For this reason, we omit some details. From configuration $(q,2^{\alpha-1})$,

we obtain that

$$[(q,2^{\alpha-1}) \nearrow (p,2^{\alpha+1})]^{\leq n}_{\alpha} \leq \sum_{t\in Q} [(q,2^{\alpha-1}) \nearrow (t,2^{\alpha})]_{\alpha-1} \cdot [(t,2^{\alpha}) \nearrow (p,2^{\alpha+1})]^{\leq n-1}_{\alpha}.$$

By the induction hypotheses and the fact we are dealing with a solution of the system, we obtain $[(q,2^{\alpha-1}) \nearrow (p,2^{\alpha+1})]^{\leq n}_{\alpha} \leq \langle (q,2^{\alpha-1}) \nearrow (p,2^{\alpha+1})\rangle^{\star}_{\alpha}$. Next, for configuration $(q,2^{\alpha})$, we obtain that

$$[(q,2^{\alpha}) \nearrow (p,2^{\alpha+1})]^{\leq n}_{\alpha} \leq \sum_{t\in Q} \Big([(q,2^{\alpha-1}) \nearrow (t,2^{\alpha})]_{\alpha-1} \cdot [(t,3\cdot 2^{\alpha-1}) \nearrow (p,2^{\alpha+1})]^{\leq n-1}_{\alpha} + [(q,2^{\alpha-1}) \searrow (t,0)]_{\alpha-1} \cdot [(t,2^{\alpha-1}) \nearrow (p,2^{\alpha+1})]^{\leq n-1}_{\alpha} \Big).$$

It follows from the induction hypotheses and the fact we deal with a solution that $[(q,2^{\alpha}) \nearrow (p,2^{\alpha+1})]^{\leq n}_{\alpha} \leq \langle (q,2^{\alpha}) \nearrow (p,2^{\alpha+1})\rangle^{\star}_{\alpha}$. Finally, for configuration $(q,3\cdot 2^{\alpha-1})$, we have

$$[(q,3\cdot 2^{\alpha-1}) \nearrow (p,2^{\alpha+1})]^{\leq n}_{\alpha} \leq \sum_{t\in Q} \Big([(q,2^{\alpha-1}) \searrow (t,0)]_{\alpha-1} \cdot [(t,2^{\alpha}) \nearrow (p,2^{\alpha+1})]^{\leq n-1}_{\alpha} \Big) + [(q,2^{\alpha-1}) \nearrow (p,2^{\alpha})]_{\alpha-1}.$$

The induction hypotheses imply that $[(q,3\cdot 2^{\alpha-1}) \nearrow (p,2^{\alpha+1})]^{\leq n}_{\alpha} \leq \langle (q,3\cdot 2^{\alpha-1}) \nearrow (p,2^{\alpha+1})\rangle^{\star}_{\alpha}$.

We have shown that the asserted solution is the least non-negative solution of the system. ◀

We now analyse the size of the equation system of Theorem 16. There are as many equations as there are variables. There are $2\cdot|Q|^2$ equations in the system where the variable of the left-hand side is indexed by 0, and, for all $\alpha \in [\![1,\beta-1]\!]$, there are $6\cdot|Q|^2$ equations in the system where the variable of the left-hand side is indexed by α. We can also show that these equations have length polynomial in $|Q|$ and $|A|$. We obtain the following result.

▶ **Lemma 17.** *The equation system of Theorem 16 has $2\cdot|Q|^2\cdot(3\beta-2)$ variables and equations. Its equations have length polynomial in $|Q|$ and $|A|$.*

Proof. The argument regarding the number of variables and equations is given above. We thus provide an analysis of the length of the equations. We analyse Equations (1)–(4) from Theorem 16. A similar analysis applies to the other equations. We need only comment on the right-hand side of each equation, as the left-hand side contains a single variable.

We start with Equation (1). If we rewrite its right-hand side as a sum of products (we substitute references to δ^I by the corresponding sum), we obtain a sum of no more than $|Q|\cdot|A|$ products of at most three variables or constants. For Equations (2)–(4), we observe that their right-hand side are respectively sums of no more than $2|Q|$ products of two variables. This ends the proof. ◀

Theorem 16 provides a system of equations that may not have a unique solution. We describe how to alter this system to have a unique solution based on the supports of the distributions assigned by σ.

We rely on the Markov chains described in Figure 5. By Theorem 16, the transition probabilities of $\mathcal{C}^{\sigma}_{\mathcal{I}}$ are reachability probabilities in these Markov chains. More precisely, the system of Theorem 16 is a collection of systems for reachability probabilities in these Markov

chains. It follows that modifying the equation system of Theorem 16 by setting all relevant probabilities to zero will ensure uniqueness of the solution.

It remains to determine how to identify the probabilities that are zero in the least solution of the system. As explained in Section 2, the probabilities that are zero only depend on the transitions (with non-zero probability) between configurations. Therefore, we need not compute the transition probabilities of the Markov chains (which would have an important computational cost, see Example 7) and need only determine the transitions qualitatively.

The idea of the procedure is to proceed gradually increasing the counter step size. First, we can study the Markov chain for counter values $\{0, 1, 2\}$, as illustrated in Figure 5a, and perform a graph-based analysis to determine which probabilities to set to zero for outgoing transitions from $Q \times \{1\}$ in $\mathcal{C}^\sigma_{\mathcal{I}}$. Then, for all $\alpha \in [\![\beta - 1]\!]$, assuming that the non-zero transition probabilities in $\mathcal{C}^\sigma_{\mathcal{I}}$ have been determined for configurations in $Q \times \{2^{\alpha-1}\}$, we can perform another graph-based analysis on the Markov chain described in Figure 5b to determine the non-zero transition probabilities from $Q \times \{2^\alpha\}$ in $\mathcal{C}^\sigma_{\mathcal{I}}$.

In this way, we obtain a procedure that runs in time polynomial in $|Q|$ and β: we perform a reachability analysis on one graph of size $3 \cdot |Q|$ for the base case and on $\beta - 1$ graphs of size $5 \cdot |Q|$ for the other cases. This analysis does not require the precise probabilities given by σ, and it is sufficient to only know which actions are chosen with positive probabilities in $Q \times I$. When given the precise probabilities, the system resulting from this procedure can be solved in polynomial time in the BSS model; by construction, its unique solution can be computed by solving β linear systems. We summarise this result in the following theorem.

▶ **Theorem 18.** *There exists an algorithm modifying the system of Theorem 16 such that (i) the least solution of the original system is the unique solution of the modified one and (ii) the algorithm runs in time polynomial in β and the representation size of $\mathcal{Q}$. This algorithms only relies on the support of the distributions in the image of σ and not the precise probabilities. The resulting system can be solved in polynomial time in the BSS model.*

Proof. We first formalise the Markov chains of Figure 5. The remainder of our argument is based on these β Markov chains. We let $\mathcal{C}_0 = (\{\bot\} \cup (Q \times \{0, 1, 2\}), \delta^I_0)$ be the Markov chain such that all states in $\{\bot\} \cup (Q \times \{0, 2\})$ are absorbing and for all $q, p \in Q$ and $u \in \{-1, 0, 1\}$, $\delta^I_0((q, 1))((p, 1 + u)) = \sum_{a \in A(q), w(q,a)=u} \sigma(q, 1)(a) \cdot \delta(q, a)(p)$ (unattributed probability goes to $\bot$). For all $\alpha \in [\![1, \beta - 1]\!]$, we let $\mathcal{C}_\alpha = (\{\bot\} \cup (Q \times \{0, 2^{\alpha-1}, 2^\alpha, 3 \cdot 2^{\alpha-1}, 2^{\alpha+1}\}), \delta^I_\alpha)$ where the states in $\{\bot\} \cup (Q \times \{0, 2^{\alpha+1}\})$ are absorbing and, for all $q, p \in Q$ and $k \in \{2^{\alpha-1}, 2^\alpha, 3 \cdot 2^{\alpha-1}\}$, we let $\delta_\alpha((q, k))((p, k - 2^{\alpha-1})) = \delta^\sigma_{\mathcal{I}}((q, 2^{\alpha-1}))((p, 0))$ and $\delta_\alpha((q, k))((p, k + 2^{\alpha-1})) = \delta^\sigma_{\mathcal{I}}((q, 2^{\alpha-1}))((p, 2^\alpha))$.

We observe that for all $p \in Q$ and all $\alpha \in [\![\beta - 1]\!]$, the subset of equations from Theorem 16 with the variables of the form $\langle (q, k) \nearrow (p, 2^{\alpha+1}) \rangle_\alpha$ (resp. $\langle (q, k) \searrow (p, 0) \rangle_\alpha$) in the left-hand side coincides with a system for reachability probabilities in $\mathcal{C}_\alpha$ for target $\{(p, 2^{\alpha+1})\}$ (resp. $\{(p, 0)\}$) when substituting variables indexed by $\alpha - 1$ by their assignment in the least solution of the system. We devise an algorithm that individually modifies every such system so it has a unique solution. Because we are dealing with systems for reachability probabilities, we obtain a system with a unique solution by setting the variables whose assignment in the least solution of the system is zero to zero (see Section 2). This set of variables can be determined using a qualitative reachability analysis of the Markov chains $\mathcal{C}_\alpha$.

We analyse the Markov chains in order, i.e., we start with $\mathcal{C}_0$, then continue with $\mathcal{C}_1$ and so on. This is necessary: the transitions with non-zero probabilities in a Markov chain $\mathcal{C}_\alpha$ with $\alpha \geq 1$ is not known beforehand, but can be inferred from the analysis of $\mathcal{C}_{\alpha-1}$. We prove the following invariant of our procedure: after processing $\mathcal{C}_\alpha$, the least non-negative solution of the modified system is the least non-negative solution of the original system

and all variables indexed by $\alpha' \leq \alpha$ have a unique valid assignment in any solution of the new system. We modify the system by adding constraints that are satisfied by the least non-negative solution of the original system. Thus, the first part of the invariant follows directly and we do not comment on it.

The transition structure of the Markov chain $\mathcal{C}_0$ can be constructed directly as follows: there exists a transition from a state $(q,1)$ to a state $(p,1+u)$ if and only if there exists an action $a \in \mathsf{supp}(\sigma(q,1))$ such that $w(q,a) = u$ and $p \in \mathsf{supp}(\delta(q,a))$ (in particular, the numerical probabilities do not matter). This yields a directed graph G_0 over $Q \times \{0,1,2\}$. For all $q,p \in Q$, we have $[(q,1) \nearrow (p,2)]_0 = 0$ (resp. $[(q,1) \searrow (p,0)]_0 = 0$) if and only if $(p,2)$ (resp. $(p,0)$) cannot be reached from $(q,1)$ in G_0, and in this case, we add $\langle (q,1) \nearrow (p,2) \rangle_0 = 0$ (resp. $\langle (q,1) \searrow (p,0) \rangle_0 = 0$) to the system. After analysing $\mathcal{C}_0$, the invariant is satisfied. Indeed, following the addition of the new equations, there is only one possible assignment of the variables indexed by 0 in any solution: all of these variables are involved in a Markov chain reachability probability system with a unique solution.

We now let $\alpha \in [\![1, \beta - 1]\!]$ and assume that $\mathcal{C}_{\alpha-1}$ has been processed. We assume that the invariant holds by induction. Through the analysis of $\mathcal{C}_{\alpha-1}$, we know which transitions of $\mathcal{C}_\alpha$ have non-zero probability because these probabilities are reachability probabilities in $\mathcal{C}_{\alpha-1}$ by Theorem 16. We construct a directed graph G_α over the state space of $\mathcal{C}_\alpha$ similarly to above. Let $s = (q,k) \in Q \times \{2^{\alpha-1}, 2^\alpha, 3 \cdot 2^{\alpha-1}\}$ and $p \in Q$. In G_α, there is an edge from s to $(p, k + 2^{\alpha-1})$ if $[(q, 2^{\alpha-1}) \nearrow (p, 2^\alpha)]_{\alpha-1} > 0$ and there is an edge from s to $(p, k - 2^{\alpha-1})$ if $[(q, 2^{\alpha-1}) \searrow (p,0)]_{\alpha-1} > 0$. Whether these probabilities are positive is known from the analysis of $\mathcal{C}_{\alpha-1}$. As above, we have $[(q,k) \nearrow (p, 2^{\alpha+1})]_\alpha = 0$ (resp. $[(q,k) \searrow (p,0)]_\alpha = 0$) if and only if $(p, 2^{\alpha+1})$ (resp. $(p,0)$) cannot be reached from (q,k) in G_α, and in this case, we add $\langle (q,k) \nearrow (p, 2^{\alpha+1}) \rangle_\alpha = 0$ (resp. $\langle (q,k) \searrow (p,0) \rangle_\alpha = 0$) to the system.

We prove that the invariant is preserved after this iteration. By induction, all variables indexed by $\alpha - 1$ have only one possible valid assignment. Given a solution of the system obtained after processing $\mathcal{C}_\alpha$, the variables indexed by α must satisfy an equation system with a unique solution, the coefficients of which are given by the unique valid assignment of the variables indexed by $\alpha - 1$. It follows that there can only be one valuation for the variables indexed by α in any solution of this system. This, in addition to the inductive hypothesis, guarantees that the invariant holds after the analysis of $\mathcal{C}_\alpha$. In the end, after analysing $\mathcal{C}_{\beta-1}$, the invariant guarantees that the resulting system has a unique solution.

To end the proof, it remains to show that the above algorithm respects the asserted complexity bounds. For all $\alpha \in [\![\beta - 1]\!]$, constructing the graph G_α takes time polynomial in the representation of $\mathcal{Q}$. Indeed, for G_0, to find all successors of a configuration $(q,1)$, it suffices to iterate over all actions $a \in \mathsf{supp}(\sigma(q,1))$ and then build on the set of successors $\mathsf{supp}(\delta(q,a))$. For the other graphs G_α, their structure is inferred from the analysis of $G_{\alpha-1}$. Each graph G_α can be analysed in polynomial time by performing a backward reachability analysis from each configuration on the right (i.e., after the arrow) of a variable indexed by α, and there are $2|Q|$ such configurations per graph. As we analyse β graphs, the overall time required to implement the procedure above respects the announced complexity bounds.

It remains to prove that the unique solution of the system provided by the procedure above can be computed in polynomial time in the BSS model. It suffices to solve linear systems for reachability probabilities in each of the Markov chains $\mathcal{C}_\beta$ for $\alpha \in [\![\beta]\!]$. This can be done in polynomial time with unit-cost arithmetic: these Markov chains have no more than $5 \cdot |Q|$ states each. ◀

4.5 Finite representations of compressed Markov chains

We have not imposed any conditions on the memoryless strategy σ: the compressed Markov chain $\mathcal{C}^\sigma_\mathcal{I}$ can be defined without assuming that σ is an OEIS or a CIS. However, for algorithmic purposes, we require that $\mathcal{C}^\sigma_\mathcal{I}$ has a finite representation that is amenable to verification algorithms. In this section, we focus on the representation of the state space of $\mathcal{C}^\sigma_\mathcal{I}$, as the results of Section 4.4 provide a finite representation of transition probabilities for each interval.

By construction, $\mathcal{C}^\sigma_\mathcal{I}$ is finite if and only if $\mathcal{I}$ is finite. Thus $\mathcal{C}^\sigma_\mathcal{I}$ can only be finite when σ is an OEIS. In the remainder of this section, our goal is to show that $\mathcal{C}^\sigma_\mathcal{I}$ has a finite representation when σ is a CIS and $\mathcal{I}$ is periodic. We assume that $r = \infty$ and that σ is a CIS that for the remainder of the section. We let ρ denote a common period of σ and $\mathcal{I}$. We let $\mathcal{J}$ be the partition of $[\![1, \rho]\!]$ induced by $\mathcal{I}$.

We claim that $\mathcal{C}^\sigma_\mathcal{I}$ is induced by a one-counter Markov chain $\mathcal{P}^\sigma_\mathcal{J} = (P_\mathcal{J}, \delta^\sigma_\mathcal{J})$ where $P_\mathcal{J} = S_\mathcal{I} \cap (\{\bot\} \cup (Q \times [\![1, \rho]\!]))$ and $\delta^\sigma_\mathcal{J}$ is described below. We first explain the interpretation of configurations before giving intuition on $\delta^\sigma_\mathcal{J}$. Let $((q, k), k')$ be a configuration of $\mathcal{C}^{\leq\infty}(\mathcal{P}^\sigma_\mathcal{J})$ such that $k' \geq 1$ or $k = \rho$ (configurations that do not satisfy these conditions will be unreachable and we ignore them). This configuration corresponds to the configuration $(q, \rho \cdot (k' - 1) + k) \in S_\mathcal{I}$. The counter value k keeps track of where in the period we are and the counter value k' indicates how many multiples of ρ the counter has exceeded. This correspondence guarantees that the configuration $((q, \rho), 0)$ of $\mathcal{C}^{\leq\infty}(\mathcal{P}^\sigma_\mathcal{J})$ represents the configuration $(q, 0) \in S_\mathcal{I}$.

Transitions are defined so that successors in $\mathcal{C}^{\leq\infty}(\mathcal{P}^\sigma_\mathcal{J})$ correspond to successors in $\mathcal{C}^\sigma_\mathcal{I}$. We formalise $\delta^\sigma_\mathcal{J} \colon P_\mathcal{J} \to \mathcal{D}(P_\mathcal{J} \times \{-1, 0, 1\})$ as follows. Like before, $\bot$ is absorbing and we give a weight of zero to its self-loop to ensure that we cannot terminate in $\bot$. In other words, we set $\delta^\sigma_\mathcal{J}(\bot)(\bot, 0) = 1$. Let $s = (q, k) \in P_\mathcal{J}$. Each transition from s in $\mathcal{C}^\sigma_\mathcal{I}$ to a state in $P_\mathcal{J}$ yields a transition with weight zero in $\mathcal{P}^\sigma_\mathcal{J}$, i.e., for all $s' \in P_\mathcal{J}$, we let $\delta^\sigma_\mathcal{J}(s)(s', 0) = \delta^\sigma_\mathcal{I}(s)(s')$. In particular, all incoming transitions of $\bot$ have weight zero. Any transition from s to a configuration $(p, 0)$ in $\mathcal{C}^\sigma_\mathcal{I}$ induces a transition from s to (p, ρ) in $\mathcal{P}^\sigma_\mathcal{J}$ with a weight of -1, i.e., we let $\delta^\sigma_\mathcal{J}(s)((p, \rho), -1) = \delta^\sigma_\mathcal{I}(s)((p, 0))$. Intuitively, in this case, we go back to the previous period. Finally, any transition from s to the configuration $(p, \rho + 1)$ in $\mathcal{C}^\sigma_\mathcal{I}$ yields a transition with a weight of 1 in $\mathcal{P}^\sigma_\mathcal{J}$ from s to $(p, 1) \in P_\mathcal{J}$ (this configuration is guaranteed to be in $S_\mathcal{I}$ because 1 is the minimum of the first interval and $\mathcal{I}$ has period ρ), i.e., we let $\delta^\sigma_\mathcal{J}(s)((p, 1), 1) = \delta^\sigma_\mathcal{I}(s)((p, \rho + 1))$. Intuitively, in this case, we have passed a multiple of ρ. We obtain a well-defined transition function with the above: for all counter values k of configurations in $P_\mathcal{J}$, the successor counter values of k are a counter value of a configuration in $P_\mathcal{J}$, 0 or $\rho + 1$, i.e., the upper and lower bound respectively of the intervals adjacent to $[\![1, \rho]\!]$ in $\mathcal{I} \cup \{[\![0]\!]\}$.

We now show that the termination probabilities in $\mathcal{C}^\sigma_\mathcal{I}$ and in $\mathcal{C}^{\leq\infty}(\mathcal{P}^\sigma_\mathcal{J})$ match from all initial configurations with the correspondence outlined previously.

▶ **Theorem 19.** *For all $(q, k) \in P_\mathcal{J} \setminus \{\bot\}$ and $k' \in \mathbb{N}$ such that $k' \geq 1$ or $k = \rho$ and all $p \in Q$, we have $\mathbb{P}^\sigma_{\mathcal{C}^\sigma_\mathcal{I}, (q, \rho\cdot(k'-1)+k)}(\mathsf{Reach}((p, 0))) = \mathbb{P}^{\sigma}_{\mathcal{C}^{\leq\infty}(\mathcal{P}^\sigma_\mathcal{J}), ((q,k),k')}(\mathsf{Term}((p, \rho)))$.*

Proof. We define an injective mapping $f \colon S_\mathcal{I} \setminus \{\bot\} \to P_\mathcal{J} \times \mathbb{N}$ such that, for any $s = (q, k) \in S_\mathcal{I}$, if k is divisible by ρ, we let $f(s) = ((q, \rho), \frac{k}{\rho})$, and otherwise, we let $f(s) = ((q, k \bmod \rho), \lfloor \frac{k}{\rho} \rfloor + 1)$. We observe that the image of f is the set of configurations $((q, k), k')$ of $\mathcal{C}^{\leq\infty}(\mathcal{P}^\sigma_\mathcal{J})$ such that $k' \geq 1$ or $k = \rho$. The configurations of $\mathcal{C}^{\leq\infty}(\mathcal{P}^\sigma_\mathcal{J})$ with a state other than $\bot$ that are not in the image of f have no incoming transitions in $\mathcal{C}^{\leq\infty}(\mathcal{P}^\sigma_\mathcal{J})$ and are

absorbing by construction. The statement of the theorem is equivalent to showing that for all $s \in S_{\mathcal{I}}$ and all $p \in Q$, we have $\mathbb{P}_{\mathcal{C}^\sigma_{\mathcal{I}},s}(\mathsf{Reach}((p,0))) = \mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{P}^\sigma_{\mathcal{J}}),f(s)}(\mathsf{Term}((p,\rho)))$.

Let $[\delta^\sigma_{\mathcal{J}}]^{\leq\infty}$ denote the transition function of $\mathcal{C}^{\leq\infty}(\mathcal{P}^\sigma_{\mathcal{J}})$. The crux of the proof is to establish that for all s, $s' \in S_{\mathcal{I}}$, we have $\delta^\sigma_{\mathcal{I}}(s)(s') = [\delta^\sigma_{\mathcal{J}}]^{\leq\infty}(f(s))(f(s'))$. To refer to this property, we say that f preserves transitions. Let $s = (q,k) \in S_{\mathcal{I}}$. If $k = 0$, we have $\delta^\sigma_{\mathcal{I}}(s)(s) = 1 = [\delta^\sigma_{\mathcal{J}}]^{\leq\infty}(f(s))(f(s)) = [\delta^\sigma_{\mathcal{J}}]^{\leq\infty}(((q,\rho),0))((q,\rho),0))$ since configurations with counter value zero are absorbing. We thus assume that $k > 0$.

We distinguish two cases below. First, we assume that $\rho = 1$, i.e., the strategy σ is counter-oblivious. It follows that for all $(p,k') \in Q \times \mathbb{N}$, we have $f((p,k')) = ((p,1),k')$. The successor counter values of k are $k-1$ and $k+1$ because $\rho = 1$. Let $p \in Q$, $u \in \{-1,1\}$ and $s' = (p, k+u)$. By definition of $\mathcal{P}^\sigma_{\mathcal{J}}$ and of $\mathcal{C}^\sigma_{\mathcal{I}}$ (in particular, its periodic structure), we have

$$[\delta^\sigma_{\mathcal{J}}]^{\leq\infty}(f(s))(f(s')) = \delta^\sigma_{\mathcal{J}}((q,1))((p,1),u) = \delta^\sigma_{\mathcal{I}}((q,1))((p,1+u)) = \delta^\sigma_{\mathcal{I}}(s)(s').$$

This ends the proof that f preserves transitions when $\rho = 1$.

Now, we assume that $\rho > 1$. First, we assume that k is divisible by ρ, i.e., that $f(s) = ((q,\rho), \frac{k}{\rho})$. Successor counter values of k are $k+1$ and $k-1$, since multiples of ρ are the maximum of their interval in $\mathcal{I}$. Let $p \in Q$. First, we consider $s' = (p, k+1)$. In this case, we have $f(s') = ((p,1), \frac{k}{\rho} + 1)$. By definition of $\mathcal{P}^\sigma_{\mathcal{J}}$ and of $\mathcal{C}^\sigma_{\mathcal{I}}$, we have

$$[\delta^\sigma_{\mathcal{J}}]^{\leq\infty}(f(s))(f(s')) = \delta^\sigma_{\mathcal{J}}((q,\rho))((p,1),1) = \delta^\sigma_{\mathcal{I}}((q,\rho))((p,\rho+1)) = \delta^\sigma_{\mathcal{I}}(s)(s').$$

Now, we consider $s' = (p, k-1)$. We obtain $f(s') = ((p,\rho-1), \frac{k}{\rho})$ and

$$[\delta^\sigma_{\mathcal{J}}]^{\leq\infty}(f(s))(f(s')) = \delta^\sigma_{\mathcal{J}}((q,\rho))((p,\rho-1),0) = \delta^\sigma_{\mathcal{I}}((q,\rho))((p,\rho-1)) = \delta^\sigma_{\mathcal{I}}(s)(s').$$

We have shown that f preserves the transitions from s whenever k is a multiple of ρ.

We now assume that k is not a multiple of ρ, and thus that $f(s) = ((q, k \bmod \rho), \lfloor \frac{k}{\rho} \rfloor + 1)$. Let $p \in Q$, k' be a successor counter value of k and $s' = (p,k')$. Let $I = [\![b^-, b^+]\!] \in \mathcal{I}$ such that $k \in I$. It follows that $k' \in [\![b^- - 1, b^+ + 1]\!]$. Since multiples of ρ are upper bounds of intervals, this implies that $k' \in [\![\lfloor \frac{k}{\rho} \rfloor \cdot \rho, (\lfloor \frac{k}{\rho} \rfloor + 1) \cdot \rho + 1]\!]$. In light of this, we distinguish four cases: $k' = \lfloor \frac{k}{\rho} \rfloor \cdot \rho$, $k' = (\lfloor \frac{k}{\rho} \rfloor + 1) \cdot \rho$, $k' = (\lfloor \frac{k}{\rho} \rfloor + 1) \cdot \rho + 1$ and finally k' is in none of the previous cases.

First, we assume that $k' = \lfloor \frac{k}{\rho} \rfloor \cdot \rho$, which implies that $f(s') = ((p,\rho), \lfloor \frac{k}{\rho} \rfloor)$. We have

$$[\delta^\sigma_{\mathcal{J}}]^{\leq\infty}(f(s))(f(s')) = \delta^\sigma_{\mathcal{J}}((q, k \bmod \rho))((p,\rho),-1) = \delta^\sigma_{\mathcal{I}}((q, k \bmod \rho))((p,0)) = \delta^\sigma_{\mathcal{I}}(s)(s').$$

Second, we assume that $k' = (\lfloor \frac{k}{\rho} \rfloor + 1) \cdot \rho$. This implies that $f(s') = ((p,\rho), \lfloor \frac{k}{\rho} \rfloor + 1)$. It holds that

$$[\delta^\sigma_{\mathcal{J}}]^{\leq\infty}(f(s))(f(s')) = \delta^\sigma_{\mathcal{J}}((q, k \bmod \rho))((p,\rho),0) = \delta^\sigma_{\mathcal{I}}((q, k \bmod \rho))((p,\rho)) = \delta^\sigma_{\mathcal{I}}(s)(s').$$

Third, we assume that $k' = (\lfloor \frac{k}{\rho} \rfloor + 1) \cdot \rho + 1$. This implies that $f(s') = ((p,1), \lfloor \frac{k}{\rho} \rfloor + 2)$. It follows that

$$[\delta^\sigma_{\mathcal{J}}]^{\leq\infty}(f(s))(f(s')) = \delta^\sigma_{\mathcal{J}}((q, k \bmod \rho))((p,1),1) = \delta^\sigma_{\mathcal{I}}((q, k \bmod \rho))((p,1)) = \delta^\sigma_{\mathcal{I}}(s)(s').$$

Finally, we assume that none of the previous cases holds. We conclude that $f(s') = ((p, k' \bmod \rho), \lfloor \frac{k}{\rho} \rfloor + 1)$. We obtain

$$\begin{aligned}
[\delta^\sigma_{\mathcal{J}}]^{\leq\infty}(f(s))(f(s')) &= \delta^\sigma_{\mathcal{J}}((q, k \bmod \rho))((p, k' \bmod \rho), 0) \\
&= \delta^\sigma_{\mathcal{I}}((q, k \bmod \rho))((p, k' \bmod \rho)) \\
&= \delta^\sigma_{\mathcal{I}}(s)(s').
\end{aligned}$$

This ends the proof that f preserves transitions.

We lift f to histories by letting, for all $h = s_1 \dots s_n \in \mathsf{Hist}(\mathcal{C}^\sigma_\mathcal{I})$ in which $\bot$ does not occur, $f(h) = f(s_1) \dots f(s_n)$ and we obtain, since f preserves transitions, $\mathbb{P}_{\mathcal{C}^\sigma_\mathcal{I},\mathsf{first}(h)}(\mathsf{Cyl}(h)) = \mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{P}^\sigma_\mathcal{J}),f(\mathsf{first}(h))}(\mathsf{Cyl}(f(h)))$. The claim of the theorem follows by writing the objectives as disjoint unions of cylinders and using the fact that f is injective. ◀

5 Verification

We present algorithms for the interval strategy verification problem based on the compressed Markov chains of Section 4. In Section 5.1, we present a polynomial time algorithm in the BSS model of computation for the verification of OEISs in bounded OC-MDPs. Sections 5.2 and 5.3 present a reduction from the verification problem for OEISs and CISs respectively to checking the validity of a universal formula in the theory of the reals. Throughout this section, we invoke the Refine and Isolate operators over interval partitions from Section 4.2.

We consider the following inputs: an OC-MDP $\mathcal{Q} = (Q, A, \delta, w)$, a counter upper bound $r \in \bar{\mathbb{N}}_{>0}$, an OEIS or CIS σ of $\mathcal{M}^{\leq r}(\mathcal{Q})$, an initial configuration $s_{\mathsf{init}} = (q_{\mathsf{init}}, k_{\mathsf{init}}) \in Q \times [\![r]\!]$, a set of targets $T \subseteq Q$ and a threshold $\theta \in [0,1] \cap \mathbb{Q}$.

We describe the algorithms in a unified fashion for both the selective termination objective $\mathsf{Term}(T)$ and the state-reachability objective $\mathsf{Reach}(T)$. We let $\Omega \in \{\mathsf{Term}(T), \mathsf{Reach}(T)\}$ denote the objective. The major difference between the algorithms for selective termination and state-reachability is with respect to the studied OC-MDP: analysing the state-reachability probabilities requires a (polynomial-time) modification of $\mathcal{Q}$ beforehand (see Theorem 12). We assume that this modification has been applied if $\Omega = \mathsf{Reach}(T)$. To further unify notation, we let $T_\Omega = T \times \{0\}$ if $\Omega = \mathsf{Term}(T)$ or $r = \infty$ and $T_\Omega = T \times \{0, r\}$ otherwise. This choice is motivated by the fact that, for all partitions $\mathcal{I}$ of $[\![1, r-1]\!]$ for which $\mathcal{C}^\sigma_\mathcal{I} = (S_\mathcal{I}, \delta^\sigma_\mathcal{I})$ is well-defined and $s_{\mathsf{init}} \in S_\mathcal{I}$, Theorems 11 and 12 ensure that $\mathbb{P}^\sigma_{\mathcal{M}^{\leq r}(\mathcal{Q}),s_{\mathsf{init}}}(\Omega) = \mathbb{P}_{\mathcal{C}^\sigma_\mathcal{I},s_{\mathsf{init}}}(\mathsf{Reach}(T_\Omega))$.

5.1 Verification in bounded one-counter Markov decision processes

We assume that $r \in \mathbb{N}_{>0}$. We provide a $\mathsf{P}^{\mathsf{PosSLP}}$ upper bound on the complexity of the OEIS verification problem in bounded OC-MDPs. We assume that $r \in \mathbb{N}_{>0}$. Let $\mathcal{I}'$ be the partition of $[\![1, r-1]\!]$ given by the description of σ. We let $\mathcal{I} = \mathsf{Refine}(\mathsf{Isolate}(\mathcal{I}', k_{\mathsf{init}}))$. It follows that σ is based on $\mathcal{I}$ and that $s_{\mathsf{init}} \in S_\mathcal{I}$ (because k_{init} is a bound of an interval in $\mathcal{I}$).

To obtain a $\mathsf{P}^{\mathsf{PosSLP}}$ complexity upper bound, we need only show that we can decide whether $\mathbb{P}^\sigma_{\mathcal{M}^{\leq r}(\mathcal{Q}),s_{\mathsf{init}}}(\Omega) \geq \theta$ in polynomial time in the BSS model [1]. In this model of computation, we can explicitly compute the transition probabilities of $\mathcal{C}^\sigma_\mathcal{I}$ in polynomial time (by Theorem 18) and use them to compute the probability of reaching T_Ω from s_{init} in $\mathcal{C}^\sigma_\mathcal{I}$. This reachability probability is exactly $\mathbb{P}^\sigma_{\mathcal{M}^{\leq r}(\mathcal{Q}),s_{\mathsf{init}}}(\Omega)$ by Theorems 11 and 12. We conclude by comparing it to θ. We obtain the following result.

▶ **Theorem 20.** *The OEIS verification problem for state-reachability and selective termination in bounded OC-MDPs is in $\mathsf{P}^{\mathsf{PosSLP}}$.*

Proof. In this proof, we reason in the BSS model of computation. Our goal is to clarify the algorithm outlined above and prove that it runs in polynomial time. We let $\mathcal{I} = \mathsf{Refine}(\mathsf{Isolate}(\mathcal{I}', k_{\mathsf{init}}))$ where $\mathcal{I}'$ is the interval partition of $[\![1, r-1]\!]$ in the representation of σ. Lemma 10 guarantees that $\mathcal{I}$ can be computed in polynomial time and has a polynomial-size representation with respect to $\mathcal{I}'$, and that $\mathcal{C}^\sigma_\mathcal{I}$ is well-defined (cf. Assumption 8). It follows from Theorem 18 that the transition probabilities of $\mathcal{C}^\sigma_\mathcal{I}$ can be computed in polynomial time.

We have $\mathbb{P}^{\sigma}_{\mathcal{M}^{\leq r}(\mathcal{Q}),s_{\mathsf{init}}}(\Omega) = \mathbb{P}_{\mathcal{C}^{\sigma}_{\mathcal{I}},s_{\mathsf{init}}}(\mathsf{Reach}(T_\Omega))$ by Theorems 11 and 12. It follows that $\mathbb{P}^{\sigma}_{\mathcal{M}^{\leq r}(\mathcal{Q}),s_{\mathsf{init}}}(\Omega)$ can be computed in polynomial time by solving a linear system for Markov chain reachability probabilities (with $|S_{\mathcal{I}}| \leq 2 \cdot |\mathcal{I}| \cdot |Q| \cdot \log_2(r)$ variables) and then can be compared to θ in constant time. We conclude that the OEIS verification problem for Ω can be solved in polynomial time in the BSS model and thus lies in $\mathsf{P}^{\mathsf{PosSLP}}$ [1]. ◀

5.2 Verifying open-ended interval strategies

We describe a co-ETR algorithm for the OEIS verification problem. Recall that co-ETR is the class of decision problems that can be reduced (in polynomial time) to checking whether a universal sentence holds in the theory of the reals and that co-ETR is included in PSPACE [22]. The algorithm of Section 5.1 provides a finer bound when dealing with bounded OC-MDPs.

We construct logic formulae in the signature of ordered fields to decide the verification problem. These formulae are also relevant for the realisability problems considered in Section 6. For this reason, we provide a formula that (essentially) depends only on $\mathcal{Q}$ and a finite partition $\mathcal{I}$ of $[\![1, r]\!]$ into intervals with respect to which compressed Markov chains are well-defined. This avoids depending only on σ which could prove problematic in a realisability setting. We fix a finite interval partition $\mathcal{I}$ of $[\![1, r-1]\!]$ satisfying Assumption 8, i.e., such that, for all $I \in \mathcal{I}$, $\log_2(|I|+1) \in \mathbb{N}$. We build a formula with respect to $\mathcal{Q}$ and $\mathcal{I}$ and show that we can answer the verification problem via this formula for all OEISs based on $\mathcal{I}$ from any initial configuration in $S_{\mathcal{I}}$. We postpone the definition of a relevant partition for σ and s_{init} to the end of the section.

Our formula uses three sets of variables. First, for all $q \in Q$, $a \in A(q)$ and $I \in \mathcal{I}$, we introduce a variable $z^I_{q,a}$ to represent $\tau(q, \min I)(a)$ for any OEIS τ based on $\mathcal{I}$. For all $I \in \mathcal{I}$, we let $\mathbf{z}^I = (z^I_{q,a})_{q \in Q, a \in A(q)}$ and let $\mathbf{z} = (\mathbf{z}^I)_{I \in \mathcal{I}}$. We let $\tau_{\mathbf{z}}$ be the formal (parametric) OEIS based on $\mathcal{I}$ defined by $\tau_{\mathbf{z}}(q, \min I)(a) = z^I_{q,a}$ for all $q \in Q$, $a \in A(q)$ and $I \in \mathcal{I}$. The notation $\tau_{\mathbf{z}}$ allows us to refer to the compressed Markov chain $\mathcal{C}^{\tau_{\mathbf{z}}}_{\mathcal{I}}$ parameterised by $\mathbf{z}$ in the following. To lighten notation, we write $\mathcal{C}^{\mathbf{z}}_{\mathcal{I}}$ instead of $\mathcal{C}^{\tau_{\mathbf{z}}}_{\mathcal{I}}$ and $\delta^{\mathbf{z}}_{\mathcal{I}}$ instead of $\delta^{\tau_{\mathbf{z}}}_{\mathcal{I}}$.

The second set of variables comes from Theorems 13 and 16 for each interval of $\mathcal{I}$ and are used to represent (and to characterise) the transition probabilities of $\mathcal{C}^{\mathbf{z}}_{\mathcal{I}}$ from configurations in $S_{\mathcal{I}} \setminus S^{\perp}_{\mathcal{I}}$. We let $\mathbf{x}$ denote the vector of all of these variables. For all configurations $s = (q, k) \in S_{\mathcal{I}} \setminus S^{\perp}_{\mathcal{I}}$ and $s' = (p, k') \in S_{\mathcal{I}} \setminus \{\perp\}$ such that k' is a successor of k, we let $x_{s,s'}$ denote the variable corresponding to $\delta^{\sigma}_{\mathcal{I}}(s)(s')$. We note that some variables represent outgoing probabilities from two configurations (see Lemma 15).

The last set of variables represents the probability of the counterpart of Ω in $\mathcal{C}^{\mathbf{z}}_{\mathcal{I}}$ from each configuration. For all $s \in S_{\mathcal{I}} \setminus S^{\perp}_{\mathcal{I}}$, we introduce a variable y_s where y_s represents $\mathbb{P}_{\mathcal{C}^{\mathbf{z}}_{\mathcal{I}},s}(\mathsf{Reach}(T_\Omega))$. We let $\mathbf{y}$ denote the vector of these variables.

We now construct, for all $s \in S_{\mathcal{I}} \setminus S^{\perp}_{\mathcal{I}}$, a quantifier-free formula such that when substituting $\mathbf{z}$ by a vector $\mathbf{z}^\star$ that yields a well-defined strategy $\tau_{\mathbf{z}^\star}$ and quantifying the other variables universally, the resulting sentence holds if and only if $\mathbb{P}^{\tau_{\mathbf{z}^\star}}_{\mathcal{M}^{\leq r}(\mathcal{Q}),s}(\Omega) \geq \theta$. We rely on universal quantification because we do not have a unique characterisation of the transition probabilities of $\mathcal{C}^{\mathbf{z}}_{\mathcal{I}}$ when $r = \infty$. We construct a quantifier-free conjunction (parameterised by the choices of the strategy) that only holds for (some) over-estimations of $\mathbb{P}^{\tau_{z^\star}}_{\mathcal{M}^{\leq r}(\mathcal{Q}),s}(\Omega)$. This allows us to check that $\mathbb{P}^{\tau_{z^\star}}_{\mathcal{M}^{\leq r}(\mathcal{Q}),s}(\Omega)$ exceeds θ by checking that all of its over-estimations do.

Our formula has two major sub-formulae. First, we define a formula depending on $\mathbf{z}$ such that the least vector satisfying it includes the transition probabilities of $\mathcal{C}^{\mathbf{z}}_{\mathcal{I}}$ from configurations to other configurations (i.e., not to $\perp$). For each $I \in \mathcal{I}$, we define $\Phi^I_\delta(\mathbf{x}, \mathbf{z}^I)$ as the conjunction of all the equations in the system characterising the transition probabilities

from $S_{\mathcal{I}} \cap (Q \times I)$ in $\mathcal{C}_{\mathcal{I}}^{\mathbf{z}}$ given by Theorems 13 and 16 (the invoked theorem depends on whether I is finite or not). We define

$$\Phi_{\delta}^{\mathcal{I}}(\mathbf{x}, \mathbf{z}) = \bigwedge_{x \in \mathbf{x}} x \geq 0 \wedge \bigwedge_{I \in \mathcal{I}} \Phi_{\delta}^{I}(\mathbf{x}, \mathbf{z}^{I}) \wedge \bigwedge_{s \in S_{\mathcal{I}} \setminus S_{\mathcal{I}}^{\perp}} \sum_{x_{s,s'} \in x_{s,\cdot}} x_{s,s'} \leq 1, \tag{6}$$

where for all $s \in S_{\mathcal{I}} \setminus S_{\mathcal{I}}^{\perp}$, $x_{s,\cdot}$ denotes the set of well-defined variables of the form $x_{s,s'}$ ($s' \in S_{\mathcal{I}}$). The first conjunction ensures that any vector satisfying $\Phi_{\delta}^{\mathcal{I}}$ is non-negative while the rightmost conjunction ensures that for all $s \in S_{\mathcal{I}} \setminus S_{\mathcal{I}}^{\perp}$, $s' \mapsto x_{s,s'}$ is a sub-probability distribution. It follows that, for all configurations $s \in S_{\mathcal{I}} \setminus S_{\mathcal{I}}^{\perp}$ and all vectors $\mathbf{x}^{\star}$ and $\mathbf{z}^{\star}$ such that $\Phi_{\delta}^{\mathcal{I}}(\mathbf{x}^{\star}, \mathbf{z}^{\star})$ holds, we can define a distribution $\delta_{\bar{x^{\star}}}(s) \in \mathcal{D}(S_{\mathcal{I}})$ such that for $s' \in S_{\mathcal{I}} \setminus \{\perp\}$, if $x_{s,s'}$ is a well-defined variable then $\delta_{\bar{x^{\star}}}(s)(s') = x_{s,s'}^{\star}$ and, otherwise, $\delta_{\bar{x^{\star}}}(s)(s') = 0$. We use these distributions in our correctness proof: they allow us to reason on Markov chains over $S_{\mathcal{I}}$.

The second block of the formula describes the probability of reaching T_{Ω} in $\mathcal{C}_{\mathcal{I}}^{\mathbf{z}}$. We consider the following formula, derived from the systems for Markov chain reachability probabilities,

$$\Phi_{\Omega}^{\mathcal{I}}(\mathbf{x}, \mathbf{y}) = \bigwedge_{s \in S_{\mathcal{I}} \setminus S_{\mathcal{I}}^{\perp}} \left(y_s \geq 0 \wedge y_s = \sum_{\substack{x_{s,s'} \in x_{s,\cdot} \\ s' \in S_{\mathcal{I}} \setminus S_{\mathcal{I}}^{\perp}}} x_{s,s'} y_{s'} + \sum_{\substack{x_{s,s'} \in x_{s,\cdot} \\ s' \in T_{\Omega}}} x_{s,s'} \right). \tag{7}$$

We now state that for all well-defined instances $\tau_{\mathbf{z}^{\star}}$ of $\tau_{\mathbf{z}}$, the conjunction $\Phi_{\delta}^{\mathcal{I}}(\mathbf{x}, \mathbf{z}^{\star}) \wedge \Phi_{\Omega}^{\mathcal{I}}(\mathbf{x}, \mathbf{y})$ only holds for over-estimations of the values represented by the variables. This mainly follows from the construction of the formulae (in particular, by Theorems 13 and 16).

▶ **Lemma 21.** *Let $\mathbf{z}^{\star}$ be a vector such that $\tau_{\mathbf{z}^{\star}}$ is a well-defined OEIS of $\mathcal{M}^{\leq r}(\mathcal{Q})$ based on $\mathcal{I}$. Let $\mathbf{x}^{\star}, \mathbf{y}^{\star}$ be vectors such that $\mathbb{R} \models \Phi_{\delta}^{\mathcal{I}}(\mathbf{x}^{\star}, \mathbf{z}^{\star}) \wedge \Phi_{\Omega}^{\mathcal{I}}(\mathbf{x}^{\star}, \mathbf{y}^{\star})$. Then, for all $s \in S_{\mathcal{I}} \setminus S_{\mathcal{I}}^{\perp}$, we have $y_s^{\star} \geq \mathbb{P}_{\mathcal{C}_{\mathcal{I}}^{\mathbf{z}^{\star}}, s}(\mathsf{Reach}(T_{\Omega}))$, and, for all $s' \in S_{\mathcal{I}} \setminus \{\perp\}$ such that $x_{s,s'}$ is a well-defined variable, $x_{s,s'}^{\star} \geq \delta_{\mathcal{I}}^{\mathbf{z}^{\star}}(s)(s')$.*

Proof. For all $s \in S_{\mathcal{I}} \setminus S_{\mathcal{I}}^{\perp}$ and all $s' \in S_{\mathcal{I}} \setminus \{\perp\}$ such that $x_{s,s'}$ is defined, we have $x_{s,s'}^{\star} \geq \delta_{\mathcal{I}}^{\mathbf{z}^{\star}}(s)(s')$ by construction of $\Phi_{\delta}^{\mathcal{I}}$ through Theorems 13 and 16. It remains to show that $y_s^{\star} \geq \mathbb{P}_{\mathcal{C}_{\mathcal{I}}^{\mathbf{z}^{\star}}, s}(\mathsf{Reach}(T_{\Omega}))$ for all $s \in S_{\mathcal{I}} \setminus S_{\mathcal{I}}^{\perp}$.

Our argument relies on the finite Markov chain $\mathcal{C}_{\mathbf{x}^{\star}} = (S_{\mathcal{I}}, \delta_{\mathbf{x}^{\star}})$ where for all $s \in S_{\mathcal{I}}^{\perp}$, we have $\delta_{\mathbf{x}^{\star}}(s)(s) = 1$ and for all $s \in S_{\mathcal{I}} \setminus S_{\mathcal{I}}^{\perp}$ and all $s' \in S_{\mathcal{I}} \setminus \{\perp\}$, we have $\delta_{\mathbf{x}^{\star}}(s)(s') = x_{s,s'}^{\star}$ whenever $x_{s,s'}$ is defined and direct the probability mass that is not assigned to a successor of s in the previous way to $\perp$.

The least vector satisfying $\Phi_{\Omega}^{\mathcal{I}}(\mathbf{x}^{\star}, \mathbf{y})$ is $(\mathbb{P}_{\mathcal{C}_{\mathbf{x}^{\star}}, s}(\mathsf{Reach}(T_{\Omega})))_{s \in S_{\mathcal{I}} \setminus S_{\mathcal{I}}^{\perp}}$. Therefore, it suffices to show that for all $s \in S_{\mathcal{I}} \setminus S_{\mathcal{I}}^{\perp}$, we have $\mathbb{P}_{\mathcal{C}_{\mathbf{x}^{\star}}, s}(\mathsf{Reach}(T_{\Omega})) \geq \mathbb{P}_{\mathcal{C}_{\mathcal{I}}^{\mathbf{z}^{\star}}, s}(\mathsf{Reach}(T_{\Omega}))$ to end the proof. It suffices to establish that for all histories $\bar{h} \in \mathsf{Hist}(\mathcal{C}_{\mathcal{I}}^{\mathbf{z}^{\star}})$ with $\mathsf{last}(\bar{h}) \in T_{\Omega}$ and no prior configuration in T_{Ω}, we have $\mathbb{P}_{\mathcal{C}_{\mathcal{I}}^{\mathbf{z}^{\star}}, \mathsf{first}(\bar{h})}(\mathsf{Cyl}(h)) \leq \mathbb{P}_{\mathcal{C}_{\mathbf{x}^{\star}}, \mathsf{first}(\bar{h})}(\mathsf{Cyl}(h))$. Let $\bar{h} \in \mathsf{Hist}(\mathcal{C}_{\mathcal{I}}^{\mathbf{z}^{\star}})$ be such a history. We assume that $\mathsf{first}(\bar{h}) \notin T_{\Omega}$, as otherwise the result is trivial. Since $\perp$ is absorbing (in both Markov chains), it follows that all states along $\bar{h}$ are configurations in $S_{\mathcal{I}}$. The desired inequality follows from $\delta_{\mathbf{x}^{\star}}(s)(s') \geq \delta_{\mathcal{I}}^{\sigma}(s)(s')$ holding for all $s, s' \in S_{\mathcal{I}} \setminus \{\perp\}$. ◀

The following theorem provides the formula we use to solve the OEIS verification problem based on the intuition given above. Its correctness follows from Lemma 21.

▶ **Theorem 22.** *Let $\mathbf{z}^\star$ be a vector such that $\tau_{\mathbf{z}^\star}$ is a well-defined OEIS. For all $s \in S_\mathcal{I} \setminus S_\mathcal{I}^\perp$, we have $\mathbb{P}^{\tau_{\mathbf{z}^\star}}_{\mathcal{M}^{\leq r}(\mathcal{Q}),s}(\Omega) \geq \theta$ if and only if $\mathbb{R} \models \forall \mathbf{x} \forall \mathbf{y}((\Phi^\mathcal{I}_\delta(\mathbf{x}, \mathbf{z}^\star) \wedge \Phi^\mathcal{I}_\Omega(\mathbf{x}, \mathbf{y})) \implies y_s \geq \theta)$.*

Proof. Let $s \in S_\mathcal{I} \setminus S_\mathcal{I}^\perp$. By Theorems 11 and 12, we have $\mathbb{P}^{\tau_{\mathbf{z}^\star}}_{\mathcal{M}^{\leq r}(\mathcal{Q}),s}(\Omega) = \mathbb{P}_{\mathcal{C}^{\mathbf{z}^\star}_\mathcal{I},s}(\mathsf{Reach}(T_\Omega))$. First, we assume that $\mathbb{P}_{\mathcal{C}^{\mathbf{z}^\star}_\mathcal{I},s}(\mathsf{Reach}(T_\Omega)) \geq \theta$. Let $\mathbf{x}^\star$ and $\mathbf{y}^\star$ such that $\mathbb{R} \models \Phi^\mathcal{I}_\delta(\mathbf{x}^\star, \mathbf{z}^\star) \wedge \Phi^\mathcal{I}_\Omega(\mathbf{x}^\star, \mathbf{y}^\star)$. By Lemma 21, we obtain $y^\star_s \geq \mathbb{P}_{\mathcal{C}^{\mathbf{z}^\star}_\mathcal{I},s}(\mathsf{Reach}(T_\Omega)) \geq \theta$. This shows the first implication.

Conversely, assume that $\mathbb{R} \models \forall \mathbf{x} \forall \mathbf{y}((\Phi^\mathcal{I}_\delta(\mathbf{x}, \mathbf{z}^\star) \wedge \Phi^\mathcal{I}_\Omega(\mathbf{x}, \mathbf{y})) \implies y_s \geq \theta)$. Let $\mathbf{x}^\star$ be the least non-negative satisfying assignment of $\mathbf{x}$ in $\Phi^\mathcal{I}_\delta(\mathbf{x}, \mathbf{z}^\star)$. The existence of $\mathbf{x}^\star$ is guaranteed by Theorems 13 and 16, which also imply that for all variables $x_{s',s''}$, we have $x^\star_{s',s''} = \delta^\sigma_\mathcal{I}(s')(s'')$. We then let $\mathbf{y}^\star$ be the least satisfying assignment of $\mathbf{y}$ in the formula with parameters $\Phi^\mathcal{I}_\Omega(\mathbf{x}^\star, \mathbf{y})$. By construction of $\Phi^\mathcal{I}_\Omega$, we conclude that $\mathbf{y}^\star$ exists and that $\mathbb{P}_{\mathcal{C}^{\mathbf{z}^\star}_\mathcal{I},s}(\mathsf{Reach}(T_\Omega)) = y^\star_s \geq \theta$. ◀

We now analyse the size of the formula of Theorem 22. We show that this formula is of size polynomial in the encoding of $\mathcal{Q}$ and the natural representation of $\mathcal{I}$, i.e., as a finite set of intervals whose bounds are described in binary. We use this to show that we can build a formula to solve the verification problem in polynomial time. This analysis is also relevant to obtain complexity bounds for realisability.

▶ **Lemma 23.** *The formula $(\Phi^\mathcal{I}_\delta(\mathbf{x}, \mathbf{z}) \wedge \Phi^\mathcal{I}_\Omega(\mathbf{x}, \mathbf{y})) \implies y_s \geq \theta$ has a number of variables and atomic sub-formulae polynomial in $|Q|$, $|A|$, $|\mathcal{I}|$ and the binary encoding of the largest integer bound in $\mathcal{I}$.*

Proof. Let $\beta = \max_{I \in \mathcal{I}, |I| < \infty} \log_2(|I| + 1)$ if there is a bounded interval in $\mathcal{I}$ and, otherwise, let $\beta = 1$. We note that $\beta \leq \log_2(b^+ + 1)$ where b^+ is the largest interval bound of $\mathcal{I}$.

First, we have, by definition of $\mathbf{z}$ and $\mathbf{y}$, $|\mathbf{z}| \leq |\mathcal{I}| \cdot |Q| \cdot |A|$ and $|\mathbf{y}| \leq |S_\mathcal{I} \setminus S_\mathcal{I}^\perp|$. By definition of $S_\mathcal{I}$, we have $|S_\mathcal{I} \setminus S_\mathcal{I}^\perp| \leq 2 \cdot \beta \cdot |\mathcal{I}| \cdot |Q|$. Second, Lemmas 14 and 17 imply that $|\mathbf{x}|$ and the number and length of the atomic sub-formulae of $\Phi^\mathcal{I}_\delta(\mathbf{x}, \mathbf{z})$ derived from Theorems 13 and 16 are polynomial in $|Q|$, $|A|$, $|\mathcal{I}|$ and β. It follows from the above that the number and length of the atomic sub-formulae of $(\Phi^\mathcal{I}_\delta(\mathbf{x}, \mathbf{z}) \wedge \Phi^\mathcal{I}_\Omega(\mathbf{x}, \mathbf{y})) \implies y_s \geq \theta$ is polynomial in $|Q|$, $|A|$, $|\mathcal{I}|$ and β. ◀

We now assume that σ is an OEIS and define the interval partition used to construct a verification formula. Let $\mathcal{I}'$ be the interval partition of $[\![1, r-1]\!]$ given in the representation of σ. We let $\mathcal{I} = \mathsf{Refine}(\mathsf{Isolate}(\mathcal{I}', k_{\mathsf{init}}))$. The partition $\mathcal{I}$ satisfies Assumption 8 and we have $s_{\mathsf{init}} \in S_\mathcal{I}$. Let $\mathbf{z}^\sigma$ denote the valuation of $\mathbf{z}$ defined by $z^I_{q,a} = \sigma(q, \min I)(a)$ for all $q \in Q$, $a \in A(q)$ and $I \in \mathcal{I}$. To decide the verification problem, we check whether the formula of Theorem 22 for s_{init} holds for this valuation $\mathbf{z}^\sigma$ of $\mathbf{z}$. We obtain the following complexity result.

▶ **Theorem 24.** *The OEIS verification problem for selective termination and state-reachability objectives is in co-ETR.*

Proof. We prove that the formula used to answer the verification problem can be constructed in polynomial time. The structure of the formula is fixed. Therefore, it can be constructed in time polynomial in its size. It follows from Lemma 23 that we can construct our formula in polynomial time if $\mathcal{I}$ admits a representation of size polynomial in the number of inputs to the verification problem.

By definition of the Refine and Isolate operators, all interval bounds of $\mathcal{I}$ are either dominated by a bound in the representation of σ or by $k_{\mathsf{init}} + 1$. Therefore, all bounds admit

a polynomial-size representation. Furthermore, Lemma 10 guarantees that when applying the refinement procedure to obtain $\mathcal{I}$, we obtain a partition of size polynomial in the size of the inputs to the verification problem. ◀

5.3 Verifying cyclic interval strategies

We provide a co-ETR algorithm for the CIS verification problem that follows the same ideas as in Section 5.2. We assume throughout this section that $r = \infty$. To analyse CISs, we use the compression approach twice. First, we use it to define the one-counter Markov chain that induces the infinite compressed Markov chain with respect to our CIS for a well-chosen periodic partition of $\mathbb{N}_{>0}$ (cf. Theorem 19). Second, we use the compression approach to reason on the obtained one-counter Markov chain.

The formulae constructed in this section are also relevant for the realisability algorithms of Section 6. We design formulae that apply to all strategies based on a given periodic partition of $\mathbb{N}_{>0}$. We let $\rho \in \mathbb{N}_{>0}$ be a period, $\mathcal{J}$ be an interval partition of $[\![1, \rho]\!]$ into intervals and let $\mathcal{I}$ be the periodic partition generated by $\mathcal{J}$. We fix a finite interval partition $\mathcal{K}$ of $\mathbb{N}_{>0}$ for the second compression. For all intervals $I \in \mathcal{J} \cup \mathcal{K}$, we assume that $\log_2(|I| + 1) \in \mathbb{N}$ to guarantee that compressed Markov chains are well-defined with respect to these partitions. We design formulae for all CISs based on $\mathcal{I}$ whose structure depends only on $\mathcal{Q}$, $\mathcal{J}$ and $\mathcal{K}$. We let $\bar{T} = (T \times \{\rho\}) \times \{0\}$ denote the target of interest in the compression of the one-counter Markov chain (see Theorems 11 and 12).

Our formula for the verification problem uses four sets of variables; we require a new set of variables comparatively to Section 5.2 for the additional compression. First, we introduce variables for the choices of strategies. For all $q \in Q$, $a \in A(q)$ and $I \in \mathcal{J}$, we introduce a variable $z^I_{q,a}$ to represent $\sigma(q, \min I)(a)$. For all $I \in \mathcal{J}$, we let $\mathbf{z}^I = (z^I_{q,a})_{q \in Q, a \in A(q)}$ and let $\mathbf{z} = (\mathbf{z}^I)_{I \in \mathcal{J}}$. We let $\tau_{\mathbf{z}}$ be the formal CIS of period ρ based on $\mathcal{I}$ defined by $\tau_{\mathbf{z}}(q, \min I)(a) = z^I_{q,a}$ for all $q \in Q$, $a \in A(q)$ and $I \in \mathcal{J}$. To lighten notation, we write $\mathcal{C}^{\mathbf{z}}_{\mathcal{I}}$ for the compressed Markov chain $\mathcal{C}^{\tau_{\mathbf{z}}}_{\mathcal{I}}$ associated to $\tau_{\mathbf{z}}$ (parameterised by $\mathbf{z}$) and $\mathcal{P}^{\mathbf{z}}_{\mathcal{J}} = (P_{\mathcal{J}}, \delta^{\tau_{\mathbf{z}}}_{\mathcal{J}})$ for the one-counter Markov chain $\mathcal{P}^{\tau_{\mathbf{z}}}_{\mathcal{J}}$ inducing $\mathcal{C}^{\mathbf{z}}_{\mathcal{I}}$ in the sense of Theorem 19. We let $P^{\top}_{\mathcal{J}} = P_{\mathcal{J}} \setminus \{\bot\}$. We let $\mathcal{C}_{\mathcal{K}}(\mathcal{P}^{\mathbf{z}}_{\mathcal{J}}) = (S_{\mathcal{K}}(P_{\mathcal{J}}), \delta_{\mathcal{K}}[\mathcal{P}^{\mathbf{z}}_{\mathcal{J}}])$ denote the compression of $\mathcal{C}^{\leq\infty}(\mathcal{P}^{\mathbf{z}}_{\mathcal{J}})$ with respect to $\mathcal{K}$.

We then introduce a new set of variables $\mathbf{v}$ for the transitions probabilities of $\mathcal{P}^{\mathbf{z}}_{\mathcal{J}}$ between configurations in $P^{\top}_{\mathcal{J}}$; these variables come from the system of Theorem 16. For any two $s, s' \in P^{\top}_{\mathcal{J}}$ and weight $u \in \{-1, 0, 1\}$, we let $v_{s,s',u}$ denote the variable corresponding to $\delta^{\tau_{\mathbf{z}}}_{\mathcal{J}}(s)(s', u)$ whenever this variable is well-defined. Third, we consider a set of variables $\mathbf{x}$ for the transitions probabilities of $\mathcal{C}_{\mathcal{K}}(\mathcal{P}^{\mathbf{z}}_{\mathcal{J}})$ taken from the systems of Theorems 13 and 16. For all $\bar{s}, \bar{s}' \in S_{\mathcal{K}}(P_{\mathcal{J}})$ such that $\bar{s} \in P^{\top}_{\mathcal{J}} \times \mathbb{N}_{>0}$, we write $x_{\bar{s},\bar{s}'}$ for the variable corresponding to $\delta_{\mathcal{K}}[\mathcal{P}^{\mathbf{z}}_{\mathcal{J}}](\bar{s})(\bar{s}')$ whenever this variable is defined. Finally, we introduce a variable $y_{\bar{s}}$ for all configurations $\bar{s} \in S_{\mathcal{K}}(P_{\mathcal{J}}) \cap (P^{\top}_{\mathcal{J}} \times \mathbb{N}_{>0})$ to represent the probability $\mathbb{P}_{\mathcal{C}_{\mathcal{K}}(\mathcal{P}^{\mathbf{z}}_{\mathcal{J}}),\bar{s}}(\mathsf{Reach}(\bar{T}))$. We let $\mathbf{y} = (y_{\bar{s}})_{\bar{s} \in S_{\mathcal{K}}(P_{\mathcal{J}}) \cap (P^{\top}_{\mathcal{J}} \times \mathbb{N}_{>0})}$.

We now formulate three sub-formulae of the formula used in our decision procedure. For all $I \in \mathcal{J}$, we let $\Psi^I_{\delta}(\mathbf{v}, \mathbf{z}^I)$ be the conjunction of the equations obtained by Theorem 16 for the outgoing transitions of $P_{\mathcal{J}} \cap (Q \times I)$ in $\mathcal{P}^{\tau_{\mathbf{z}}}_{\mathcal{J}}$. Similarly to Equation (6), we define a formula for the transitions of $\mathcal{P}^{\mathbf{z}}_{\mathcal{J}}$ by

$$\Psi^{\mathcal{J}}_{\delta}(\mathbf{v}, \mathbf{z}) = \bigwedge_{v \in \mathbf{v}} v \geq 0 \wedge \bigwedge_{I \in \mathcal{J}} \Psi^I_{\delta}(\mathbf{v}, \mathbf{z}^I) \wedge \bigwedge_{s \in P_{\mathcal{J}} \setminus \{\bot\}} \sum_{v_{s,s',u} \in v_{s,\cdot,\cdot}} v_{s,s',u} \leq 1. \tag{8}$$

We then construct the counterpart $\Phi^{\mathcal{K}}_{\delta}(\mathbf{x}, \mathbf{v})$ of the formula of Equation (6) for the compressed Markov chain $\mathcal{C}_{\mathcal{K}}(\mathcal{P}^{\sigma}_{\mathcal{J}})$. In this case, the sub-formulae derived from the systems

of Theorem 13 and Theorem 16 for each interval of $\mathcal{K}$ depend on $\mathbf{v}$ instead of $\mathbf{z}$. We also build a counterpart $\Phi^{\mathcal{K}}_{\Omega}(\mathbf{x},\mathbf{y})$ of the formula given in Equation (7) for $\mathcal{C}_{\mathcal{K}}(\mathcal{P}^{\sigma}_{\mathcal{J}})$ with respect to the target $\bar{T}$.

To decide the verification problem, we rely on a formula similar to that of Theorem 22: we check that over-estimations of the probability of interest exceed the threshold θ. To validate this approach, we establish a counterpart of Lemma 21 for the conjunction $\Psi^{\mathcal{J}}_{\delta}(\mathbf{v},\mathbf{z}^\star) \wedge \Phi^{\mathcal{K}}_{\delta}(\mathbf{x},\mathbf{v}) \wedge \Phi^{\mathcal{K}}_{\Omega}(\mathbf{x},\mathbf{y})$ given a vector $\mathbf{z}^\star$ such that $\tau_{\mathbf{z}^\star}$ is a well-defined CIS based on $\mathcal{I}$.

▶ **Lemma 25.** *Let $\mathbf{z}^\star$ be a vector such that $\tau_{\mathbf{z}^\star}$ is a well-defined CIS of $\mathcal{M}^{\leq\infty}(\mathcal{Q})$ based on $\mathcal{I}$. Let $\mathbf{v}^\star$, $\mathbf{x}^\star$ and $\mathbf{y}^\star$ be vectors such that $\mathbb{R} \models \Psi^{\mathcal{J}}_{\delta}(\mathbf{v}^\star,\mathbf{z}^\star) \wedge \Phi^{\mathcal{K}}_{\delta}(\mathbf{x}^\star,\mathbf{v}^\star) \wedge \Phi^{\mathcal{K}}_{\Omega}(\mathbf{x}^\star,\mathbf{y}^\star)$. Then, it holds that*

1. *for all $\bar{s} \in S_{\mathcal{K}}(P_{\mathcal{J}}) \cap (P^{\top}_{\mathcal{J}} \times \mathbb{N}_{>0})$, we have $y^\star_{\bar{s}} \geq \mathbb{P}_{\mathcal{C}_{\mathcal{K}}(\mathcal{P}^{\mathbf{z}^\star}_{\mathcal{J}}),\bar{s}}(\mathsf{Reach}(\bar{T}))$;*
2. *for all $\bar{s} \in S_{\mathcal{K}}(P_{\mathcal{J}}) \cap (P^{\top}_{\mathcal{J}} \times \mathbb{N}_{>0})$ and all $\bar{s}' \in S_{\mathcal{K}}(P_{\mathcal{J}})$ such that $x_{\bar{s},\bar{s}'}$ is defined, we have $x^\star_{\bar{s},\bar{s}'} \geq \delta_{\mathcal{K}}[\mathcal{P}^{\mathbf{z}^\star}_{\mathcal{J}}](\bar{s})(\bar{s}')$;*
3. *for all s, $s' \in P_{\mathcal{J}} \setminus \{\bot\}$ and $u \in \{-1,0,1\}$ such that $v_{s,s',u}$ is defined, we have $v^\star_{s,s',u} \geq \delta^{\mathbf{z}^\star}_{\mathcal{J}}(s)(s',u)$.*

Proof. Item 3 follows directly from the construction of $\Psi^{\mathcal{J}}_{\delta}$ based on Theorem 16. To prove Items 1 and 2, we consider the one-counter Markov chain $\mathcal{P}_{\mathbf{v}^\star} = (P_{\mathcal{J}}, \delta_{\mathbf{v}^\star})$ where for all s, $s' \in P_{\mathcal{J}}$ and all $u \in \{-1,0,1\}$, $\delta_{\mathbf{v}^\star}(s)(s',u) = v^\star_{s,s',u}$ whenever the variable $v_{s,s',u}$ is defined and any probability that is not assigned in this way is attributed to $\delta_{\mathbf{v}^\star}(s)(\bot,0)$. We show that in the compression $\mathcal{C}_{\mathcal{K}}(\mathcal{P}_{\mathbf{v}^\star}) = (S_{\mathcal{K}}(P_{\mathcal{J}}), \delta_{\mathcal{K}}[\mathcal{P}_{\mathbf{v}^\star}])$ of $\mathcal{P}_{\mathbf{v}^\star}$ with respect to $\mathcal{K}$, we have the two following properties:

a. for all $\bar{s},\bar{s}' \in S_{\mathcal{K}}(P_{\mathcal{J}}) \cap (P^{\top}_{\mathcal{J}} \times \mathbb{N}_{>0})$, $\delta_{\mathcal{K}}[\mathcal{P}_{\mathbf{v}^\star}](\bar{s})(\bar{s}') \geq \delta_{\mathcal{K}}[\mathcal{P}^{\sigma}_{\mathcal{J}}](\bar{s})(\bar{s}')$;
b. for all $\bar{s} \in S_{\mathcal{K}}(P_{\mathcal{J}}) \cap (P^{\top}_{\mathcal{J}} \times \mathbb{N}_{>0})$, we have $\mathbb{P}_{\mathcal{P}_{\mathcal{K}}(\mathcal{P}_{\mathbf{v}^\star}),\bar{s}}(\mathsf{Reach}(\bar{T})) \geq \mathbb{P}_{\mathcal{C}_{\mathcal{K}}(\mathcal{P}^{\mathbf{z}^\star}_{\mathcal{J}}),\bar{s}}(\mathsf{Reach}(\bar{T}))$.

These two properties along with Lemma 21 (with respect to the parameterised formula $\Phi^{\mathcal{K}}_{\delta}(\mathbf{x},\mathbf{v}^\star) \wedge \Phi^{\mathcal{K}}_{\Omega}(\mathbf{x},\mathbf{y})$) yield Items 1 and 2.

We first prove Item a. Let $\bar{s},\bar{s}' \in S_{\mathcal{K}}(P_{\mathcal{J}}) \setminus \{\bot\}$. Let s, $s \in P_{\mathcal{J}}$ and $k,k' \in \mathbb{N}$ such that $\bar{s} = (s,k)$ and $\bar{s}' = (s',k')$. If k' is not a successor counter value of k with respect to $\mathcal{K}$, then we have $\delta_{\mathcal{K}}[\mathcal{P}_{\mathbf{v}^\star}](\bar{s})(\bar{s}') = \delta_{\mathcal{K}}[\mathcal{P}^{\mathbf{z}^\star}_{\mathcal{J}}](\bar{s})(\bar{s}') = 0$. Otherwise, $\delta_{\mathcal{K}}[\mathcal{P}^{\mathbf{z}^\star}_{\mathcal{J}}](\bar{s})(\bar{s}')$ is the probability of reaching $\bar{s}'$ from $\bar{s}$ in $\mathcal{C}^{\leq\infty}(\mathcal{P}^{\mathbf{z}^\star}_{\mathcal{J}})$ without visiting another configuration with a successor value of k beforehand. Along the relevant plays for this probability, there are no $\bot$ configurations. Since $\delta_{\mathcal{K}}[\mathcal{P}_{\mathbf{v}^\star}](\bar{s})(\bar{s}')$ is similarly defined, the desired inequality follows from Item 3.

Item a implies Item b, as there are no $\bot$ configuration that can occur on plays ending in $\bar{T}$ (of which all states are absorbing). ◀

We obtain an adaptation of Theorem 22 for CISs via Lemma 25. We use the correspondence between configurations of $\mathcal{P}^{\mathbf{z}}_{\mathcal{J}}$ and $\mathcal{C}_{\mathcal{K}}(\mathcal{P}^{\mathbf{z}}_{\mathcal{J}})$ established in Theorem 19 in this result.

▶ **Theorem 26.** *Let $\mathbf{z}^\star$ be a vector such that $\tau_{\mathbf{z}^\star}$ is a well-defined CIS of $\mathcal{M}^{\leq\infty}(\mathcal{Q})$ based on $\mathcal{I}$. For all $\bar{s} = ((q,k),k') \in S_{\mathcal{K}}(P_{\mathcal{J}}) \cap (P^{\top}_{\mathcal{J}} \times \mathbb{N}_{>0})$, we have $\mathbb{P}^{\tau_{\mathbf{z}^\star}}_{\mathcal{M}^{\leq\infty}(\mathcal{Q}),(q,(k'-1)\cdot\rho+k)}(\mathsf{Reach}(\bar{T})) \geq \theta$ if and only if $\mathbb{R} \models \forall\mathbf{x}\,\forall\mathbf{y}\forall\mathbf{v}((\Psi^{\mathcal{J}}_{\delta}(\mathbf{v},\mathbf{z}^\star) \wedge \Phi^{\mathcal{K}}_{\delta}(\mathbf{x},\mathbf{v}) \wedge \Phi^{\mathcal{K}}_{\Omega}(\mathbf{x},\mathbf{y})) \implies y_{\bar{s}} \geq \theta)$.*

Proof. Let $\bar{s} = ((q,k),k') \in S_{\mathcal{K}}(P_{\mathcal{J}}) \cap (P^{\top}_{\mathcal{J}} \times \mathbb{N}_{>0})$. By Theorems 19 and 11, we have $\mathbb{P}^{\tau_{\mathbf{z}^\star}}_{\mathcal{M}^{\leq\infty}(\mathcal{Q}),(q,(k'-1)\rho+k)}(\mathsf{Reach}(\bar{T})) = \mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{P}^{\mathbf{z}^\star}_{\mathcal{J}}),\bar{s}}(\mathsf{Reach}(\bar{T})) = \mathbb{P}_{\mathcal{C}_{\mathcal{K}}(\mathcal{P}^{\mathbf{z}^\star}_{\mathcal{J}}),\bar{s}}(\mathsf{Reach}(\bar{T}))$

If $\mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{P}^{\sigma}_{\mathcal{J}}),\bar{s}}(\mathsf{Reach}(\bar{T})) \geq \theta$, then we have $\mathbb{R} \models \forall\mathbf{x}\,\forall\mathbf{y}\forall\mathbf{v}((\Psi^{\mathcal{J}}_{\delta}(\mathbf{v},\mathbf{z}^\star) \wedge \Phi^{\mathcal{K}}_{\delta}(\mathbf{x},\mathbf{v}) \wedge \Phi^{\mathcal{K}}_{\Omega}(\mathbf{x},\mathbf{y})) \implies y_{\bar{s}} \geq \theta)$ by Lemma 25.

Conversely, assume that $\mathbb{R} \models \forall\mathbf{x}\,\forall\mathbf{y}\forall\mathbf{v}((\Psi^{\mathcal{J}}_{\delta}(\mathbf{v},\mathbf{z}^\star) \wedge \Phi^{\mathcal{K}}_{\delta}(\mathbf{x},\mathbf{v}) \wedge \Phi^{\mathcal{K}}_{\Omega}(\mathbf{x},\mathbf{y})) \implies y_{\bar{s}} \geq \theta)$. Let $\mathbf{v}^\star$ be the least vector satisfying $\Psi^{\mathcal{J}}_{\delta}(\mathbf{v},\mathbf{z}^\star)$, $\mathbf{x}^\star$ be the least vector satisfying $\Phi^{\mathcal{K}}_{\delta}(\mathbf{x},\mathbf{v}^\star)$ and $\mathbf{y}^\star$ be the least vector satisfying $\Phi^{\mathcal{K}}_{\Omega}(\mathbf{x}^\star,\mathbf{y})$. By construction of these three

formulae (and Theorems 13 and 16), these vectors are well-defined and we obtain $y^\star_{\bar{s}} = \mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{P}^\sigma_{\mathcal{J}}),\bar{s}}(\mathsf{Reach}(\bar{T})) \geq \theta$. ◀

We now study the size of the formula of Theorem 26 for our complexity analysis. We obtain a conclusion similar to that provided by Lemma 23.

▶ **Lemma 27.** *The formula* $(\Psi^{\mathcal{J}}_\delta(\mathbf{v},\mathbf{z}) \wedge \Phi^{\mathcal{K}}_\delta(\mathbf{x},\mathbf{v}) \wedge \Phi^{\mathcal{K}}_\Omega(\mathbf{x},\mathbf{y})) \implies y_{\bar{s}} \geq \theta$ *has a number of variables and atomic sub-formulae polynomial in* $|Q|$, $|A|$, $|\mathcal{J}|$, $|\mathcal{K}|$, *the binary encoding of* ρ *and the binary encoding of the largest integer bound in* $\mathcal{K}$. *The length of its atomic sub-formulae is polynomial in the same parameters.*

Proof. Lemma 23 implies that the number of variables and atomic formulae of $\Phi^{\mathcal{K}}_\delta(\mathbf{x},\mathbf{v}) \wedge \Phi^{\mathcal{K}}_\Omega(\mathbf{x},\mathbf{y})$, well as the length of these atomic formulae, is polynomial in $|P^\top_{\mathcal{J}}|$, $|\mathcal{K}|$ and the binary encoding of the largest bound in $\mathcal{K}$ (actions are not relevant; remark that we deal with a Markov chain). By construction of $P^\top_{\mathcal{J}}$, we have $|P^\top_{\mathcal{J}}| \leq 2 \cdot \log_2(\rho+1) \cdot |\mathcal{J}| \cdot |Q|$ (interval bounds of $\mathcal{J}$ are at most ρ). Regarding $\Psi^{\mathcal{J}}_\delta(\mathbf{v},\mathbf{z})$, it suffices to adapt the analysis performed in the proof of Lemma 23 for the formula $\Phi^{\mathcal{T}}_\delta$ to obtain the desired bounds. ◀

We now assume that the input strategy σ is a CIS. We close the section by explaining how to construct $\mathcal{J}$ and $\mathcal{K}$ in polynomial time from the representation of σ and s_{init} to prove that the verification problem is in co-ETR. Let $\mathcal{J}'$ denote the partition of $[\![1,\rho]\!]$ given in the representation of σ. We let $\mathcal{J} = \mathsf{Refine}(\mathsf{Isolate}(\mathcal{J}, k_{\mathsf{init}} \bmod \rho))$. For the partition for the second compression, we let $\mathcal{K} = \mathsf{Refine}([\![1, \lfloor \frac{k_{\mathsf{init}}}{\rho} \rfloor]\!]) \cup \{[\![\lfloor \frac{k_{\mathsf{init}}}{\rho} \rfloor + 1, +\infty]\!]\}$ of $\mathbb{N}_{>0}$. We observe that the counterpart of s_{init} in $S_{\mathcal{K}}(P_{\mathcal{J}})$ (in the sense of Theorem 19) is guaranteed to exist: if $k_{\mathsf{init}} \bmod \rho = 0$, then we have $((q_{\mathsf{init}},\rho), \frac{k_{\mathsf{init}}}{\rho}) \in S_{\mathcal{K}}(P_{\mathcal{J}})$, and, otherwise, we have $((q_{\mathsf{init}}, k_{\mathsf{init}} \bmod \rho), \lfloor \frac{k_{\mathsf{init}}}{\rho} \rfloor + 1) \in S_{\mathcal{K}}(P_{\mathcal{J}})$. We let $\mathbf{z}^\sigma$ denote the substitution of $\mathbf{z}$ such that $z^I_{q,a}$ is set to $\sigma(q, \min I)(a)$ for all $q \in Q$, $a \in A(q)$ and $I \in \mathcal{J}$. With the partitions $\mathcal{J}$ and $\mathcal{K}$ given above and the formula of Theorem 26 with respect to the counterpart of s_{init} in $S_{\mathcal{K}}(P_{\mathcal{J}})$ and the parameter $\mathbf{z}^\sigma$, we can decide our instance of the verification problem. We thus obtain the following complexity result.

▶ **Theorem 28.** *The CIS verification problem for selective termination and reachability objectives is in co-ETR.*

Proof. We observe, in the same way as in the proof of Theorem 24, that Lemmas 10 and 27 imply that the formula of Theorem 26 can be constructed in polynomial time for the partitions described above. ◀

6 Fixed-interval and parameterised realisability

We provide complexity upper bounds for the fixed-interval and parameterised realisability problems for interval strategies. Our algorithms are built on the verification techniques presented in Section 5. We first provide a technical result for analysing the complexity of the parameterised realisability problem in Section 6.1. In Section 6.2, we study our realisability problems in bounded OC-MDPs. We consider OEISs in general, i.e., we provide an approach applicable for both the bounded and unbounded setting, in Section 6.3. We close the section with CISs in Section 6.4.

We use similar approaches for all settings. For fixed-interval realisability for pure strategies, we non-deterministically construct strategies and verify them. This approach is not viable for fixed-interval realisability for randomised strategies; instead, we quantify existentially over strategy variables in the logical formulae used for verification. For the parameterised

realisability problem, we build on our algorithms for the fixed-interval case. The main idea is use non-determinism to find an interval partition compatible with the input parameters and then use fixed-interval algorithms with this partition to answer our problem. All complexity bounds provided in this section are in PSPACE.

We consider the following inputs for the whole section: an OC-MDP $\mathcal{Q} = (Q, A, \delta, w)$, a counter upper bound $r \in \bar{\mathbb{N}}_{>0}$, an initial configuration $s_{\mathsf{init}} = (q_{\mathsf{init}}, k_{\mathsf{init}}) \in Q \times [\![r]\!]$, a set of targets $T \subseteq Q$, an objective $\Omega \in \{\mathsf{Reach}(T), \mathsf{Term}(T)\}$ and a threshold $\theta \in [0,1] \cap \mathbb{Q}$. We specify the other inputs below. As in Section 5, we assume that we work with the modified OC-MDP of Theorem 12 if $\Omega = \mathsf{Reach}(T)$.

6.1 Parameters and compatible interval partitions

We present algorithms for parameterised interval strategy realisability problems that rely on non-determinism to test all interval partitions compatible with the input parameters to check that there is a well-performing strategy based on one of these interval partitions. We state a result implying that all of these interval partition admit a representation that is polynomial in the encoding of the input parameters.

▶ **Lemma 29.** *Let $d \in \mathbb{N}_{>0}$, $n \in \mathbb{N}_{>0}$ and $k \in \bar{\mathbb{N}}$. Let $\mathcal{I}$ be an interval partition of $[\![1,k]\!]$ such that $|\mathcal{I}| \leq d$ and for all bounded $I \in \mathcal{I}$, $|I| \leq n$. Then $\mathcal{I}$ can be explicitly represented in space $\mathcal{O}(d \cdot (\log_2(d) + \log_2(n)))$.*

Proof. We can represent each interval $[\![b^-, b^+]\!] \in \mathcal{I}$ by the pair (b^-, b^+) (where b^+ can be $+\infty$). We prove that each finite interval bound in these pairs is at most $n \cdot d$.

First, assume that $k \in \mathbb{N}$. In this case, the interval $[\![1,k]\!]$ is the union of at most d sets of at most n elements (by the assumption on $\mathcal{I}$). It follows that $k \leq d \cdot n$, and thus, that all finite interval bounds of $\mathcal{I}$ are no more than $d \cdot n$.

Second, assume that $k = \infty$. Let I_∞ denote the unbounded interval in $\mathcal{I}$. It holds (by the same reasoning as above) that the bounds of all intervals in $\mathcal{I} \setminus \{I_\infty\}$ are no more than $(d-1) \cdot n$. This implies that $\min I_\infty - 1 \leq (d-1) \cdot n$. We obtain that in this second case, all finite interval bounds in $\mathcal{I}$ are no more than $(d-1) \cdot n + 1 \leq d \cdot n$.

We conclude that, in both cases, we can represent $\mathcal{I}$ using no more than d pairs of numbers whose binary encoding is in $\mathcal{O}(\log_2(d) + \log_2(n))$. ◀

Lemma 29 implies that, under the assumption that the parameter d for the number of intervals is given in unary, the interval partitions that are compatible with the parameters admit a polynomial-size representation with respect to the inputs to the parameterised interval strategy realisability problems.

6.2 Realisability in bounded one-counter Markov decision processes

Assume that $r \in \mathbb{N}_{>0}$. We provide algorithms for the fixed-interval and parameterised OEIS realisability problems in bounded OC-MDPs. We first discuss the variants of these problems for pure strategies, and then discuss the variants for randomised strategies below.

First, we consider the fixed-interval pure OEIS realisability problem. Fix an input interval partition $\mathcal{I}'$ of $[\![1, r-1]\!]$. We obtain a straightforward non-deterministic algorithm: we guess a pure interval strategy σ based on the partition $\mathcal{I}'$ and then use our verification algorithm for OEISs in bounded OC-MDPs as a $\mathsf{P}^{\mathsf{PosSLP}}$ sub-procedure (Theorem 20) to check whether $\mathbb{P}^{\sigma}_{\mathcal{M}^{\leq r}(\mathcal{Q}), s_{\mathsf{init}}}(\Omega) \geq \theta$. This realisability algorithm runs in non-deterministic polynomial time with a PosSLP oracle: we non-deterministically choose $d \cdot |Q|$ actions, i.e.,

one per state-interval pair, and then run a deterministic polynomial-time algorithm with a PosSLP oracle. This yields an $\mathsf{NP}^{\mathsf{PosSLP}}$ upper bound for this problem.

For the parameterised pure OEIS realisability problem in bounded OC-MDPs, we proceed similarly. Let $d \in \mathbb{N}_{>0}$ and $n \in \mathbb{N}_{>0}$ respectively denote the input parameters for the number and size of intervals. We non-deterministically guess an interval partition $\mathcal{I}'$ of $[\![1, r-1]\!]$ that is compatible with d and n (these partitions can be represented in polynomial space by Lemma 29), guess a pure strategy based on $\mathcal{I}'$ and verify it. In this case, by adapting the analysis made above, we also obtain an $\mathsf{NP}^{\mathsf{PosSLP}}$ upper complexity bound. We summarise the above upper bounds in the following theorem.

▶ **Theorem 30.** *The fixed-interval and parameterised pure OEIS realisability problems for selective termination and state-reachability objectives in bounded OC-MDPs are in* $\mathsf{NP}^{\mathsf{PosSLP}}$.

We now consider the fixed-interval and parameterised randomised OEIS realisability problem and describe an $\mathsf{NP}^{\mathsf{ETR}}$ algorithm. We start with the fixed-interval realisability problem. Let $\mathcal{I}' = (I_j)_{j \in [\![1,d]\!]}$ denote an input interval partition of $[\![1, r-1]\!]$. Prefacing the formula of Theorem 22 with existential quantifiers for the strategy probabilities yields a polynomial-space procedure (cf. Section 6.3). We provide an alternative approach for bounded OC-MDPs which yields a more precise bound.

The key is to rely on a unique characterisation of the transition and reachability probabilities in the compressed Markov chain that we consider. Theorem 18 provides the means to do this: it provides systems whose only solution contains the transition probabilities of a compressed Markov chain. These systems also indicate which transitions have positive probability in the compressed Markov chain. We can thus refine the reachability probability system (described in the formula in Equation (7)) to have a unique solution.

To adequately refine systems using Theorem 18, we must know the supports of the distributions assigned by the considered strategy. For this, we use non-determinism; we guess the action supports for each state-interval pair. We then construct an existential formula that is dependent on these supports. This formula holds if and only if there exists a strategy witnessing a positive answer to the fixed-interval randomised OEIS realisability problem that uses these supports.

We let $T_\Omega = T \times \{0\}$ if $\Omega = \mathsf{Term}(T)$ and $T_\Omega = T \times \{0, r\}$ if $\Omega = \mathsf{Reach}(T)$. Let $\mathcal{I} = \mathsf{Refine}(\mathsf{Isolate}(\mathcal{I}', k_{\mathsf{init}}))$. We note that $s_{\mathsf{init}} \in S_{\mathcal{I}}$. For all $j \in [\![1, d]\!]$, we let $\mathcal{I}_j = \{I \in \mathcal{I} \mid I \subseteq I_j\}$. We use the same variables as the formula of Theorem 22 presented in Section 5.2. More precisely, we have a variable vector $\mathbf{z}$ for the probabilities assigned by strategies and a vector $\mathbf{z}^I$ for all $I \in \mathcal{I}$ for the strategy probabilities specific to I. We let $\tau_{\mathbf{z}}$ denote a formal strategy given by $\tau_{\mathbf{z}}(q, \min I)(a) = z^I_{q,a}$ for all $q \in Q$, $a \in A(q)$ and $I \in \mathcal{I}$. We then have $\mathbf{x}$ for the transition probabilities of the compressed Markov chain $\mathcal{C}^{\mathbf{z}}_{\mathcal{I}}$ and $\mathbf{y}$ for the probability of reaching T_Ω from each configuration in $S_{\mathcal{I}} \setminus S^{\perp}_{\mathcal{I}}$.

We call functions $\mathcal{B} \colon Q \times [\![1, d]\!] \to 2^A \setminus \{\emptyset\}$ such that for all $q \in Q$ and $j \in [\![1, d]\!]$, the inclusion $\mathcal{B}(q, j) \subseteq A(q)$ holds *support-assigning functions*. An OEIS σ based on $(I_j)_{j \in [\![1,d]\!]}$ is *compatible* with $\mathcal{B}$ if for all $q \in Q$ and $j \in [\![1, d]\!]$, we have $\mathsf{supp}(\sigma(q, \min I_j)) = \mathcal{B}(q, j)$.

We now define the required sub-formulae used in our algorithm. The first formula checks that the substitution of $\mathbf{z}$ results in an interpretation of the symbolic strategy $\tau_{\mathbf{z}}$ that is based on $\mathcal{I}'$ and is compatible with $\mathcal{B}$. For each $j \in [\![1, d]\!]$, we fix $I^\star_j \in \mathcal{I}_j$. We define the formula $\Phi^{\mathcal{I},\mathcal{I}',\mathcal{B}}_{\sigma}(\mathbf{z})$ by

$$\bigwedge_{j \in [\![1,d]\!]} \left(\bigwedge_{q \in Q} \left(\bigwedge_{a \in \mathcal{B}(q,j)} z^{I^\star_j}_{q,a} > 0 \wedge \bigwedge_{a \notin \mathcal{B}(q,j)} z^{I^\star_j}_{q,a} = 0 \wedge \sum_{a \in A(q)} z^{I^\star_j}_{q,a} = 1 \right) \wedge \bigwedge_{I \in \mathcal{I}_j} \mathbf{z}^I = \mathbf{z}^{I^\star_j} \right). \quad (9)$$

The other sub-formulae are built under the assumption that we consider an interpretation of $\tau_{\mathbf{z}}$ with the supports described by $\mathcal{B}$. The second sub-formula is a parallel of the formula of Equation (6). For each $I \in \mathcal{I}$, we let $\Phi_\delta^{I,\mathcal{I}',\mathcal{B}}(\mathbf{x}, \mathbf{z}^I)$ be the conjunction of the equations in the system with a unique solution obtained from Theorem 18 for the transitions from $S_{\mathcal{I}} \cap (Q \times I)$ in $\mathcal{C}_{\mathcal{I}}^{\mathbf{z}}$. We let $\Phi_\delta^{\mathcal{I},\mathcal{I}',\mathcal{B}}(\mathbf{x}, \mathbf{z}) = \bigwedge_{I \in \mathcal{I}} \Phi_\delta^{I,\mathcal{I}',\mathcal{B}}(\mathbf{x}, \mathbf{z}^I)$. From these equations, we can deduce the transition structure of $\mathcal{C}_{\mathcal{I}}^{\mathbf{z}}$ and construct a linear system with a unique solution describing the probability of reaching T_Ω in $\mathcal{C}_{\mathcal{I}}^{\mathbf{z}}$; we let $\Phi_\Omega^{\mathcal{I},\mathcal{I}',\mathcal{B}}(\mathbf{x}, \mathbf{y})$ denote the conjunction of the equations of this system. Using the fact that these last two formulae have unique satisfying assignments for a valuation of $\mathbf{z}$ satisfying $\Phi_\sigma^{\mathcal{I},\mathcal{I}',\mathcal{B}}(\mathbf{z})$, we obtain the following theorem.

▶ **Theorem 31.** *Let $\mathcal{B} \colon Q \times [\![1, d]\!] \to 2^A \setminus \{\emptyset\}$ be a support-assigning function and $s \in S_{\mathcal{I}} \setminus S_{\mathcal{I}}^{\perp}$. There exists a strategy σ based on $\mathcal{I}'$ that is compatible with $\mathcal{B}$ such that $\mathbb{P}^{\sigma}_{\mathcal{M}^{\leq r}(\mathcal{Q}),s}(\Omega) \geq \theta$ if and only if $\mathbb{R} \models \exists \mathbf{z} \exists \mathbf{x} \exists \mathbf{y} (\Phi_\sigma^{\mathcal{I},\mathcal{I}',\mathcal{B}}(\mathbf{z}) \wedge \Phi_\delta^{\mathcal{I},\mathcal{I}',\mathcal{B}}(\mathbf{x}, \mathbf{z}) \wedge \Phi_\Omega^{\mathcal{I},\mathcal{I}',\mathcal{B}}(\mathbf{x}, \mathbf{y}) \wedge y_s \geq \theta)$.*

Proof. Assume that there exists a strategy σ based on $\mathcal{I}'$ compatible with $\mathcal{B}$ such that $\mathbb{P}^{\sigma}_{\mathcal{M}^{\leq r}(\mathcal{Q}),s}(\Omega) \geq \theta$. Let $\mathbf{z}^\sigma$ denote the valuation of $\mathbf{z}$ given by $z^I_{q,a} = \sigma(q, \min I)(a)$ for all $q \in Q$, $a \in A(q)$ and $I \in \mathcal{I}$. It is easy to see that $\Phi_\sigma^{\mathcal{I},\mathcal{I}',\mathcal{B}}(\mathbf{z}^\sigma)$ holds by compatibility of σ with $\mathcal{B}$. By construction of $\Phi_\delta^{\mathcal{I},\mathcal{I}',\mathcal{B}}(\mathbf{x}, \mathbf{z})$ (via Theorem 18), there is a unique vector $\mathbf{x}^\star$ such that $\Phi_\delta^{\mathcal{I},\mathcal{I}',\mathcal{B}}(\mathbf{x}^\star, \mathbf{z}^\sigma)$ holds which contains the transition probabilities of $\mathcal{C}_{\mathcal{I}}^{\sigma}$. In turn, this implies that there is a unique vector $\mathbf{y}^\star$ such that $\Phi_\Omega^{\mathcal{I},\mathcal{I}',\mathcal{B}}(\mathbf{x}^\star, \mathbf{y}^\star)$ holds, and this vector is $(\mathbb{P}_{\mathcal{C}_{\mathcal{I}}^{\sigma},s'}(\mathsf{Reach}(T_\Omega)))_{s' \in S_{\mathcal{I}} \setminus S_{\mathcal{I}}^{\perp}}$. By Theorems 11 and 12, we obtain that $y^\star_s = \mathbb{P}_{\mathcal{C}_{\mathcal{I}}^{\sigma},s}(\mathsf{Reach}(T_\Omega)) = \mathbb{P}^{\sigma}_{\mathcal{M}^{\leq r}(\mathcal{Q}),s}(\Omega) \geq \theta$.

Conversely, let $\mathbf{z}^\star$, $\mathbf{x}^\star$ and $\mathbf{y}^\star$ witnessing that the existential formula above holds and define $\sigma = \tau_{\mathbf{z}^\star}$. The strategy σ is well-defined and compatible with $\mathcal{B}$ because $\Phi_\sigma^{\mathcal{I},\mathcal{I}',\mathcal{B}}(\mathbf{z}^\star)$ holds. Furthermore, by construction of the formulae $\Phi_\delta^{\mathcal{I},\mathcal{I}',\mathcal{B}}$ and $\Phi_\Omega^{\mathcal{I},\mathcal{I}',\mathcal{B}}$, we deduce that $y^\star_s = \mathbb{P}_{\mathcal{C}_{\mathcal{I}}^{\sigma},s}(\mathsf{Reach}(T_\Omega)) \geq \theta$. We conclude that $\mathbb{P}^{\sigma}_{\mathcal{M}^{\leq r}(\mathcal{Q}),s}(\Omega) = \mathbb{P}_{\mathcal{C}_{\mathcal{I}}^{\sigma},s}(\mathsf{Reach}(T_\Omega)) \geq \theta$ by Theorems 11 and 12. ◀

To decide the fixed-interval randomised OEIS realisability problem, it suffices to check that there exists a support-assigning function $\mathcal{B}$ (using non-determinism) such that the formula of Theorem 31 for holds for s_{init} (by construction of $\mathcal{I}$, $s_{\mathsf{init}} \in S_{\mathcal{I}}$). We thus obtain an $\mathsf{NP}^{\mathsf{ETR}}$ upper bound for this variant of the fixed-interval realisability problem.

For the parameterised realisability problem, we obtain an $\mathsf{NP}^{\mathsf{ETR}}$ upper bound by altering the fixed-interval algorithm slightly. In this case, we use non-determinism to guess an interval partition $\mathcal{I}'$ that is compatible with the input parameters and a support-assigning function. We then check the validity of the formula of Theorem 31 for the interval partition $\mathcal{I} = \mathsf{Refine}(\mathsf{Isolate}(\mathcal{I}', k_{\mathsf{init}}))$ and the initial configuration s_{init}. We obtain the following result.

▶ **Theorem 32.** *The fixed-interval and parameterised randomised OEIS realisability problems for selective termination and state-reachability objectives in bounded OC-MDPs are in $\mathsf{NP}^{\mathsf{ETR}}$.*

Proof. We present a unified argument for both the fixed-interval and parameterised realisability problems. The only difference between our complexity analysis for these two problems is how the interval partition $\mathcal{I}'$ is obtained. In the fixed-interval case, the interval partition $\mathcal{I}'$ is a part of the input. In the parameterised case, $\mathcal{I}'$ is of size polynomial in the input parameters by Lemma 29 (recall that the parameter bounding the number of intervals is assumed to be given in unary).

We let $\mathcal{I} = \mathsf{Refine}(\mathsf{Isolate}(\mathcal{I}', k_{\mathsf{init}}))$ and $\mathcal{B} \colon Q \times [\![1, |\mathcal{I}'|]\!] \to 2^A \setminus \{\emptyset\}$ be a support-assigning function. Lemma 10 guarantees that $\mathcal{I}$ has a representation of size polynomial in that

of $\mathcal{I}'$ and k_{init}. The support-assigning function $\mathcal{B}$ can be explicitly represented in space polynomial in $|Q|$, $|A|$ and $|\mathcal{I}'|$. Therefore, to prove that the $\mathsf{NP}^{\mathsf{ETR}}$ upper bound holds, it remains to prove that the formula $\Phi_\sigma^{\mathcal{I},\mathcal{I}',\mathcal{B}}(\mathbf{z}) \wedge \Phi_\delta^{\mathcal{I},\mathcal{I}',\mathcal{B}}(\mathbf{x},\mathbf{z}) \wedge \Phi_\Omega^{\mathcal{I},\mathcal{I}',\mathcal{B}}(\mathbf{x},\mathbf{y}) \wedge y_{s_{\mathsf{init}}} \geq \theta$ can be constructed in deterministic polynomial time from $\mathcal{I}'$, $\mathcal{I}$, $\mathcal{B}$ and the other inputs to the considered realisability problem.

Lemma 23 implies that the number of variables in this formula is polynomial in $|Q|$, $|A|$ and $|\mathcal{I}|$, and that the formulae $\Phi_\delta^{\mathcal{I},\mathcal{I}',\mathcal{B}}(\mathbf{x},\mathbf{z})$ and $\Phi_\Omega^{\mathcal{I},\mathcal{I}',\mathcal{B}}(\mathbf{x},\mathbf{y})$ have size polynomial in $|Q|$, $|A|$ and $|\mathcal{I}|$. Indeed, for these two formulae, we observe that they can be derived from the original formulae $\Phi_\delta^{\mathcal{I}}$ and $\Phi_\Omega^{\mathcal{I}}$ for verification by taking their conjunction with atomic propositions requiring that some variables in $\mathbf{x}$ and $\mathbf{y}$ are equal to zero (these variables can be identified in polynomial time through the algorithm of Theorem 18 and with a reachability analysis of the compressed Markov chain). Finally, we observe that, in the formula $\Phi_\sigma^{\mathcal{I},\mathcal{I}',\mathcal{B}}(\mathbf{z})$, there are no more than $|\mathcal{I}'| \cdot (2 \cdot |Q| + |\mathcal{I}| \cdot |A| \cdot |Q|)$ atomic formulae of length in $\mathcal{O}(|A|)$. We have thus shown that the formula $\Phi_\sigma^{\mathcal{I},\mathcal{I}',\mathcal{B}}(\mathbf{z}) \wedge \Phi_\delta^{\mathcal{I},\mathcal{I}',\mathcal{B}}(\mathbf{x},\mathbf{z}) \wedge \Phi_\Omega^{\mathcal{I},\mathcal{I}',\mathcal{B}}(\mathbf{x},\mathbf{y}) \wedge y_{s_{\mathsf{init}}} \geq \theta$ can be constructed in polynomial time. ◀

6.3 Realisability for open-ended interval strategies

We now consider the fixed-interval and parameterised realisability problems for OEISs in general OC-MDPs. For pure strategies, we can adapt the approach of Section 6.2. In the fixed-interval case, we guess a pure OEIS based on the input interval partition of the set of counter values and verify it. In the parameterised case, we guess an interval partition compatible with the input parameters (its representation is of size polynomial in the representation of the parameters by Lemma 29) and a pure OEIS based on it, then verify it. Theorem 24 thus implies that the fixed-interval and parameterised realisability problems for pure OEISs can be solved by a non-deterministic polynomial-time algorithm that uses a co-ETR oracle (which is the same as using an ETR oracle). We obtain the following upper bound.

▶ **Theorem 33.** *The fixed-interval and parameterised pure OEIS realisability problems for selective termination and state-reachability objectives are in* $\mathsf{NP}^{\mathsf{ETR}}$.

We now consider the variant of our realisability problems for randomised OEISs. First, we discuss the fixed-interval problem and let $\mathcal{I}' = (I_j)_{j \in [\![1,d]\!]}$ be an input partition. Let $\mathcal{I} = \mathsf{Refine}(\mathsf{Isolate}(\mathcal{I}', k_{\mathsf{init}}))$. We let, for all $j \in [\![1,d]\!]$, $\mathcal{I}_j = \{I \in \mathcal{I} \mid I \subseteq I_j\}$. Next, we consider the sets of variables $\mathbf{z}$, $\mathbf{x}$ and $\mathbf{y}$ and the formulae $\Phi_\delta^{\mathcal{I}}(\mathbf{x},\mathbf{z})$ and $\Phi_\Omega^{\mathcal{I}}(\mathbf{x},\mathbf{y})$ from Section 5.2. We introduce a new formula to constrain the variables $\mathbf{z}$ similarly to the formula of Equation (9): we let

$$\Phi_\sigma^{\mathcal{I},\mathcal{I}'}(\mathbf{z}) = \bigwedge_{j \in [\![1,d]\!]} \bigwedge_{I \in \mathcal{I}_j} \left(\bigwedge_{q \in Q} \left(\bigwedge_{a \in A(q)} z_{q,a}^I \geq 0 \wedge \sum_{a \in A(q)} z_{q,a}^I = 1 \right) \wedge \bigwedge_{I' \in \mathcal{I}_j} \mathbf{z}^I = \mathbf{z}^{I'} \right). \quad (10)$$

Any vector $\mathbf{z}^\star$ satisfying $\Phi_\sigma^{\mathcal{I},\mathcal{I}'}(\mathbf{z}^\star)$ defines a strategy $\tau_{\mathbf{z}^\star}$ based on $\mathcal{I}'$ and any such strategy induces such a vector. We obtain the following corollary of Theorem 22.

▶ **Theorem 34.** *Let $s \in S_{\mathcal{I}} \setminus S_{\mathcal{I}}^{\perp}$. There exists an OEIS σ based on the partition $\mathcal{I}'$ such that $\mathbb{P}^\sigma_{\mathcal{M}^{\leq r}(\mathcal{Q}),s}(\Omega) \geq \theta$ if and only if $\mathbb{R} \models \exists \mathbf{z} \forall \mathbf{x} \forall \mathbf{y} (\Phi_\sigma^{\mathcal{I},\mathcal{I}'}(\mathbf{z}) \wedge ((\Phi_\delta^{\mathcal{I}}(\mathbf{x},\mathbf{z}) \wedge \Phi_\Omega^{\mathcal{I}}(\mathbf{x},\mathbf{y})) \implies y_s \geq \theta))$.*

We obtain that the fixed-interval realisability problem for randomised OEISs can be reduced to deciding a sentence in $\mathbb{R}$ with two blocks of quantifiers. This shows that the problem is decidable in polynomial space [5, Rmk. 13.10]. For the parameterised randomised

OEIS realisability problem, we obtain an NPSPACE = PSPACE [44] upper bound through the following algorithm: we use non-determinism to obtain a partition $\mathcal{I}'$ of $[\![1, r-1]\!]$ compatible with the input parameters and then check the validity of the formula of Theorem 34 for this partition. We obtain the following.

▶ **Theorem 35.** *The fixed-interval and parameterised randomised OEIS realisability problems for selective termination and state-reachability objectives are in PSPACE.*

Proof. Let $\mathcal{I}'$ denote the input interval partition of $[\![1, r-1]\!]$ if we consider the fixed-interval realisability problem or the partition obtained using non-determinism if we consider the parameterised realisability problem. In the latter, the representation of $\mathcal{I}'$ is of size polynomial in that of the input parameters by Lemma 29.

Whether a formula with two blocks of quantifiers holds in the theory of the reals can be decided in polynomial space [5, Rmk. 13.10]. Therefore, to obtain the claim of the theorem, it suffices to show that the formula $\Phi_\sigma^{\mathcal{I},\mathcal{I}'}(\mathbf{z}) \wedge ((\Phi_\delta^{\mathcal{I}}(\mathbf{x},\mathbf{z}) \wedge \Phi_\Omega^{\mathcal{I}}(\mathbf{x},\mathbf{y})) \implies y_s \geq \theta)$ can be constructed in polynomial time with respect to the representation of $\mathcal{Q}$ and $\mathcal{I}'$.

Lemma 10 implies that $\mathcal{I} = \mathsf{Refine}(\mathsf{Isolate}(\mathcal{I}', k_{\mathsf{init}}))$ admits a representation of size polynomial in the representation of the inputs to the fixed-interval realisability problem. It follows from Lemma 23 that the sub-formula $(\Phi_\delta^{\mathcal{I}}(\mathbf{x},\mathbf{z}) \wedge \Phi_\Omega^{\mathcal{I}}(\mathbf{x},\mathbf{y})) \implies y_s \geq \theta$ can be constructed in polynomial time. For $\Phi_\sigma^{\mathcal{I},\mathcal{I}'}(\mathbf{z})$, we observe that it is a conjunction of no more than $|\mathcal{I}| \cdot (|Q| \cdot (|A|+1) + |\mathcal{I}| \cdot |Q| \cdot |A|)$ atomic formulae of length in $\mathcal{O}(|A|)$. ◀

6.4 Realisability for cyclic interval strategies

We now consider the fixed-interval and parameterised realisability problems for CISs. We assume that $r = \infty$ for the remainder of the section. We adapt the techniques of Section 6.3.

First, we provide non-deterministic algorithms for the variants of these problems for pure CISs. In the fixed-interval case, it suffices to non-deterministically select an action for each state of the OC-MDP and interval from the input interval partition and then verify the resulting CIS. We now consider the parameterised case. Let $d \in \mathbb{N}_{>0}$ and $n \in \mathbb{N}_{>0}$ respectively denote the parameter bounding the number of intervals and the size of intervals for the desirable interval partitions. To solve the parameterised realisability problem, we guess a period $\rho \leq d \cdot n$, an interval partition $\mathcal{J}'$ of $[\![1, \rho]\!]$ that is compatible with d and n and actions for all pairs in $Q \times \mathcal{J}'$, then verify the obtained CIS. Theorem 28, along with Lemma 29 in the parameterised case, imply that both of these problems can be solved in non-deterministic polynomial time with an ETR oracle.

▶ **Theorem 36.** *The fixed-interval and parameterised pure CIS realisability problems for selective termination and state-reachability objectives are in $\mathsf{NP}^{\mathsf{ETR}}$.*

We now consider the problem variants for randomised strategies. As before, we first focus on the fixed-interval case. Let $\rho \in \mathbb{N}_{>0}$ denote the input period and $\mathcal{J}' = (I_j)_{j \in [\![1,d]\!]}$ denote the input interval partition of $[\![1, \rho]\!]$. We define $\mathcal{J} = \mathsf{Refine}(\mathsf{Isolate}(\mathcal{J}', k_{\mathsf{init}} \bmod \rho))$ of $[\![1, \rho]\!]$. We let, for all $j \in [\![1, d]\!]$, $\mathcal{J}_j = \{I \in \mathcal{J} \mid I \subseteq I_j\}$. We also let $\mathcal{K} = \mathsf{Refine}([\![1, \lfloor \frac{k_{\mathsf{init}}}{\rho} \rfloor]\!]) \cup \{[\![\lfloor \frac{k_{\mathsf{init}}}{\rho} \rfloor + 1, \infty]\!]\}$. These choices guarantee that the counterpart in the sense of Theorem 19 of s_{init} in $S_{\mathcal{K}}(P_{\mathcal{J}})$ exists.

Next, we reintroduce the sets of variables $\mathbf{z}$, $\mathbf{v}$, $\mathbf{x}$ and $\mathbf{y}$ and the formulae $\Psi_\delta^{\mathcal{J}}(\mathbf{v},\mathbf{z})$, $\Phi_\delta^{\mathcal{K}}(\mathbf{x},\mathbf{v})$ and $\Phi_\Omega^{\mathcal{K}}(\mathbf{x},\mathbf{y})$ from Section 5.3. We formulate an adaptation of the formulae of

Equations (9) and (10) with respect to $\mathcal{J}$: we let

$$\Phi_\sigma^{\mathcal{J},\mathcal{J}'}(\mathbf{z}) = \bigwedge_{j\in[\![1,d]\!]} \bigwedge_{I\in\mathcal{J}_j} \left(\bigwedge_{q\in Q} \left(\bigwedge_{a\in A(q)} z_{q,a}^I \geq 0 \wedge \sum_{a\in A(q)} z_{q,a}^I = 1 \right) \wedge \bigwedge_{I'\in\mathcal{J}_j} \mathbf{z}^I = \mathbf{z}^{I'} \right). \quad (11)$$

Once again, we obtain a natural correspondence between vectors $\mathbf{z}^\star$ satisfying $\Phi_\sigma^{\mathcal{J},\mathcal{J}'}(\mathbf{z}^\star)$ and the CISs considered for our realisability problem. The following theorem is therefore implied by Theorem 26.

▶ **Theorem 37.** *Let $\bar{s} \in S_{\mathcal{K}}(P_{\mathcal{J}}) \cap (P_{\mathcal{J}} \times \mathbb{N}_{>0})$. There exists a CIS σ based on the periodic partition generated by $\mathcal{J}'$ such that $\mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{P}_{\mathcal{J}}^\sigma),\bar{s}}(\mathsf{Reach}(\bar{T})) \geq \theta$ if and only if $\mathbb{R} \models \exists\mathbf{z}\forall\mathbf{x}\forall\mathbf{y}\forall\mathbf{v}(\Phi_\sigma^{\mathcal{J},\mathcal{J}'}(\mathbf{z}) \wedge ((\Psi_\delta^{\mathcal{J}}(\mathbf{v},\mathbf{z}) \wedge \Phi_\delta^{\mathcal{K}}(\mathbf{x},\mathbf{v}) \wedge \Phi_\Omega^{\mathcal{K}}(\mathbf{x},\mathbf{y})) \implies y_{\bar{s}} \geq \theta))$.*

We thus obtain that the fixed-interval realisability problem for randomised CISs is reducible in polynomial time to deciding whether a sentence with two quantifier blocks holds in the theory of the reals. We obtain a PSPACE upper bound [5, Rmk. 13.10]. For the parameterised case, we obtain an NPSPACE = PSPACE [44] upper bound; we non-deterministically guess an interval partition $\mathcal{J}'$ as we have done for the parameterised pure CIS realisability problem, then reduce to checking the validity of the formula of Theorem 37 as in the fixed-interval case. The following theorem summarises our complexity bounds.

▶ **Theorem 38.** *The fixed-interval and parameterised randomised CIS realisability problems for selective termination and state-reachability objectives are in PSPACE.*

Proof. This theorem follows from an adaptation of the proof of Theorem 35. The major difference, in this case, is that we refer to Lemma 27 instead of Lemma 23 for bounds on the size of the sentence used in to solve the realisability problem. ◀

7 Lower bounds

We present lower complexity bounds for the interval strategy verification problem, the fixed-interval realisability problem and the parameterised realisability problem. In Section 7.1, we prove the square-root-sum hardness of all variants of these problems. We show that our interval strategy realisability problems are NP-hard when considering selective termination in Section 7.2.

7.1 Square-root-sum hardness

The square-root-sum problem consists in comparing a sum of square roots of natural numbers to some integer bound. Formally, the inputs to the problems are natural numbers $x_1, \ldots, x_n$ and a natural number y. We formalise the *square-root-sum problem* as the problem of determining whether $\sum_{i=1}^n \sqrt{x_i} \geq y$.

▶ Remark 39. The square-root-sum problem is typically formulated as having to decide whether $\sum_{i=1}^n \sqrt{x_i} \leq y$, i.e., with the opposite inequality. For the sake of illustrating the hardness of a problem, both problems can be seen as equally suitable.

We argue this by briefly showing that an efficient solution to either variant of the problem yields an efficient solution to the other one. We observe that these two variants are almost the complement of one another. The only case in which the two problems have the same solution for the same inputs is when $\sum_{i=1}^n \sqrt{x_i} = y$. Deciding whether $\sum_{i=1}^n \sqrt{x_i} = y$ can be done in polynomial time [11]. Therefore, an efficient decision procedure for one variant of the square-root-sum problem would entail an efficient one for the other. ◁

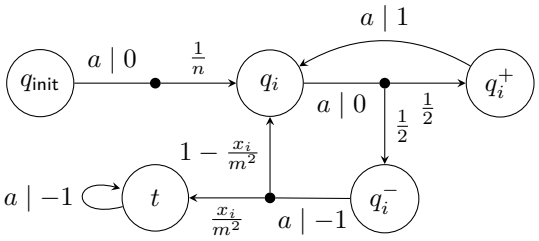


Figure 6 A fragment of $\mathcal{Q}_{\bar{x}}$. Transition probabilities are well-defined: q_{init} has n successors and its outgoing transition share the same probabilities and, for the states q_i^-, we have $x_i \leq m$.

The square-root-sum problem is not known to be solvable in polynomial time in the Turing model of computation. It is known that the square-root-sum problem can be solved in polynomial time in the BSS model [47]. In particular, the square-root-sum problem is in $\mathsf{P}^{\mathsf{PosSLP}}$ [1] and thus in the counting hierarchy.

Our hardness result rely on an existing reduction from the square-root-sum problem to (a variant of) the verification for one-counter Markov chains from [28]. We recall this reduction (adapted to our formalism) in Section 7.1.1 and derive square-root-sum hardness for our problems in unbounded OC-MDPs from it. In Section 7.1.2, we adapt the same reduction to the bounded setting by proving that we can, in polynomial time, compute a large enough counter bound so the bounded one-counter Markov chain approximates the one used in the unbounded reduction well enough for the reduction to still be valid.

7.1.1 Unbounded one-counter Markov decision processes

We show that the reduction from [28] from the square-root-sum problem to a verification problem for selective termination in one-counter Markov chains can be used to obtain lower complexity bounds for our interval strategy problems in unbounded OC-MDPs. We fix inputs $x_1, \ldots, x_n$ and y to the square-root-sum problem, and let $m = \max_{1 \leq i \leq n} x_i$ and $\bar{x} = (x_1, \ldots, x_n)$. In [28], the authors provide a one-counter Markov chain based on $\bar{x}$ such that the probability of terminating in a given state t is $\frac{1}{nm} \sum_{i=1}^{n} \sqrt{x_i}$. We reformulate this one-counter Markov chain as an equivalent OC-MDP $\mathcal{Q}_{\bar{x}}$ with one action.

We depict the fragment of $\mathcal{Q}_{\bar{x}}$ that is associated with x_i in Figure 6. Formally, we define $\mathcal{Q}_{\bar{x}} = (Q_{\bar{x}}, \{a\}, \delta_{\bar{x}}, w_{\bar{x}})$ where $Q_{\bar{x}} = \{q_{\mathsf{init}}, t\} \cup \bigcup_{1 \leq x \leq n} \{q_i, q_i^+, q_i^-\}$ and, for all $i \in [\![1, n]\!]$, transitions and weights to and from the state q_i, q_i^+, q_i^- and t match those in the illustration. For any $r \in \bar{\mathbb{N}}_{>0}$, we let $\mathcal{C}^{\leq r}(\mathcal{Q}_{\bar{x}})$ denote the Markov chain induced by the sole strategy of $\mathcal{M}^{\leq r}(\mathcal{Q}_{\bar{x}})$. We have the following theorem (it can also be seen as a corollary of Theorem 13).

▶ **Theorem 40** ([28]). *We have* $\mathbb{P}_{\mathcal{C}^{\leq \infty}(\mathcal{Q}_{\bar{x}}),(q_{\mathsf{init}},1)}(\mathsf{Term}(t)) = \frac{1}{nm} \sum_{i=1}^{n} \sqrt{x_i}$ *and, for all* $1 \leq i \leq n$, $\mathbb{P}_{\mathcal{C}^{\leq \infty}(\mathcal{Q}_{\bar{x}}),(q_i,1)}(\mathsf{Term}(t)) = \frac{1}{m} \sqrt{x_i}$.

Due to the structure of $\mathcal{Q}_{\bar{x}}$, reaching t and terminating in t are equivalent. Thus, Theorem 40 implies that we can reduce the square-root sum instance fixed above to the verification problem for selective termination and state-reachability on $\mathcal{Q}_{\bar{x}}$ for the unique (counter-oblivious) strategy of $\mathcal{M}^{\leq \infty}(\mathcal{Q}_{\bar{x}})$, which is both an OEIS and a CIS, and the threshold $\theta = \frac{y}{nm}$. Furthermore, as there is only a single strategy, the answer to the verification problem is the same as the answer to the realisability problem for (pure or randomised) counter-oblivious strategies, which is a special case of the fixed-interval and parameterised interval strategy realisability problems for OEISs and CISs.

▶ **Theorem 41.** *The interval strategy verification, fixed-interval realisability and parameterised realisability problems for state-reachability and selective termination are square-root-sum-hard in unbounded OC-MDPs.*

Proof. The OC-MDP $\mathcal{Q}_{\bar{x}}$ and the threshold $\theta = \frac{y}{nm}$ can be computed in polynomial time. Therefore, we need only comment on the correctness of the reduction.

By Theorem 40 and due to the structure of $\mathcal{Q}_{\bar{x}}$, we have $\mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{Q}_{\bar{x}}),(s_{\mathsf{init}},1)}(\mathsf{Reach}(t)) = \mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{Q}_{\bar{x}}),(s_{\mathsf{init}},1)}(\mathsf{Term}(t)) = \frac{1}{nm}\sum_{i=1}^{n}\sqrt{x_i}$. Let $\Omega \in \{\mathsf{Term}(t), \mathsf{Reach}(t)\}$. We clearly have $\mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{Q}_{\bar{x}}),(s_{\mathsf{init}},1)}(\Omega) \geq \theta$ if and only if $\sum_{i=1}^{n}\sqrt{x_i} \geq y$ in light of the above, and thus the reduction is correct. ◀

7.1.2 Bounded one-counter Markov decision processes

We now establish the square-root sum hardness for the interval strategy verification, fixed-interval realisability and parameterised realisability problems in bounded OC-MDPs. Intuitively, in most cases, the reduction consists in adding a counter upper bound to the reduction of Section 7.1.1. We fix inputs $x_1, \ldots, x_n$ and y to the square-root-sum problem for the remainder of the section and let $m = \max_{1\leq i\leq n} x_i$ and $\bar{x} = (x_1, \ldots, x_n)$. Let $\theta = \frac{y}{nm}$ denote the threshold used for the reduction. We assume that for all $i \in [\![1, n]\!]$, $x_i \neq 0$, and thus that $m \geq 1$.

We aim to determine a bound $r \in \mathbb{N}_{>0}$ such that $\mathbb{P}_{\mathcal{C}^{\leq r}(\mathcal{Q}_{\bar{x}}),(q_{\mathsf{init}},1)}(\mathsf{Term}(t)) \geq \theta$ if and only if $\sum_{i=1}^{n}\sqrt{x_i} \geq y$. For all $r \in \mathbb{N}_{>0}$, let $\varepsilon_r = \mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{Q}_{\bar{x}}),(q_{\mathsf{init}},1)}(\mathsf{Reach}(Q \times \{r\}))$. Using Theorem 40, we obtain that for all $r \in \mathbb{N}_{>0}$, we have

$$\mathbb{P}_{\mathcal{C}^{\leq r}(\mathcal{Q}_{\bar{x}}),(q_{\mathsf{init}},1)}(\mathsf{Term}(t)) = \mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{Q}_{\bar{x}}),(q_{\mathsf{init}},1)}(\mathsf{Term}(t)) - \varepsilon_r = \frac{1}{nm}\sum_{i=1}^{n}\sqrt{x_i} - \varepsilon_r.$$

Therefore, we require a bound $r \in \mathbb{N}_{>0}$ with a polynomial-size representation such that the positive error term ε_r above is small enough to ensure the correctness of the reduction.

There is a particular case for which it is clear that no suitable r exists: whenever $\sum_{i=1}^{n}\sqrt{x_i} = y$ holds. Since this equality can be decided in polynomial time [11], we consider a reduction that is conditioned on this equality. First, we check if the equality holds in polynomial time. If it does, we reduce to a fixed positive instance of the considered interval strategy problem (in bounded OC-MDPs). Otherwise, we mirror the reduction of the unbounded case with a well-chosen counter upper bound r.

From here, we assume that $\sum_{i=1}^{n}\sqrt{x_i} \neq y$ and try to determine a suitable bound r. We first formulate a lower bound on the distance $|\sum_{i=1}^{n}\sqrt{x_i} - y|$. We adapt [47, Lemma 3] to suit our formulation of the square-root-sum problem. For the sake of completeness, we provide an adapted proof in Appendix A.

▶ **Lemma 42.** *Let λ be the sum of the bit-sizes of $x_1, \ldots, x_n$ and y. If $\sum_{i=1}^{n}\sqrt{x_i} \neq y$, then $|\sum_{i=1}^{n}\sqrt{x_i} - y| > 2^{-2^n(\lambda+1)}$.*

Lemma 42 implies that our reduction is correct whenever we ensure that $n \cdot m \cdot \varepsilon_r \leq 2^{-2^n(\lambda+1)}$ where λ denotes the sum of bit-sizes of the inputs to the square-root-sum problem. We claim that choosing $r = 2^n m \cdot (\lambda+1) + nm^2 + 1$ is sufficient. This bound can be computed in polynomial time as n is the number of inputs. To show that this choice of r is suitable, we first bound ε_r by an explicit function in r and m.

▶ **Lemma 43.** *For all $r \in \mathbb{N}_{>0}$, $\varepsilon_r \leq (\frac{m}{m+1})^{r-1}$.*

Proof. We have $\varepsilon_1 = 1 = (\frac{m}{m+1})^0$, and thus the inequality holds trivially for $r = 1$. For the remainder of the proof, we prove properties that hold for all $r \geq 2$ to obtain the general result.

For all $r \geq 2$ and $i \in [\![1, n]\!]$, let $\varepsilon_r^{(i)} = \mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{Q}_{\bar{x}}),(q_i,1)}(\mathsf{Reach}((q_i, r)))$ denote the probability of hitting counter value r from $(q_i, 1)$. For all $r \geq 2$, we have $\varepsilon_r = \frac{1}{n}\sum_{i=1}^{n} \varepsilon_r^{(i)}$ due to the topology of $\mathcal{Q}_{\bar{x}}$. To obtain the lemma, it suffices to show that for all $i \in [\![1, n]\!]$, we have $\varepsilon_r^{(i)} \leq (\frac{m}{m+1})^{r-1}$. We fix $i \in [\![1, n]\!]$.

For all $r \geq 2$, we let $\eta_r^{(i)} = \mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{Q}_{\bar{x}}),(q_i,r-1)}(\mathsf{Reach}((q_i, r)))$ denote the probability of reaching counter value r from $(q_i, r-1)$. To conclude this proof, we establish three statements. First, we show that for all $r \geq 2$, we have $\varepsilon_{r+1}^{(i)} = \varepsilon_r^{(i)} \cdot \eta_{r+1}^{(i)}$. Second, we show that for all $r \geq 2$, we have $\eta_r^{(i)} \leq \frac{m}{m+1}$. Finally, we combine these two properties to conclude using an inductive argument.

For all $r \in \mathbb{N}_{>0}$ and $k \in [\![1, r]\!]$, we let $\mathcal{H}_{k\to r}$ denote the set of histories of $\mathcal{M}^{\leq\infty}(\mathcal{Q}_{\bar{x}})$ starting in (q_i, k) and ending in (q_i, r) with only one occurrence of this last configuration. The sets $\mathcal{H}_{k\to r}$ are prefix-free; the cylinders of their elements are pairwise disjoint.

We now show the first claim. Let $r \geq 2$. All histories of $\mathcal{H}_{1\to r}$ can be written as the concatenation of a history of $\mathcal{H}_{1\to r-1}$ and a history of $\mathcal{H}_{r-1\to r}$. We obtain

$$\begin{aligned}
\varepsilon_r^{(i)} &= \mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{Q}_{\bar{x}}),(q_i,1)}(\mathsf{Cyl}(\mathcal{H}_{1\to r})) \\
&= \sum_{h_1 \in \mathcal{H}_{1\to r-1}} \sum_{h_2 \in \mathcal{H}_{r-1\to r}} \mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{Q}_{\bar{x}}),(q_i,1)}(\mathsf{Cyl}(h_1 \cdot h_2)) \\
&= \left(\sum_{h_1 \in \mathcal{H}_{1\to r-1}} \mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{Q}_{\bar{x}}),(q_i,1)}(\mathsf{Cyl}(h_1))\right) \cdot \left(\sum_{h_2 \in \mathcal{H}_{r-1\to r}} \mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{Q}_{\bar{x}}),(q_i,r-1)}(\mathsf{Cyl}(h_2))\right) \\
&= \varepsilon_{r-1}^{(i)} \cdot \eta_r^{(i)}.
\end{aligned}$$

This proves the first claim.

For the second claim, we show that the sequence $(\eta_r^{(i)})_{r\geq 2}$ is increasing and convergent, and that $\lim_{r\to\infty} \eta_r^{(i)} \leq \frac{m}{m+1}$. Let us prove that $(\eta_r^{(i)})_{r\geq 2}$ is increasing. Let $r \geq 2$. We consider the mapping $f_{+1}\colon \mathcal{H}_{r-1\to r} \to \mathcal{H}_{r\to r+1}$ that increases all counter values along a history by 1. This mapping is injective. Furthermore, for all $h \in \mathcal{H}_{r-1\to r}$, we have $\mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{Q}_{\bar{x}}),(q_i,r-1)}(\mathsf{Cyl}(h)) = \mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{Q}_{\bar{x}}),(q_i,r)}(\mathsf{Cyl}(f_{+1}(h)))$. It follows that

$$\eta_r^{(i)} = \mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{Q}_{\bar{x}}),(q_i,r-1)}(\mathsf{Cyl}(\mathcal{H}_{r-1\to r})) \leq \mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{Q}_{\bar{x}}),(q_i,r)}(\mathsf{Cyl}(\mathcal{H}_{r\to r+1})) = \eta_{r+1}^{(i)}.$$

This shows that $(\eta_r^{(i)})_{r\geq 2}$ is increasing.

The sequence $(\eta_r^{(i)})_{r\geq 2}$ is bounded and increasing, thus it converges. We prove that $\lim_{r\to\infty} \eta_r^{(i)} = \frac{m}{m+\sqrt{x_i}}$. To this end, we establish an inductive relation on the elements of this sequence: we prove that for all $r \geq 2$, we have $\eta_{r+1}^{(i)} = \frac{1}{2} + (1 - \frac{x_i}{m^2})\eta_r^{(i)} \cdot \eta_{r+1}^{(i)}$. Let $r \geq 2$. By separating histories for which a counter increment occurs first from those for which a counter decrement occurs first, we obtain that

$$\eta_{r+1}^{(i)} = \frac{1}{2} + \frac{1}{2} \cdot \left(1 - \frac{x_i}{m^2}\right) \cdot \mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{Q}_{\bar{x}}),(q_i,r-1)}(\mathsf{Cyl}(\mathcal{H}_{r-1\to r+1})).$$

We obtain $\eta_{r+1}^{(i)} = \frac{1}{2} + (1 - \frac{x_i}{m^2})\eta_{r-1}^{(i)} \cdot \eta_r^{(i)}$ by observing that elements of $\mathcal{H}_{r-1\to r+1}$ can be written as concatenations of elements of $\mathcal{H}_{r-1\to r}$ and $\mathcal{H}_{r\to r+1}$ and following the same reasoning as for $\varepsilon_r^{(i)}$ above.

By taking the limits on both sides of the above inductive relation, we obtain that $\lim_{r\to\infty} \eta_r^{(i)} = \frac{1}{2} + (1 - \frac{x_i}{m^2}) \cdot (\lim_{r\to\infty} \eta_r^{(i)})^2$. If $m = 1$, then $x_i = 1$ (inputs are positive) and we directly obtain $\lim_{r\to\infty} \eta_r^{(i)} = \frac{1}{2} = \frac{m}{m+\sqrt{x_i}}$. Otherwise, if $m \geq 2$, we have $\frac{x_i}{m^2} < 1$ and we can solve a quadratic equation to deduce that $\lim_{r\to\infty} \eta_r^{(i)} \in \{\frac{m}{m+\sqrt{x_i}}, \frac{m}{m-\sqrt{x_i}}\}$. It follows from $\frac{m}{m-\sqrt{x_i}} > 1$ and $(\eta_r^{(i)})_{r\geq 2}$ being a sequence of probabilities that $\lim_{r\to\infty} \eta_r^{(i)} = \frac{m}{m+\sqrt{x_i}}$. To end the proof of the second claim, we observe that since $(\eta_r^{(i)})_{r\geq 2}$ is increasing and $x_i \geq 1$, we have, for all $r \geq 2$, $\eta_r^{(i)} \leq \frac{m}{m+\sqrt{x_i}} \leq \frac{m}{m+1}$.

We now combine the two claims to provide an inductive proof that $\varepsilon_r^{(i)} \leq (\frac{m}{m+1})^{r-1}$ for all $r \geq 2$. For $r = 2$, we have $\varepsilon_2^{(i)} = \frac{1}{2} \leq \frac{m}{m+1}$ (because $m \geq 1$). We now assume that $\varepsilon_r^{(i)} \leq (\frac{m}{m+1})^{r-1}$ holds. Via the two claims, we conclude that $\varepsilon_{r+1}^{(i)} = \varepsilon_r^{(i)} \cdot \eta_{r+1}^{(i)} \leq (\frac{m}{m+1})^r$. ◀

The following result is a technical inequality required to show that the candidate for r for the reduction is well-chosen. We separate it from the main proof for the sake of clarity.

▶ **Lemma 44.** *It holds that* $\frac{1}{\log_2(\frac{m+1}{m})} \leq m$.

Proof. The inequality above is equivalent to $1+\frac{1}{m} \geq 2^{\frac{1}{m}}$. To prove this equivalent formulation, we show that the function $f\colon [1, +\infty[\to \mathbb{R}\colon z \mapsto 1 + \frac{1}{z} - 2^{\frac{1}{z}}$ is non-negative. Let $z_0 = -\log_2(\ln(2))^{-1} > 1$. We show that f is increasing on $[1, z_0]$ and decreasing on $[z_0, +\infty[$. This property implies that f is non-negative. Indeed, on the one hand, we have $f(1) = 0$, implying that f is non-negative on $[1, z_0]$. On the other hand, because $\lim_{z\to+\infty} f(z) = 0$, f is necessarily non-negative on the interval $[z_0, +\infty[$.

We study the sign of the derivative of f to determine its intervals of monotonicity. We define $g\colon [1, +\infty[\to \mathbb{R}$ such that, for all $z \in [1, +\infty[$, $g(z) = \ln(2) \cdot 2^{\frac{1}{z}} - 1$. For all $z \in [1, +\infty[$, we have $f'(z) = \frac{1}{z^2} g(z)$. We obtain that for all $z \in [1, +\infty[$, the sign of $f'(z)$ depends only on the sign of $g(z)$. The function g is a decreasing function, because $]1, +\infty[\to \mathbb{R}\colon z \mapsto 2^{\frac{1}{z}}$ is a restriction of the composition of the decreasing function $]0, +\infty[\to \mathbb{R}\colon z \mapsto \frac{1}{z}$ and the increasing function $]0, +\infty[\to \mathbb{R}\colon z \mapsto 2^z$. Because $g(1) = 2\ln(2) - 1 > 0$ and $g(z_0) = 0$, it follows that f' is positive on the interval $]1, z_0]$ and negative on the interval $]z_0, +\infty[$. This implies the desired property for f, ending the proof. ◀

We can now show that choosing $r = 2^n m \cdot (\lambda + 1) + nm^2 + 1$ (where λ is the sum of bit-sizes of the inputs to our square-root sum instance) is sufficient to achieve the precision given by Lemma 42 that ensures the validity of the reduction.

▶ **Lemma 45.** *Let* $\lambda \in \mathbb{N}_{>0}$. *For all* $r \geq 2^n m \cdot (\lambda + 1) + nm^2 + 1$, *we have* $\varepsilon_r \leq \frac{1}{nm} 2^{-2^n(\lambda+1)}$.

Proof. We claim that it is sufficient to show that, for all $r \geq 2^n m \cdot (\lambda + 1) + nm^2 + 1$,

$$\left(\frac{m}{m+1}\right)^{r-1} \leq 2^{-2^n(\lambda+1)-nm}. \tag{12}$$

We observe that m, $n \in \mathbb{N}_{>0}$ implies that $2^{-nm} \leq \frac{1}{nm}$. Combining this with Lemma 43 implies that, for all $r \in \mathbb{N}_{>0}$ such that Equation (12) holds,

$$\varepsilon_r \leq \left(\frac{m}{m+1}\right)^{r-1} \leq 2^{-2^n(\lambda+1)-nm} \leq \frac{1}{nm} 2^{-2^n(\lambda+1)}.$$

This guarantees that establishing Equation (12) for the relevant upper bounds r is sufficient.

For all $r \in \mathbb{N}_{>0}$, Equation (12) is equivalent (by applying $\log_2$ on both sides then using algebraic manipulations) to

$$r - 1 \geq \frac{2^n(\lambda+1) + nm}{\log_2\left(\frac{m+1}{m}\right)}. \tag{13}$$

By Lemma 44, Equation (13) is guaranteed to hold whenever $r - 1 \geq m(2^n(\lambda+1) + nm)$, and thus the same applies to Equation (12). This ends the proof of this lemma. ◀

We can now formulate and prove our square-root-sum hardness result for our OEIS-related problems in bounded OC-MDPs.

▶ **Theorem 46.** *The interval strategy verification, fixed-interval realisability and parameterised realisability problems for state-reachability and selective termination are square-root-sum-hard in bounded OC-MDPs.*

Proof. The reduction only differ slightly between the three considered problem; we discuss the verification problem and comment on the additional steps for the other two problems below. The reduction is the same for state reachability and selective termination, and thus we only mention the target in the following without specifying the objective. We consider inputs $x_1, \ldots, x_n$ and y to the square-root-sum problem. We assume that for all $i \in [\![1, n]\!]$, $x_i \neq 0$. Let $m = \max_{1 \leq i \leq n} x_i$ and $\bar{x} = (x_1, \ldots, x_n)$. We describe the reduction and prove its correctness.

First, we check in polynomial time whether $\sum_{i=1}^{n} \sqrt{x_i} = y$. If this equality holds, we construct an OC-MDP $\mathcal{Q}$ with a single state q and a single action a where the self-loop of q labelled by a has weight -1. We reduce our instance of the square-root-sum problem to the verification problem on $\mathcal{Q}$ with counter upper bound $r = 2$, initial configuration $(q, 1)$, target $\{q\}$ and threshold $\theta = 1$. The reduction is trivially correct in this case and is in polynomial time.

If $\sum_{i=1}^{n} \sqrt{x_i} \neq y$, we construct the OC-MDP $\mathcal{Q}_{\bar{x}}$. Let $r = 2^n m \cdot (\lambda+1) + nm^2 + 1$ where λ is the sum of bit-sizes of the inputs of the square-root-sum instance. We reduce our instance of the square-root-sum problem to the verification problem on $\mathcal{Q}_{\bar{x}}$ with counter upper bound r, initial configuration $(q_{\mathsf{init}}, 1)$, target $\{t\}$ and threshold $\theta = \frac{y}{nm}$. This reduction is in polynomial time, and thus it remains to prove its correctness.

Let $\varepsilon_r = \mathbb{P}_{\mathcal{C}^{\leq\infty}(\mathcal{Q}_{\bar{x}}),(q_{\mathsf{init}},1)}(\mathsf{Reach}(Q \times \{r\}))$. It follows from Theorem 40 that we must show that $\frac{1}{nm}\sum_{i=1}^{n} \sqrt{x_i} - \varepsilon_r \geq \theta$ if and only if $\sum_{i=1}^{n} \sqrt{x_i} \geq y$. It is direct that $\frac{1}{nm}\sum_{i=1}^{n} \sqrt{x_i} - \varepsilon_r \geq \theta$ implies $\sum_{i=1}^{n} \sqrt{x_i} \geq y$. We prove the converse implication. Assume that $\sum_{i=1}^{n} \sqrt{x_i} \geq y$. By Lemmas 42 and 45, it holds that $\varepsilon_r \leq \frac{1}{nm}|\sum_{i=1}^{n} \sqrt{x_i} - y| = \frac{1}{nm}(\sum_{i=1}^{n} \sqrt{x_i} - y)$. We obtain that $\frac{1}{nm}\sum_{i=1}^{n} \sqrt{x_i} - \varepsilon_r \geq \frac{1}{nm} \cdot y = \theta$. This shows that the reduction is correct.

It remains to comment on how to adapt the above reduction to the fixed-interval and parameterised realisability problems. Instead of specifying the strategy as an input, we specify the interval partition $\mathcal{I} = [\![1, r-1]\!]$ for the fixed-interval case, and the parameters $d = 1$ for the number of intervals and $n = r - 1$ for the size of intervals in the parameterised case. These inputs are such that we check the existence of a well-performing counter-oblivious strategy (with respect to the threshold θ specified above). ◀

7.2 NP-hardness of fixed-interval and parameterised realisability

The goal of this section is to prove that the realisability problem for counter-oblivious strategies is NP-hard for the selective termination objective. This implies the NP-hardness of the fixed-interval and parameterised realisability problems: counter-oblivious strategies are

single-interval OEISs and are also CISs with a period of one. We prove this hardness result by a reduction from the problem of deciding if a directed graph has a Hamiltonian cycle.

Let $G = (V, E)$ denote a directed graph where V is a finite set of vertices and $E \subseteq V^2$ is a set of edges. A *Hamiltonian cycle* of G is a simple cycle $v_0 v_1 \dots v_n$ such that $n = |V|$, i.e., a cycle that passes through all vertices exactly once, except the first vertex which is visited twice. Deciding whether G has a Hamiltonian cycle is NP-complete (e.g., [30]).

We sketch a reduction from the problem of deciding if a graph has a Hamiltonian cycle to the counter-oblivious strategy realisability problem for targeted termination. Let $G = (V, E)$ be a directed graph. We fix an initial vertex v_{init}. We derive an OC-MDP $\mathcal{Q}$ with deterministic transitions from G by adding vertices and redirecting transitions. We add a copy $v'_{\mathsf{init}} \notin V$ of the initial vertex and a fresh absorbing state $q \notin V$. All incoming transitions of v_{init} are redirected to v'_{init} and the only successor of v'_{init} is set to be q. All transitions are given a weight of -1. We assume a counter upper bound $r \in \{|V| + 1, \infty\}$ that exceeds the initial counter value chosen below.

We claim that there is a Hamiltonian cycle in G if and only if there is a strategy guaranteeing (almost-)sure termination in v'_{init} from the initial configuration $(v_{\mathsf{init}}, |V|)$. Intuitively, all cycles of G from v_{init} with k edges (i.e., $k + 1$ vertices) are equivalent to a history from $(v_{\mathsf{init}}, |V|)$ to $(v'_{\mathsf{init}}, |V| - k)$ in $\mathcal{M}^{\leq r}(\mathcal{Q})$. If there is a Hamiltonian cycle in G, because it is simple, we obtain a history of length $|V|$ in $\mathcal{M}^{\leq r}(\mathcal{Q})$ that can be obtained via a pure counter-oblivious strategy, and this strategy provides a positive answer to the realisability problem. For the converse, we observe that any other simple cycle from v_{init} of G yields a history terminating in the additional state q. Therefore, if there is no Hamiltonian cycle in G, all (randomised) counter-oblivious strategies have a history consistent with them that either terminates in q if v'_{init} is reached in under $|V|$ steps or terminates in V.

▶ **Theorem 47.** *The problem of deciding whether there exists a counter-oblivious (pure or randomised) strategy ensuring almost-sure selective termination is NP-hard. In particular, the fixed-interval and parameterised realisability problems for selective termination are NP-hard.*

Proof. We provide a reduction from the NP-complete problem of deciding whether a directed graph contains a Hamiltonian cycle. We fix a directed graph $G = (V, E)$ and an initial vertex $v_{\mathsf{init}} \in V$ for the remainder of the proof.

We consider $\mathcal{Q} = (Q, A, \delta, w)$ such that $Q = V \cup \{v'_{\mathsf{init}}, q\}$ (where $v'_{\mathsf{init}}, q \notin V$), $A = V$. The transition function is deterministic: we view it as a function $\delta \colon Q \times A \to Q$. We formalise δ as follows. First, for all $(v, v') \in E$ such that $v' \neq v_{\mathsf{init}}$, we let $\delta(v, v') = v'$. Second, for all $v \in V$ such that $(v, v_{\mathsf{init}}) \in E$, we let $\delta(v, v_{\mathsf{init}}) = v'_{\mathsf{init}}$. Finally, for all $v \in V$, we let $\delta(v'_{\mathsf{init}}, v) = \delta(q, v) = q$. All weights are -1. Recall that counter-oblivious strategies can be seen as functions $\sigma \colon Q \to A$.

We show that the three following assertions are equivalent:

1. there exists a Hamiltonian cycle of G;
2. there exists a pure counter-oblivious strategy σ of $\mathcal{Q}$ such that $\mathbb{P}^{\sigma}_{(v_{\mathsf{init}}, |V|)}(\mathsf{Term}(v'_{\mathsf{init}})) = 1$;
3. there exists a counter-oblivious strategy σ of $\mathcal{Q}$ such that $\mathbb{P}^{\sigma}_{(v_{\mathsf{init}}, |V|)}(\mathsf{Term}(v'_{\mathsf{init}})) = 1$;

We prove that Item 1 implies Item 2 and that Item 3 implies Item 1. The implication from Item 2 to Item 3 is direct.

We assume that there exists a Hamiltonian cycle $v_0 v_1 \dots v_{|V|}$ of G. Assume without loss of generality that $v_0 = v_{\mathsf{init}}$. It is easy to see that the pure counter-oblivious strategy σ such that $\sigma(v_\ell) = v_{\ell+1}$ for all $\ell \in [\![|V| - 1]\!]$ ensures that $\mathbb{P}^{\sigma}_{(v_{\mathsf{init}}, |V|)}(\mathsf{Term}(v'_{\mathsf{init}})) = 1$. The strategy σ is well-defined because $v_0 v_1 \dots v_{|V|-1}$ is a simple path. This shows that Item 1 implies Item 2.

We now prove the contrapositive of the implication from Item 3 to Item 1. Assume that there is no Hamiltonian cycle in G. Let σ be a counter-oblivious strategy. We show that termination occurs in a state other than v'_{init} with positive probability. If $\mathbb{P}^{\sigma}_{(v_{\mathsf{init}},|V|)}(\mathsf{Term}(v'_{\mathsf{init}})) = 0$, then the claim is direct. We assume that $\mathbb{P}^{\sigma}_{(v_{\mathsf{init}},|V|)}(\mathsf{Term}(v'_{\mathsf{init}})) > 0$. Thus, there exists a history $h = (v_0, |V|)v_1(v_1, |V| - 1)\dots(v_{|V|-1}, 1)v_{\mathsf{init}}(v'_{\mathsf{init}}, 0)$ consistent with σ such that $v_0 = v_{\mathsf{init}}$. Since there are no Hamiltonian cycles in G and h induces a cycle of G, there must be a state other than v_{init} (due to the structure of $\mathcal{Q}$) that is repeated in this induced cycle. Let $0 < \ell < \ell' < |V|$ such that $v_\ell = v_{\ell'}$. It is easy to see that the history starting in $(v_0, |V|)$ that follows h up to index ℓ' and then loops in the cycle between v_ℓ and $v_{\ell'}$ until termination is consistent with σ and therefore $\mathbb{P}^{\sigma}_{(v_{\mathsf{init}},|V|)}(\mathsf{Term}(v'_{\mathsf{init}})) < 1$. This ends the proof that Item 3 implies Item 1.

To conclude, we note that $\mathcal{Q}$ can be constructed in polynomial time. This ends our NP-hardness proof. ◀

References

1 Eric Allender, Peter Bürgisser, Johan Kjeldgaard-Pedersen, and Peter Bro Miltersen. On the complexity of numerical analysis. *SIAM Journal on Computing*, 38(5):1987–2006, 2009. doi:10.1137/070697926.

2 Pranav Ashok, Mathias Jackermeier, Jan Kretínský, Christoph Weinhuber, Maximilian Weininger, and Mayank Yadav. dtControl 2.0: Explainable strategy representation via decision tree learning steered by experts. In Jan Friso Groote and Kim Guldstrand Larsen, editors, *Proceedings (Part II) of the 27th International Conference on Tools and Algorithms for the Construction and Analysis of Systems, TACAS 2021, Held as Part of ETAPS 2021, Luxemburg City, Luxemburg, March 27–April 1, 2021*, volume 12652 of *Lecture Notes in Computer Science*, pages 326–345. Springer, 2021. doi:10.1007/978-3-030-72013-1_17.

3 Christel Baier and Joost-Pieter Katoen. *Principles of model checking*. MIT Press, 2008.

4 Nikhil Balaji, Stefan Kiefer, Petr Novotný, Guillermo A. Pérez, and Mahsa Shirmohammadi. On the complexity of value iteration. In Christel Baier, Ioannis Chatzigiannakis, Paola Flocchini, and Stefano Leonardi, editors, *Proceedings of the 46th International Colloquium on Automata, Languages, and Programming, ICALP 2019, Patras, Greece, July 9–12, 2019*, volume 132 of *LIPIcs*, pages 102:1–102:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. URL: https://doi.org/10.4230/LIPIcs.ICALP.2019.102, doi:10.4230/LIPICS.ICALP.2019.102.

5 Saugata Basu, Richard Pollack, and Marie-Françoise Roy. *Algorithms in Real Algebraic Geometry*. 1431-1550. Springer, 2nd edition, 2006. doi:10.1007/3-540-33099-2.

6 Noam Berger, Nevin Kapur, Leonard J. Schulman, and Vijay V. Vazirani. Solvency games. In Ramesh Hariharan, Madhavan Mukund, and V. Vinay, editors, *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2008, December 9-11, 2008, Bangalore, India*, volume 2 of *LIPIcs*, pages 61–72. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2008. URL: https://doi.org/10.4230/LIPIcs.FSTTCS.2008.1741, doi:10.4230/LIPICS.FSTTCS.2008.1741.

7 Kark-Josef Bierth. An expected average reward criterion. *Stochastic processes and their applications*, 26:123–140, 1987.

8 Frantisek Blahoudek, Tomás Brázdil, Petr Novotný, Melkior Ornik, Pranay Thangeda, and Ufuk Topcu. Qualitative controller synthesis for consumption Markov decision processes. In Shuvendu K. Lahiri and Chao Wang, editors, *Proceedings (Part II) of the 32nd International Conference on Computer Aided Verification, CAV 2020, Los Angeles, CA, USA, July 21–24, 2020*, volume 12225 of *Lecture Notes in Computer Science*, pages 421–447. Springer, 2020. doi:10.1007/978-3-030-53291-8_22.

9 Roderick Bloem, Krishnendu Chatterjee, and Barbara Jobstmann. Graph games and reactive synthesis. In Edmund M. Clarke, Thomas A. Henzinger, Helmut Veith, and

Roderick Bloem, editors, *Handbook of Model Checking*, pages 921–962. Springer, 2018. doi:10.1007/978-3-319-10575-8_27.

10 Lenore Blum, Mike Shub, and Steve Smale. Over the real numbers: NP-completeness, recursive functions and universal machines. *Bulletin of the American Mathematical Society*, 21(1), 1989.

11 Allan Borodin, Ronald Fagin, John E. Hopcroft, and Martin Tompa. Decreasing the nesting depth of expressions involving square roots. *Journal of Symbolic Computation*, 1(2):169–188, 1985. doi:10.1016/S0747-7171(85)80013-4.

12 Patricia Bouyer, Stéphane Le Roux, Youssouf Oualhadj, Mickael Randour, and Pierre Vandenhove. Games where you can play optimally with arena-independent finite memory. *Logical Methods in Computer Science*, 18(1), 2022. doi:10.46298/lmcs-18(1:11)2022.

13 Patricia Bouyer, Youssouf Oualhadj, Mickael Randour, and Pierre Vandenhove. Arena-independent finite-memory determinacy in stochastic games. *Log. Methods Comput. Sci.*, 19(4), 2023. URL: https://doi.org/10.46298/lmcs-19(4:18)2023, doi:10.46298/LMCS-19(4:18)2023.

14 Patricia Bouyer, Mickael Randour, and Pierre Vandenhove. The true colors of memory: A tour of chromatic-memory strategies in zero-sum games on graphs (invited talk). In Anuj Dawar and Venkatesan Guruswami, editors, *Proceedings of the 42nd IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2022, IIT Madras, Chennai, India, December 18–20, 2022*, volume 250 of *LIPIcs*, pages 3:1–3:18. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2022. doi:10.4230/LIPIcs.FSTTCS.2022.3.

15 Tomás Brázdil, Václav Brozek, Kousha Etessami, and Antonín Kucera. Approximating the termination value of one-counter MDPs and stochastic games. *Information and Computation*, 222:121–138, 2013. URL: https://doi.org/10.1016/j.ic.2012.01.008, doi:10.1016/J.IC.2012.01.008.

16 Tomás Brázdil, Václav Brozek, Kousha Etessami, Antonín Kucera, and Dominik Wojtczak. One-counter Markov decision processes. In Moses Charikar, editor, *Proceedings of the Twenty-First Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2010, Austin, Texas, USA, January 17-19, 2010*, pages 863–874. SIAM, 2010. doi:10.1137/1.9781611973075.70.

17 Tomás Brázdil, Krishnendu Chatterjee, Martin Chmelik, Andreas Fellner, and Jan Kretínský. Counterexample explanation by learning small strategies in Markov decision processes. In Daniel Kroening and Corina S. Pasareanu, editors, *Proceedings (Part I) of the 27th International Conference on Computer Aided Verification, CAV 2015, San Francisco, CA, USA, July 18–24, 2015*, volume 9206 of *Lecture Notes in Computer Science*, pages 158–177. Springer, 2015. doi:10.1007/978-3-319-21690-4_10.

18 Tomás Brázdil, Krishnendu Chatterjee, Jan Kretínský, and Viktor Toman. Strategy representation by decision trees in reactive synthesis. In Dirk Beyer and Marieke Huisman, editors, *Proceedings (Part I) of the 24th International Conference on Tools and Algorithms for the Construction and Analysis of Systems, TACAS 2018, Held as Part of ETAPS 2018, Thessaloniki, Greece, April 14–20, 2018*, volume 10805 of *Lecture Notes in Computer Science*, pages 385–407. Springer, 2018. doi:10.1007/978-3-319-89960-2_21.

19 Tomás Brázdil, Krishnendu Chatterjee, Antonín Kucera, and Petr Novotný. Efficient controller synthesis for consumption games with multiple resource types. In P. Madhusudan and Sanjit A. Seshia, editors, *Computer Aided Verification - 24th International Conference, CAV 2012, Berkeley, CA, USA, July 7-13, 2012 Proceedings*, volume 7358 of *Lecture Notes in Computer Science*, pages 23–38. Springer, 2012. doi:10.1007/978-3-642-31424-7_8.

20 Tomás Brázdil, Stefan Kiefer, and Antonín Kucera. Efficient analysis of probabilistic programs with an unbounded counter. In Ganesh Gopalakrishnan and Shaz Qadeer, editors, *Computer Aided Verification - 23rd International Conference, CAV 2011, Snowbird, UT, USA, July 14-20, 2011. Proceedings*, volume 6806 of *Lecture Notes in Computer Science*, pages 208–224. Springer, 2011. doi:10.1007/978-3-642-22110-1_18.

21 Thomas Brihaye, Aline Goeminne, James C. A. Main, and Mickael Randour. Reachability games and friends: A journey through the lens of memory and complexity (invited talk).

In Patricia Bouyer and Srikanth Srinivasan, editors, *Proceedings of the 43rd IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2023, IIIT Hyderabad, Telangana, India, December 18–20, 2023*, volume 284 of *LIPIcs*, pages 1:1–1:26. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2023. URL: https://doi.org/10.4230/LIPIcs.FSTTCS.2023.1, doi:10.4230/LIPICS.FSTTCS.2023.1.

22 John F. Canny. Some algebraic and geometric computations in PSPACE. In Janos Simon, editor, *Proceedings of the 20th Annual ACM Symposium on Theory of Computing, STOC 1988, Chicago, Illinois, USA, May 2–4, 1988*, pages 460–467. ACM, 1988. doi:10.1145/62212.62257.

23 Krishnendu Chatterjee, Marcin Jurdzinski, and Thomas A. Henzinger. Quantitative stochastic parity games. In J. Ian Munro, editor, *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2004, New Orleans, Louisiana, USA, January 11–14, 2004*, pages 121–130. SIAM, 2004. URL: http://dl.acm.org/citation.cfm?id=982792.982808.

24 Krishnendu Chatterjee, Mickael Randour, and Jean-François Raskin. Strategy synthesis for multi-dimensional quantitative objectives. *Acta Informatica*, 51(3-4):129–163, 2014. doi:10.1007/s00236-013-0182-6.

25 Florent Delgrange, Joost-Pieter Katoen, Tim Quatmann, and Mickael Randour. Simple strategies in multi-objective MDPs. In Armin Biere and David Parker, editors, *Proceedings (Part I) of the 26th International Conference on Tools and Algorithms for the Construction and Analysis of Systems, TACAS 2020, Held as Part of ETAPS 2020, Dublin, Ireland, April 25–30, 2020*, volume 12078 of *Lecture Notes in Computer Science*, pages 346–364. Springer, 2020. doi:10.1007/978-3-030-45190-5_19.

26 Rick Durrett. *Probability: Theory and Examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 5th edition, 2019. doi:10.1017/9781108591034.

27 Kousha Etessami, Marta Z. Kwiatkowska, Moshe Y. Vardi, and Mihalis Yannakakis. Multi-objective model checking of Markov decision processes. *Logical Methods in Computer Science*, 4(4), 2008. doi:10.2168/LMCS-4(4:8)2008.

28 Kousha Etessami, Dominik Wojtczak, and Mihalis Yannakakis. Quasi-birth-death processes, tree-like QBDs, probabilistic 1-counter automata, and pushdown systems. *Performance Evaluation*, 67(9):837–857, 2010. URL: https://doi.org/10.1016/j.peva.2009.12.009, doi:10.1016/J.PEVA.2009.12.009.

29 Nathanaël Fijalkow, Nathalie Bertrand, Patricia Bouyer-Decitre, Romain Brenguier, Arnaud Carayol, John Fearnley, Hugo Gimbert, Florian Horn, Rasmus Ibsen-Jensen, Nicolas Markey, Benjamin Monmege, Petr Novotný, Mickael Randour, Ocan Sankur, Sylvain Schmitz, Olivier Serre, and Mateusz Skomra. Games on graphs. *CoRR*, abs/2305.10546, 2023. doi:10.48550/arXiv.2305.10546.

30 M. R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman, 1979.

31 Hugo Gimbert. Pure stationary optimal strategies in Markov decision processes. In Wolfgang Thomas and Pascal Weil, editors, *Proceedings of the 24th Annual Symposium on Theoretical Aspects of Computer Science, STACS 2007, Aachen, Germany, February 22–24, 2007*, volume 4393, pages 200–211. Springer, 2007. doi:10.1007/978-3-540-70918-3_18.

32 Alston Scott Householder. *The Numerical Treatment of a Single Nonlinear Equation*. McGraw-Hill, 1970.

33 Florian Jüngermann, Jan Kretínský, and Maximilian Weininger. Algebraically explainable controllers: decision trees and support vector machines join forces. *International Journal on Software Tools for Technology Transfer*, 25(3):249–266, 2023. URL: https://doi.org/10.1007/s10009-023-00716-z, doi:10.1007/S10009-023-00716-Z.

34 Stefan Kiefer, Richard Mayr, Mahsa Shirmohammadi, Patrick Totzke, and Dominik Wojtczak. How to play in infinite MDPs (invited talk). In Artur Czumaj, Anuj Dawar, and Emanuela

Merelli, editors, *Proceedings of the 47th International Colloquium on Automata, Languages, and Programming, ICALP 2020, Saarbrücken, Germany, July 8–11, 2020*, volume 168 of *LIPIcs*, pages 3:1–3:18. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. URL: https://doi.org/10.4230/LIPIcs.ICALP.2020.3, doi:10.4230/LIPICS.ICALP.2020.3.

35 Antonín Kucera, Javier Esparza, and Richard Mayr. Model checking probabilistic pushdown automata. *Logical Methods in Computer Science*, 2(1), 2006. doi:10.2168/LMCS-2(1:2)2006.

36 Serge Lang. *Algebra*, volume 211 of *Graduate Texts in Mathematics*. Springer, revised third edition edition, 2002. URL: https://www.springer.com/gp/book/9780387953854.

37 James C. A. Main and Mickael Randour. Different strokes in randomised strategies: Revisiting Kuhn's theorem under finite-memory assumptions. *Information and Computation*, 301:105229, 2024. URL: https://doi.org/10.1016/j.ic.2024.105229, doi:10.1016/J.IC.2024.105229.

38 James C. A. Main and Mickael Randour. Mixing any cocktail with limited ingredients: On the structure of payoff sets in multi-objective MDPs and its impact on randomised strategies. *CoRR*, abs/2502.18296, 2025. doi:10.48550/arXiv.2502.18296.

39 Donald Ornstein. On the existence of stationary optimal strategies. *Proceedings of the American Mathematical Society*, 20(2):563–569, 1969. URL: http://www.jstor.org/stable/2035700.

40 Joël Ouaknine and James Worrell. Positivity problems for low-order linear recurrence sequences. In Chandra Chekuri, editor, *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014, Portland, Oregon, USA, January 5–7, 2014*, pages 366–379. SIAM, 2014. doi:10.1137/1.9781611973402.27.

41 Jakob Piribauer and Christel Baier. Positivity-hardness results on Markov decision processes. *TheoretiCS*, 3, 2024. URL: https://doi.org/10.46298/theoretics.24.9, doi:10.46298/THEORETICS.24.9.

42 Mickael Randour. Automated synthesis of reliable and efficient systems through game theory: A case study. In *Proc. of ECCS 2012*, Springer Proceedings in Complexity XVII, pages 731–738. Springer, 2013. doi:10.1007/978-3-319-00395-5_90.

43 Mickael Randour, Jean-François Raskin, and Ocan Sankur. Percentile queries in multi-dimensional Markov decision processes. *Formal methods in system design*, 50(2-3):207–248, 2017. doi:10.1007/s10703-016-0262-7.

44 Walter J. Savitch. Relationships between nondeterministic and deterministic tape complexities. *Journal of Computer and System Sciences*, 4(2):177–192, 1970. doi:10.1016/S0022-0000(70)80006-X.

45 Lloyd S. Shapley. Stochastic games. *Proceedings of the National Academy of Sciences*, 39(10):1095–1100, 1953. URL: https://www.pnas.org/content/39/10/1095, arXiv:https://www.pnas.org/content/39/10/1095.full.pdf, doi:10.1073/pnas.39.10.1095.

46 Richard S. Sutton and Andrew G. Barto. *Reinforcement Learning: An Introduction*. MIT Press, 2018.

47 Prasoon Tiwari. A problem that is easier to solve on the unit-cost algebraic RAM. *Journal of Complexity*, 8(4):393–397, 1992. doi:10.1016/0885-064X(92)90003-T.

A Bounds on sufficient approximations for square-root sum

The goal of this section is to prove Lemma 42. Let $x_1, \ldots, x_n \in \mathbb{N}$ and $y \in \mathbb{N}$ be fixed for the remainder of this section. This section has three parts. First, we recall some field-theoretic notions that are required for the proof. We refer the reader to [36] for a reference on field theory. Second, we show that all roots of the minimal polynomial of $\sum_{i=1}^{n} \sqrt{x_i} - y$ are of a certain form. We end with a proof of Lemma 42.

Field-theoretic background. A complex number is *algebraic* (over $\mathbb{Q}$) if it is the root of a polynomial with rational coefficients. An *algebraic extension* of $\mathbb{Q}$ is a field K such that $\mathbb{Q} \subseteq K \subseteq \mathbb{C}$ such that all elements of K are algebraic. Let $K \subseteq L \subseteq \mathbb{C}$ be algebraic extensions of $\mathbb{Q}$. We write L/K as shorthand to mean that L is an extension of K. The *minimum polynomial* of $\alpha \in L$ over K is the unique monic polynomial with coefficients in K of minimum degree that has α as a root. An algebraic number is an *algebraic integer* if its minimal polynomial over $\mathbb{Q}$ has integer coefficients. Algebraic integers form a sub-ring of the algebraic closure of $\mathbb{Q}$.

Given algebraic numbers $\alpha_1, \ldots, \alpha_\ell$, we let $K(\alpha_1, \ldots, \alpha_\ell)$ be the smallest algebraic extension of K containing $\alpha_1, \ldots, \alpha_\ell$. The *degree* of the extension L/K, denoted by $[L : K]$, is the dimension of L as a vector space over K, and if $L = K(\alpha)$, $[L : K]$ is the degree of the minimal polynomial of α over K. Degrees of successive extensions multiply, in the sense that, given F/L, it holds that $[F : K] = [F : L] \cdot [L : K]$.

An extension L/K is *Galois* if and only if any embedding of K in the algebraic closure of $\mathbb{Q}$ induces an automorphism of K. Given a polynomial $P \in K[X]$, the *splitting field* of P is the smallest algebraic extension of K that contains all of the complex roots of P. If L/K is of finite degree, then L/K is Galois if and only if it is the splitting field of a polynomial in $K[X]$. If L/K is Galois, the Galois group $\mathsf{Gal}(L/K)$ of L/K is the group formed by the (field) automorphisms of L whose restriction to K is the identity function over K. The order of the Galois group of a Galois extension is the degree of the extension.

Minimal polynomials. The following lemma provides a set that is guaranteed to contain all roots of the minimal polynomial of $\sum_{i=1}^{n} \sqrt{x_i} - y$. Through this lemma, we can bound the coefficients of the minimal polynomial of $\sum_{i=1}^{n} \sqrt{x_i} - y$, which is the crux of the proof of Lemma 42.

▶ **Lemma 48.** *Let $\beta = \sum_{i=1}^{n} \sqrt{x_i} - y$. The minimal polynomial of β over $\mathbb{Q}$ has at most 2^n roots and all are included in the set $\{\sum_{i=1}^{n} (-1)^{b_i} \sqrt{x_i} - y \mid (b_1, \ldots, b_n) \in \{0,1\}^n\}$.*

Proof. Let P_β denote the minimum polynomial of β and let $K = \mathbb{Q}(\sqrt{x_1}, \ldots, \sqrt{x_n})$. To bound the number of roots of P_β, it suffices to bound its degree, i.e., $[\mathbb{Q}(\beta) : \mathbb{Q}]$. We have $\mathbb{Q}(\beta) \subseteq K$ because $\beta \in K$ by definition of K. It follows that $[\mathbb{Q}(\beta) : \mathbb{Q}]$ is a divisor of $[K : \mathbb{Q}]$. We have that $[K : \mathbb{Q}]$ is at most 2^n because

$$[K : \mathbb{Q}] = \prod_{0 \leq i \leq n-1} [\mathbb{Q}(\sqrt{x_1}, \ldots, \sqrt{x_{i+1}}) : \mathbb{Q}(\sqrt{x_1}, \ldots, \sqrt{x_i})]$$

and the degrees in the product are one or two; for all $1 \leq i \leq n$, $\sqrt{x_i}$ is a root of $X^2 - x_i$.

We now show that all roots of P_β are in $\{\sum_{i=1}^{n} (-1)^{b_i} \sqrt{x_i} - y \mid (b_1, \ldots, b_n) \in \{0,1\}^n\}$. First, we note that K is the splitting field of the polynomial $\prod_{i=1}^{n} (X^2 - x_i)$. Therefore, $K/\mathbb{Q}$ is Galois (as its degree is finite).

We determine the Galois group of $K/\mathbb{Q}$. Let $R \subseteq \{\sqrt{x_1}, \ldots, \sqrt{x_n}\}$ be a minimal set such that $K = \mathbb{Q}(R)$. Assume that $R = \{\sqrt{x_1}, \ldots, \sqrt{x_{n'}}\}$. For all $1 \leq i \leq n'$, $[K :$

$\mathbb{Q}(R \setminus \{\sqrt{x_i}\})] = 2$ and $K/\mathbb{Q}(R \setminus \{\sqrt{x_i}\})$ is Galois. It follows that the group $\mathsf{Gal}(K/\mathbb{Q})$ contains, for all $1 \leq i \leq n'$, the automorphism of K in $\mathsf{Gal}(K/\mathbb{Q}(R \setminus \{\sqrt{x_i}\}))$ that is such that $\sqrt{x_i} \mapsto -\sqrt{x_i}$ that leaves other elements of R unchanged. These different automorphisms commute. It follows that these automorphisms generate the Galois group $\mathsf{Gal}(K/\mathbb{Q})$ whose order is $2^{n'}$.

Let L denote the splitting field of P_β (thus $L/\mathbb{Q}$ is Galois). It holds that $L \subseteq K$, because $K/\mathbb{Q}$ is Galois and P_β has a root in K. On the one hand, elements of $\mathsf{Gal}(L/\mathbb{Q})$ are the restrictions of elements of $\mathsf{Gal}(K/\mathbb{Q})$. On the other hand, the action of $\mathsf{Gal}(L/\mathbb{Q})$ on the set of roots of P_β is transitive (because P_β is irreducible in $\mathbb{Q}[X]$). It follows that the roots are all of the claimed form. ◀

Proof of Lemma 42. We now provide a proof of Lemma 42.

▶ **Lemma 42.** *Let λ be the sum of the bit-sizes of $x_1, \ldots, x_n$ and y. If $\sum_{i=1}^{n} \sqrt{x_i} \neq y$, then $|\sum_{i=1}^{n} \sqrt{x_i} - y| > 2^{-2^n(\lambda+1)}$.*

Proof. Assume that $\sum_{i=1}^{n} \sqrt{x_i} \neq y$. Let P_β denote the minimal polynomial of $\beta = \sum_{i=1}^{n} \sqrt{x_i} - y$. It has integer coefficients, because β is an algebraic integer. Indeed, square roots of integers are algebraic integers and algebraic integers are a sub-ring of the algebraic closure of $\mathbb{Q}$. We bound the coefficients of P_β to conclude via the following result: the non-zero roots of a non-zero polynomial with k-bit integer coefficients are greater than 2^{-k} in absolute value [32, 47].

Let d denote the degree of P_β. We have $d \leq 2^n$ by Lemma 48. By the same result, it follows that the roots of P_β are of absolute value at most $\sum_{i=1}^{n} x_i + y < 2^\lambda$. From the decomposition of P_β in linear factors, we obtain that its coefficients are sums of at most 2^d products of roots of P_β, and thus are strictly less than $2^d \cdot (2^\lambda)^d = 2^{d(\lambda+1)}$ in absolute value, i.e., their bit-size is at most $d(\lambda+1)$. We obtain that $|\beta| > 2^{-d(\lambda+1)} \geq 2^{-2^n(\lambda+1)}$. ◀