On Piecewise Affine Reachability with Bellman Operators

Anton Varonka
TU Wien, Austria

Kazuki Watanabe
National Institute of Informatics, Japan
The Graduate University for Advanced Studies (SOKENDAI), Hayama, Japan

Abstract

A piecewise affine map is one of the simplest mathematical objects exhibiting complex dynamics. The reachability problem of piecewise affine maps is as follows: Given two vectors $\boldsymbol{s}, \boldsymbol{t} \in \mathbb{Q}^d$ and a piecewise affine map $f\colon \mathbb{Q}^d \to \mathbb{Q}^d$, is there $n \in \mathbb{N}$ such that $f^n(\boldsymbol{s}) = \boldsymbol{t}$? Koiran, Cosnard, and Garzon show that the reachability problem of piecewise affine maps is undecidable even in dimension 2.

Most of the recent progress has been focused on decision procedures for one-dimensional piecewise affine maps, where the reachability problem has been shown to be decidable for some subclasses. However, the general undecidability discouraged research into positive results in arbitrary dimension.

In this work, we investigate a rich subclass of piecewise affine maps arising as Bellman operators of Markov decision processes (MDPs). We consider the reachability problem restricted to this subclass and examine its decidability in arbitrary dimensions. We establish that the reachability problem for Bellman operators is decidable in any dimension under either of the following conditions: (i) the target vector $\boldsymbol{t}$ is not the fixed point of the operator f; or (ii) the initial and target vectors $\boldsymbol{s}$ and $\boldsymbol{t}$ are comparable with respect to the componentwise order. Furthermore, we show that the reachability problem for two-dimensional Bellman operators is decidable for arbitrary $\boldsymbol{s}, \boldsymbol{t} \in \mathbb{Q}^d$, in contrast to the known undecidability of reachability for general piecewise affine maps.

2012 ACM Subject Classification Theory of computation → Abstract machines; Mathematics of computing → Markov processes; Theory of computation → Program verification

Keywords and phrases piecewise affine map, reachability, value iteration, Markov decision process, Bellman operator

Digital Object Identifier 10.4230/LIPIcs.MFCS.2025.92

Funding The authors were supported by the ASPIRE grant No. JPMJAP2301, JST.
This paper is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement No. 10103444).
Anton Varonka: A. Varonka gratefully acknowledges the support of the ERC consolidator grant ARTIST 101002685.

Acknowledgements We would like to thank the anonymous reviewers for their valuable comments and suggestions.
We further thank Ichiro Hasuo and Laura Kovács for their helpful feedback.

1 Introduction

Solving reachability problems is central to formal verification, but it is challenging, as evidenced by a body of work that has been pursued for decades. Specifically, the reachability problem that we are interested in asks the following question: Given two vectors $\boldsymbol{s}, \boldsymbol{t}$ and a map f, is there $n \in \mathbb{N}$ such that $f^n(\boldsymbol{s}) = \boldsymbol{t}$? One of the seminal results is given by

Kannan and Lipton [17,18]. They answer the decidability question in the affirmative – by presenting a novel polynomial-time algorithm – for *affine maps* f and vectors $\boldsymbol{s}, \boldsymbol{t}$ with rational coefficients.[1]

Unfortunately, the reachability problem becomes undecidable by slightly extending the class of maps beyond the affine maps – *piecewise affine maps (PAMs).* Koiran et al. [20] show that the reachability problem for PAMs is undecidable even in the two-dimensional space, which witnesses the significant difficulty of the problem, compared to that of affine maps. Specifically, a PAM f on the domain $\mathcal{D}$ is a function with the property that for some family $\{P_1, \ldots, P_k\}$ of sets such that $P_1 \cup \cdots \cup P_k = \mathcal{D}$, the restriction of f to each P_i is an affine function. We consider piecewise affine maps over the domain $\mathcal{D} = [0,1]^d$, where $[0,1]$ is the unit interval. Each of the finitely many *pieces* $P_1, \ldots, P_k$ is defined by a conjunction of finitely many linear inequalities.

▶ **Example 1.** Consider an example of a PAM in the dimension $d = 2$. Let $f : [0,1] \times [0,1] \to [0,1] \times [0,1]$ be defined by $f(x_1, x_2) = (x_1', x_2')$ with

$$x_1' = \frac{1}{2}x_2 + \frac{1}{3}, \qquad x_2' = \begin{cases} \frac{1}{2}x_1 + \frac{1}{2}, & \text{if } x_1 \geq x_2, \\ \frac{1}{4}x_1 + \frac{1}{4}x_2 + \frac{1}{2}, & \text{if } x_1 < x_2. \end{cases}$$

Here, P_1 is defined as $\{\boldsymbol{x} = (x_1, x_2) : x_1 \geq x_2\}$ and P_2 as $\{\boldsymbol{x} = (x_1, x_2) : x_1 < x_2\}$, where according to the definition $P_1 \cup P_2 = \mathcal{D}$.

In fact, piecewise affine maps characterise a very rich mathematical object [3,7]. The reachability problem for PAMs is thus one of a series of challenging problems whose crux lies in the unpredictable behaviour of iterative maps and corresponding discrete-time dynamical systems, see also [3,11,19,27].

A natural question might be: *When does the reachability problem for PAMs become decidable?* Indeed, the reachability problem for PAMs on a unit interval (dimension one) has received attention in recent research [8,13,21], where classes of PAMs with decidable reachability problems have been found while fostering new elaborate techniques. Yet, the decidability of the general one-dimensional reachability problem for PAMs remains open, even for maps defined with two pieces.

Our Approach. In this work, we propose an orthogonal approach to investigate the challenges behind the reachability problem for PAMs, focusing not on the restriction of the *dimension* or the number of *pieces*, but on the restriction of the *structure* of PAMs. Specifically, we consider the reachability problem for the *Bellman operators* on *Markov decision processes (MDPs)* – Bellman operators are, in fact, PAMs that have been studied mostly in the context of software verification or reinforcement learning [1,4,24]. For instance, the PAM in Example 1 is a Bellman operator.

MDPs are a standard probabilistic model for systems with uncertainties, and the least fixed points of the Bellman operators $\Phi\colon [0,1]^d \to [0,1]^d$ represent the "optimal" reachability probability to the specific target state. Here, optimal means that the maximum reachability probability is induced by a *scheduler* that resolves the non-deterministic behaviour on MDPs. We formulate our target problem as follows: Given two vectors $\boldsymbol{s}, \boldsymbol{t} \in [0,1]^d$ and a Bellman operator $\Phi\colon [0,1]^d \to [0,1]^d$, is there $n \in \mathbb{N}$ such that $\Phi^n(\boldsymbol{s}) = \boldsymbol{t}$? We refer to this problem as the *reachability problem for Bellman operators* (or BOR, for Bellman Operator Reachability).

[1] The results by Kannan and Lipton hold for linear maps. Since the reachability problem is decidable in arbitrary dimension, a standard technique extends it to the affine maps by encoding a d-dimensional affine map as a linear map in dimension $d + 1$; see also [27, Section 5].

Under a reasonable assumption, from any vector $\boldsymbol{s}$, the sequence $\langle \Phi^n(\boldsymbol{s}) \rangle_{n \in \mathbb{N}}$ converges to the unique fixed point $\mu\Phi$. Here, the unique fixed point is precisely the vector of optimal reachability probabilities from each state to the target state. The iterative procedure where $\mu\Phi$ is approximated by applying Φ is referred to as *value iteration* and is widely studied [1, 2, 10]. Note that the unique fixed point is computable in polynomial time by linear programming [1], so for our problem we can assume that we know the unique fixed point $\mu\Phi$ a priori. Nevertheless, even a convergent sequence does not generally reach $\mu\Phi$ at any n. In fact, the question of the reachability to the fixed point (in finite time) is known in both theoretical computer science and software verification communities. A case in point is the discussion in [20], where this question is explicitly asked for one-dimensional PAMs (see also [6]). Furthermore, in a value iteration survey [10] the same property is listed. While the authors observe that fixed points are not reachable in general, they do not discuss the decidability aspect. In the present paper, we investigate the decidability of reaching $\boldsymbol{t} = \mu\Phi$ as part of our problem.

Contributions. We present some decidability results for our target problem under a condition that ensures the existence of the unique fixed point $\mu\Phi$ of the Bellman operators [15]. First, we show that the reachability problem for Bellman operators is decidable for any dimension if the target vector $\boldsymbol{t}$ does not coincide with $\mu\Phi$ ($\boldsymbol{t} \neq \mu\Phi$). This is true because from any vector $\boldsymbol{s}$, the iteration of the Bellman operator converges to $\mu\Phi$. It becomes rather non-trivial when $\boldsymbol{t} = \mu\Phi$, that is, for the reachability problem to the unique fixed point $\mu\Phi$. We show that the reachability problem for Bellman operators when $\boldsymbol{t} = \mu\Phi$ is decidable for any dimension if $\boldsymbol{s}$ is comparable to $\mu\Phi$, that is, either $\boldsymbol{s} \leq \mu\Phi$ or $\mu\Phi \leq \boldsymbol{s}$ holds for the componentwise order. The crux is to show that eventually only "optimal" actions are chosen, and we reduce the reachability problem to a simple qualitative reachability problem that can be shown decidable.

Finally, we address the remaining case: $\boldsymbol{t} = \mu\Phi$ and $\boldsymbol{s}$ is incomparable to $\mu\Phi$. In dimension 2, we show an algorithmic procedure also for this case – finding the last piece of the puzzle – the reachability problem for two-dimensional Bellman operators is thus decidable. Our argument is based on analysing the equivalent problem for matrix semigroups, and our proof exploits the existence of a total order on the lines induced by actions. To the best of our knowledge, this is the first result to give a reasonably large class of PAMs for which the reachability problem is decidable in the two-dimensional case.

Organization. We outline our paper as follows.

- In Section 2, we formally define the main objects of our study, MDPs and Bellman operators, and recall some known properties.
- In Section 3, we show that the reachability problem for Bellman operators is decidable when either $\boldsymbol{t} \neq \mu\Phi$ (Proposition 12) or $\boldsymbol{s}$ is comparable to $\boldsymbol{t} = \mu\Phi$ (Theorem 22). Importantly, the result holds for arbitrary dimension.
- In Section 4, we prove that the reachability problem for Bellman operators is decidable in the two-dimensional case (Theorem 31), by presenting an algorithm that solves the remaining case ($\boldsymbol{s}$ is incomparable to $\boldsymbol{t} = \mu\Phi$).
- In Section 5, we discuss our result with related work, and list some future directions.

2 Background

We first recall some definitions and properties for Markov decision processes (MDPs) and their Bellman operators, which are necessary for our development. We then define our target problem, namely the piecewise affine reachability problem with Bellman operators.

2.1 Preliminary

▶ **Definition 2** (MDP [24]). *An MDP $\mathcal{M}$ is a tuple $(S, Act, \mathbb{P})$ such that (i) S is a finite non-empty set of* states*; (ii) Act is an indexed family $(Act_s)_{s\in S}$ of finite sets of* actions *such that the set Act_s and $Act_{s'}$ of actions on s and s' are disjoint for any $s, s' \in S$; and (iii) $\mathbb{P}$ is the transition probability function $\mathbb{P}(s, \alpha, \_) \in S \to [0, 1] \cap \mathbb{Q}$ with finite support that satisfies $\sum_{s'\in S} \mathbb{P}(s, \alpha, s') = 1$, for any $s \in S$ and $\alpha \in Act_s$.*

We refer to the support of $\mathbb{P}(s, \alpha, \_)$ by $\mathrm{supp}(s, \alpha)$. We fix a *target state* t and assume that t is a sink, i.e., $Act_t = \emptyset$.

▶ **Example 3.** We present an MDP $\mathcal{M} = (S, Act, \mathbb{P})$, where (i) $S = \{s_1, s_2, s_3, t\}$; (ii) $Act_{s_1} := \{\alpha\}$, $Act_{s_2} := \{\beta_1, \beta_2\}$, and $Act_{s_3} = Act_t := \emptyset$; and (iii) $\mathbb{P}$ is defined by

$$\begin{aligned}
&\mathbb{P}(s_1, \alpha, s_2) := 1/2, \quad \mathbb{P}(s_1, \alpha, s_3) := 1/6, \quad \mathbb{P}(s_1, \alpha, t) := 1/3,\\
&\mathbb{P}(s_2, \beta_1, s_1) := 1/2, \quad \mathbb{P}(s_2, \beta_1, t) := 1/2,\\
&\mathbb{P}(s_2, \beta_2, s_1) := 1/4, \quad \mathbb{P}(s_2, \beta_2, s_2) := 1/4, \quad \mathbb{P}(s_2, \beta_2, t) := 1/2.
\end{aligned}$$

Given states $s, s' \in S$, a *path* π from s to s' is a sequence $\pi := (s_1, \ldots, s_m)$ such that $s_i \in S\backslash\{t\}$ for any $i \in [1, m-1]$, $s_1 = s$, and $s_m = s'$. We denote the set of paths from s to s' by $\mathrm{Path}(s, s')$. A *scheduler* is a function $\sigma \colon S^+ \to \cup_{s\in S} Act_s$ such that $\sigma(s_1 \cdots s_m) \in Act_{s_m}$. As deterministic schedulers suffice for the reachability objective [1], we further consider the set Σ of all deterministic schedulers. A scheduler is *positional* if for any $s_1 \cdots s_m \cdot s$ and $s'_1 \cdots s'_n \cdot s$, the actions $\sigma(s_1 \cdots s_m \cdot s)$ and $\sigma(s'_1 \cdots s'_n \cdot s)$ coincide. For a path $\pi := (s_1, \ldots, s_m)$ and a scheduler $\sigma \in \Sigma$, define $\mathbb{P}^\sigma(\pi) := \prod_{i\in[1,m-1]} \mathbb{P}(s_i, \sigma(\pi_i), s_{i+1})$, where $\pi_i = (s_1, \ldots, s_i)$.

▶ **Definition 4** (reachability probability). *Given a scheduler σ, and $s \in S$, the* reachability probability *$\mathbb{P}^\sigma(s \models \Diamond t)$ under σ is defined by $\mathbb{P}^\sigma(s \models \Diamond t) := \sum_{\pi\in \mathrm{Path}(s,t)} \mathbb{P}^\sigma(\pi)$.*
The optimal reachability probability *is defined as $p_s := \sup_{\sigma\in\Sigma} \mathbb{P}^\sigma(s \models \Diamond t) \in \mathbb{Q}$.*

We write $\boldsymbol{p}^*$ for the vector $(p_s)_{s\in S\backslash\{t\}}$ indexed by states $S\backslash\{t\}$. The optimal reachability probabilities are in fact achievable by a positional scheduler.

▶ **Proposition 5** (e.g. [1]). *There exists an optimal positional scheduler $\sigma_{pos} \in \Sigma$ such that $\mathbb{P}^{\sigma_{pos}}(s \models \Diamond t) = p_s$ holds for all $s \in S$.*

Value Iteration (VI) [1, 24] is a standard technique to approximate the vector of optimal reachability probabilities $\boldsymbol{p}^*$. Specifically, VI applies the *Bellman operator* Φ to the current approximation for each iteration step.

Let $S_d = \{s_1, \cdots, s_d\}$ be the set of all non-target states whose optimal reachability probability is positive. We note in passing that the set S_d can be computed using graph reachability techniques [1, Chapter 10.6.1].

For each $s \in S_d$, we associate a polynomial of degree 1, *a linear polynomial* of α, with each action $\alpha \in Act_s$. This polynomial $L_\alpha \in \mathbb{Q}[\boldsymbol{x}]$ is defined as

$$L_\alpha(\boldsymbol{x}) = \sum_{s'\in S_d} \mathbb{P}(s, \alpha, s')x_{s'} + \mathbb{P}(s, \alpha, t).$$

▶ **Definition 6** (Bellman operator). *The* Bellman operator $\Phi\colon [0,1]^d \to [0,1]^d$ *is defined by* $\Phi(\boldsymbol{x})_s := \max_{\alpha\in Act_s} L_\alpha(\boldsymbol{x})$ *for each* $\boldsymbol{x} = (x_s)_{s\in S_d} \in [0,1]^d$ *and* $s \in S_d$*.*

▶ **Example 7.** The Bellman operator Φ of the MDP given in Example 3 is given by

$$\Phi(\boldsymbol{x})_{s_1} := 1/2 \cdot x_2 + 1/3, \quad \Phi(\boldsymbol{x})_{s_2} := \max\big(1/2 \cdot x_1 + 1/2,\ 1/4 \cdot x_1 + 1/4 \cdot x_2 + 1/2\big),$$

where $S_d = \{s_1, s_2\}$. The Bellman operator Φ is indeed the PAM f given in Example 1.

By an abuse of notation, we write Act_i for the set Act_{s_i} of actions on $s_i \in S_d$, and the optimal reachability probability p_i for p_{s_i}. We also restrict $\boldsymbol{p}^*$ to the vector over S_d and write $\boldsymbol{p}^* = (p_1, \ldots, p_d)$ when it is clear from the context. Each action $\alpha \in Act_i$ is associated with a set $\mathrm{succ}(\alpha)$ of successor states defined as $\mathrm{succ}(\alpha) := \mathrm{supp}(s_i, \alpha) \cap S_d$. We emphasise that the successor set, together with transition probabilities $\mathbb{P}(s, \alpha, s')$, $s' \in S_d$, is a probabilistic *sub*distribution.

We define the partial order $\leq$ on vectors in $\mathbb{R}^d$ by $\boldsymbol{u} \leq \boldsymbol{v}$ if $u_i \leq v_i$ holds for each $i \in [1, d]$. We refer to vectors $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^d$ as *comparable* if either $\boldsymbol{u} \geq \boldsymbol{v}$ or $\boldsymbol{u} \leq \boldsymbol{v}$ holds. Otherwise, the vectors are incomparable, denoted $\boldsymbol{u} \bowtie \boldsymbol{v}$.

Let $||\cdot||_\infty$ denote the ℓ^∞*-norm*, or the max-*norm*, defined by $||\boldsymbol{x}||_\infty := \max(|x_1|, \ldots, |x_d|)$ for a vector $\boldsymbol{x} = (x_1, \ldots, x_d)$. In the sequel, the notation $||\boldsymbol{x}||$ stands for $||\boldsymbol{x}||_\infty$. We further define the ℓ^∞-metric, i.e., the distance between two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ is $d(\boldsymbol{x}, \boldsymbol{y}) := \max_i(|x_i - y_i|)$.

The set $[0,1]^d$ is a complete lattice with the componentwise ordering and the Bellman operator is ω-continuous, that is, it preserves joins of ascending sequences. Due to this, the iterative update by the Bellman operator Φ from the bottom vector $\mathbf{0} := (0, \cdots, 0) \in [0,1]^d$, which is called VI (from $\mathbf{0}$), converges to the least fixed point $\mu\Phi$, which is the optimal reachability probabilities $\boldsymbol{p}^* = (p_1, \cdots, p_d)$.

▶ **Proposition 8** ([1, 10, 24]). *The sequence* $\langle \Phi^n(\mathbf{0})\rangle_{n\in\mathbb{N}}$ *is monotonically increasing and converges to the optimal reachability probabilities* $\boldsymbol{p}^*$*.*

By the Kleene fixed-point theorem, we can further see that the iteration by Φ converges to $\boldsymbol{p}^*$ from any initial vector if Φ has a unique fixed point.

▶ **Proposition 9** ([15]). *Assume* Φ *has a unique fixed point. The sequence* $\langle \Phi^n(\boldsymbol{x})\rangle_{n\in\mathbb{N}}$ *converges to the unique fixed point* $\boldsymbol{p}^*$ *from any initial* $\boldsymbol{x} \in [0,1]^d$*.*

2.2 Target problem

We begin by recalling the value iteration algorithm. As an iterative procedure, VI boils down to repeatedly applying the Bellman operator Φ starting from a certain vector $\boldsymbol{s} \in [0,1]^d$ (usually, $\boldsymbol{s} = \mathbf{0}$) and converging to the least fixed point $\mu\Phi (= \boldsymbol{p}^*)$ – this becomes the unique fixed point in our setting.

Observe that every Bellman operator Φ is indeed a piecewise affine map on the domain $[0,1]^d$. On close inspection, for every $\boldsymbol{x} \in [0,1]^d$, the value of $\Phi(\boldsymbol{x})$ is computed as the maximum of finitely many affine functions $\phi_1, \ldots, \phi_k : [0,1]^d \to [0,1]^d$ evaluated at $\boldsymbol{x}$. Let $P_i \subseteq [0,1]^d$ be the set of points $\boldsymbol{x} \in [0,1]^d$, for which $\phi_i(\boldsymbol{x}) \geq \phi_j(\boldsymbol{x})$ for each $j \neq i$. Two observations are straightforward: 1) each P_i is defined by a conjunction of linear inequalities; 2) every $\boldsymbol{x}$ belongs to at least one set P_i. Moreover, Φ is a well-defined function, which can be observed from $\boldsymbol{x} \in P_i \cap P_j$ implying $\phi_i(\boldsymbol{x}) = \phi_j(\boldsymbol{x})$. Therefore, Φ is a PAM on $[0,1]^d$.

In this paper, we investigate the specialisation of the piecewise affine reachability problem to Bellman operators. A standard assumption that ensures the uniqueness of fixed points for Bellman operators is the absence of *end components* in MDPs [15].

▶ **Definition 10** (end component [12]). *Let $\mathcal{M} = (S, Act, \mathbb{P})$ be an MDP. A pair (S', Act') such that $\emptyset \neq S' \subseteq S$ and $\emptyset \neq Act' \subseteq \cup_{s \in S'} Act_s$ is an* end component *if (i) for all $s \in S'$ and $\alpha \in Act' \cap Act_s$, $\mathrm{supp}(s, \alpha) \subseteq S'$; and (ii) the directed graph that is induced by (S', Act') is strongly connected.*

Note that according to our definitions, the target state of $\mathcal{M}$ is not an end component, as there is no action defined at it. We also eliminate all states from which the target state cannot be reached by considering the subset S_d.

Problem BOR (Bellman Operator Reachability). Let $\boldsymbol{s}, \boldsymbol{t} \in [0,1]^d \cap \mathbb{Q}^d$ and $\Phi : [0,1]^d \to [0,1]^d$ be a Bellman operator of an MDP $\mathcal{M}$ with no end components. Does there exist $n \in \mathbb{N}$ such that $\Phi^n(\boldsymbol{s}) = \boldsymbol{t}$?

In the sequel, we always assume that an MDP $\mathcal{M}$ has no end components and hence its Bellman operator Φ has a unique fixed point. Such MDPs remain expressive and exhibit highly non-trivial behaviors, making them an important subject of study in probabilistic verification; see e.g. [2, 9, 15, 16, 25]. Notably, the presence of end components in a given MDP can be checked in P [12]. Moreover, [15] describes a reduction that, for an arbitrary MDP, constructs an MDP with the same least fixed point and without end components; the reduction never increases the dimension.

3 Bellman Operator Reachability in Arbitrary Dimension

In this section, we will discuss the BOR problem without restricting the dimension d. Recall that for the BOR problem, a Bellman operator Φ has a unique fixed point $\mu\Phi$. It will be instrumental to split the discussion of decidability depending on how the initial and target vectors $\boldsymbol{s}$ and $\boldsymbol{t}$ compare to $\mu\Phi$ with respect to the componentwise order on $\mathbb{Q}^d$.

3.1 Target vector that is not the fixed point

First, we show that BOR is decidable when $\boldsymbol{t}$ is not the unique fixed point $\mu\Phi$.

▶ **Lemma 11.** *Let $\Phi : [0,1]^d \to [0,1]^d$ be a Bellman operator. Consider an arbitrary vector $\boldsymbol{x} \in [0,1]^d$ and let $\delta := ||\boldsymbol{x} - \mu\Phi||$. We have $||\Phi(\boldsymbol{x}) - \mu\Phi|| \leq \delta$.*

Proof. It holds $\boldsymbol{x} \leq \mu\Phi + \boldsymbol{\delta}$, where $\boldsymbol{\delta} := (\delta, \dots, \delta)$. For any action α in state i, we have

$$\begin{aligned} L_\alpha(\mu\Phi + \boldsymbol{\delta}) &= \sum_{j \in S_d} \mathbb{P}(i, \alpha, j) \cdot (\mu\Phi_j + \delta) + \mathbb{P}(i, \alpha, t) \\ &= \sum_{j \in S_d} \mathbb{P}(i, \alpha, j) \cdot \mu\Phi_j + \mathbb{P}(i, \alpha, t) + \sum_{j \in S_d} \mathbb{P}(i, \alpha, j) \cdot \delta \\ &= L_\alpha(\mu\Phi) + \sum_{j \in S_d} \mathbb{P}(i, \alpha, j) \cdot \delta \leq L_\alpha(\mu\Phi) + \delta. \end{aligned}$$

Let $\tilde{\alpha} := \arg\max_{\alpha \in Act_i} L_\alpha(\mu\Phi + \boldsymbol{\delta})$. Then, $\Phi(\mu\Phi + \boldsymbol{\delta})_i = L_{\tilde{\alpha}}(\mu\Phi + \boldsymbol{\delta}) \leq \mu\Phi_i + \delta$.

From monotonicity of Φ we conclude $\Phi(\boldsymbol{x})_i \leq \Phi(\mu\Phi + \boldsymbol{\delta})_i \leq \mu\Phi_i + \delta$. ◀

▶ **Proposition 12.** *Let $\Phi : [0,1]^d \to [0,1]^d$ be a Bellman operator of an MDP with no end components, and $\boldsymbol{s} \in [0,1]^d$ be an arbitrary initial vector. For every $\boldsymbol{t} \in [0,1]^d$ with $\boldsymbol{t} \neq \mu\Phi$, there exists an effectively computable bound N such that*

$$\Phi^n(\boldsymbol{s}) = \boldsymbol{t} \quad \Rightarrow \quad \Phi^n(\boldsymbol{s}) = \boldsymbol{t} \text{ for some } n \leq N.$$

Proof. Fix vectors $\boldsymbol{s}, \boldsymbol{t} \in [0,1]^d$ assuming $\boldsymbol{t} \neq \mu\Phi$, where $\mu\Phi$ is the unique fixed point of Φ. We crucially use the convergence properties of the interval iteration algorithm [15]. Let $\mathbf{1} := (1,\dots,1)$ be the greatest element of the lattice $[0,1]^d$. Choose a convergence threshold $\varepsilon > 0$. From [15, Theorem 2] we have $||\Phi^N(\mathbf{0}) - \Phi^N(\mathbf{1})|| < \varepsilon$ for some $N \leq A\lceil \frac{\log \varepsilon}{\log(1-B^A)} \rceil$, where constants A, B only depend on $\mathcal{M}$ and can be computed directly from its representation.

From the monotonicity of Φ we have $\Phi^N(\mathbf{0}) \leq \Phi^N(\boldsymbol{s}) \leq \Phi^N(\mathbf{1})$ for any $\boldsymbol{s} \in [0,1]^d$. We also have $\Phi^N(\mathbf{0}) \leq \mu\Phi \leq \Phi^N(\mathbf{1})$, and so $||\Phi^N(\boldsymbol{s}) - \mu\Phi|| < \varepsilon$.

Recall that we can compute the vector $\mu\Phi$ exactly. Now choose the threshold $\varepsilon := ||\boldsymbol{t} - \mu\Phi||$, and let N_ε be the previously discussed bound for this threshold. We have $||\Phi^{N_\varepsilon}(\boldsymbol{s}) - \mu\Phi|| < ||\boldsymbol{t} - \mu\Phi||$ and hence, from Lemma 11, $||\Phi^n(\boldsymbol{s}) - \mu\Phi|| < ||\boldsymbol{t} - \mu\Phi||$ for every $n \geq N_\varepsilon$. Therefore, either $\boldsymbol{t} = \Phi^n(\boldsymbol{s})$ for some $n < N_\varepsilon$, or $\boldsymbol{t}$ is not reachable from $\boldsymbol{s}$ under iteratively applying Φ. The first condition can be checked in finite time since N_ε is effectively bounded. ◀

From Proposition 12 we immediately derive an algorithmic procedure for the BOR problem instances $(\Phi, \boldsymbol{s}, \boldsymbol{t})$, where $\boldsymbol{t} \neq \mu\Phi$. For an instance like this, it suffices to compute the bound N as above, and to test whether $\Phi^n(\boldsymbol{s})$ is equal to $\boldsymbol{t}$ for some $n \leq N$.

We now move on to the case $\boldsymbol{t} = \mu\Phi$. In the sequel, it will be important to differentiate between two types of actions. These types are defined based on preserving the probabilities of the unique fixed point $\boldsymbol{t} = (t_1, \dots, t_d)$.

▶ **Definition 13.** *An action α available in state i is* tight*, if $t_i = L_\alpha(t_1, \dots, t_d)$, where L_α is the linear polynomial of action α. An action α is* leaking *in state i, if it is not tight.*

3.2 Initial vector below the fixed point

We now assume that $\boldsymbol{s} \leq \boldsymbol{t}$ and $\boldsymbol{t} = \mu\Phi$. Note that $\Phi^n(\boldsymbol{s}) \leq \boldsymbol{t}$ holds for all $n \geq 0$.

▶ **Lemma 14.** *Let $\boldsymbol{x} \leq \boldsymbol{t}$ be a $[0,1]^d$-vector and let $\alpha \in Act_i$ be the action chosen in state i when the Bellman operator is applied in $\boldsymbol{x}$, i.e. $\Phi(\boldsymbol{x})_i = L_\alpha(\boldsymbol{x})$.*

$\Phi(\boldsymbol{x})_i = t_i$ holds if and only if α is tight and for each $j \in \operatorname{succ}(\alpha)$ we have $x_j = t_j$.

Proof. One implication follows from directly applying the definitions. Now consider the $(\Rightarrow)$ implication. For every action β, we have $L_\beta(\boldsymbol{x}) \leq L_\beta(\boldsymbol{t})$. Hence, for a leaking β, this implies $L_\beta(\boldsymbol{x}) < t_i$. Since for α we have $L_\alpha(\boldsymbol{x}) = t_i$, it must be tight. Assume further that there exists $j \in \operatorname{succ}(\alpha)$ such that $x_j < t_j$. Then, $L_\alpha(\boldsymbol{x}) \leq L_\alpha(t_1, \dots, t_{j-1}, x_j, t_{j+1}, \dots, t_d) < L_\alpha(\boldsymbol{t}) = t_i$. This again contradicts $L_\alpha(\boldsymbol{x}) = t_i$, hence $x_j = t_j$ holds for all $j \in \operatorname{succ}(\alpha)$. ◀

From probabilities to $\{-1, 0\}$. The reasoning of Lemma 14 can be extended. Intuitively, we can abstract away from the actual probabilities in vectors $\Phi^n(\boldsymbol{s})$, $n \geq 0$. This succeeds by *only keeping track of whether these probabilities are different from probabilities in $\boldsymbol{t}$*. This abstraction makes the space of the BOR problem finite, provided $\boldsymbol{s} \leq \mu\Phi$.

Formally, we introduce a sign abstraction $f : [0,1]^d \to \{-1,0\}^d$ by associating a sign vector $\boldsymbol{\varepsilon} = f(\boldsymbol{x}) = (\varepsilon_1, \dots, \varepsilon_d)$ with every vector $\boldsymbol{x}$ such that $\boldsymbol{x} \leq \boldsymbol{t}$: $\varepsilon_i = \begin{cases} 0, & x_i = t_i, \\ -1, & \text{otherwise.} \end{cases}$
According to this definition, $f(\boldsymbol{t}) = \mathbf{0}$. Lemma 14 can now be read as follows: $\Phi(\boldsymbol{x}) = \boldsymbol{t}$ holds if and only if there exists a choice of tight actions $(\alpha_1, \dots, \alpha_d)$ in $\boldsymbol{x}$ (whose sign vector is $\boldsymbol{\varepsilon}$) such that for each state $s \in \operatorname{succ} \alpha_i$, we have $\varepsilon_s = 0$.

We further prove that the successor of $f(\boldsymbol{x})$ with respect to the Bellman operator is well-defined. This would allow to only consider the evolution of sign vectors later on.

▶ **Lemma 15.** *Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be two vectors satisfying $\boldsymbol{x} \le \boldsymbol{t}$ and $\boldsymbol{y} \le \boldsymbol{t}$. Provided $f(\boldsymbol{x}) = f(\boldsymbol{y})$, we have $f(\Phi(\boldsymbol{x})) = f(\Phi(\boldsymbol{y}))$.*

Proof. Let $\boldsymbol{\varepsilon}' = (\varepsilon_1', \dots, \varepsilon_d')$ be the abstraction of $\Phi(\boldsymbol{x})$, that is, $\boldsymbol{\varepsilon}' = f(\Phi(\boldsymbol{x}))$.

First, notice that a leaking action chosen in state i at $\boldsymbol{z} = (z_1, \dots, z_d)$ always implies $L_\alpha(\boldsymbol{z}) < t_i$. Second, recall that having $j \in \mathrm{succ}(\alpha)$ with $z_j < t_j$ implies $L_\alpha(\boldsymbol{z}) < t_i$. Hence, $\Phi(\boldsymbol{z})_i = t_i$ if and only if there exists an action $\alpha \in Act_i$ such that $L_\alpha(\boldsymbol{z}) = t_i$. This action is necessarily tight. Using the vocabulary of the sign abstraction, we state $f(\Phi(\boldsymbol{z}))_i = 0$ holds if and only if there exists a tight action $\alpha \in Act_i$ such that $L_\alpha(\boldsymbol{z}) = t_i$. We summarise these observations as

$$\varepsilon_i' = \max_{\substack{\alpha \in Act_i \\ \alpha \text{ tight}}} \min_{j \in \mathrm{succ}(\alpha)} \varepsilon_j, \tag{1}$$

and thus make sure that ε_i' does not depend on the actual values in $\boldsymbol{x}$ and $\boldsymbol{y}$, as soon as those two vectors have the same sign abstraction. ◀

▶ **Proposition 16.** *There exists an algorithmic procedure for the BOR problem instances $(\Phi, \boldsymbol{s}, \boldsymbol{t})$, where $\boldsymbol{s} \le \mu\Phi$ and $\boldsymbol{t} = \mu\Phi$.*

Proof. It is easy to observe that the space $\{-1, 0\}^d$ of possible sign vectors is finite. Given $\boldsymbol{s}$, we compute the abstraction $f(\boldsymbol{s})$ and ask whether $\mathbf{0}$ is reached by iteratively applying the map $\boldsymbol{\varepsilon} \mapsto \boldsymbol{\varepsilon}'$ as defined by Equation (1). Once an already explored vector occurs in the sequence $\langle \boldsymbol{\varepsilon} = f(\boldsymbol{s}), \boldsymbol{\varepsilon}', \boldsymbol{\varepsilon}'', \dots \rangle$, we can stop. This happens in at most $2^d - 1$ iterations. In this finite sequence, $\mathbf{0}$ occurs if and only if $\boldsymbol{t}$ is reached by iterating Φ starting from $\boldsymbol{s}$. This is due to Lemma 15. ◀

3.3 Initial vector above the fixed point

Next assumption we are going to work with is $\boldsymbol{s} \ge \boldsymbol{t} = \mu\Phi$. The main result of this subsection is the following proposition.

▶ **Proposition 17.** *There exists an algorithmic procedure for the BOR problem instances $(\Phi, \boldsymbol{s}, \boldsymbol{t})$, where $\boldsymbol{s} \ge \mu\Phi$ and $\boldsymbol{t} = \mu\Phi$.*

The procedure for this case is more intricate than in Section 3.2. This is due to the new phenomenon that occurs for sequences initialised with $\boldsymbol{s} \ge \boldsymbol{t}$. An iteration of the Bellman operator can choose a leaking action $\beta \in Act_i$ over all tight actions available in state s_i. Intuitively, this happens if the successor states $\mathrm{succ}(\beta)$ have probabilities significantly greater than optimal – enough to compensate for the "leakage" $t_i - L_\beta(\boldsymbol{t})$. In other words, $\Phi(\boldsymbol{x})_i > L_\alpha(\boldsymbol{x})$ might hold for all tight $\alpha \in Act_i$, unlike in the case $\boldsymbol{x} \le \boldsymbol{t}$ (cf. Lemma 14).

▶ **Example 18.** Consider the MDP $\mathcal{M}_1$ in Figure 1. It has $S = \{s_1, s_2, s_3, s_4, t\}$ where the "missing" probabilistic transitions lead to s_4. Moreover, Act_4 comprises a single action with $\mathbb{P}(s_4, \cdot, s_4) = 1$. We omit s_4 and transitions to/from it, for simplicity of presentation.

There is one tight action α and one leaking action β in s_1. Let $\boldsymbol{s} := (1, 1/3, 2/3)$. Clearly, $\boldsymbol{s} > \mu\Phi = (7/12, 1/4, 1/4)$. The tight action α is chosen for the first iteration, and $\Phi(\boldsymbol{s}) = (13/18, 2/3, 1/4)$. Next, the leaking action β is chosen and $\Phi^2(\boldsymbol{s}) = (9/12, 1/4, 1/4)$, and finally $\Phi^3(\boldsymbol{s}) = \boldsymbol{t}$ by choosing the tight action α.

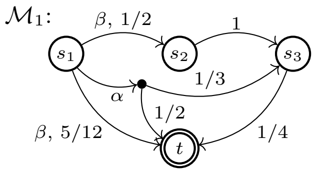


Figure 1 An MDP $\mathcal{M}_1$.

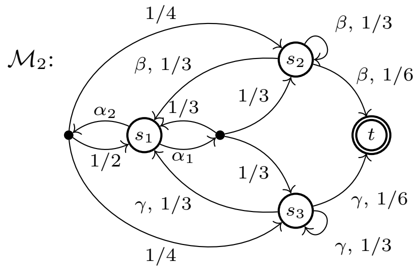


Figure 2 An MDP $\mathcal{M}_2$.

Only tight actions eventually. However, we show that in a convergent sequence, the actions chosen by the Bellman operator Φ are all tight, after some number of iterations.

▶ **Lemma 19.** *Let $\Phi : [0,1]^d \to [0,1]^d$ be a Bellman operator such that $\mu\Phi = \boldsymbol{t}$. There exists a δ-neighbourhood of the fixed point*

$$U_\delta(\boldsymbol{t}) = \{\boldsymbol{x} \in [0,1]^d : d(\boldsymbol{x}, \boldsymbol{t}) < \delta\}$$

such that for every $\boldsymbol{x} \in U_\delta(\boldsymbol{t})$, the vector $\Phi(\boldsymbol{x})$ is obtained by applying only tight actions. That is,

$$\Phi(\boldsymbol{x}) = (L_{\alpha_1}(\boldsymbol{x}), \ldots, L_{\alpha_d}(\boldsymbol{x}))\,,$$

where each $\alpha_i \in Act_i$, $1 \le i \le d$, is tight.

Proof. Since the transition probabilities in $\mathcal{M}$ are rational numbers, one argues that $t_1, \ldots, t_d$ are rational, too. We consider the set of rational numbers that consists of $t_1, \ldots, t_d$, along with $L_\alpha(\boldsymbol{t})$ for each action α. It is sufficient to consider leaking actions, by definition. Let D be the least common denominator of the numbers in the aforedescribed set.

Let $\delta := \frac{1}{2D}$ and pick $\boldsymbol{x} \in U_\delta(\boldsymbol{t})$. Further let $A = (\alpha_1, \ldots, \alpha_d)$ be the actions chosen at $\boldsymbol{x}$ by the Bellman operator Φ. We show that each action in A is tight. Denote by f_A the effect of applying A, i.e.,

$$f_A(\boldsymbol{t}) = (L_{\alpha_1}(\boldsymbol{t}), \ldots, L_{\alpha_d}(\boldsymbol{t})) \quad \text{and} \quad f_A(\boldsymbol{x}) = \Phi(\boldsymbol{x}).$$

We notice that f_A is 1-Lipschitz (which, in fact, holds for any choice A of actions). Indeed, let us consider arbitrary $\boldsymbol{u}, \boldsymbol{v} \in [0,1]^d$. We have

$$\begin{aligned} ||f_A(\boldsymbol{u}) - f_A(\boldsymbol{v})|| &= \max_{1\le i\le d} |L_{\alpha_i}(\boldsymbol{u}) - L_{\alpha_i}(\boldsymbol{v})| \\ &\le \max_{1\le i\le d} \sum_{j\in\mathrm{succ}(\alpha_i)} \mathbb{P}(i, \alpha_i, j) \cdot |u_j - v_j| \\ &\le \max_{1\le i\le d} \left(1 \cdot \max_{j\in\mathrm{succ}\,\alpha_i} |u_j - v_j|\right) = \max_{1\le i\le d} |u_i - v_i| = ||\boldsymbol{u} - \boldsymbol{v}||. \end{aligned}$$

Utilising the 1-Lipschitz property, we observe $||f_A(\boldsymbol{t}) - \Phi(\boldsymbol{x})|| \le ||\boldsymbol{t} - \boldsymbol{x}||$. Moreover, we have $||\Phi(\boldsymbol{x}) - \boldsymbol{t}|| \le ||\boldsymbol{x} - \boldsymbol{t}||$ due to Lemma 11. Therefore,

$$||f_A(\boldsymbol{t}) - \boldsymbol{t}|| \le ||f_A(\boldsymbol{t}) - \Phi(\boldsymbol{x})|| + ||\Phi(\boldsymbol{x}) - \boldsymbol{t}|| \le ||\boldsymbol{t} - \boldsymbol{x}|| + ||\boldsymbol{x} - \boldsymbol{t}|| < \delta + \delta = \frac{1}{D}. \tag{2}$$

On the other hand, $||f_A(\boldsymbol{t}) - \boldsymbol{t}|| = \max_{1\le i\le d} |L_{\alpha_i}(\boldsymbol{t}) - t_i|$. By the definition of D, we know that $L_{\alpha_i}(\boldsymbol{t}) \ne t_i$ implies $|L_{\alpha_i}(\boldsymbol{t}) - t_i| \ge \frac{1}{D}$. We conclude from Equation (2) that $f_A(\boldsymbol{t}) = \boldsymbol{t}$. Equivalently, all actions chosen at $\boldsymbol{x}$ are tight. ◀

The vector sequence $\boldsymbol{s}, \Phi(\boldsymbol{s}), \Phi^2(\boldsymbol{s}), \ldots$ converges to $\boldsymbol{t}$ due to Proposition 9. Therefore, it reaches a $\frac{1}{2D}$-neighbourhood of $\boldsymbol{t}$ after finitely many steps. Furthermore, an upper bound on the number of necessary steps can be computed as in Proposition 12.

▶ **Corollary 20.** *For an arbitrary initial vector $\boldsymbol{s} \in [0,1]^d$, there exists an effectively computable $N \in \mathbb{N}$ such that in the sequence $\left(\Phi^n(\boldsymbol{s})\right)_{n\in\mathbb{N}}$ for every $n \geq N$, $\Phi^{n+1}(\boldsymbol{s})$ is obtained by applying only tight actions to $\Phi^n(\boldsymbol{s})$.*

We point out that the argument used in the proof of Lemma 19, bears resemblance to (and is inspired by) the proof of [15, Theorem 3]. However, the eventual optimality of tight actions, which we establish here, is not a matter of discussion in [15] or any other work we know.

From probabilities to $\{0, 1\}$. We now extend the sign abstraction from the previous section to encompass vectors above $\boldsymbol{t}$. Let $f : [0,1]^d \to \{0,1\}^d$ be defined by associating a sign vector $\boldsymbol{\varepsilon} = f(\boldsymbol{x}) = (\varepsilon_1, \ldots, \varepsilon_d)$ with each vector $\boldsymbol{x} \geq \boldsymbol{t}$: $\varepsilon_i = \begin{cases} 0, & x_i = t_i, \\ 1, & \text{otherwise.} \end{cases}$

Conforming to this definition is $f(\boldsymbol{t}) = \mathbf{0}$. As before, we show that f is well-defined.

▶ **Lemma 21.** *Let δ be chosen as above to guarantee that Φ only picks tight actions in the neighbourhood $U_\delta(\boldsymbol{t})$. Consider vectors $\boldsymbol{x}, \boldsymbol{y} \in [0,1]^d$ satisfying $\boldsymbol{t} \leq \boldsymbol{x}, \boldsymbol{y} \leq \boldsymbol{t} + (\delta, \ldots, \delta)$.*

Provided $f(\boldsymbol{x}) = f(\boldsymbol{y})$, we have $f(\Phi(\boldsymbol{x})) = f(\Phi(\boldsymbol{y}))$.

Proof. The assumption $\boldsymbol{x}, \boldsymbol{y} \in U_\delta(\boldsymbol{t})$ is necessary to shake off the effect of the leaking actions (cf. the assumptions of Lemma 15). Let $\boldsymbol{z} = (z_1, \ldots, z_d)$ be an arbitrary vector in $U_\delta(\boldsymbol{t})$. This way we guarantee that for every i, there is a tight action $\alpha \in Act_i$ with $\Phi(\boldsymbol{z})_i = L_\alpha(\boldsymbol{z})$. It is not hard to see now that having $j \in \mathrm{succ}(\alpha)$ with $z_j > t_j$ implies $L_\alpha(\boldsymbol{z}) > t_i$. Therefore, $\Phi(\boldsymbol{z})_i = t_i$ if and only if there exists no tight action $\alpha \in Act_i$ that depends on $j \in \mathrm{succ}(\alpha)$ with $z_j > t_j$. Equivalently, we have

$$\varepsilon'_i = \max_{\substack{\alpha \in Act_i \\ \alpha \text{ tight}}} \max_{j \in \mathrm{succ}(\alpha)} \varepsilon_j, \tag{3}$$

where $\boldsymbol{\varepsilon}' = (\varepsilon'_1, \ldots, \varepsilon'_d)$ is the abstraction of $\Phi(\boldsymbol{x})$, that is, $\boldsymbol{\varepsilon}' = f(\Phi(\boldsymbol{x}))$.

Therefore, the abstraction vectors of $\Phi(\boldsymbol{x})$ and $\Phi(\boldsymbol{y})$ do not depend on the actual values in $\boldsymbol{x}$ and $\boldsymbol{y}$, but only on their abstractions $f(\boldsymbol{x}) = f(\boldsymbol{y})$. ◀

The abstraction f is thus well-defined for vectors in a certain neighbourhood of $\boldsymbol{t}$.

Proof of Proposition 17. Starting at $\boldsymbol{s}$ above the fixed point, $\boldsymbol{s} \geq \boldsymbol{t}$, we first compute the bound N of Corollary 20. If $\Phi^n(\boldsymbol{s}) = \boldsymbol{t}$ holds for some $n < N$, we terminate with a positive answer to the BOR problem. Otherwise, we continue with the abstraction argument. We compute $f(\Phi^N(\boldsymbol{s}))$ and ask whether $\mathbf{0}$ is reached by iteratively applying the map $\boldsymbol{\varepsilon} \mapsto \boldsymbol{\varepsilon}'$ as defined by Equation (3). In at most $2^d - 1$ iterations we either reach $\mathbf{0}$, or discover an ever-repeating vector subsequence that does not contain $\mathbf{0}$. The rest follows from Lemma 21. ◀

Combining Propositions 16 and 17, we obtain the following theorem.

▶ **Theorem 22.** *There exists an algorithmic procedure that solves all BOR problem instances $(\Phi, \boldsymbol{s}, \boldsymbol{t})$ with $\boldsymbol{t} = \mu\Phi$ and $\boldsymbol{s}\ \{\leq, \geq\}\ \mu\Phi$.*

3.4 Initial and target vectors are incomparable

We keep assuming $\boldsymbol{t} = \mu\Phi$ and consider the remaining case, that is, the case when $\boldsymbol{s}$ and $\boldsymbol{t}$ are two incomparable vectors, denoted $\boldsymbol{s} \bowtie \boldsymbol{t}$.

We can assume $\boldsymbol{s} \in U_\delta(\boldsymbol{t})$ as defined in Lemma 19. Clearly, starting with an arbitrary incomparable vector, we can apply Φ up to the pre-computed power N, reaching either a comparable vector (potentially including $\boldsymbol{t}$ itself), or the δ-neighbourhood of $\boldsymbol{t}$. In the latter case, set $\boldsymbol{s}$ to be the first vector inside the neighbourhood. Hence, for every $\boldsymbol{x}$ discussed below, the vector $\Phi(\boldsymbol{x})$ is obtained by applying only tight actions.

Positionality. Even after eventually adhering to tight actions, the behaviour of Φ's iterations is not described by a single linear transformation. We recall from Proposition 5 that positionality is prominently sufficient for achieving optimal probabilities *in the limit* [1]. However, Example 23 shows a more subtle behaviour for our reachability problem.

▶ **Example 23.** In Figure 2, we present an MDP $\mathcal{M}_2$ with $S = \{s_1, s_2, s_3, s_4, t\}$ where the "missing" probabilistic transitions lead to $s_4 \notin S_d$. We omit s_4 and transitions to/from it, for simplicity. Let $Act_1 = \{\alpha_1, \alpha_2\}$, $Act_2 = \{\beta\}$, $Act_3 = \{\gamma\}$ be the actions available in the states of S_3. Observe that $\mu\Phi = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ and all actions are tight, including both α_1, α_2.

Let $\boldsymbol{s} = (0, \frac{5}{6}, \frac{5}{6})$. First, α_1 is chosen over α_2 in $\boldsymbol{s}$, and so $\Phi(\boldsymbol{s}) = (\frac{5}{9}, \frac{4}{9}, \frac{4}{9})$. However, in the next iteration, $\alpha_2(\frac{5}{9}, \frac{4}{9}, \frac{4}{9}) = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2}) \geq (\frac{13}{27}, \frac{1}{2}, \frac{1}{2}) = \alpha_1(\frac{5}{9}, \frac{4}{9}, \frac{4}{9})$ yielding $\Phi^2(\boldsymbol{s}) = \mu\Phi$.

Spectacularly, none of two positional schedulers reaches $\mu\Phi$, which can be proved using [17].

With positionality out of question, we need to study schedulers that switch actions over time.

Matrix semigroups. In the sequel, we take a matrix perspective on BOR by introducing a $d \times d$–matrix for every tuple of tight actions. The set of such tuples is finite, and thus we argue that the behaviour of Bellman operator iterations from $\boldsymbol{s}$ is governed by multiplying the vector $\boldsymbol{s}$ with elements of a semigroup $\mathcal{S}$ generated by finitely many matrices $M_1, \ldots, M_k$.

We associate with every tight action $\alpha \in Act_i$ a row vector $(\mathbb{P}(s_i, \alpha, s_1), \ldots, \mathbb{P}(s_i, \alpha, s_d))$ of probabilities for going to the states of S_d, as well as a scalar $\mathbb{P}(s_i, \alpha, t)$ for reaching t.

For each state s_i, let $\mathcal{F}_i$ denote the set of row vectors for tight actions in Act_i. Notice that all row vectors only have non-negative entries; furthermore, the sum of elements in each row vector is at most 1. We will further refer to matrices with all rows satisfying these properties as *substochastic*.

▶ **Definition 24.** *A family $\mathcal{F} \subset \mathbb{Q}^{d\times d}$ of matrices is called a* product family *if $\mathcal{F}$ consists of all possible matrices with i-th row from $\mathcal{F}_i$ for all $i \in \{1, \ldots, d\}$.*

We further let $\mathcal{S} := \langle \mathcal{F} \rangle$ be the semigroup generated by $\mathcal{F}$.

We introduce the map f by $\boldsymbol{x} \mapsto \boldsymbol{x} - \mu\Phi$.

The Product Family Reachability (PFR) Problem. Let $\mathcal{F} = \{M_1, \ldots, M_k\}$ be a product family of substochastic matrices, and $\boldsymbol{\varepsilon} \in [-1, 1]^d \cap \mathbb{Q}^d$. We define $F : \mathbb{Q}^d \to \mathbb{Q}^d$ by $F(\boldsymbol{v}) \coloneqq \max_{1\leq i\leq k} (M_i \cdot \boldsymbol{v})$. Then the *PFR problem* asks: Does there exist $n \geq 0$ such that $F^n(\boldsymbol{\varepsilon}) = \boldsymbol{0}$?

Note in passing that max over vectors is taken with respect to the partial order $\leq$. The operator F is thus well-defined, that is, there exists a matrix $M \in \mathcal{F}$ such that $F(\boldsymbol{v}) = M \cdot \boldsymbol{v}$. It holds indeed that $F(\boldsymbol{v})_i \geq \boldsymbol{w} \cdot \boldsymbol{v}$ for any $\boldsymbol{w} \in \mathcal{F}_i$, $1 \leq i \leq k$.

▶ **Proposition 25.** *Every d–dimensional instance of BOR with $\boldsymbol{t} = \mu\Phi$ and $\boldsymbol{s} \in U_\delta(\boldsymbol{t})$ (thus guaranteeing that only tight actions are used) is equivalent to a d–dimensional instance of the PFR problem.*

Proof. Let $\mathcal{M}$ be an MDP whose Bellman operator Φ has a unique fixed point $\mu\Phi \in [0,1]^d$. Fix $\boldsymbol{t} = \mu\Phi$ and $\boldsymbol{s} \in U_\delta(\boldsymbol{t})$, where the neighbourhood $U_\delta(\boldsymbol{t})$ is as in Lemma 19. We will introduce a product family of substochastic matrices in $\mathbb{Q}^{d\times d}$ with a corresponding operator F and we will prove that there exists $n \in \mathbb{N}$ such that $\Phi^n(\boldsymbol{s}) = \boldsymbol{t}$ if and only if $F^n(f(\boldsymbol{s})) = \boldsymbol{0}$.

Let $\mathcal{F}_i$ be the set of all row vectors for actions in Act_i of $\mathcal{M}$, for each $i \in \{1,\dots,d\}$. Then, a finite product family $\mathcal{F} = \{M_1,\dots,M_k\} \subset \mathbb{Q}^{d\times d}$ obtained from these sets only contains substochastic matrices.

Let $A = (\alpha_1,\dots,\alpha_d)$ be a tuple of actions in $\mathcal{M}$, where $\alpha_i \in Act_i$ for each i. Consider its corresponding matrix $M \in \mathcal{F}$. Formally, $(M)_{i,j} = \mathbb{P}(s_i,\alpha_i,s_j)$. Let $L_A(\boldsymbol{x})$ denote the vector $(L_{\alpha_1}(\boldsymbol{x}),\dots,L_{\alpha_d}(\boldsymbol{x}))^\top$. By definition of a row vector of an action, we have $L_A(\boldsymbol{x}) = M\cdot\boldsymbol{x}+\mathbb{P}_A$, where $\mathbb{P}_A = (\mathbb{P}(s_1,\alpha_1,t),\dots,\mathbb{P}(s_d,\alpha_d,t))^\top$. In particular, $\mu\Phi = M\cdot\mu\Phi + \mathbb{P}_A$ since every action in A is tight. As an intermediate step, we prove the claim below.

▷ Claim. The following equalities hold for every $\boldsymbol{x} \in U_\delta(\boldsymbol{t})$:
- $f(L_A(\boldsymbol{x})) = M\cdot f(\boldsymbol{x})$,
- $f(\Phi(\boldsymbol{x})) = F(f(\boldsymbol{x}))$,
- and $f(\Phi^n(\boldsymbol{x})) = F^n(f(\boldsymbol{x}))$ for every $n \geq 1$.

Proof (of the Claim). Use linearity of L_A and the equality $L_A(\mu\Phi) = \mu\Phi$ to obtain

$$M\cdot f(\boldsymbol{x}) = M\cdot(\boldsymbol{x}-\mu\Phi) = M\cdot\boldsymbol{x} - M\cdot\mu\Phi = M\cdot\boldsymbol{x} - (\mu\Phi - \mathbb{P}_A) = M\cdot\boldsymbol{x} - \mu\Phi + \mathbb{P}_A$$
$$\text{and}\qquad f(L_A(\boldsymbol{x})) = L_A(\boldsymbol{x}) - \mu\Phi = M\cdot\boldsymbol{x} + \mathbb{P}_A - \mu\Phi.$$

The first statement thus holds. The second statement follows from $\Phi(\boldsymbol{x}) = \max_A L_A(\boldsymbol{x})$. Indeed, $f(\Phi(\boldsymbol{x})) = f(\max_A L_A(\boldsymbol{x})) = \max_A(f(L_A(\boldsymbol{x}))) = \max_{M\in\mathcal{F}} M\cdot f(\boldsymbol{x}) = F(f(\boldsymbol{x}))$.

The second statement further serves both as the base case (with argument $\Phi(\boldsymbol{x})$) and the induction step (with argument $\Phi^{n+1}(\boldsymbol{x})$) to prove the final statement. ◁

Now, set $\boldsymbol{\varepsilon} := f(\boldsymbol{s}) = \boldsymbol{s} - \mu\Phi$ in the PFR problem. Clearly, $\boldsymbol{\varepsilon} \in [-1,1]^d \cap \mathbb{Q}^d$. The equality $\Phi^n(\boldsymbol{s}) = \mu\Phi$ holds if and only if $f(\Phi^n(\boldsymbol{s})) = \boldsymbol{0}$ if and only if $F^n(\boldsymbol{\varepsilon}) = \boldsymbol{0}$. ◀

Example 23 (revisited). The MDP $\mathcal{M}_2$ in Figure 2 has a product family $\mathcal{F} = \{M_1, M_2\}$. The two matrices correspond to the action tuples $A_1 = (\alpha_1,\beta,\gamma)$ and $A_2 = (\alpha_2,\beta,\gamma)$:

$$M_1 = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 0 \\ 1/3 & 0 & 1/3 \end{pmatrix}, \qquad M_2 = \begin{pmatrix} 1/2 & 1/4 & 1/4 \\ 1/3 & 1/3 & 0 \\ 1/3 & 0 & 1/3 \end{pmatrix}.$$

We have $\boldsymbol{\varepsilon} := f(\boldsymbol{s}) = \boldsymbol{s} - \mu\Phi = (-1/2, 1/3, 1/3)$, and $F^2(\boldsymbol{\varepsilon}) = \max_{1\leq i,j\leq 2}(M_j M_i \boldsymbol{\varepsilon}) = M_2 M_1 \boldsymbol{\varepsilon} = \boldsymbol{0}$. Indeed, $\Phi^2(\boldsymbol{s}) = \mu\Phi$. Meanwhile, $M_1^n\cdot\boldsymbol{\varepsilon} \neq \boldsymbol{0}$ and $M_2^n\cdot\boldsymbol{\varepsilon} \neq \boldsymbol{0}$ for all n.

▶ Remark 26. An instance of the PFR problem is a "yes" instance if and only if

$$\exists n.\ (\exists M = M_{i_n}\dots M_{i_1}.\ M\cdot\boldsymbol{\varepsilon} = \boldsymbol{0})\ \wedge\ (\forall M' = M_{j_n}\dots M_{j_1}.\ M'\cdot\boldsymbol{\varepsilon} \leq \boldsymbol{0}),$$

where $i_1,\dots,i_n,j_1,\dots,j_n \in \{1,\dots,k\}$. Finding a semigroup element $M \in \mathcal{S}$ that satisfies $M\cdot\boldsymbol{\varepsilon} = \boldsymbol{0}$, together with previously discussed techniques, is sufficient to answer the BOR problem. If $M\cdot\boldsymbol{\varepsilon} = M_{i_n}\dots M_{i_1}\cdot\boldsymbol{\varepsilon} = \boldsymbol{0}$, then $F^n(\boldsymbol{\varepsilon}) \geq 0$ and hence $\Phi^n(\boldsymbol{s}) \geq \boldsymbol{t}$. Then, we can apply the complete algorithm from Proposition 17. Notice that this does not necessarily imply that $(\Phi,\boldsymbol{s},\boldsymbol{t})$ is a positive BOR instance.

However, deciding whether a matrix $M \in \mathcal{S}$ with $M \cdot \boldsymbol{\varepsilon} = \mathbf{0}$ exists is per se an undecidable problem for general matrices [3]. There, the so-called *vector reachability problem* for a matrix semigroup $\mathcal{S} = \langle M_1, \dots, M_k \rangle$ asks for given $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{Q}^d$, whether there exists $M \in \mathcal{S}$ such that $M \cdot \boldsymbol{x} = \boldsymbol{y}$.

Unique tight actions. We conclude this section with the discussion of the PFR Problem restricted to semigroups generated by a single matrix. This corresponds to the assumption that there is a unique tight action in every state $s_i \in S_d$. The decidability for this restriction is not surprising and, in particular, follows from the algorithmic procedure of [18].

Nevertheless, we employ simple linear algebra techniques to provide an alternative proof for our version of the problem (importantly, the target vector is zero).

▶ **Proposition 27.** *Let $M \in \mathbb{Q}^d$ be a substochastic matrix and $\boldsymbol{\varepsilon} \in [-1,1]^d \cap \mathbb{Q}^d$ an arbitrary vector. If there exists $n \geq 0$ such that $M^n \cdot \boldsymbol{\varepsilon} = \mathbf{0}$, then there exists such $n \leq d$. Whether there exists $n \geq 0$ such that $M^n \cdot \boldsymbol{\varepsilon} = \mathbf{0}$ can thus be answered algorithmically.*

Proof. Let $\ker M$ denote the kernel of M, the set of all vectors $\boldsymbol{x}$ such that $M \cdot \boldsymbol{x} = \mathbf{0}$.

Clearly, for all $n \geq 0$, $\ker M^n \subseteq \ker M^{n+1}$. Furthermore, notice that $\ker M^n = \ker M^{n+1}$ implies $\ker M^{n+k} = \ker M^n$ for all $k \geq 0$. Assume, for sake of contradiction, that there exists n and $k \geq 2$ such that $\ker M^n = \cdots = \ker M^{n+k-1} \subsetneq \ker M^{n+k}$. Then consider $\boldsymbol{x} \in \ker M^{n+k} \setminus \ker M^n$. It holds $M^{k-1}\boldsymbol{x} \in \ker M^{n+1}$ and, by assumption, $M^{k-1}\boldsymbol{x} \in \ker M^n$. Then, however, $\boldsymbol{x} \in \ker M^{n+k-1} \setminus \ker M^n$, a contradiction that proves that, for any n, either $\ker M^n \subsetneq \ker M^{n+1}$, or $\ker M^n = \ker M^{n+1} = \cdots = \ker M^{n+k} = \dots$ holds.

Recall that for every n, $\ker M^n$ is a linear subspace of $\mathbb{R}^d$. Therefore, either $\dim \ker M^{n+1} > \dim \ker M^n$, or $\dim \ker M^n = \cdots = \dim \ker M^{n+k} = \dots$. Since $\dim \ker M^{n+i}$ is bounded from above by d, the first statement follows.

Finally, it suffices to check $M^n \cdot \boldsymbol{\varepsilon} \neq \mathbf{0}$ for $n \leq d$ in order to deduce that the equality does not hold for any n. This gives a complete algorithm to solve the $\mathbf{0}$-reachability problem. ◀

The reachability problem for Bellman operators is thus completely solved for MDPs in which a positional scheduler with all tight actions is unique. Equivalently, it is solved for $\mathcal{S} = \langle M \rangle$.

However, we cannot employ the previous argument for $\mathcal{S} = \langle M_1, M_2 \rangle$, since the implication $\ker(M_1^k) = \ker(M_1^{k+1}) \Rightarrow \ker(M_2 \cdot M_1^k) = \ker(M_2 \cdot M_1^{k+1})$ does not hold there. For instance, in Example 23, $\ker M_2 M_1 \neq \ker M_2 M_1^2$. The decidability remains open for general semigroups – that is, when there are states with non-unique tight actions.

4 A Decidable Case $d = 2$

In this section, we show that the PFR problem is decidable in dimension $d = 2$. Following Proposition 25, this will suffice for the decidability of the BOR problem in $d = 2$.

The key property that helps us establish the decidability is the existence of a total order associated with the (two-dimensional) row vectors of actions. By arguing about this order, we show that the sequence $\boldsymbol{\varepsilon}, F(\boldsymbol{\varepsilon}), F^2(\boldsymbol{\varepsilon}), \dots$ either has a vector comparable with $\mathbf{0}$ among its first terms – or never reaches $\mathbf{0}$.

Actions in $d = 2$ and lines. We further consider an MDP $\mathcal{M}$ with $d = 2$. Let $Act_1 = \{\alpha_1, \dots, \alpha_k\}$ and $Act_2 = \{\beta_1, \dots, \beta_\ell\}$ be the sets of tight actions available in two states $S_d = \{s_1, s_2\}$ of $\mathcal{M}$. Recall that we associate the sets of row vectors $\mathcal{F}_1$ and $\mathcal{F}_2$ with actions in Act_1 and Act_2, respectively. An action is a *zero* α- or β-action if its row vector is a zero vector $(0, 0)$. Denote by Act_i^* the subset of all non-zero actions in Act_i, for each $i \in \{1, 2\}$.

We now discuss lines that correspond to the actions of MDPs with $d = 2$. For our proof, it is important to identify the actions whose lines have the greatest/least slope. With a mild abuse of notation, we denote by $\alpha_i(x_1, x_2)$ the homogeneous part of the linear polynomial of α_i (similarly for β_j). This is exactly the dot product of action's row vector with (x_1, x_2), or $\alpha_i(x_1, x_2) := \mathbb{P}(s_1, \alpha_i, s_1)x_1 + \mathbb{P}(s_1, \alpha_i, s_2)x_2$.

For each non-zero action γ, the set of points (x_1, x_2) with $\gamma(x_1, x_2) = 0$ is a line orthogonal to the row vector of γ. We denote the *angle* between this line and the positive direction of the x-axis by $\measuredangle\gamma$. The angles of actions in $Act^* = Act_1^* \cup Act_2^*$ are numbers in $[\frac{\pi}{2}, \pi]$ and thus are totally ordered.

We denote by α_{lo} an α-action with the greatest angle. The sign of $\alpha_{lo}(x_1, x_2)$ is well-defined, that is, it is independent of the choice of action among those with the greatest angle. Similarly, α_{hi} is an α-action with the least angle. Actions β_{lo} and β_{hi} are defined analogously.

We now resume the matrix argumentation. The row vectors in $\mathcal{F}_1$ and $\mathcal{F}_2$ define the product family $\{M_{1,1}, \dots, M_{k,\ell}\} \subset \mathbb{Q}^{2\times 2}$. That is, $M_{i,j} \cdot \boldsymbol{v} = \big(\alpha_i(\boldsymbol{v}), \beta_j(\boldsymbol{v})\big)^\top$ holds for all $\alpha_i \in Act_1$, $\beta_j \in Act_2$, and $\boldsymbol{v} \in \mathbb{Q}^2$.

▶ **Definition 28.** *We define $F^{-1}(\mathbf{0})$ as the set of all vectors $\boldsymbol{\varepsilon}$ such that $F(\boldsymbol{\varepsilon}) = \mathbf{0}$.*

▶ Remark 29. The set $F^{-1}(\mathbf{0})$ is entirely contained in the union $\cup_{i,j} \ker M_{i,j}$ of kernels.

One immediate consequence thereof is that at least one of the matrices is singular. In other words, there exists $M_{i,j}$ with $\ker M_{i,j} \neq \{\mathbf{0}\}$. Otherwise, no vector $\boldsymbol{x} \neq \mu\Phi$ ever reaches $\mu\Phi$.

▶ **Lemma 30.** *Given a product family $\{M_{1,1}, \dots, M_{k,\ell}\} \subset \mathbb{Q}^{2\times 2}$ of substochastic matrices and a vector $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2)$. The map F is defined as above by*

$$F(\boldsymbol{v}) = \max_{1\le i\le k, 1\le j\le \ell} (M_{i,j} \cdot \boldsymbol{v}).$$

Exactly one of the two statements holds:

1. *$F^n(\boldsymbol{\varepsilon}) \neq \mathbf{0}$ for all n.*
2. *$F^2(\boldsymbol{\varepsilon})$ is comparable with $\mathbf{0}$.*

Proof. We can assume $\boldsymbol{\varepsilon} \bowtie \mathbf{0}$, otherwise 2. holds trivially. Let Q_i, $i \in \{1, 2, 3, 4\}$, be a coordinate plane quadrant. Here, each Q_i is a closed set. For example, $Q_1 = \{\boldsymbol{x} = (x_1, x_2) : x_1 \ge 0 \wedge x_2 \ge 0\}$. Let $\operatorname{int} Q_1 = \{\boldsymbol{x} = (x_1, x_2) : x_1 > 0 \wedge x_2 > 0\}$ be the interior of Q_1, similarly for other quadrants. Note that $\{\operatorname{int} Q_1, \operatorname{int} Q_3, Q_2 \setminus \{0\}, Q_4\}$ is a partition π of $\mathbb{R}^2$.

A matrix kernel $\ker M_{i,j} \subseteq \mathbb{R}^2$ is either a singleton $\{\mathbf{0}\}$, a line through the origin, or the entire ambient space. Observe that either there exist both a zero α- and a zero β-action; or $F^{-1}(\mathbf{0})$ is a union of finitely many lines. In the latter case, both $\operatorname{int} Q_1$ and $\operatorname{int} Q_3$ do not intersect $F^{-1}(\mathbf{0})$ because no one-dimensional kernel intersects $\operatorname{int} Q_1$ or $\operatorname{int} Q_3$ – all matrices are non-negative. From Remark 29, we have $F^{-1}(\mathbf{0}) \subseteq Q_2 \cup Q_4$ for this case.

The discussion below is driven by the question

"When are the vectors $F(\boldsymbol{\varepsilon})$ and $F^2(\boldsymbol{\varepsilon})$ incomparable with $\mathbf{0}$?".

We perform a case distinction by comparing the angles $\measuredangle\alpha_{lo}$ and $\measuredangle\beta_{lo}$. Crucially,

$$F(\boldsymbol{v}) = \left(\max_{1\le i\le k} \alpha_i(\boldsymbol{v}), \max_{1\le j\le \ell} \beta_j(\boldsymbol{v})\right). \tag{4}$$

Without loss of generality, let $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2)$ be such that $\varepsilon_1 < 0$ and $\varepsilon_2 > 0$, i.e., $\boldsymbol{\varepsilon} \in \operatorname{int} Q_2$.

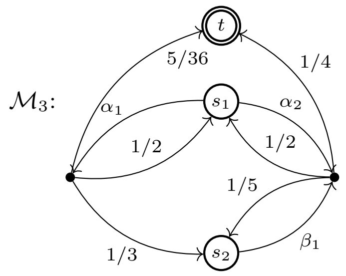
1. $Act_1^* = \varnothing$ or $Act_2^* = \varnothing$.
If all α-actions are zero, then $F(\boldsymbol{\varepsilon}) = (0, x_2)$ and thus comparable with $\mathbf{0}$. Analogously, if all β-actions are zero, then $F(\boldsymbol{\varepsilon}) = (x_1, 0)$ is a vector comparable with $\mathbf{0}$. In the rest of the case distinction, we assume that the sets Act_1^*, Act_2^* are *both non-empty*.
2. $\measuredangle\alpha_{lo} = \measuredangle\beta_{lo}$. We first consider $\boldsymbol{\varepsilon}$ with $\alpha_{lo}(\boldsymbol{\varepsilon}) < 0$ and show that $F(\boldsymbol{\varepsilon})$ is comparable. Indeed, *if there are no zero actions*, we have $\alpha_i(\boldsymbol{\varepsilon}) < 0$ for all i. Similarly, all β-actions yield negative values, and so $F(\boldsymbol{\varepsilon})$ is a strictly negative vector by (4). *Otherwise*, at least one zero action is chosen, resulting in a vector with a zero entry (hence comparable with $\mathbf{0}$). Now, if $\alpha_{lo}(\boldsymbol{\varepsilon}) \geq 0$, then so is $\beta_{lo}(\boldsymbol{\varepsilon}) \geq 0$. It follows immediately that $F(\boldsymbol{\varepsilon}) \geq \mathbf{0}$.
3. $\measuredangle\beta_{lo} > \measuredangle\alpha_{lo}$. There are two subcases based on the existence of zero actions.
   a. $Act_1^* \subsetneq Act_1$ (there exists a zero α-action). If $\alpha_{lo}(\boldsymbol{\varepsilon}) \geq 0$, then $F(\boldsymbol{\varepsilon}) \geq \mathbf{0}$. Otherwise, a zero α-action is chosen, hence $F(\boldsymbol{\varepsilon})$ is comparable with $\mathbf{0}$ from the argument of Case 1.
   b. $Act_1^* = Act_1$ (there are no zero α-actions). We deduce from $\measuredangle\beta_{lo} > \measuredangle\alpha_{lo}$ that

$$\boldsymbol{x} \in Q_2 \setminus \{0\} \quad \Rightarrow \quad \max_i \alpha_i(\boldsymbol{x}) < 0 \ \vee \ \max_j \beta_j(\boldsymbol{x}) > 0. \tag{5}$$

   We show that $F^n(\boldsymbol{\varepsilon}) \notin Q_4$ for all $n \geq 0$. Assume towards a contradiction that m is the smallest integer such that $F^m(\boldsymbol{\varepsilon}) \in Q_4$. Consider $y := F^{m-1}(\boldsymbol{\varepsilon})$. Since $F(\boldsymbol{y})_1 \geq 0$, we have $\alpha_{lo}(\boldsymbol{y}) \geq 0$. This implies $\boldsymbol{y} \in Q_2 \cup \operatorname{int} Q_1$. But $\boldsymbol{y} \in \operatorname{int} Q_1$ means $\beta_{lo}(\boldsymbol{y}) > 0$ and hence $F(\boldsymbol{y})_2 > 0$, contradicting $F(\boldsymbol{y}) \in Q_4$. We have deduced $\boldsymbol{y} \in Q_2$. Now, from (5) we either have $\max_i \alpha_i(\boldsymbol{y}) < 0$ or $\max_j \beta_j(\boldsymbol{y}) > 0$. The former implies $F(\boldsymbol{y})_1 < 0$ whereas the latter implies $F(\boldsymbol{y})_2 > 0$. Either of two contradicts $F(\boldsymbol{y}) \in Q_4$. Our assumption was wrong, and $F^n(\varepsilon) \notin Q_4$ for all $n \geq 0$.
4. $\measuredangle\beta_{lo} < \measuredangle\alpha_{lo}$.
   a. $Act_2^* \subsetneq Act_2$ (there exists a zero β-action). If $\beta_{lo}(\boldsymbol{\varepsilon}) \geq 0$, then $F(\boldsymbol{\varepsilon}) \geq \mathbf{0}$. Otherwise, a zero β-action is chosen, hence $F(\boldsymbol{\varepsilon})$ is comparable with $\mathbf{0}$.
   b. $Act_2^* = Act_2$ (there are no zero β-actions). A new phenomenon happens now: there might exist $n > 0$ such that an incomparable vector $F^n(\boldsymbol{\varepsilon})$ is in Q_4. However, observe first that $\alpha_{lo}(\boldsymbol{\varepsilon}) \leq 0$ implies $\beta_{lo}(\boldsymbol{\varepsilon}) < 0$ and hence the vector $F(\boldsymbol{\varepsilon})$ is comparable. Moreover, if $\beta_{lo}(\boldsymbol{\varepsilon}) \geq 0$, then similarly $\alpha_{lo}(\boldsymbol{\varepsilon}) > 0$ and hence, $F(\boldsymbol{\varepsilon}) \geq \mathbf{0}$. We move on to the case when $\alpha_{lo}(\boldsymbol{\varepsilon}) > 0$ and $\beta_{lo}(\boldsymbol{\varepsilon}) < 0$.
   Indeed, we then have $F(\boldsymbol{\varepsilon}) \in Q_4 \setminus \{\mathbf{0}\}$. In the discussion that follows we analyse how $F^2(\boldsymbol{\varepsilon})$ depends on $F(\boldsymbol{\varepsilon})$. From now on, we focus our attention on actions α_{hi} and β_{hi}, *rather than* α_{lo} *and* β_{lo}. In particular, if α_{hi} evaluates to a non-negative number at $F(\boldsymbol{\varepsilon}) \in Q_4$, then $F^2(\boldsymbol{\varepsilon})_1 \geq 0$. We keep assuming that no zero β-actions exist.
      i. $\measuredangle\beta_{hi} = \measuredangle\alpha_{hi}$. Consider first the case $\alpha_{hi}(F(\boldsymbol{\varepsilon})) < 0$. Then $F^2(\boldsymbol{\varepsilon}) \leq \mathbf{0}$ because $\beta_i(F(\boldsymbol{\varepsilon})) < 0$ holds for all $\beta_i \in Act_2$. See also Case 2. Now, if $\alpha_{hi}(F(\boldsymbol{\varepsilon})) \geq 0$, then so is $\beta_{hi}(F(\boldsymbol{\varepsilon})) \geq 0$. It follows immediately that $F^2(\boldsymbol{\varepsilon}) \geq \mathbf{0}$.
      ii. $\measuredangle\beta_{hi} > \measuredangle\alpha_{hi}$. Similarly to the Case 3b, we observe

$$\boldsymbol{x} \in Q_4 \setminus \{0\} \quad \Rightarrow \quad \max_i \alpha_i(\boldsymbol{x}) > 0 \ \vee \ \max_j \beta_j(\boldsymbol{x}) < 0.$$

      We have $F^{-1}(\mathbf{0}) \cap Q_4 = \{\mathbf{0}\}$, hence $F^{-1}(\mathbf{0}) \subset Q_2$. Now if for some $\boldsymbol{x} \in Q_2 \setminus \{0\}$, we have $(F(\boldsymbol{x}))_1 = 0$, then $\alpha_{lo}(\boldsymbol{x}) \leq 0$. Then $\beta_j(\boldsymbol{x}) < 0$ for all j, implying $F(\boldsymbol{x}) \neq \mathbf{0}$. Therefore, $F^{-1}(\mathbf{0}) = \{\mathbf{0}\}$ and we have $F^n(\boldsymbol{\varepsilon}) \neq \mathbf{0}$ for all $n \geq \mathbf{0}$.



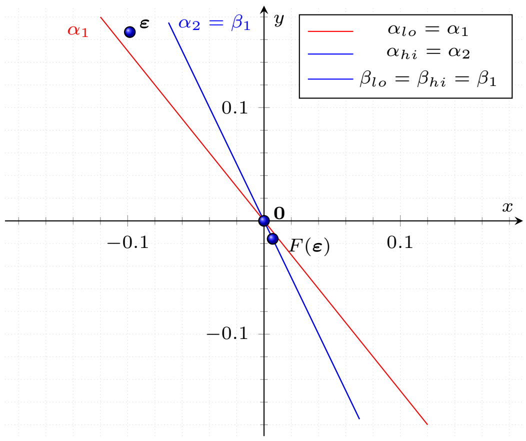


Figure 3 An MDP $\mathcal{M}_3$ of Example 32.

Figure 4 $\mathcal{M}_3$ illustrating case 4.b.i.

iii. $\measuredangle\beta_{hi} < \measuredangle\alpha_{hi}$. Assume first that a zero α-action exists. In this case, either $\alpha_{hi}(F(\boldsymbol{\varepsilon})) > 0$ and $\beta_{hi}(F(\boldsymbol{\varepsilon})) > 0$ follows as well; or $\alpha_{hi}(F(\boldsymbol{\varepsilon})) \leq 0$. Regardless, $F^2(\boldsymbol{\varepsilon})$ is a comparable vector (positive; or having 0 in the first component). We now proceed under the assumption that no zero (α- or β-) actions exist. Recall also that we keep assuming $\measuredangle\beta_{lo} < \measuredangle\alpha_{lo}$ as well as $\measuredangle\beta_{hi} < \measuredangle\alpha_{hi}$. This suffices to prove that $F^{-1}(\mathbf{0}) = \{\mathbf{0}\}$. Assume the opposite.

- Let $\boldsymbol{x} \in F^{-1}(\mathbf{0}) \cap Q_2 \setminus \{0\}$. It necessarily holds $\alpha_{lo}(\boldsymbol{x}) = 0$. Hence, $\beta_j(\boldsymbol{x}) < 0$ for all $1 \leq j \leq \ell$, and $(F(\boldsymbol{x}))_2 < 0$. This contradicts $\boldsymbol{x} \in F^{-1}(\mathbf{0})$.
- Let $\boldsymbol{x} \in F^{-1}(\mathbf{0}) \cap Q_4 \setminus \{0\}$. We necessarily have $\beta_{hi}(\boldsymbol{x}) = 0$. Hence, $\alpha_i(\boldsymbol{x}) < 0$ for all $1 \leq i \leq k$, and $(F(\boldsymbol{x}))_1 < 0$. This clearly contradicts $\boldsymbol{x} \in F^{-1}(\mathbf{0})$.

After deducing $F^{-1}(\mathbf{0}) = \{\mathbf{0}\}$, we have $F^n(\boldsymbol{\varepsilon}) \neq \mathbf{0}$ for all n also in this case.

In most cases we were able to show $F(\boldsymbol{\varepsilon})$ is comparable with $\mathbf{0}$. This implies $F^2(\boldsymbol{\varepsilon})$ is comparable with $\mathbf{0}$, too. In all other cases we either directly showed that $F^2(\boldsymbol{\varepsilon})$ is comparable with $\mathbf{0}$, or derived $F^n(\boldsymbol{\varepsilon}) \neq \mathbf{0}$ for all n. ◀

With Lemma 30, if $\Phi^2(\boldsymbol{s}) \bowtie \mu\Phi$ for $\boldsymbol{s} \in U_\delta(\mu\Phi)$, then the answer to the BOR problem is guaranteed negative. Otherwise, we can exploit the procedure of Theorem 22. We derive the following result for Bellman operators in $d = 2$ with arbitrarily many pieces.

▶ **Theorem 31.** *The BOR problem is decidable with $d = 2$.*

▶ **Example 32.** We provide an example $\mathcal{M}_3$ with $d = 2$, where $\boldsymbol{\varepsilon}, F(\boldsymbol{\varepsilon})$ are vectors incomparable with $\mathbf{0}$, and $F^2(\boldsymbol{\varepsilon}) = \mathbf{0}$. Here, $\alpha_1(x_1, x_2) = \frac{1}{2}x_1 + \frac{1}{3}x_2$, $\alpha_2(x_1, x_2) = \beta_1(x_1, x_2) = \frac{1}{2}x_1 + \frac{1}{5}x_2$. That is, the example fits into the Case 4b.i. Let $\boldsymbol{\varepsilon} = (-\frac{31}{315}, \frac{1}{6})$. Then, $\alpha_1(\boldsymbol{\varepsilon}) > \alpha_2(\boldsymbol{\varepsilon})$; we get $F(\boldsymbol{\varepsilon}) = (\frac{2}{315}, -\frac{1}{63})$. Further, $\alpha_2(F(\boldsymbol{\varepsilon})) > \alpha_1(F(\boldsymbol{\varepsilon}))$ and $F^2(\boldsymbol{\varepsilon}) = (0, 0)$.

5 Related Work and Discussion

In our work, we have outlined a series of phenomena inherent to the iterative application of Bellman operators. Crucial was the fact that the update coefficients were non-negative.

Notice that any PAM on the domain $[0, 1]^d$ can be represented as a nested min-max of its affine components [14, 23]. This supports the relevance of PAMs with updates defined by the maximum of affine pieces, Bellman operators of MDPs being among them. Generalising our results beyond Bellman operators of MDPs using monotonicity and fixed-point convergence is a subject of our future work.

It is worth pointing out that we capture a large class of PAMs for which the assumptions of the known techniques do not hold. In proving decidability for $d = 2$, we do not impose the restrictions used in works on the one-dimensional version of the problem. Bellman operators are, in general, neither injective as in [13] nor complete [8], nor even surjective.

Our techniques and results can be applied to other problems about PAMs. Consider, in particular, the universally quantified version of BOR: given a vector $\boldsymbol{t} \in [0,1]^d \cap \mathbb{Q}$ and a Bellman operator $\Phi : [0,1]^d \to [0,1]^d$, does there exist $n \in \mathbb{N}$ *for every* $\boldsymbol{s} \in [0,1]^d \cap \mathbb{Q}$ such that $\Phi^n(\boldsymbol{s}) = \boldsymbol{t}$? This is the mortality problem, known to be undecidable for general PAMs in dimension 2 [5,6]. For Bellman operators, however, it is equivalent to solving the BOR problem for $\boldsymbol{s} = \mathbf{0}$ and for $\boldsymbol{s} = \mathbf{1}$. Indeed, both BOR instances are "yes"-instances if and only if every instance with $\mathbf{0} \leq \boldsymbol{s} \leq \mathbf{1}$ is a "yes". We can answer this using our novel algorithm (Theorem 22), hence mortality for Bellman operators is decidable in all dimensions.

Without restrictions on the dimension d, the decidability of the BOR problem remains open. Notably, if there exist states with multiple tight actions, then the dynamics of the sequence $\langle F^n(\boldsymbol{\varepsilon})\rangle_{n\in\mathbb{N}}$ as defined in Section 3.4 is intricate for $\boldsymbol{\varepsilon} \bowtie \mathbf{0}$. For more works discussing the iterative dynamics of map F see [26,30]. We highlight that $F^n(\boldsymbol{\varepsilon})$ does not have a closed form, in contrast to $M^n\boldsymbol{\varepsilon}$ for a fixed matrix M. While the entries of the vector $M^n\boldsymbol{\varepsilon}$ are terms of linear recurrence sequences, the class of such sequences is not closed under max [29]. The behaviour of entries in $F^n(\boldsymbol{\varepsilon})$ is even subtler than that, as witnessed by the non-positionality discussion and Example 23.

We emphasise that due to the undecidability of the vector reachability problem for matrix semigroups [3], following Remark 26 we need to argue about ultimate non-positivity of $F^n(\boldsymbol{\varepsilon})$. However, even for one M, deciding whether $(M^n\boldsymbol{\varepsilon})_1 \leq 0$ for some n is equivalent to the positivity problem [22], open in dimension $d > 5$. This hardness carries on to the stochastic matrices [28]. Moreover, the question whether there exists n such that $M^n\boldsymbol{\varepsilon} \leq \mathbf{0}$ corresponds to the polyhedron-hitting problem, and is Diophantine-hard [11] for general matrices M.

References

1. Christel Baier and Joost-Pieter Katoen. *Principles of model checking.* MIT Press, 2008.
2. Christel Baier, Joachim Klein, Linda Leuschner, David Parker, and Sascha Wunderlich. Ensuring the reliability of your model checker: Interval iteration for Markov decision processes. In *CAV (1)*, volume 10426 of *Lecture Notes in Computer Science*, pages 160–180. Springer, 2017. doi:10.1007/978-3-319-63387-9_8.
3. Paul Bell and Igor Potapov. On undecidability bounds for matrix decision problems. *Theor. Comput. Sci.*, 391(1-2):3–13, 2008. doi:10.1016/J.TCS.2007.10.025.
4. Richard Bellman. A Markovian decision process. *Journal of Mathematics and Mechanics*, 6(5):679–684, 1957. URL: http://www.jstor.org/stable/24900506.
5. Amir M. Ben-Amram. Mortality of iterated piecewise affine functions over the integers: Decidability and complexity. *Comput.*, 4(1):19–56, 2015. doi:10.3233/COM-150032.
6. Vincent D. Blondel, Olivier Bournez, Pascal Koiran, Christos H. Papadimitriou, and John N. Tsitsiklis. Deciding stability and mortality of piecewise affine dynamical systems. *Theor. Comput. Sci.*, 255(1-2):687–696, 2001. doi:10.1016/S0304-3975(00)00399-6.
7. Vincent D. Blondel and John N. Tsitsiklis. A survey of computational complexity results in systems and control. *Autom.*, 36(9):1249–1274, 2000. doi:10.1016/S0005-1098(00)00050-9.
8. Olivier Bournez, Oleksiy Kurganskyy, and Igor Potapov. Reachability problems for one-dimensional piecewise affine maps. *Int. J. Found. Comput. Sci.*, 29(4):529–549, 2018. doi:10.1142/S0129054118410046.
9. Tomás Brázdil, Krishnendu Chatterjee, Martin Chmelik, Vojtech Forejt, Jan Kretínský, Marta Z. Kwiatkowska, David Parker, and Mateusz Ujma. Verification of Markov decision

processes using learning algorithms. In *ATVA*, volume 8837 of *Lecture Notes in Computer Science*, pages 98–114. Springer, 2014. doi:10.1007/978-3-319-11936-6_8.

10 Krishnendu Chatterjee and Thomas A. Henzinger. Value iteration. In *25 Years of Model Checking*, volume 5000 of *Lecture Notes in Computer Science*, pages 107–138. Springer, 2008. doi:10.1007/978-3-540-69850-0_7.

11 Ventsislav Chonev, Joël Ouaknine, and James Worrell. The polyhedron-hitting problem. In *SODA*, pages 940–956. SIAM, 2015. doi:10.1137/1.9781611973730.64.

12 Luca de Alfaro. *Formal verification of probabilistic systems*. PhD thesis, Stanford University, USA, 1997. URL: https://searchworks.stanford.edu/view/3910936.

13 Faraz Ghahremani, Edon Kelmendi, and Joël Ouaknine. Reachability in injective piecewise affine maps. In *LICS*, pages 1–11. IEEE, 2023. doi:10.1109/LICS56636.2023.10175723.

14 V. V. Gorokhovik, O. I. Zorko, and G. Birkhoff. Piecewise affine functions and polyhedral sets. *Optimization*, 31(3):209–221, 1994.

15 Serge Haddad and Benjamin Monmege. Interval iteration algorithm for MDPs and IMDPs. *Theor. Comput. Sci.*, 735:111–131, 2018. doi:10.1016/J.TCS.2016.12.003.

16 Arnd Hartmanns and Benjamin Lucien Kaminski. Optimistic value iteration. In *CAV (2)*, volume 12225 of *Lecture Notes in Computer Science*, pages 488–511. Springer, 2020. doi:10.1007/978-3-030-53291-8_26.

17 R. Kannan and R. J. Lipton. Polynomial-Time Algorithm for the Orbit Problem. *J. ACM*, 33(4):808–821, August 1986. doi:10.1145/6490.6496.

18 Ravindran Kannan and Richard J. Lipton. The orbit problem is decidable. In *STOC*, pages 252–261. ACM, 1980. doi:10.1145/800141.804673.

19 Toghrul Karimov, Edon Kelmendi, Joël Ouaknine, and James Worrell. What's decidable about discrete linear dynamical systems? In *Principles of Systems Design*, volume 13660 of *Lecture Notes in Computer Science*, pages 21–38. Springer, 2022. doi:10.1007/978-3-031-22337-2_2.

20 Pascal Koiran, Michel Cosnard, and Max Garzon. Computability with low-dimensional dynamical systems. *Theor. Comput. Sci.*, 132(1):113–128, September 1994. doi:10.1016/0304-3975(94)90229-1.

21 Bart Kuijpers. Deciding the point-to-fixed-point problem for skew tent maps on an interval. *J. Comput. Syst. Sci.*, 115:113–120, 2021. doi:10.1016/J.JCSS.2020.07.005.

22 Joël Ouaknine and James Worrell. Positivity problems for low-order linear recurrence sequences. In *SODA*, pages 366–379. SIAM, 2014. doi:10.1137/1.9781611973402.27.

23 Sergei Ovchinnikov. Max-min representation of piecewise linear functions. *Beiträge zur Algebra und Geometrie*, 43(1):297–302, 2002. URL: http://eudml.org/doc/225460.

24 Martin L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley Series in Probability and Statistics. Wiley, 1994. doi:10.1002/9780470316887.

25 Tim Quatmann and Joost-Pieter Katoen. Sound value iteration. In *CAV (1)*, volume 10981 of *Lecture Notes in Computer Science*, pages 643–661. Springer, 2018. doi:10.1007/978-3-319-96145-3_37.

26 Karel Sladký. Bounds on discrete dynamic programming recursions. i. models with non-negative matrices. *Kybernetika*, 16(6):(526)–547, 1980. URL: http://eudml.org/doc/28460.

27 Ashish Tiwari. Termination of linear programs. In *CAV*, volume 3114 of *Lecture Notes in Computer Science*, pages 70–82. Springer, 2004. doi:10.1007/978-3-540-27813-9_6.

28 Mihir Vahanwala. Skolem and positivity completeness of ergodic Markov chains. *Inf. Process. Lett.*, 186:106481, 2024. doi:10.1016/J.IPL.2024.106481.

29 Gerco van Heerdt, Justin Hsu, Joël Ouaknine, and Alexandra Silva. Convex language semantics for nondeterministic probabilistic automata. In *ICTAC*, volume 11187 of *Lecture Notes in Computer Science*, pages 472–492. Springer, 2018. doi:10.1007/978-3-030-02508-3_25.

30 W.H.M. Zijm. Generalized eigenvectors and sets of nonnegative matrices. *Linear Algebra and its Applications*, 59:91–113, 1984. doi:10.1016/0024-3795(84)90161-7.