HAL open science

Contributions on complexity bounds for Deterministic Partially Observed Markov Decision Process

Cyrille Vessaire, Jean-Philippe Chancelier, Michel de Lara, Pierre Carpentier, Alejandro Rodríguez-Martínez

▶ **To cite this version:**

Cyrille Vessaire, Jean-Philippe Chancelier, Michel de Lara, Pierre Carpentier, Alejandro Rodríguez-Martínez. Contributions on complexity bounds for Deterministic Partially Observed Markov Decision Process. ENPC - École des Ponts ParisTech. 2022. hal-03902089

HAL Id: hal-03902089

https://hal.science/hal-03902089v1

Submitted on 18 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Contributions on complexity bounds for Deterministic Partially Observed Markov Decision Process

Cyrille Vessaire[*], Jean-Philippe Chancelier[*], Michel De Lara[*],
Pierre Carpentier[†], Alejandro Rodríguez-Martínez[‡]

January 18, 2023

Abstract

Markov Decision Processes (MDPs) form a versatile framework used to model a wide range of optimization problems. The MDP model consists of sets of states, actions, time steps, rewards, and probability transitions. When in a given state and at a given time, the decision maker's action generates a reward and determines the state at the next time step according to the probability transition function. However, MDPs assume that the decision maker knows the state of the controlled dynamical system. Hence, when one needs to optimize controlled dynamical systems under partial observation, one often turns toward the formalism of Partially Observed Markov Decision Processes (POMDP). POMDPs are often untractable in the general case as Dynamic Programming suffers from the curse of dimensionality. Instead of focusing on the general POMDPs, we present a subclass where transitions and observations mappings are deterministic: Deterministic Partially Observed Markov Decision Processes (DET-POMDP). That subclass of problems has been studied by (Littman, 1996) and (Bonet, 2009). It was first considered as a limit case of POMDPs by Littman, mainly used to illustrate the complexity of POMDPs when considering as few sources of uncertainties as possible. In this paper, we improve on Littman's complexity bounds. We then introduce and study an even simpler class: Separated DET-POMDPs and give some new complexity bounds for this class. This new class of problems uses a property of the dynamics and observation to push back the curse of dimensionality.

1 Introduction

Markov Decision Processes (MDPs) form a versatile framework used to model a wide range of optimization problems. Indeed, one often uses the formalism of MDPs to optimize controlled dynamical systems. It is very popular in both optimal control and machine learning community, as it can be used to model complex real-life problems (see the survey (White, 1993) for common applications). Moreover, it provides the mathematical foundations for Reinforcement Learning (see (Sutton and Barto, 2018)), and algorithms such as Policy Iteration and Dynamic Programming can efficiently solve MDPs.

In the MDP framework, a decision maker can sequentially act upon a controlled dynamical system and get some rewards. The MDP model consists of sets of states, actions, time steps, rewards, and probability transitions. When in a given state and at a given time, the decision maker's action generates a reward and determines the state at the next time step according to the probability transition function.

However, MDPs assume that the decision maker knows the state of the controlled dynamical system. Hence, when one needs to optimize controlled dynamical systems under partial observation, one often turns toward the formalism of Partially Observed Markov Decision Processes (POMDP). An extensive literature exists on POMDPs, most of which focuses on the infinite horizon case. POMDPs can be applied to numerous fields, from medical models (as in (Steimle et al., 2021)) to robotics (as in (Pajarinen and Kyrki, 2017)) to name a few. Algorithms based on Dynamic Programming (see (Bellman, 1957)) have been designed to exploit specific structures in POMDPs in order to solve this difficult class of problems. They

[*]CERMICS, Ecole des Ponts, Marne-la-Vallée, France
[†]UMA, ENSTA Paris, Institut Polytechnique de Paris, Palaiseau, France
[‡]IAM, TotalEnergies SE, Pau, France

do so by first reformulating the problem through the use of beliefs (probability distributions over the state space), as in (Smallwood and Sondik, 1973). One such algorithm is Sarsop, described in (Kurniawati et al., 2008). However, Pomdps are often untractable in the general case as Dynamic Programming suffers from the curse of dimensionality. Indeed, working with beliefs implies working on the space of distributions over the state space, which is, by nature, an infinite space.

Yet not all Pomdps suffer equally from the curse of dimensionality. Indeed, instead of focusing on the general Pomdps, we present a subclass where transitions and observations mappings are deterministic: Deterministic Partially Observed Markov Decision Processes (Det-Pomdp). That subclass of problems has been studied by (Littman, 1996) and (Bonet, 2009). It was first considered as a limit case of Pomdps by Littman, mainly used to illustrate the complexity of Pomdps when considering as few sources of uncertainties as possible. For Bonet, Det-Pomdps became of interest after some applications were found. He presented examples in (Bonet, 2009, §2), such as the navigation of a robot in a partially observed terrain.

In this paper, we improve on Littman's complexity bounds. We then introduce and study an even simpler class: Separated Det-Pomdps. This new class of problems uses a property of the dynamics and observation to push back the curse of dimensionality.

The paper is organized as follows. First, in §2, we present a general formulation of Det-Pomdp. Second, in §3, we present Dynamic Programming on beliefs for Det-Pomdps with constraints, and present complexity bounds. Third, in §4, we introduce a subclass of Det-Pomdp, Separated Det-Pomdp. Finally, in §5 we illustrate Separated Det-Pomdp with a toy problem: emptying a tank containing water when considering partial observation of the level of water in the tank. Meanwhile, in Appendix A.1, we present technical lemmata and considerations on pushforward measures. Finally, in Appendix A.2, we present complements on Separated Det-Pomdps.

We now detail our main contributions. In §3, we improve Littman (1996) bound on the cardinality of the set of reachable beliefs for Det-Pomdps (see Theorem 4). This new bound comes from a new representation of the belief dynamics in Det-Pomdps using the notion of *pushforward measure* (see Lemma 7). In §4, we introduce a subclass of Det-Pomdps, Separated Det-Pomdps. As shown in Theorem 13, the interest of Separated Det-Pomdps is that they further push back the curse of dimensionality for Dynamic Programming with beliefs (see Theorem 13). Moreover, this last bound is tight (see Proposition 16).

2 Formulation of Deterministic Partially Observed Markov Decision Processes

A Det-Pomdp is a particular case of Pomdps, itself an extension of Markov Decision Processes (Mdps). Backgrounds on Mdps can be found in Puterman (1994), whereas backgrounds on Pomdps can be found in Bertsekas and Shreve (1978). As with Mdps, the model consists of a dynamical system, defined thanks to states, controls (also called actions), transitions and time steps. At each time-step, the decision maker (also called the agent) chooses a given action, which generates a random reward depending on the state of the system and on the time. The state then transits to its next random value. However, in the case of Det-Pomdp (and Pomdp), the decision maker has only partial knowledge of the state of the dynamical system. Instead, he has access to functions of the state and controls: the *observations*. For Det-Pomdps, the transitions and observations are given by deterministic evolution and observation functions.

First, we present the ingredients of a Det-Pomdp. Second, we present the formulation of a Det-Pomdp optimization problem.

Ingredients of a Det-Pomdp A Det-Pomdp is defined by the tuple

$$\mathcal{D} = \left(\mathcal{T}, \mathbb{U}, \mathbb{O}, \mathbb{X}, \{\mathcal{L}_t\}_{t\in\mathcal{T}\setminus\{T\}}, \{f_t\}_{t\in\mathcal{T}\setminus\{T\}}, \{\mathcal{U}_t^{ad}\}_{t\in\mathcal{T}\setminus\{T\}}, \{h_t\}_{t\in\mathcal{T}}\right) , \tag{1}$$

which we now detail[1].

The set $\mathcal{T} = [\![0, T]\!]$[2] is the set of time-steps, where the positive integer $T \in \mathbb{N}\setminus\{0\}$ is colloquially known as the horizon. The set $\mathbb{U}$ is the set of controls the decision maker can choose from. The set $\mathbb{O}$ is the set

[1] For simplicity, we assume that $\mathbb{U}$, $\mathbb{O}$ and $\mathbb{X}$ are not indexed by time

[2] Let t and t' be two integers, with $t' \geq t$. The set $\{t, t+1, \ldots, t'\}$ is denoted by $[\![t, t']\!]$.

of observations available to the decision maker. The set $\mathbb{X}$ is the set of states. The collection $\{\mathcal{L}_t\}_{t\in\mathcal{T}\setminus\{T\}}$ is the collection of instantaneous costs functions: for all time $t \in \mathcal{T}\setminus\{T\}$, $\mathcal{L}_t : \mathbb{X}\times\mathbb{U} \to \mathbb{R}\cup\{+\infty\}$. Moreover, the final cost function $\mathcal{L}_T$ is by convention denoted by $\mathcal{K} : \mathbb{X} \to \mathbb{R}\cup\{+\infty\}$. The collection $\{f_t\}_{t\in\mathcal{T}\setminus\{T\}}$ is the collection of evolution functions: for all time $t \in \mathcal{T}\setminus\{T\}$, $f_t : \mathbb{X}\times\mathbb{U} \to \mathbb{X}$. They define the transitions of the dynamical system. The collection $\{\mathcal{U}_t^{ad}\}_{t\in\mathcal{T}\setminus\{T\}}$ is the collection of admissibility constraints: for all time $t \in \mathcal{T}\setminus\{T\}$, $\mathcal{U}_t^{ad} : \mathbb{X} \rightrightarrows \mathbb{U}$ is a set-valued mapping from $\mathbb{X}$ to $\mathbb{U}$, that is, for all state $x \in \mathbb{X}$, the admissible controls at time t are given by the subset of $\mathbb{U}$, $\mathcal{U}_t^{ad}(x)$. $\{h_t\}_{t\in\mathcal{T}}$ is the collection of observation functions. The initial observation is given by the mapping $h_0 : \mathbb{X} \to \mathbb{O}$ whereas, for all time $t \in \mathcal{T}\setminus\{0\}$, the observations are given by the mappings $h_t : \mathbb{X}\times\mathbb{U} \to \mathbb{O}$.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, where Ω is the set of possible outcomes, $\mathcal{F}$ is a σ-field over Ω and $\mathbb{P}$ is a probability measure on Ω. We denote by $\mathbb{E}$ the mathematical expectation operator.

In this paper, we only consider Det-Pomdps which satisfy the following finite sets assumption.

Assumption 1 (Finite sets assumption)**.** *The sets of possible outcomes Ω, of states $\mathbb{X}$, of controls $\mathbb{U}$, and observations $\mathbb{O}$ have finite cardinality. Moreover, we consider a finite number of timesteps, i.e. the horizon is finite: $T < +\infty$.*

As we consider finite sets, we introduce a notation for the set of probability distributions on finite sets. Let $\mathbb{Y}$ be a finite set. We denote by $\Delta(\mathbb{Y})$ the set of probability distributions on $\mathbb{Y}$. The set $\Delta(\mathbb{Y})$ is in bijection with the simplex $\Delta_{|\mathbb{Y}|}$ of dimension[3] $|\mathbb{Y}|$ (hence the notation).

We now present the formulation of the optimization problem which we study in this paper.

Formulation of a Det-Pomdp optimization problem A finite-horizon Det-Pomdp optimization problem is formulated as follows

$$\mathcal{V}^\star(b_0) = \min_{\boldsymbol{X},\boldsymbol{O},\boldsymbol{U}} \mathbb{E}\Big[\sum_{t=0}^{T-1} \mathcal{L}_t(\boldsymbol{X}_t, \boldsymbol{U}_t) + \mathcal{K}(\boldsymbol{X}_T)\Big] \tag{2a}$$

$$s.t.\ \ \mathbb{P}_{\boldsymbol{X}_0} = b_0\ , \tag{2b}$$

$$\boldsymbol{X}_{t+1} = f_t(\boldsymbol{X}_t, \boldsymbol{U}_t)\ ,\ \ \forall t \in \mathcal{T}\setminus\{T\}\ , \tag{2c}$$

$$\boldsymbol{O}_0 = h_0(\boldsymbol{X}_0)\ , \tag{2d}$$

$$\boldsymbol{O}_{t+1} = h_{t+1}(\boldsymbol{X}_{t+1}, \boldsymbol{U}_t)\ ,\ \ \forall t \in \mathcal{T}\setminus\{T\}\ , \tag{2e}$$

$$\boldsymbol{U}_t \in \mathcal{U}_t^{ad}(\boldsymbol{X}_t)\ ,\ \ \forall t \in \mathcal{T}\setminus\{T\}\ , \tag{2f}$$

$$\sigma(\boldsymbol{U}_t) \subset \sigma(\boldsymbol{O}_0, \ldots, \boldsymbol{O}_t, \boldsymbol{U}_0, \ldots, \boldsymbol{U}_{t-1})\ ,\ \ \forall t \in \mathcal{T}\setminus\{T\}\ , \tag{2g}$$

where we denote by $\mathcal{V}^\star(b_0)$ the optimal value of Problem (2), that is, the optimal value of the Det-Pomdp optimization problem when the initial probability distribution of the state is given by the initial belief $b_0 \in \Delta(\mathbb{X})$. In Problem (2), there are three processes $\boldsymbol{X} = \big\{\boldsymbol{X}_t\big\}_{t\in\mathcal{T}}$, $\boldsymbol{U} = \big\{\boldsymbol{U}_t\big\}_{t\in\mathcal{T}\setminus\{T\}}$ and $\boldsymbol{O} = \big\{\boldsymbol{O}_t\big\}_{t\in\mathcal{T}}$. For all time $t \in \mathcal{T}$, $\boldsymbol{X}_t : \Omega \to \mathbb{X}$ and $\boldsymbol{O}_t : \Omega \to \mathbb{O}$ are random variables representing respectively the state and the observation variables of the system at time t, and for all time $t \in \mathcal{T}\setminus\{T\}$, $\boldsymbol{U}_t : \Omega \to \mathbb{U}$ are random variables representing the controls at time t.

The optimization criterion of Problem (2) is given by Equation (2a). In this paper, we only consider the minimization of the expected value in the finite horizon case.

We now detail the constraints of the optimization Problem (2). First, Equation (2b) is the *initialization* constraint. As the initial state is not fully known, we instead use the probability distribution $b_0 \in \Delta(\mathbb{X})$ of the initial state of the system for the initialization. Second, Equation (2c) is called the *state evolution* equation of the system. It is defined thanks to the dynamics which describe the evolution of the states of the controlled dynamical system. Third, Equations (2d) and (2e) define the *observations* of the system available at each time step. Fourth, Equation (2f) is called the *admissibility constraints*: it defines which controls can be applied at each time step. Note that the proper formulation of the admissibility constraints should contain an added quantification, "$\forall\omega \in \Omega$", which we omit in this paper as the set Ω is finite, and we can always assume that $\mathbb{P}(\omega) > 0$ for all $\omega \in \Omega$. Finally, Equation (2g) is the *non-anticipativity* constraint: it defines the information available to the decision maker before choosing

[3]The cardinality of a finite set is the number of its elements and is denoted by $|\cdot|$.

a control at each time step. As all sets Ω, $\mathbb{X}$, $\mathbb{U}$ and $\mathbb{O}$ are assumed to be finite by Assumption 1, all mappings with domain Ω are random variables and Equation (2a) is well defined because $\mathcal{L}_t$ and $\mathcal{K}$ takes their values in $\mathbb{R} \cup \{+\infty\}$, hence the optimization Problem (2) is well defined.

3 Complexity analysis of Dynamic Programming for Det-Pomdps

In §3.1, we present Dynamic Programming for Det-Pomdps. Then, in §3.2 we study its *complexity*, i.e. the number of "operations" necessary to solve Problem (2). In §3.3, we present a new representation of beliefs as pushforward measures, that will be used to prove the complexity results.

3.1 Dynamic Programming for Det-Pomdp

We now present Dynamic Programming Equations with beliefs for Problem (2). As a Det-Pomdp is a Pomdp, all the results and numerical methods that apply to Pomdps are carried over to Det-Pomdps. Notably, it is possible to write Dynamic Programming equations for a finite horizon problem associated with a Pomdp. To do so, it is classical to formulate a belief-Mdp where the state is a probability distribution over the state space, called belief (see (Bertsekas and Shreve, 1978) for details on the assumptions for general Pomdps). Here, we detail this methodology for the specific Det-Pomdp case, and extend it to tackle cases with admissibility constraints on the controls.

First, in §3.1.1, we formally define sets and mappings which are necessary for the formulation of the belief-Mdp. Second, in §3.1.2, we present the Dynamic Programming equations for the resulting belief-Mdp.

3.1.1 Beliefs in Det-Pomdp

First, we present the set of beliefs. Second, we present the mappings necessary for the formulation of the belief-Mdp, notably the beliefs dynamics.

Sets for the beliefs. The dynamic programming equation for Det-Pomdps is formulated using states in the set $\Delta(\mathbb{X})$, the probability distributions over the "initial" state space $\mathbb{X}$, which are called beliefs. However, the beliefs dynamics, as described later in Equation (9), may lead to a null measure over the space $\mathbb{X}$ when considering some combination of observations and controls which are in contradiction with each other. As we want to be able to compose belief dynamics, we combine $\Delta(\mathbb{X})$ and the null measure over $\mathbb{X}$ as follows.

We introduce an extended state set $\overline{\mathbb{X}}$, obtained as the union of the original set $\mathbb{X}$ with an extra element, denoted by ∂ ($\partial \notin \mathbb{X}$), which is used as the support of the null measure over $\mathbb{X}$.

$$\overline{\mathbb{X}} = \mathbb{X} \cup \{\partial\} \ . \tag{3}$$

We denote by $\mathbb{B}$ the subset of $\Delta(\overline{\mathbb{X}})$ defined by

$$\mathbb{B} = \Delta(\mathbb{X}) \cup \{\delta_\partial\} \ , \tag{4}$$

where we identify the set $\Delta(\mathbb{X})$ with $\{\mu \in \Delta(\overline{\mathbb{X}}) \,|\, \mathrm{supp}(\mu) \subset \mathbb{X}\}$ and where $\delta_\partial \in \Delta(\overline{\mathbb{X}})$ is the discrete probability measure on $\overline{\mathbb{X}}$ concentrated on ∂, that is $\delta_\partial(\{\partial\}) = 1$, and where the mapping "supp" is the support of a nonnegative measure. For any nonnegative measure μ on the finite set $\mathbb{Y}$, we have

$$\mathrm{supp}(\mu) = \big\{ y \in \mathbb{Y} \,\big|\, \mu(\{y\}) > 0 \big\} \ . \tag{5}$$

We call the probability measure δ_∂ the *cemetery* belief as we will see in Equation (9) that the belief dynamics, when reaching the belief state δ_∂, remains in δ_∂ forever. A probability measure $\nu \in \Delta(\overline{\mathbb{X}})$ is represented, in some equations, by the ordered pair $\big(\nu_{|\mathbb{X}}, \nu(\partial)\big)$, where $\nu_{|\mathbb{X}}$ is a nonnegative measure on the set $\mathbb{X}$ and $\nu(\partial) \in \mathbb{R}_+$.

Now that the set of beliefs $\mathbb{B}$ is defined, we present the beliefs dynamics.

Beliefs dynamics. In order to define the beliefs dynamics, we introduce, for each $t \in \mathcal{T} \setminus \{T\}$ two mappings, $Q_{t+1} : \mathbb{B} \times \mathbb{U} \times \mathbb{O} \to [0,1]$ and $\tau_t : \mathbb{B} \times \mathbb{U} \times \mathbb{O} \to \mathbb{B}$. They are defined using partial mappings, defined as follows.

Let $\mathbb{A}$, $\mathbb{D}$, $\mathbb{F}$ and $\mathbb{G}$ be sets. Let $g : \mathbb{A} \times \mathbb{D} \to \mathbb{F}, (a,d) \mapsto g(a,d)$ be a mapping. We denote by g^d the mapping

$$g^d : \mathbb{A} \to \mathbb{F} \ , \ \ a \mapsto g(a,d) \ , \tag{6}$$

i.e. the mapping $g(\cdot, d)$ obtained from g by setting its second variable to a fixed value $d \in \mathbb{D}$. When considering mappings with n inputs, we extend this notation to the last $n-1$ inputs using a Cartesian product over the last $n-1$ sets. For example, let $g : \mathbb{A} \times \mathbb{D} \times \mathbb{F} \to \mathbb{G}$. We denote by $g^{(d,f)} = g(\cdot, d, f)$ the mapping $g^{(d,f)} : \mathbb{A} \to \mathbb{G}, a \mapsto g(a,d,f)$.

Now, the mapping Q_{t+1} gives the probability of observing o at time $t+1$ when applying control u on the dynamical system when considering belief b at time t, and is given by

$$\forall t \in \mathcal{T} \setminus \{T\} \ , \ \ Q_{t+1} : (b,u,o) \ni \mathbb{B} \times \mathbb{U} \times \mathbb{O} \mapsto b\big((h^u_{t+1} \circ f^u_t)^{-1}(o)\big) \ , \tag{7}$$

where $f^u_t(\cdot)$ and $h^u_t(\cdot)$ are partial mapping that follow the notation defined in Equation (6):

$$\forall u \in \mathbb{U}, \quad f^u_t : \mathbb{X} \to \mathbb{X} \ , \ \ x \mapsto f_t(x,u) \ , \quad \text{and } \forall u \in \mathbb{U}, \quad h^u_t : \mathbb{X} \to \mathbb{O} \ , \ \ x \mapsto h_t(x,u) \ ,$$

and where $b\big((h^u_{t+1} \circ f^u_t)^{-1}(o)\big)$ is the probability of the set $(h^u_{t+1} \circ f^u_t)^{-1}(o)$ with respect to the probability distribution b. Note that, we always have that

$$Q_{t+1}(\delta_\partial, u, o) = \delta_\partial\big((h^u_{t+1} \circ f^u_t)^{-1}(o)\big) = 0 \ , \tag{8}$$

as $(h^u_{t+1} \circ f^u_t)^{-1}(o)$ is always a subset of $\mathbb{X}$ and thus has a null intersection with $\{\partial\}$.

For all time $t \in \mathcal{T} \setminus \{T\}$, the mapping τ_t gives the evolution of the beliefs when applying control u on the dynamical system when considering belief b at time t and observing o at time $t+1$, and is given by

$$\forall y \in \mathbb{X} \ , \ \ \tau_t(b,u,o)(y) = \begin{cases} \dfrac{b\big((f^u_t)^{-1}(y)\big)}{Q_{t+1}(b,u,o)} & \text{if } Q_{t+1}(b,u,o) \neq 0, \text{ and } y \in \big(h^u_{t+1}\big)^{-1}(o) \ , \\ 0 & \text{otherwise,} \end{cases} \tag{9a}$$

$$\tau_t(b,u,o)(\partial) = 1 - \tau_t(b,u,o)(\mathbb{X}) \ . \tag{9b}$$

Hence, δ_∂ is used as a last resort belief, which appears when it is not possible to observe o after applying control u to any state of the support of belief b. Indeed, δ_∂ is used to ensure that the mappings τ_t are well defined for all beliefs, controls and observations.

Using the sequences of mappings $\{Q_t\}_{t \in \mathcal{T} \setminus \{0\}}$ and $\{\tau_t\}_{t \in \mathcal{T} \setminus \{T\}}$, we have a properly defined belief-MDP, which can be solved by Dynamic Programming.

3.1.2 Dynamic Programming Equations for Det-Pomdp

In the case of POMDP (without constraints on the controls), Dynamic Programming equations with beliefs as new states were first given in (Åström, 1965). More general cases (still without explicit constraints on the controls) are treated in Bertsekas and Shreve (1978, Chapter 10) and in Bertsekas (2000, Chapter 4). Dynamic Programming Equations for DET-POMDP can be obtained as a special case of Dynamic Programming for POMDP. They are given in Equations (10a) and (10b) together with the expression of the beliefs dynamics $\{\tau_t\}_{t \in \mathcal{T} \setminus \{T\}}$ (see Equation (9)) in the case where there are no constraints on the controls in (Littman, 1996). In (Bertsekas and Shreve, 1978) the proof that beliefs are *statistics sufficient for controls* was made for POMDPs without any admissibility constraint. We thus cannot directly apply this result on Problem (2), as it contains Constraint (2f). We extend the classical results by (Bertsekas and Shreve, 1978) in Proposition 1 in order to tackle such constraints. We identify an admissibility set for beliefs of the form $\mathcal{U}^{\mathrm{b,ad}}(b) = \bigcap_{x \in \mathrm{supp}(b)} \mathcal{U}^{ad}(x)$. Note that we use an upper index b to distinguish admissibility sets for beliefs from admissibility sets for states. Also note that, as far as we know, the first Dynamic Programming equations using such sets $\mathcal{U}^{\mathrm{b,ad}}(b)$ were given in (Geffner and Bonet, 1998, §5) with no explicit proof.

Proposition 1. *Consider a* Det-Pomdp *optimization problem given by Problem* (2) *which satisfies the finite sets Assumption 1. Let* $\mathbb{B} = \Delta(\mathbb{X}) \cup \{\delta_\partial\}$*, as defined in Equation* (4) *and consider the sequence of value functions* $(V_t : \mathbb{B} \to \mathbb{R} \cup \{+\infty\})_{t\in\mathcal{T}}$ *defined by the following backward induction. First, for all* $t \in \mathcal{T}$*, we have that* $V_t(\delta_\partial) = 0$*. Second, we have that*

$$V_T : b \in \Delta(\mathbb{X}) \mapsto \sum_{x\in\mathbb{X}} b(x)\mathcal{K}(x) \ , \tag{10a}$$

$$V_t : b \in \Delta(\mathbb{X}) \mapsto \min_{u\in\mathcal{U}_t^{\mathrm{b,ad}}(b)} \Big(\sum_{x\in\mathbb{X}} b(x)\mathcal{L}_t(x,u) + \sum_{o\in\mathbb{O}} Q_{t+1}(b,u,o)V_{t+1}\big(\tau_t(b,u,o)\big)\Big) \ , \tag{10b}$$

where $\mathcal{U}_t^{\mathrm{b,ad}}(b) = \bigcap_{x\in\mathrm{supp}(b)} \mathcal{U}_t^{ad}(x)$*.*

Then, the optimal value of Problem (2) *and the value of the function* V_0 *at the initial belief* b_0 *are equal, that is,* $V_0(b_0) = \mathcal{V}^\star(b_0)$*. Moreover, a policy* $\pi = (\pi_0, \dots, \pi_{T-1})$*, defined by a sequence of mappings* $\pi_t : \mathbb{B} \to \mathbb{U}$*, which minimizes the right-hand side of Equation* (10b) *for each* b *and* t *is an optimal policy of Problem* (2)*: the controls given by* $\boldsymbol{U}_t = \pi_t(\boldsymbol{B}_t)$ *(where* $\boldsymbol{B}_t$ *is computed thanks to the recursion* $\boldsymbol{B}_{t+1} = \tau_t(\boldsymbol{B}_t, \boldsymbol{U}_t, \boldsymbol{O}_{t+1})$*, with* $\boldsymbol{B}_0 = b_0$*) are optimal controls of Problem* (2)*.*

Proof. We present a sketch of proof of Proposition 1.

First, we rewrite Problem (2) as an equivalent problem, without constraint (2f) by adding characteristic functions of the constraints to the instantaneous costs. The equivalent problem then follows the framework of (Bertsekas and Shreve, 1978).

Second, we apply the results of (Bertsekas and Shreve, 1978) to the reformulated problem, and obtain associated Dynamic Programming equations.

Third, the Dynamic Programming equations which solve the equivalent problem are equivalent to Equations (10) presented in Proposition 1, thus concluding that Equation (10) gives the solution of Problem (2) as formulated in Proposition 1. This step is a bit technical, but is otherwise straightforward and does not present any major difficulty. □

Now that we have presented Dynamic Programming equations on beliefs, we present the complexity of Dynamic Programming.

3.2 Dynamic Programming complexity for Det-Pomdps

According to Proposition 1, we can solve Problem (2) by computing $V_0(b_0)$ by means of Equations (10). Solving Dynamic Programming equations (10) implies that we are able to numerically evaluate the value functions at each reachable belief starting from b_0. Thus, we introduce the subsets of reachable beliefs starting from b_0. We start by formally defining the set of reachable beliefs, before we present our first complexity result on Dynamic Programming for Det-Pomdp.

The set of reachable beliefs $\mathbb{B}^{\mathrm{R},\mathcal{D}}$ is defined as follows. Note that we use the upper index $\mathcal{D}$ to recall that we consider the set of reachable beliefs of a Det-Pomdp defined by the data tuple $\mathcal{D}$, in Equation (1), whereas the upper index R stands for reachable.

Definition 2. *Let* $b_0 \in \Delta(\mathbb{X})$ *be given and consider the sequence* $\{\mathbb{B}_t^{\mathrm{R},\mathcal{D}}\}_{t\in\mathcal{T}}$ *of subsets of the set of beliefs* $\mathbb{B} = \Delta(\mathbb{X}) \cup \{\delta_\partial\}$ *defined by the induction*

$$\mathbb{B}_0^{\mathrm{R},\mathcal{D}}(b_0) = \{b_0\} \quad \textit{and} \quad \forall t \in \mathcal{T} \setminus \{T\} \ , \ \mathbb{B}_{t+1}^{\mathrm{R},\mathcal{D}}(b_0) = \tau_t\big(\mathbb{B}_t^{\mathrm{R},\mathcal{D}}(b_0), \mathbb{U}, \mathbb{O}\big) \ , \tag{11}$$

where τ_t *is defined in Equation* (9)*. For any* $t \in \mathcal{T}$*, the subset* $\mathbb{B}_t^{\mathrm{R},\mathcal{D}}(b_0) \subset \mathbb{B}$ *is called the set of* reachable beliefs a time t *starting from initial belief* b_0*.*

Moreover, we denote by $\mathbb{B}_{[\![t,t']\!]}^{\mathrm{R},\mathcal{D}}(b_0)$ *the union, for* t'' *in the time interval* $[\![t,t']\!]$*,* $t < t'$*, of the reachable beliefs at time* t'' *starting from the initial belief* $b_0 \in \Delta(\mathbb{X})$*, that is,*

$$\forall (t,t') \in \mathcal{T}^2 \ , \ t < t' \ , \ \mathbb{B}_{[\![t,t']\!]}^{\mathrm{R},\mathcal{D}}(b_0) = \bigcup_{t''=t}^{t'} \mathbb{B}_{t''}^{\mathrm{R},\mathcal{D}}(b_0) \ . \tag{12}$$

The set $\mathbb{B}_{[\![1,T]\!]}^{\mathrm{R},\mathcal{D}}(b_0)$ *is called the* set of reachable beliefs *from the initial belief* b_0*.*

Note that, under Assumption 1, the set $\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)$ is finite.

We now present a classical complexity result for Dynamic Programming algorithm (which we call Dp Algorithm in the rest of this paper).

Proposition 3. *Consider a* Det-Pomdp *optimization problem given by Problem* (2) *which satisfies the finite sets Assumption 1. Let $b_0 \in \Delta(\mathbb{X})$. Then, a standard* Dp *Algorithm (numerically) solves Problem* (2)*, and its complexity is $O(|\mathcal{T}||\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)||\mathbb{U}||\mathbb{O}|)$, where the set of reachable beliefs $\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)$ is defined in Equation* (12)*.*

Proof. First, as we consider that Assumption 1 holds, note that $\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)$ is finite and we can apply Proposition 1 on Problem (2). We hence solve Problem (2) by computing value functions given by Equations (10).

For a given time $t \in \mathcal{T} \setminus \{T\}$ and reachable belief $b \in \mathbb{B}^{\mathrm{R},\mathcal{D}}_t(b_0)$, we compute the value function V_t by evaluating the next value for each control $u \in \mathbb{U}$ and each resulting observations. We hence need $\sum_{t\in\mathcal{T}} |\mathbb{B}^{\mathrm{R},\mathcal{D}}_t(b_0)||\mathbb{U}||\mathbb{O}|$ operations to solve Problem (2). Then, since for all time $t \in \mathcal{T}$, $t > 0$, $\mathbb{B}^{\mathrm{R},\mathcal{D}}_t(b_0) \subset \mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)$ (see Equation (12)), we have $|\mathbb{B}^{\mathrm{R},\mathcal{D}}_t(b_0)| \leq |\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)|$. Moreover, we also have $\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0) \neq \emptyset$ (there is always at least one belief in $\mathbb{B}^{\mathrm{R},\mathcal{D}}_1(b_0)$, as for a given control $u \in \mathbb{U}$ and an observation $o \in \mathbb{O}$, $\tau_0(b_0, u, o) \in \mathbb{B}^{\mathrm{R},\mathcal{D}}_1(b_0) \subset \mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)$) and $\mathbb{B}^{\mathrm{R},\mathcal{D}}_0(b_0) = \{b_0\}$, hence $|\mathbb{B}^{\mathrm{R},\mathcal{D}}_0(b_0)| \leq |\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)|$.

Hence, $\sum_{t\in\mathcal{T}} |\mathbb{B}^{\mathrm{R},\mathcal{D}}_t(b_0)||\mathbb{U}||\mathbb{O}| \leq |\mathcal{T}||\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)||\mathbb{U}||\mathbb{O}|$, and thus we can solve Problem (2) in $O(|\mathcal{T}||\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)||\mathbb{U}||\mathbb{O}|)$ operations. □

In order to apply Proposition 3 on Problem (2) and to get complexity bounds on the Dp Algorithm, we now study the set of reachable beliefs $\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)$, more specifically, we give bounds on its cardinality.

Theorem 4. *Consider a* Det-Pomdp *optimization problem given by Problem* (2) *which satisfies the finite sets Assumption 1, and such that $|\mathbb{U}| > 1$. For all initial belief $b_0 \in \Delta(\mathbb{X})$, the cardinality of the set of reachable beliefs starting from b_0, defined in Equation* (12)*, satisfies the following bound*

$$\left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)\right| \leq \min\left((1+|\mathbb{X}|)^{|\mathrm{supp}(b_0)|}, 1 + |\mathrm{supp}(b_0)||\mathbb{U}|^{|\mathcal{T}|}\right) . \tag{13}$$

Proof. A sketch of proof is postponed to §3.3, as it relies on a new representation of the belief dynamics presented in §3.3. The complete proof can be found in Appendix §A.1.3. □

In Theorem 4, we gave a bound on the cardinality of the set $\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,t]\!]}(b_0)$ which improves a previous result we now recall.

Littman presents in (Littman, 1996, Lemma 6.1) a bound on the set of reachable beliefs starting from belief $b_0 \in \Delta(\mathbb{X})$:

$$\forall t \in \mathcal{T} , \ \left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![0,t]\!]}(b_0)\right| \leq (1+|\mathbb{X}|)^{|\mathbb{X}|} . \tag{14}$$

Equation (13) is an improvement on the bound given in Equation (14) which takes into account the support of the initial belief b_0: indeed, as $b_0 \in \Delta(\mathbb{X})$ and $|\mathrm{supp}(b_0)| \leq |\mathbb{X}|$, Equation (13) is tighter than Equation (14).

Using Equation (13), we obtain that the number of reachable beliefs of a Det-Pomdp is finite even when considering the case of an infinite horizon. Indeed, the first inequality in Equation (13) is well defined even in the infinite horizon case.

A direct consequence of Proposition 3 and Theorem 4 is that the complexity of the Dp Algorithm is $O\big(|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)||\mathcal{T}||\mathbb{U}||\mathbb{O}|\big)$, i.e. in $O\left(\min\left((1+|\mathbb{X}|)^{|\mathrm{supp}(b_0)|}, 1 + |\mathrm{supp}(b_0)||\mathbb{U}|^{|\mathcal{T}|}\right)|\mathcal{T}||\mathbb{U}||\mathbb{O}|\right)$.

Remark 5. *As a side note, we can remark that we could also use Theorem 4 to characterize the complexity of a general* Pomdp*. Indeed, we can reformulate any finite* Pomdp *with independent noises on the dynamics $\{w_t\}_{t\in\mathcal{T}\setminus\{T\}}$ and independent noises on the observations $\{v_t\}_{t\in\mathcal{T}}$ and admissibility constraints of the form $\mathcal{U}^{ad} : \mathbb{X} \rightrightarrows \mathbb{U}$ as a finite* Det-Pomdp*. To do so, we expand the state of the* Pomdp *with the realization of all noises. We model the problem as though the realization of the noises are predetermined, but the decision maker does not know the noises in advance. We then obtain a* Det-Pomdp*, with states $\mathbb{X}'$, controls $\mathbb{U}$ and observations $\mathbb{O}$. However, such reformulation leads to a drastic increase in the dimension of the state and the cardinality of the support of the initial belief. Indeed, the initial*

belief contains all possible values of the initial state and all the possible realizations of noises, i.e. its cardinality is multiplied by a factor $|\mathbb{V}|^{T+1} \times |\mathbb{W}|^{T}$*, with* $\mathbb{V}$ *the set of noises on the observation and* $\mathbb{W}$ *the set of noises on the dynamics. Hence, we are doubly penalized when considering the bound presented in Theorem 4: we both increase* $|\mathbb{X}|$ *and* $|\mathrm{supp}(b_0)|$*. This reinforces the point on the difficulty of solving* POMDPs *as even the ones with simple structures are far more difficult than similar sized* DET-POMDPs*.*

3.3 Belief dynamics as pushforward measures

Here, we expose another representation of the beliefs evolution functions $\{\tau_t\}_{t\in\mathcal{T}\setminus\{T\}}$ defined in Equation (9), used in the proof of Theorem 4. First, we recall the notion of *pushforward measures* when considering finite sets. Second, we introduce the mappings necessary for the new representation. We then present in Lemma 7 the representation of the belief dynamics as pushforward measures.

Definition 6. *Consider two finite sets* $\mathbb{A}$ *and* $\mathbb{D}$ *and a mapping* $h : \mathbb{A} \to \mathbb{D}$*. The* pushforward measure *(or the* image-measure*) of a probability measure* $\mu \in \Delta(\mathbb{A})$ *on the set* $\mathbb{A}$ *by the mapping* h *is the probability measure* $h_\star\mu \in \Delta(\mathbb{D})$ *on the set* $\mathbb{D}$ *defined by*

$$\forall d \in \mathbb{D}\ ,\ \ (h_\star\mu)(d) = \mu\big(h^{-1}(d)\big) = \sum_{a\in\mathbb{A}, h(a)=d} \mu(a)\ . \tag{15}$$

We also denote by $h_\star$ *the mapping from* $\Delta(\mathbb{A})$ *to* $\Delta(\mathbb{D})$ *such that* $h_\star(\mu) = h_\star\mu$*.*

Before presenting Lemma 7, we first introduce some mappings: $F_t^{u,o}$, and $\mathcal{R}$. For each pair $(u, o) \in \mathbb{U} \times \mathbb{O}$, and each $t \in \mathcal{T} \setminus \{T\}$, we denote by $F_t^{u,o}$ the self-mapping on the extended state set $\overline{\mathbb{X}} = \mathbb{X} \cup \{\partial\}$ (defined in Equation (3)), defined by:

$$F_t^{u,o} : \overline{\mathbb{X}} \to \overline{\mathbb{X}}\ ,\ \ \overline{x} \mapsto \begin{cases} f_t^u(\overline{x}) & \text{if} \quad \overline{x} \neq \partial \text{ and } f_t^u(\overline{x}) \in \big(h_{t+1}^u\big)^{-1}(o)\ , \\ \partial & \text{otherwise.} \end{cases} \tag{16}$$

The mapping $F_t^{u,o}$ hence applies the dynamics f_t, as defined in Problem (2), given control u, and only keeps the resulting state if it is consistent with observation o. Meanwhile, the renormalization mapping $\mathcal{R} : \Delta(\overline{\mathbb{X}}) \to \Delta(\overline{\mathbb{X}})$ is defined by

$$\mathcal{R} : \nu \in \Delta(\overline{\mathbb{X}}) \mapsto \begin{cases} \big(\frac{1}{\nu(\mathbb{X})}\nu_{|\mathbb{X}}, 0\big) & \text{if } \nu(\mathbb{X}) \neq 0\ , \\ \delta_\partial & \text{if } \nu(\mathbb{X}) = 0\ . \end{cases} \tag{17}$$

We now express the belief dynamics as pushforward measures.

Lemma 7. *Let* $(u, o) \in \mathbb{U} \times \mathbb{O}$ *be given, and let* $t \in \mathcal{T} \setminus \{T\}$*. We have that*

$$\forall b \in \mathbb{B}\ ,\ \ \tau_t(b, u, o) = \mathcal{R} \circ (F_t^{u,o})_\star(b)\ , \tag{18}$$

where the pushforward $(F_t^{u,o})_\star(b)$ *follows Notation* (15)*.*

Proof. The proof is detailed in Appendix A.1. □

This new representation is of interest as, for all time $t \in \mathcal{T} \setminus \{T\}$, the composition of belief dynamics τ_t is given by the pushforward measure of the composition of mappings $F_t^{u,o}$ for the relevant pairs $(u, o) \in \mathbb{U} \times \mathbb{O}$. Indeed, when considering a composition of belief dynamics, we can factorize the renormalization mapping $\mathcal{R}$. We thus apply the renormalization mapping $\mathcal{R}$ to the composition of the pushforward measures, which is the pushforward measure of the composition of mappings $F_t^{u,o}$. There is therefore an equivalence between studying the composition for time $t \in \mathcal{T} \setminus \{T\}$ of the belief dynamics τ_t and the composition, for the relevant pairs $(u, o) \in \mathbb{U} \times \mathbb{O}$, of the mappings $F_t^{u,o}$. Notably, we use this representation to bound the cardinality of the set of reachable beliefs, and thus study the complexity of Dynamic Programming for DET-POMDP.

To do so, we introduce notations for sets and mappings.

Notation for sets and mappings. For any given sets $\mathbb{Y}$ and $\mathbb{V}$, we denote by $\mathbb{L}(\mathbb{Y};\mathbb{V}) = \mathbb{V}^{\mathbb{Y}}$ the set of mappings from $\mathbb{Y}$ to $\mathbb{V}$.

- For all $\mathbb{G} \subset \mathbb{L}(\mathbb{Y};\mathbb{V})$, $Y \subset \mathbb{Y}$, $B \subset \Delta(\mathbb{Y})$ and $b \in \Delta(\mathbb{Y})$ we introduce the notations $\mathbb{G}(Y)$, and $\mathbb{G}_\star(B)$, and $\mathbb{G}_\star(b)$ for the sets defined by

$$\mathbb{G}_\star(b) = \mathbb{G}_\star(\{b\}) \subset \Delta(\mathbb{V}) \ . \tag{19a}$$

- Given two subsets $\mathbb{G}'$ and $\mathbb{G}''$ of $\mathbb{L}(\mathbb{Y};\mathbb{Y})$ we introduce the subset $\mathbb{G}' \circ \mathbb{G}''$ defined by

$$\mathbb{G}' \circ \mathbb{G}'' = \left\{ g' \circ g'' \,\middle|\, g' \in \mathbb{G}' \text{ and } g'' \in \mathbb{G}'' \right\} \subset \mathbb{L}(\mathbb{Y};\mathbb{Y}) \ . \tag{19b}$$

- For any sequence $\{\mathbb{G}_k\}_{k\in\mathbb{N}}$, with $\mathbb{G}_k \subset \mathbb{L}(\mathbb{Y};\mathbb{Y})$ for all $k \in \mathbb{N}$, we introduce for any $k \in \mathbb{N}$ the subsets $\mathbb{G}_{0:k}$ defined by

$$\forall k \in \mathbb{N} \ , \ \ \mathbb{G}_{0:k} = \mathbb{G}_k \circ \mathbb{G}_{k-1} \circ \cdots \circ \mathbb{G}_0 \subset \mathbb{L}(\mathbb{Y};\mathbb{Y}) \ . \tag{19c}$$

For a fixed value of $u \in \mathbb{U}$, and $o \in \mathbb{O}$, for all $t \in \mathcal{T} \setminus \{T\}$, we have obtained in Lemma 7 that $\tau_t(\cdot, u, o) = \mathcal{R} \circ (F_t^{u,o})_\star$. Now, for each $t \in \mathcal{T}$, we introduce the sets

$$\mathbb{T}_t^{\mathcal{D}} = \left\{ \tau_t(\cdot, u, o) \,\middle|\, u \in \mathbb{U}, o \in \mathbb{O} \right\} \subset \mathbb{L}(\mathbb{B};\mathbb{B}) \ , \tag{20}$$

$$\mathbb{F}_t^{\mathcal{D}} = \left\{ F_t^{u,o} \,\middle|\, u \in \mathbb{U}, o \in \mathbb{O} \right\} \subset \mathbb{L}(\overline{\mathbb{X}};\overline{\mathbb{X}}) \ , \tag{21}$$

$$\mathbb{F}^{\mathcal{D}} = \bigcup_{t\in\mathcal{T}\setminus\{T\}} \mathbb{F}_{0:t}^{\mathcal{D}} \ , \tag{22}$$

where the composition of sets of mapping is given by Notation (19b) and (19c). Note that[4] $\mathbb{F}_{0:t}^{\mathcal{D}} \neq \mathbb{F}_{[\![0,t]\!]}^{\mathcal{D}}$. Moreover, we call $\mathbb{F}^{\mathcal{D}}$, defined by Equation (22), the *set of pushforwards of the* Det-Pomdp defined by Equation (2).

Lemma 8. *Let $b_0 \in \Delta(\mathbb{X})$. We have that*

$$\forall t \in \mathcal{T} \setminus \{0\} \ , \ \ \mathbb{B}_t^{\mathrm{R},\mathcal{D}}(b_0) = \mathbb{T}_{0:t-1}^{\mathcal{D}}(b_0) = \mathcal{R} \circ (\mathbb{F}^{\mathcal{D}})_\star(b_0) \ , \tag{23}$$

$$\mathbb{T}^{\mathcal{D}} = \bigcup_{t\in\mathcal{T}\setminus\{T\}} \mathbb{T}_{0:t}^{\mathcal{D}} = \mathcal{R} \circ (\mathbb{F}_{0:t}^{\mathcal{D}})_\star \ , \tag{24}$$

i.e.

$$\mathbb{B}_{[\![1,T]\!]}^{\mathrm{R},\mathcal{D}}(b_0) = \bigcup_{t\in\mathcal{T}\setminus\{T\}} \mathbb{T}_{0:t}^{\mathcal{D}}(b_0) = \mathcal{R} \circ (\mathbb{F}^{\mathcal{D}})_\star(b_0) \ . \tag{25}$$

Proof. The proof is detailed in Appendix A.1. □

Lemmata 7 and 8 are illustrated in Figures 1 and 2. A direct application of Lemma 8 is that there is an equivalence between studying the cardinality of $\mathbb{B}_{[\![1,T]\!]}^{\mathrm{R},\mathcal{D}}(b_0)$ and studying the cardinality of $(\mathbb{F}^{\mathcal{D}})_\star(b_0)$.

We now present the postponed sketch of proof of Theorem 4. A detailed proof can be found in Appendix §A.1.3.

Sketch of proof of Theorem 4. Let $b_0 \in \Delta(\mathbb{X})$ be given.

By Lemma 8, we have that $\mathbb{B}_{[\![1,T]\!]}^{\mathrm{R},\mathcal{D}}(b_0) = \mathcal{R} \circ (\mathbb{F}^{\mathcal{D}})_\star(b_0)$.

The first inequality $|\mathbb{B}_{[\![1,T]\!]}^{\mathrm{R},\mathcal{D}}(b_0)| \leq (1 + |\mathbb{X}|)^{|\mathrm{supp}(b_0)|}$ comes from the fact that $\left|(\mathbb{F}^{\mathcal{D}})_\star(b_0)\right|$ is bounded by the number of mappings from $\mathrm{supp}(b_0)$ to $\overline{\mathbb{X}}$, as shown in Lemma 21.

Meanwhile, the second inequality $\left|\mathbb{B}_{[\![1,T]\!]}^{\mathrm{R},\mathcal{D}}(b_0)\right| \leq 1 + |\mathrm{supp}(b_0)||\mathbb{U}|^{|\mathcal{T}|}$ comes from the fact that, for all time and controls $(t, u) \in \mathcal{T}\setminus\{T\}\times\mathbb{U}$, and for any belief $b \in \Delta(\mathbb{X})$, we have that $\sum_{o\in\mathbb{O}} \left|\mathrm{supp}\big((F_t^{u,o})_\star b\big)\right| \leq \left|\mathrm{supp}\big(b\big)\right|$ by Lemma 24. Therefore, for a given sequence of controls $u_{0:t} \in \mathbb{U}^{t+1}$, there can be at most $|\mathrm{supp}(b_0)|$ resulting beliefs (see Lemma 23). As there are at most $|\mathbb{U}|^{|\mathcal{T}|}$ such sequences $u_{0:t}$, $t \in \mathcal{T} \setminus \{T\}$, this leads to $\left|\mathbb{B}_{[\![1,T]\!]}^{\mathrm{R},\mathcal{D}}(b_0)\right| \leq 1 + |\mathrm{supp}(b_0)||\mathbb{U}|^{|\mathcal{T}|}$. □

[4] $\mathbb{F}_{0:t}^{\mathcal{D}}$ is the set of compositions of mappings $F_{t'}^{u,o}$ from time $t' = 0$ to time $t' = t$ for all controls $u \in \mathbb{U}$ and observation $o \in \mathbb{O}$, while the set $\mathbb{F}_{[\![0,t]\!]}^{\mathcal{D}}$ is the set of all mappings $F_t^{u,o}$ between time 0 and time t.

$$
\begin{tikzcd}
\Delta(\mathbb{X}) \arrow[r, "\tau_t^{u,o}"] \arrow[d, "b", "(b,0)"'] & \mathbb{B} = \Delta(\mathbb{X}) \cup \{\delta_\partial\} \\
\Delta(\overline{\mathbb{X}}) \arrow[r, "\left(F_t^{u,o}\right)_\star"'] & \Delta(\overline{\mathbb{X}}) \arrow[u, "\mathcal{R}\big((b'_{|\mathbb{X}}, \underbrace{b'(\partial)}_{\in\mathbb{R}})\big)"']
\end{tikzcd}
$$

Figure 1: Illustration of the beliefs dynamics as pushforward measures

$$
\begin{tikzcd}
\Delta(\mathbb{X}) \arrow[rr, "\tau_{t+1}^{u',o'} \circ \tau_t^{u,o}"] \arrow[d, "b"] & & \mathbb{B} = \Delta(\mathbb{X}) \cup \{\delta_\partial\} \\
\Delta(\overline{\mathbb{X}}) \arrow[r, "\left(F_t^{u,o}\right)_\star"'] & \Delta(\overline{\mathbb{X}}) \arrow[r, "\left(F_{t+1}^{u',o'}\right)_\star"'] & \Delta(\overline{\mathbb{X}}) \arrow[u, "\mathcal{R}"']
\end{tikzcd}
$$

$$
\underbrace{\Delta(\overline{\mathbb{X}}) \longrightarrow \Delta(\overline{\mathbb{X}}) \longrightarrow \Delta(\overline{\mathbb{X}})}_{= \left(\underbrace{F_{t+1}^{u',o'} \circ F_t^{u,o}}_{\in \overline{\mathbb{X}}^{\overline{\mathbb{X}}}}\right)_\star}
$$

Figure 2: Illustration of the composition of belief dynamics as pushforward measures.

We now present the subclass of Separated Deterministic Partially Observed Markov Decision Processes (Separated Det-Pomdp), which is simpler than Det-Pomdp.

4 Separated Det-Pomdp and complexity

In this section, we introduce a subclass of Det-Pomdps: Separated Det-Pomdps. First, we define this subclass in §4.1. Second, in §4.2, we present an improved bound on the cardinality of the set of reachable beliefs for Separated Det-Pomdps compared to Det-Pomdps. Third, in §4.3, we show that the improved bound is tight.

4.1 Definition of (∂)-separated mapping set and Separated Det-Pomdp

Let us first define separated mapping sets.

Definition 9. *Let $\mathbb{Y}_1$ and $\mathbb{Y}_2$ be two given sets. A set $\mathbb{G} \subset \mathbb{L}(\mathbb{Y}_1; \mathbb{Y}_2)$ of mappings from $\mathbb{Y}_1$ to $\mathbb{Y}_2$ is called a* separated mapping set *if*

$$
\forall (g_1, g_2) \in \mathbb{G}^2 \ , \ \ \forall y \in \mathbb{Y}_1 \ , \ \ \Big(g_1(y) = g_2(y) \Rightarrow g_1 = g_2\Big) \ .
$$

A separated mapping set $\mathbb{G} \subset \mathbb{L}(\mathbb{Y}_1; \mathbb{Y}_2)$ is hence a set of mappings where all pairs of mappings are either different everywhere, or equal everywhere. Otherwise stated, all the evaluation mappings on set $\mathbb{G}$ (i.e. the mappings $\mathbb{G} \to \mathbb{Y}_2, g \mapsto g(y)$, for a fixed $y \in \mathbb{Y}_1$) are injective for all $y \in \mathbb{Y}_1$. For example, let $\mathbb{Y}_1 = [\![1, n]\!]$ and $\mathbb{Y}_2 = \mathbb{R}$. Then, $\mathbb{G} \subset \mathbb{R}^{\mathbb{Y}_1}$ is identified with $G \subset \mathbb{R}^n$, and $\mathbb{G}$ is a Separated mapping set if and only if the projections of G along each axis are injective.

In the special case where $\mathbb{Y}_1 = \mathbb{Y}_2 = \overline{\mathbb{X}}$, with the extended set $\overline{\mathbb{X}} = \mathbb{X} \cup \{\partial\}$ defined in Equation (3), we want to extend the above notion of separated mapping set to tackle the added point ∂ in a specific way. We thus introduce the notion of (∂)-separation for a pair of self-mappings on the set $\overline{\mathbb{X}}$ and the notion of (∂)-separated mapping set.

Definition 10. *Let $\overline{\mathbb{X}} = \mathbb{X} \cup \{\partial\}$. A pair $(g_1, g_2) \in \mathbb{L}(\overline{\mathbb{X}}; \overline{\mathbb{X}})$ of self-mappings on the set $\overline{\mathbb{X}}$ is (∂)-separated if the restriction of the pair (g_1, g_2) to the set $g_1^{-1}(\mathbb{X}) \cap (g_2)^{-1}(\mathbb{X})$ is separated. Moreover, a set $\mathbb{G}$ of self-mappings on the set $\overline{\mathbb{X}}$ is called a (∂)-separated mapping set if all pairs of mappings $(g_1, g_2) \in \mathbb{G}^2$ are (∂)-separated.*

Definition 11. *A Separated* Det-Pomdp *is a* Det-Pomdp *such that the set of pushforward of the* Det-Pomdp $\mathbb{F}^{\mathcal{D}}$*, defined in Equation* (22)*, is a (∂)-separated mapping set.*

Otherwise stated, for a Separated Det-Pomdp, if two sequences of controls lead to the same state when starting in state x, then applying the two sequences of controls to another state x' either leads to the same state, or at least one sequence of controls leads to the cemetery point ∂.

We now present a link between the notion of separated mapping set and the notion of Separated Det-Pomdp. This allows us to propose a sufficient condition in order to ensure that a Det-Pomdp is a Separated Det-Pomdp.

Proposition 12. *If the set* $\bigcup_{t\in\mathcal{T}\setminus\{T\}} f_{0:t}^{\mathbb{U}^{t+1}} = \{f_{0:t}^{u_{0:t}} \,|\, \forall t \in \mathcal{T}\setminus\{T\}, \forall u_{0:t} \in \mathbb{U}^{t+1}\}$ *of the composition of the evolution functions of Problem* (2) *is a separated mapping set, as defined if Definition 9, then Problem* (2) *is a Separated* Det-Pomdp.

Proof. The proof of Proposition 12 is a direct consequence of Corollary 26. The detailed proof is found in Appendix A.2. □

Note that the observation mappings $\{h_t\}_{t\in\mathcal{T}\setminus\{T\}}$ do not play any role in Proposition 12.

Now that we have defined the subclass of Separated Det-Pomdps, we present a bound on the cardinality of the set of reachable beliefs for this subclass.

4.2 Complexity analysis of Separated Det-Pomdp

We now present the main interest of Separated Det-Pomdp when compared to Det-Pomdp, namely that the bound on cardinality of the set of reachable beliefs is lowered from $(1+|\mathbb{X}|)^{|\mathrm{supp}(b_0)|}$ to $1+\left(2^{|\mathrm{supp}(b_0)|}-|\mathrm{supp}(b_0)|\right)|\mathbb{X}|$.

Theorem 13. *Consider a Separated* Det-Pomdp *optimization problem given by Problem* (2) *which satisfies the finite sets Assumption 1. For all initial belief* $b_0 \in \Delta(\mathbb{X})$*, the cardinality of the set* $\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)$ *of reachable beliefs starting from* b_0 *satisfies the following bound*

$$\left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)\right| \leq 1+\left(2^{|\mathrm{supp}(b_0)|}-|\mathrm{supp}(b_0)|\right)|\mathbb{X}| \,. \tag{26}$$

Proof. The proof is detailed in Appendix A.2. □

We have therefore an improved complexity of the Dp Algorithm for Separated Det-Pomdp compared with standard Det-Pomdp.

Corollary 14. *Consider a Separated* Det-Pomdp *optimization problem given by Problem* (2) *which satisfies the finite sets Assumption 1. Then, the* Dp *Algorithm numerically solves Problem* (2) *by Dynamic Programming and its complexity is*

$$O\left(\min\left(1+\left(2^{|\mathrm{supp}(b_0)|}-|\mathrm{supp}(b_0)|\right)|\mathbb{X}|, 1+|\mathrm{supp}(b_0)||\mathbb{U}|^{|\mathcal{T}|}\right)|\mathcal{T}||\mathbb{U}||\mathbb{O}|\right) \,.$$

Proof. By Proposition 3, the Dp Algorithm solves Problem (2)and its complexity is $O\left(|\mathcal{T}||\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)||\mathbb{U}||\mathbb{O}|\right)$. Then, by Theorem 13, we have $\left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)\right| \leq 1+\left(2^{|\mathrm{supp}(b_0)|}-|\mathrm{supp}(b_0)|\right)|\mathbb{X}|$ and, by Theorem 4, we have that $\left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)\right| \leq 1+|\mathrm{supp}(b_0)||\mathbb{U}|^{|\mathcal{T}|}$. □

As the bound presented in Theorem 13 depends on the states that can be reached when starting from states in the support of the initial belief, we can obviously improve the bound when the support of the belief belongs to a subset of $\mathbb{X}$ stable by the dynamics $\{f_t\}_{t\in\mathcal{T}}$.

Remark 15. *Assuming that Problem* (2) *is a Separated* Det-Pomdp*, that Assumption 1 holds, that* $|\mathrm{supp}(b_0)| > 1$*, that the evolution functions* $\{f_t\}_{t\in\mathcal{T}\setminus\{T\}}$ *of Problem* (2) *satisfy the property that there exists a subset* $A \subset \mathbb{X}$ *such that, for all time* $t \in \mathcal{T}\setminus\{T\}$*,* $f_t(A,\mathbb{U}) \subset A$*. Assume that* $\mathrm{supp}(b_0) \subset A$*. Then the bound presented in Theorem 13 can be improved as*

$$\left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)\right| \leq 1+\left(2^{|\mathrm{supp}(b_0)|}-|\mathrm{supp}(b_0)|\right)|A| \,. \tag{27}$$

Now that we have a better bound than with non-separated Det-Pomdps, the question is whether it is tight or not. We now show that it is.

4.3 Existence of Separated Det-Pomdps with tight bound

In Theorem 13, we have given an improved bound on the cardinality of the set of reachable beliefs for Separated Det-Pomdp compared with standard Det-Pomdp. We now prove that the bound is tight.

Proposition 16. *There exist a Separated* Det-Pomdp *such that equality is obtained in Equation* (26)*, that is,*

$$\left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)\right| = 1 + \left(2^{|\mathrm{supp}(b_0)|} - |\mathrm{supp}(b_0)|\right)|\mathbb{X}| \ . \tag{28}$$

Proof. We exhibit a simple Separated Det-Pomdp for which the set of reachable beliefs $\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)$ satisfies Equation (28). Following the framework of §2, let $\mathbb{X} = \{x_1, x_2, x_3\}$ consists of three distinct states, $\mathbb{O} = \{\bar{o}_1, \bar{o}_2\}$ of two distinct observations, and $\mathbb{U} = \{\bar{u}_1, \bar{u}_2\}$ of two distinct controls. The evolution functions are defined as $\forall x \in \mathbb{X}$, $f(x, \bar{u}_1) = x$, and $\forall i \in \{1, 2, 3\}$, $f(x_i, \bar{u}_2) = x_{\mathrm{mod}(i,3)+1}$, where $\mathrm{mod}(i, 3)$ is the remainder of the euclidean division of i by 3. Finally, the observation mapping is given by $h(x, u) = \begin{cases} \bar{o}_2 & \text{if } x = x_3 \text{ and } u = \bar{u}_1 \ , \\ \bar{o}_1 & \text{otherwise} \ . \end{cases}$.

We show in Figure 3 the mappings $F^{(u,o)}$ defined in Equation (16) for this simple case, and we illustrate the dynamics and observation functions in Figure 4.

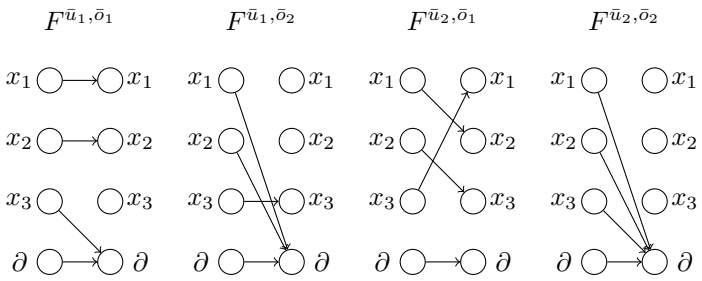

Figure 3: Representation of the $F^{(u,o)}$ mappings in the case of §4.3

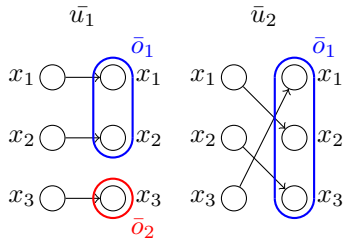

Figure 4: Representation of the dynamics and the observations depending on the control of the case of §4.3

By adding a cost function $\mathcal{L}$, a horizon $T > 0$ and admissibility constraints $\mathcal{U}^{ad} : x \rightrightarrows \mathbb{U}$, the resulting problem has all the ingredients of a Det-Pomdp (as presented in §2), where Assumption 1 holds.

We now prove that the resulting Det-Pomdp is a Separated Det-Pomdp. For that purpose, we enumerate all the possible results of the dynamics before applying Proposition 12. For this purpose, let us consider a sequence of controls $(u_1, \ldots, u_t) \in \mathbb{U}^t$. By denoting $f^{u_{1:t}}$ the compositions of dynamics (i.e. $f^{u_{1:t}}(x) = f^{u_t} \circ \cdots \circ f^{u_1}(x)$), we have, for all $i \in [\![1, 3]\!]$, $f^{u_{1:t}}(x_i) = x_{\mathrm{mod}(i+\gamma(u_{1:t})-1,3)+1}$, where γ is the function that counts the number of times $\bar{u}_2$ appears in a sequence of controls. The function γ is defined as $\gamma : \mathbb{U}^t \to \mathbb{N}$, $u_{1:t} \mapsto \left|\{u_i, i \in [\![1, t]\!] \,|\, u_i = \bar{u}_2\}\right|$.

The set $\{f^{u_{1:t}} \,|\, u_{1:t} \in \mathbb{U}^t\}$ is thus such that, for all sequences of controls $(u_{1:t}, u'_{1:t'}) \in \mathbb{U}^t \times \mathbb{U}^{t'}$, if there is a state $x \in \mathbb{X}$ such that $f^{u_{1:t}}(x) = f^{u'_{1:t'}}(x)$, then for any state $x' \in \mathbb{X}$, $f^{u_{1:t}}(x') = f^{u'_{1:t'}}(x')$. Hence, the set $\cup_{t \in \mathcal{T} \setminus \{T\}} f^{\mathbb{U}^{t+1}}_{0:t}$ is a separated mapping set. By Proposition 12, the optimization problem is hence a Separated Det-Pomdp.

We now chose an initial belief b_0 such that $\mathrm{supp}(b_0) = \{x_1, x_2\}$, for which we can compute explicitly the reachable beliefs. We can apply Theorem 13 with such initial belief. Therefore, according to Equation (26), there can be at most 7 reachable beliefs (including δ_∂). In Table 1, we enumerate all possible supports of the reachable beliefs when starting with belief b_0 .

We have therefore 7 different supports for the reachable beliefs, hence at least 7 beliefs in the set of reachable beliefs starting from b_0. As Equation (26) states that there can be at most 7 reachable beliefs, we obtain that we have exactly 7 reachable beliefs and thus Equation (28) is obtained. □

Note that, while the proof of Proposition 16 was made with a Separated Det-Pomdp with $|\mathbb{X}| = 3$, we can generate a Separated Det-Pomdp such that equality is obtained in Equation (26) for any $|\mathbb{X}| = n$, $n \geq 3$. We need once again that $\mathbb{X} = \{x_i\}_{i \in [\![1,n]\!]}$ consists of n distinct states, $\mathbb{O} = \{\bar{o}_1, \bar{o}_2\}$

Mapping applied	Support of resulting belief
$F^{\bar{u}_1,\bar{o}_1}$	$\{x_1, x_2\}$
$F^{\bar{u}_2,\bar{o}_1}$	$\{x_2, x_3\}$
$F^{\bar{u}_2,\bar{o}_1} \circ F^{\bar{u}_2,\bar{o}_1}$	$\{x_3, x_1\}$
$F^{\bar{u}_1,\bar{o}_2} \circ F^{\bar{u}_2,\bar{o}_1}$	$\{x_3\}$
$F^{\bar{u}_2,\bar{o}_1} \circ F^{\bar{u}_1,\bar{o}_2} \circ F^{\bar{u}_2,\bar{o}_1}$	$\{x_1\}$
$F^{\bar{u}_2,\bar{o}_1} \circ F^{\bar{u}_2,\bar{o}_1} \circ F^{\bar{u}_1,\bar{o}_2} \circ F^{\bar{u}_2,\bar{o}_1}$	$\{x_2\}$
$F^{\bar{u}_1,\bar{o}_2}$	$\{\partial\}$

Table 1: Resulting support when applying given mappings to the initial belief b_0 with $\text{supp}(b_0) = \{x_1, x_2\}$

of two distinct observations and $\mathbb{U} = \{\bar{u}_1, \bar{u}_2\}$ of two distinct controls. Then, the dynamics is given by $\forall x \in \mathbb{X}$, $f(x, \bar{u}_1) = x$, and $\forall i \in [\![1, n]\!]$, $f(x_i, \bar{u}_2) = x_{\text{mod}(i,n)+1}$. Finally, the observation mapping is given by $h(x, u) = \begin{cases} \bar{o}_2 & \text{if } x = x_n \text{ and } u = \bar{u}_1 \ , \\ \bar{o}_1 & \text{otherwise} \ . \end{cases}$

Now that we have presented the subclass of Separated Det-Pomdps, we give a numerical illustration.

5 Numerical application on a toy example of Separated Det-Pomdp

In this section, we present a simple one-dimensional illustration of Separated Det-Pomdp. We consider that we empty a tank while minimizing an associated cost, as illustrated in Figure 5. The state is one-dimensional and consists in the volume of water present in the tank. The control is also one-dimensional and is the amount of water that the decision maker removes during one time step. The decision maker has access at time t to partial observation, as he only knows that the volume of water in the tank is between two quantized levels.

5.1 A partially observed tank as a Separated Det-Pomdp

More precisely, the problem is the following.

- The state x consists of a discrete volume of water in the tank, with $x \in \mathbb{X} = \{x^{(1)}, x^{(2)}, \dots, x^{(n)}\} \subset \mathbb{R}_+$ of finite cardinality n.
- The observation o consists of a discrete level of water in the tank, with $o \in \mathbb{O} = \{o^{(1)}, o^{(2)}, \dots, o^{(m)}\} \subset \mathbb{R}_+$ of finite cardinality m.
- The control u consists of a discrete volume of water to be removed, with $u \in \mathbb{U} = \{u^{(1)}, u^{(2)}, \dots, u^{(d)}\} \subset \mathbb{R}_+$ of finite cardinality d.
- The unitary price of water at each time $t \in \mathcal{T} \setminus \{T\}$ is given by $c_t \in \mathbb{R}$.

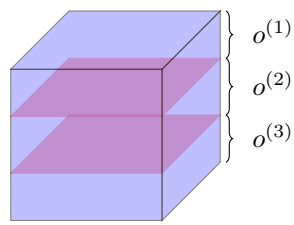

Figure 5: Illustration of the water tank "quantum" of observation ($m = 3$)

Optimization problem. We now adapt the Problem (2) to the tank case presented above:

$$\min_{\boldsymbol{X},\boldsymbol{U},\boldsymbol{O}} \mathbb{E}\Big[\sum_{t=0}^{T-1} c_t \boldsymbol{U}_t\Big] \tag{29a}$$

$$s.t.\ \mathbb{P}_{\boldsymbol{X}_0} = b_0 \ , \tag{29b}$$

$$\boldsymbol{X}_{t+1} = \boldsymbol{X}_t - \boldsymbol{U}_t \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , \tag{29c}$$

$$\boldsymbol{U}_t \in \{u^{(i)} \in \mathbb{U} \,|\, u^{(i)} \leq \boldsymbol{X}_t\} \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , \tag{29d}$$

$$\boldsymbol{O}_t = \max\{o^{(j)} \in \mathbb{O} \,|\, \boldsymbol{X}_t \geq o^{(j)}\} \ , \ \ \forall t \in \mathcal{T} \ , \tag{29e}$$

$$\sigma(\boldsymbol{U}_t) \subset \sigma\left(\boldsymbol{O}_0, \ldots, \boldsymbol{O}_t, \boldsymbol{U}_0, \ldots, \boldsymbol{U}_{t-1}\right) \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ . \tag{29f}$$

Equation (29a) represents the objective function of the tank problem, i.e. Equation (2a) of Problem (2). The instantaneous cost function at time t is defined as $\mathcal{L}_t(u_t) = c_t u_t$, and hence only depends on the controls. The evolution function corresponds to emptying the tank and is given by $f : (x, u) \mapsto x - u$, which gives Equation (29c). The observation function h is given by a piecewise constant function which does not depend on the controls u: $h(x) = \max\{o^{(i)} \,|\, x \geq o^{(i)}\}$.

This leads to equation (29e), which is the implementation of (2e). The admissibility set of the tank problem is given by $\mathcal{U}^{ad}(\boldsymbol{X}_t) = [0, \boldsymbol{X}_t]$ (see Equation (29d)). It ensures that we cannot remove more water than what is in the tank. Note that this could be a problem as we do not observe $\boldsymbol{X}_t$.

Problem (29) has the same form as Problem (2). It is therefore a Det-Pomdp and all the relevant results presented in §3.1 hence apply.

Associated beliefs dynamics τ. Let $(b, u, o) \in \mathbb{B} \times \mathbb{U} \times \mathbb{O}$, with $\mathbb{B} = \Delta(\mathbb{X}) \cup \{\delta_\partial\}$, as defined in Equation (4). As the evolution functions and observation functions are stationary, the belief dynamics are also stationary.

We note $\mathbb{I}(o) \subset \mathbb{X}$, the set of states compatible with the observation o, i.e.

$$\mathbb{I}(o) = \{x \in \mathbb{X} \,|\, h(x) = o\} \ . \tag{30}$$

By Equation (29c), we have $(f^u)^{-1}(y) = y + u$. Moreover, we have, by the definition of $\mathbb{I}(o)$ in Equation (30), that $\left(h^u\right)^{-1}(o) = \mathbb{I}(o)$. Hence, the function Q in (7) is here

$$Q : \mathbb{B} \times \mathbb{U} \times \mathbb{O} \to [0, 1] \, , (b, u, o) \mapsto \sum_{x \in \mathbb{I}(o) - \{u\}} b(x) \ ,$$

and Equation (9) gives

$$\tau(b, u, o)(y) = \begin{cases} \dfrac{b(y+u)}{\displaystyle\sum_{x' \in \mathbb{I}(o) - \{u\}} b(x')} & \text{if} \quad y \in \mathbb{I}(o) - \{u\} \ , \\ 0 & \text{if} \quad y \notin \mathbb{I}(o) - \{u\} \ , \end{cases} \qquad \text{where } \mathbb{I}(o) - \{u\} \text{ is defined in Equation (30).}$$

Bellman equations for the partially observed tank problem. As Problem (29) is a Det-Pomdp and the finite sets Assumption 1 holds, we can apply Proposition 1. Equations (10a) and (10b) are here

$$V_T : \mathbb{B}_T^{\mathrm{R},\mathcal{D}}(b_0) \to \mathbb{R} \ , \ \ b \mapsto 0 \tag{31a}$$

$$V_t : \mathbb{B}_t^{\mathrm{R},\mathcal{D}}(b_0) \to \mathbb{R} \ , \ \ b \mapsto \min_{u \leq \min_{x \in \mathrm{supp}(b)} x} \Big(c_t u + \sum_{o \in \mathbb{O}} \sum_{x - u \in [\underline{o}, \overline{o}]} b(x) V_{t+1}\big(\tau(b, u, o)\big)\Big) \ . \tag{31b}$$

Indeed, the intersection $\mathcal{U}_t^{\mathrm{b,ad}}(b) = \bigcap_{x \in \mathrm{supp}(b)} \mathcal{U}_t^{ad}(x)$ is $\{u^{(i)} \in \mathbb{U} \,|\, u \leq \min_{x \in \mathrm{supp}(b)} x\}$, as the admissibility set is given by Equation (29d), and as

$$\{u^{(i)} \in \mathbb{U} \,|\, u^{(i)} \leq x^{(j)}\} \cap \{u^{(i)} \in \mathbb{U} \,|\, u^{(i)} \leq x^{(k)}\} = \{u^{(i)} \in \mathbb{U} \,|\, u^{(i)} \leq \min\left(x^{(j)}, x^{(k)}\right)\} \ .$$

The partially observed tank problem as a Separated Det-Pomdp. The tank Det-Pomdp is a Separated Det-Pomdp as a direct consequence of Corollary 29, in Appendix A.2. Indeed, Corollary 29 states that if the evolution functions f_t of a Det-Pomdp are linear, then it is a Separated Det-Pomdp. As the evolution function f of the partially observed tank is indeed linear, the tank Det-Pomdp is a Separated Det-Pomdp.

5.2 Numerical results

We now present some numerical results for the tank problem described by Problem (29).

Presentation of the instances We made a numerical application with the following parameters:

- $\mathbb{X} = [\![0, 300]\!]$,
- $\mathbb{U} = [\![0, 9]\!]$,
- $\mathbb{O} = \{0, 1, 20, 40, 60, 80, 100, 120, 140, 160, 180, 200, 220, 240, 260, 280, 300\}$,
- $\mathcal{T} = [\![0, 100]\!]$,
- $\text{supp}(b_0) = [\![260, 300]\!]$, with a randomly generated probability distribution over that support, the distribution used is detailed in Figure 6.

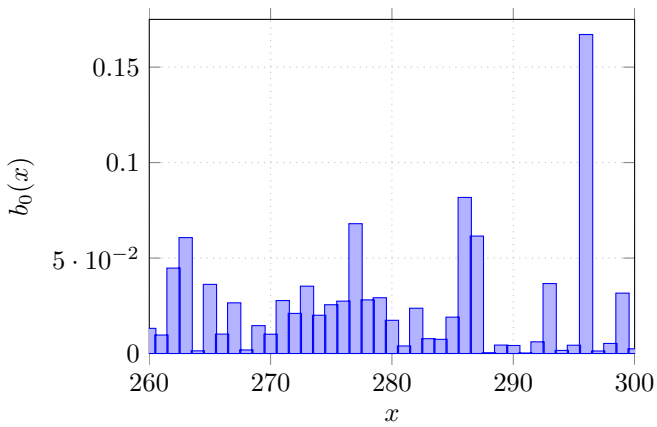

Figure 6: Probability distribution used as the initial belief b_0 for the numerical applications

When considering the initial belief b_0 presented in Figure 6 and a "true" (unknown) initial state of $x_0 = 290$ (used to simulate the observation process depending on the policy), we obtain the tank water volume represented in Figure 7.

Moreover, we have a set of reachable beliefs $\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![0,100]\!]}$ such that $|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![0,100]\!]}| = 64,400$. We therefore do not display value functions, as they are defined on sets with large cardinality.

We also made a second numerical application where the observation $\mathbb{O}$ is changed to:

- $\mathbb{O} = \{1, 6, 11, 51, 101, 151, 201, 251\}$

When considering the new observations set and the same initial belief and initial state, we obtain a trajectory represented in Figure 8.

Figures 7 and 8 both illustrate some properties of Det-Pomdps.

- First, in both cases, we see that the support of the beliefs decreases with time (the vertical red slices are non increasing).
- Second, we remark that such a decrease is due to the observations. Indeed, in Problem (29), the observation function ensures that the support of the beliefs must belong to intervals $[\underline{o}_t, \overline{o}_t]$ when we observe o_t (see Equation (30)). Thus, the supports of the beliefs are reduced along the limit of

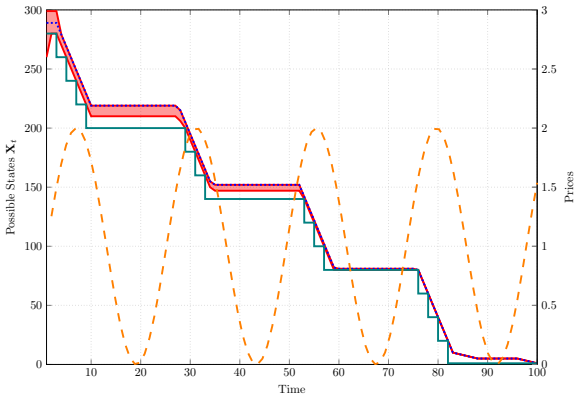

Figure 7: Representation of a trajectory of the volume of water in the tank when applying the optimal controls when considering the first set of observations. A vertical slice at time t of the red area represents the support of the belief held at time t, the dotted blue curve represents the trajectory of the "true" state, the piecewise constant green curve is the observation we have access to at time t, and the dashed orange curve represents the periodic prices.

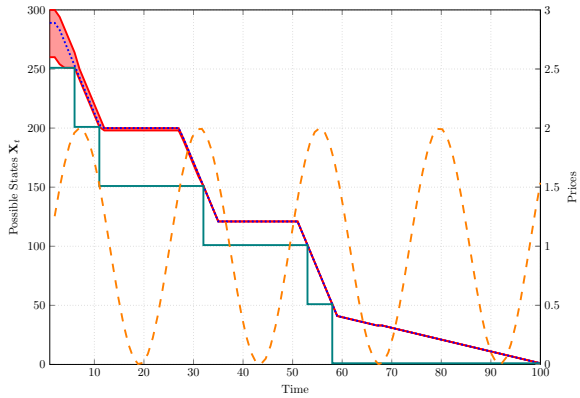

Figure 8: Representation of a trajectory of the volume of water in the tank when applying the optimal controls when considering the second set of observations. A vertical slice at time t of the red area represents the support of the belief held at time t, the dotted blue curve represents the trajectory of the "true" state, the piecewise constant green curve is the observation we have access to at time t, and the dashed orange curve represents the periodic prices.

those intervals, as is more easily seen in Figure 8 between time $t = 1$ to $t = 6$ (we apply a control, i.e. removing some water, and we see that the lower part of the support remains at the observation value until time $t = 7$, which is when we change observation and we see that the upper bound of the support gets just beneath the previous observation, i.e. at $x = 249$).

- Third, we remark that, as could be expected, the optimal policy consists of removing water when prices are high, and stopping when prices are low.
- Fourth, we remark that, despite having fewer observations in the second case, the optimal trajectory in the second case reaches a deterministic belief (i.e. such that $|\text{supp}(b)| = 1$) much sooner in Figure 8 compared to Figure 7 (at time $t = 33$ for the second case and time $t = 53$ for the first case). Having more observations hence does not guarantee to remove ambiguities at a faster rate.

We now present the computation time of the Dp Algorithm and compare it to another algorithm, Sarsop.

Comparison with Sarsop. In this paragraph, we focus on the comparison with Sarsop, first introduced in (Kurniawati et al., 2008). We used the Julia implementation of this algorithm, with the POMDPs package API. The following results were obtained on a computer equipped with a Core i7-8665U and 32 GB of memory, using Julia v1.7.3, POMDPs v0.9.3 and Sarsop v0.5.5.

However, we must first warn the reader that Sarsop is an algorithm that solves an infinite horizon Pomdp. We hence reformulate the finite horizon Det-Pomdp as an infinite time Pomdp by extending the state with the time variable. Such reformulation leads to a much bigger problem in terms of data and size of the state space, which heavily penalizes Sarsop. Hence, the reformulation prevents any fair comparison of computation times. We still present some computation time in Table 2.

Note that, for each instance where the computation did not stop (i.e. those without a "$>$" symbol in the computation time column) due to hitting the memory limit of the computer, Sarsop and the Dp Algorithm found the same value, hence Sarsop indeed converged toward the optimal solution of Problem (29).

$\lvert\mathbb{X}\rvert$	$\lvert\mathbb{U}\rvert$	$\lvert\mathbb{O}\rvert$	$\lvert\text{supp}(b_0)\rvert$	T	Sarsop computation time (s)	Dp Algorithm computation time (s)
11	2	3	2	20	0.376	0.002
21	2	5	2	25	0.16	0.003
51	5	5	2	100	24.9	0.20
51	5	5	4	100	27.2	1.20
51	5	5	6	100	29.4	3.03
101	5	5	2	200	359	0.96
101	5	5	10	200	1930	32.2
101	10	5	10	200	1069	78.2
201	5	5	10	200	3506	62.1
201	10	5	10	200	15618	309
201	5	5	20	200	3652	225
201	10	6	20	200	33562	497
301	5	6	10	200	4638	86.8
301	10	6	10	300	> 38000 ($>$ 19217s of iterations)	762

Table 2: Computation time of different instances of both Sarsop and the Dp Algorithm

6 Conclusion

In this paper, we have presented a subclass of Pomdps, Separated Det-Pomdps, which has properties that contribute to push back the curse of dimensionality for Dynamic Programming. Indeed, we have shown that the conditions on the dynamics for Separated Det-Pomdp improve the bound on the cardinality of the set of the reachable beliefs: the bound is reduced from $\left(1+|\mathbb{X}|\right)^{|\text{supp}(b_0)|}$ (in the case of Det-Pomdp, see Theorem 4) to $2^{|\text{supp}(b_0)|}|\mathbb{X}|$ (Theorem 13), as presented in Table 3. This tighter bound allows Dynamic Programming algorithms to efficiently solve Separated Det-Pomdp problems, especially when considering small supports of the initial state distributions. Moreover, the bound is tight (see Proposition 16).

The Separated Det-Pomdp class is, therefore, an interesting framework for some problems as only a fraction of the number of beliefs needs to be considered, in comparison with Det-Pomdp or Pomdp. The Separated Det-Pomdps are therefore tractable with larger instances than regular Pomdps or Det-Pomdps.

Class	Infinite horizon bound	Finite horizon bound
Det-Pomdp	$(1+\lvert\mathbb{X}\rvert)^{\lvert\mathbb{X}\rvert}$ (Littman, 1996)	$\min\left((1+\lvert\mathbb{X}\rvert)^{\lvert\mathbb{X}\rvert}, \left(\lvert\mathbb{U}\rvert\lvert\mathbb{O}\rvert\right)^{\lvert\mathcal{T}\rvert}\right)$
Det-Pomdp improved bounds	$(1+\lvert\mathbb{X}\rvert)^{\lvert\text{supp}(b_0)\rvert}$ Theorem 4	$\min\left((1+\lvert\mathbb{X}\rvert)^{\lvert\text{supp}(b_0)\rvert}, 1+\lvert\text{supp}(b_0)\rvert\lvert\mathbb{U}\rvert^{\lvert\mathcal{T}\rvert}\right)$ Theorem 4
Separated Det-Pomdp	$1+\left(2^{\lvert\text{supp}(b_0)\rvert}-\lvert\text{supp}(b_0)\rvert\right)\lvert\mathbb{X}\rvert$ Theorem 13	$\min\Big(1+\left(2^{\lvert\text{supp}(b_0)\rvert}-\lvert\text{supp}(b_0)\rvert\right)\lvert\mathbb{X}\rvert,$ $1+\lvert\text{supp}(b_0)\rvert\lvert\mathbb{U}\rvert^{\lvert\mathcal{T}\rvert}\Big)$ Corollary 14

Table 3: Summary of the bounds depending on the class of problem

A Appendix

First, in §A.1, we present technical lemmata used to prove bounds on the cardinality of the sets of reachable beliefs. Second, in §A.2, we present complementary results on (∂)-separated mappings sets.

A.1 Technical lemmata

In this subsection, we present technical lemmata used in the proofs of Theorem 4. We first introduce in §A.1.1 the notions of forward and backward mappings. Second, in §A.1.2, we present properties on the composition and pushforward measures by those forward and backward mapping. Third, in §A.1.3, we present properties on the cardinality of sets of forward and backward mappings used notably in the proof of Theorem 4.

A.1.1 Forward and backward mappings

For any subset $X \subset \mathbb{X}$, we introduce the notion of X-forward and X-backward mappings. Given a mapping $h : \mathbb{X} \to \mathbb{X}$ and a subset $X \subset \mathbb{X}$, we define a mapping $h_{\overrightarrow{X}}$: from $\overline{\mathbb{X}}$ to $\overline{\mathbb{X}}$, called a *X-forward mapping*, as follows

$$h_{\overrightarrow{X}} : x \in \overline{\mathbb{X}} \mapsto \begin{cases} h(\overline{x}) & \text{if} \quad x \in \mathbb{X} \quad \text{and} \quad h(x) \in X \ , \\ \partial & \text{if} \quad x = \partial \quad \text{or} \quad h(x) \notin X \ . \end{cases} \tag{32}$$

We call $h_{\overrightarrow{X}} : \overline{\mathbb{X}} \to \overline{\mathbb{X}}$ a X-forward mapping as we have $h_{\overrightarrow{X}}(\overline{\mathbb{X}}) \subset X \cup \{\partial\}$. X-forward imposes a constraint on the codomain (set of destinations): we only keep the values that belong to X, whereas the others are sent to ∂. The set X is thus a subset of the codomain of h.

We also introduce the *X-backward mapping* $h_{\overleftarrow{X}} : \overline{\mathbb{X}} \to \overline{\mathbb{X}}$, defined by

$$h_{\overleftarrow{X}} : x \in \overline{\mathbb{X}} \mapsto \begin{cases} h(x) & \text{if} \quad x \in X \ , \\ \partial & \text{otherwise.} \end{cases} \tag{33}$$

We call $h_{\overleftarrow{X}} : \overline{\mathbb{X}} \to \overline{\mathbb{X}}$ a X-backward mapping as we have $h_{\overleftarrow{X}}(X) \subset \mathbb{X}$, and $h_{\overleftarrow{X}}\big(\overline{\mathbb{X}} \setminus X\big) = \{\partial\}$. X-backward imposes a constraint on the domain (set of departures): we only keep the values whose inputs are in X, whereas the others are sent to ∂. The set X is thus a subset of the domain of h.

It is straightforward to check that we have

$$\forall X \subset \mathbb{X} \ , \ h_{\overrightarrow{X}} = h_{\overleftarrow{h^{-1}(X)}} \ , \tag{34a}$$

$$\forall X \subset \mathbb{X} \ , \ h_{\overrightarrow{X}} = h_{\overrightarrow{X \cap \mathrm{Im}(h)}} \ , \tag{34b}$$

where Im is the image of a mapping, that is $\mathrm{Im}(h) = h(\mathbb{X})$. A forward mapping can hence be rewritten as a backward mapping. The reverse is not true, as we have

$$h_{\overleftarrow{X}} = h_{\overrightarrow{h(X)}} \Leftrightarrow h^{-1}\big(h(X)\big) = X \ .$$

A.1.2 Results on pushforward measures by forward and backward mappings sets

We now present properties of the composition of pushforward measures of forward and backward mappings.

Definition 17. *Let $\mathbb{M} \subset \mathbb{L}(\mathbb{X};\mathbb{X})$ be a subset of self mappings on the set $\mathbb{X}$. We say that $\mathbb{G} \subset \mathbb{L}(\overline{\mathbb{X}};\overline{\mathbb{X}})$ is an* $\big(\mathbb{M}, \overleftarrow{\mathbb{X}}\big)$-mappings set *(resp. an* $\big(\mathbb{M}, \overrightarrow{\mathbb{X}}\big)$-mappings set*) if it satisfies the following property*

$$\mathbb{G} \subset \big\{h_{\overleftarrow{X}} \,\big|\, h \in \mathbb{M} \textit{ and } X \subset \mathbb{X}\big\} \ , \tag{35a}$$

$$\big(\textit{resp.} \quad \mathbb{G} \subset \big\{h_{\overrightarrow{X}} \,\big|\, h \in \mathbb{M} \textit{ and } X \subset \mathbb{X}\big\}\big) \ , \tag{35b}$$

where $h_{\overleftarrow{X}}$ (resp. $h_{\overrightarrow{X}}$) is defined in Equation (33) *(resp. Equation* (32)*). When $\mathbb{M} = \mathbb{L}(\mathbb{X};\mathbb{X})$, a* $\big(\mathbb{M}, \overleftarrow{\mathbb{X}}\big)$*-mappings set (resp. an* $\big(\mathbb{M}, \overrightarrow{\mathbb{X}}\big)$*-mappings set) is just named a* $\big(\overleftarrow{\mathbb{X}}\big)$-mappingsset *(resp. an* $\big(\overrightarrow{\mathbb{X}}\big)$-mappings set*).*

We obtain the following properties.

- If $\mathbb{G}$ is an $\big(\mathbb{M}, \overrightarrow{\mathbb{X}}\big)$-mappings set, then $\mathbb{G}$ is an $\big(\mathbb{M}, \overleftarrow{\mathbb{X}}\big)$-mappings set (using Equality (34a)).

- $(\overleftarrow{\mathbb{X}})$-mappings sets are stable by composition, as we easily obtain that

$$h'_{\overleftarrow{X'}} \circ h_{\overleftarrow{X}} = (h' \circ h)_{\overleftarrow{X \cap h^{-1}(X')}} \ . \tag{36}$$

- Let $\mathbb{G}$ be an $(\overleftarrow{\mathbb{X}})$-mappings set and consider, for any $X \subset \mathbb{X}$, the subset $\mathbb{G}_{\overleftarrow{X}}$ of $\mathbb{G}$ defined by

$$\mathbb{G}_{\overleftarrow{X}} = \big\{ g \in \mathbb{G} \,\big|\, \exists h \in \mathbb{L}(\mathbb{X};\mathbb{X}), g = h_{\overleftarrow{X}} \big\} \ . \tag{37}$$

  Then, for any belief $b_0 \in \Delta(\mathbb{X})$, we have

$$\big(\mathcal{R} \circ (\mathbb{G}_{\overleftarrow{X \cap \mathrm{supp}(b_0)}})_\star\big)(b_0) = \big(\mathcal{R} \circ (\mathbb{G}_{\overleftarrow{X}})_\star\big)(b_0) \ . \tag{38}$$

  The Equation (38) is a consequence of the following Lemma 18. Indeed, assuming Lemma 18, the expression of $\big(\mathcal{R} \circ (\mathbb{G}_{\overleftarrow{X}})_\star\big)(b_0)$ given by Equation (39b) only depends on the restriction of the measure b_0 to the subset X – which coincides with the restriction of the measure b_0 to the subset $X \cap \mathrm{supp}(b_0)$ – as the measure b_0 is null outside its support.

Lemma 18. *Let X be a subset of $\mathbb{X}$. The mappings $\mathcal{R} \circ (h_{\overleftarrow{X}})_\star$ and $\mathcal{R} \circ (h_{\overrightarrow{X}})_\star$ in $\mathbb{L}(\Delta(\overline{\mathbb{X}});\mathbb{B})$, where the pushforward measure is defined in Equation* (15)*, and the mapping $\mathcal{R}$ is defined in Equation* (17)*, have the following expressions for all $\nu \in \Delta(\overline{\mathbb{X}})$:*

$$\big(\mathcal{R} \circ (h_{\overrightarrow{X}})_\star\big)(\nu) = \begin{cases} \Big[\overline{x} \in \overline{\mathbb{X}} \mapsto \dfrac{\nu\big(h^{-1}(\overline{x})\big)\mathbf{1}_X(\overline{x})}{\nu\big(h^{-1}(X)\big)}\Big] & \textit{if} \quad \nu\big(h^{-1}(X)\big) \neq 0 \ , \\ \delta_\partial & \textit{otherwise,} \end{cases} \tag{39a}$$

and

$$\big(\mathcal{R} \circ (h_{\overleftarrow{X}})_\star\big)(\nu) = \begin{cases} \Big[\overline{x} \in \overline{\mathbb{X}} \mapsto \dfrac{\nu\big(h^{-1}(\overline{x}) \cap X\big)}{\nu\big(h^{-1}(\mathbb{X}) \cap X\big)}\Big] & \textit{if} \quad \nu\big(h^{-1}(\mathbb{X}) \cap X\big) \neq 0 \ , \\ \delta_\partial & \textit{otherwise.} \end{cases} \tag{39b}$$

Proof. For any probability measure ν on the finite set $\overline{\mathbb{X}}$, it is straightforward, using the definition of pushforward measure in Equation (15), to obtain that the pushforward of the measure ν through the mapping $h_{\overrightarrow{X}}$, as defined in Equation (32), is given by

$$\begin{aligned} &(h_{\overrightarrow{X}})_\star \nu : \overline{\mathbb{X}} \to \mathbb{R}_+ \\ &\qquad \overline{y} \mapsto \nu\big((h_{\overrightarrow{X}})^{-1}(\overline{y})\big) = \begin{cases} \nu\big(h^{-1}(\overline{y})\big) & \text{if} \quad \overline{y} \in X \ , \\ \Big(1 - \nu\big(h^{-1}(X)\big)\Big) & \text{if} \quad \overline{y} = \partial \ , \\ 0 & \text{if} \quad \overline{y} \neq \partial \text{ and } \overline{y} \notin X \ . \end{cases} \end{aligned} \tag{40}$$

Thus, we obtain that

$$\forall x \in \mathbb{X} \ , \ \big((h_{\overrightarrow{X}})_\star \nu\big)_{|\mathbb{X}}(x) = \nu\big(h^{-1}(x)\big)\mathbf{1}_X(x) \ , \tag{41}$$

and that

$$\big((h_{\overrightarrow{X}})_\star \nu\big)(\mathbb{X}) = \sum_{x \in \mathbb{X}} \nu\big(h^{-1}(x)\big)\mathbf{1}_X(x) = \nu\big(h^{-1}(X)\big) \ . \tag{42}$$

Hence, using the definition of $\mathcal{R}$ in Equation (17), the result follows from Equation (39a). The proof of Equation (39b) is very similar and left to the reader. □

The composition of self-mappings of the form $\mathcal{R} \circ (h_{\overrightarrow{X}})_\star$ can also be written without resorting to multiple renormalizations. Instead, we only need to renormalize the composition of the pushforward measures, as shown below.

Lemma 19. *Assume that h and h' are self-mappings on the finite set $\mathbb{X}$. Then, for any subsets X and X' of $\mathbb{X}$, we have the following composition equalities*

$$\mathcal{R} \circ (h_{\overrightarrow{X}})_\star \circ \mathcal{R} \circ (h'_{\overrightarrow{X'}})_\star = \mathcal{R} \circ (h_{\overrightarrow{X}} \circ h'_{\overrightarrow{X'}})_\star \ , \tag{43a}$$

$$\mathcal{R} \circ (h_{\overleftarrow{X}})_\star \circ \mathcal{R} \circ (h'_{\overleftarrow{X'}})_\star = \mathcal{R} \circ (h_{\overleftarrow{X}} \circ h'_{\overleftarrow{X'}})_\star \ . \tag{43b}$$

Proof. We just prove Equation (43a) as the proof follows the same lines for Equation (43b). As a preliminary, we remark that the mapping $\mathcal{R} \circ (h_{\overrightarrow{X}})_\star$ is defined on the nonnegative measures on the set $\overline{\mathbb{X}}$ and not just on probability measures. Now, given $\mu \in \Delta(\overline{\mathbb{X}})$, we consider the nonnegative measure $\mu' = (\mu_{|\mathbb{X}}, 0)$. The two nonnegative measures μ and μ' coincide on the set $\mathbb{X}$. Thus using the expression of $\mathcal{R} \circ (h_{\overrightarrow{X}})_\star$ in Equation (39a) and the fact that $X \subset \mathbb{X}$, we obtain that $\mathcal{R} \circ (h_{\overrightarrow{X}})_\star(\mu) = \mathcal{R} \circ (h_{\overrightarrow{X}})_\star(\mu_{|\mathbb{X}}, 0)$.

Now, let $\nu \in \Delta(\overline{\mathbb{X}})$ be given. We denote by $\nu' \in \Delta(\overline{\mathbb{X}})$ the probability measure $\nu' = (h'_{\overrightarrow{X'}})_\star \nu$. We consider two cases: either $\nu'(\mathbb{X}) \neq 0$, or $\nu'(\mathbb{X}) = 0$.

First case. We assume that $\nu'(\mathbb{X}) \neq 0$. Then, we successively have

$$\begin{aligned}
\mathcal{R} \circ (h_{\overrightarrow{X}})_\star \circ \mathcal{R} \circ (h'_{\overrightarrow{X'}})_\star \nu &= \mathcal{R} \circ (h_{\overrightarrow{X}})_\star \circ \mathcal{R}(\nu') && \text{(by replacing } (h'_{\overrightarrow{X'}})_\star \nu \text{ by } \nu') \\
&= \mathcal{R} \circ (h_{\overrightarrow{X}})_\star \big(\frac{1}{\nu'(\mathbb{X})} \nu'_{|\mathbb{X}}, 0 \big) && \text{(using } \mathcal{R} \text{ definition in (17), with } \nu'(\mathbb{X}) \neq 0) \\
&= \mathcal{R} \circ (h_{\overrightarrow{X}})_\star \big(\frac{1}{\nu'(\mathbb{X})} (\nu'_{|\mathbb{X}}, 0) \big) && \text{(factorizing by } \tfrac{1}{\nu'(\overline{\mathbb{X}})}) \\
&= \mathcal{R} \Big(\frac{1}{\nu'(\mathbb{X})} (h_{\overrightarrow{X}})_\star \big(\nu'_{|\mathbb{X}}, 0 \big) \Big) && \text{(as } (h_{\overrightarrow{X}})_\star \text{ is 1-positively homogeneous)} \\
&= \mathcal{R} \big((h_{\overrightarrow{X}})_\star \big(\nu'_{|\mathbb{X}}, 0 \big) \big) && \text{(as } \mathcal{R} \text{ is 0-positively homogeneous)} \\
&= \mathcal{R} \big((h_{\overrightarrow{X}})_\star (\nu') \big) && \text{(using the preliminary part)} \\
&= \mathcal{R} \circ (h_{\overrightarrow{X}})_\star \circ (h'_{\overrightarrow{X'}})_\star \nu && \text{(as } \nu' = (h'_{\overrightarrow{X'}})_\star \nu) \\
&= \mathcal{R} \circ (h_{\overrightarrow{X}} \circ h'_{\overrightarrow{X'}})_\star (\nu) \ . && \text{(as } f_\star \circ h_\star = (f \circ h)_\star)
\end{aligned}$$

Second case. We assume that $\nu'(\mathbb{X}) = 0$. Then, we have that $\nu' = \delta_\partial$ as $\nu' \in \Delta(\overline{\mathbb{X}})$, and we obtain

$$\begin{aligned}
\mathcal{R} \circ (h_{\overrightarrow{X}})_\star \circ \mathcal{R} \circ (h'_{\overrightarrow{X'}})_\star \nu &= \mathcal{R} \circ (h_{\overrightarrow{X}})_\star \circ \mathcal{R}(\delta_\partial) && \text{(by replacing } (h'_{\overrightarrow{X'}})_\star \nu \text{ by } \nu' = \delta_\partial) \\
&= \mathcal{R} \circ (h_{\overrightarrow{X}})_\star (\delta_\partial) && \text{(as } \mathcal{R}(\delta_\partial) = \delta_\partial) \\
&= \mathcal{R} \circ (h_{\overrightarrow{X}})_\star \circ (h'_{\overrightarrow{X'}})_\star \nu && \text{(by replacing } \delta_\partial = \nu' \text{ by } (h'_{\overrightarrow{X'}})_\star \nu) \\
&= \mathcal{R} \circ (h_{\overrightarrow{X}} \circ h'_{\overrightarrow{X'}})_\star (\nu) \ .
\end{aligned}$$

Hence, in both cases, we obtain Equation (43a). □

Now that we have exposed technical lemmata on the composition and renormalization of $(\overrightarrow{\mathbb{X}})$-mappings and $(\overleftarrow{\mathbb{X}})$-mappings, we present lemmata on the cardinality of sets of pushforward measures, notably the cardinality of pushforward measures by $(\overrightarrow{\mathbb{X}})$-mappings and $(\overleftarrow{\mathbb{X}})$-mappings.

A.1.3 Results on the cardinality of sets of pushforward measures

We now present results on the cardinality of sets of forward and backward mappings.

Lemma 20. *Let $\{\mathbb{G}_k\}_{k \in \mathbb{N}}$ be a given sequence where, for each $k \in \mathbb{N}$, the set $\mathbb{G}_k \subset \mathbb{L}(\overline{\mathbb{X}}; \overline{\mathbb{X}})$ is a finite set of self-mappings on the set $\overline{\mathbb{X}}$. The sets $\mathbb{G}_k$, for all $k \in \mathbb{N}$, are assumed to be either all $(\overrightarrow{\mathbb{X}})$-mappings sets or all $(\overleftarrow{\mathbb{X}})$-mappings sets. We define the sequence $\{\Phi_k\}_{k \in \mathbb{N}}$, where, for each $k \in \mathbb{N}$, the set $\Phi_k \subset \mathbb{L}(\Delta(\overline{\mathbb{X}}); \Delta(\overline{\mathbb{X}}))$ is a finite set of self-mappings on the set $\overline{\mathbb{X}}$ given by*

$$\forall k \in \mathbb{N} \ , \ \ \Phi_k = \mathcal{R} \circ (\mathbb{G}_k)_\star \ . \tag{44}$$

Then, for any $b_0 \in \Delta(\mathbb{X})$, we have the following bound

$$\forall n \in \mathbb{N}\ , \quad \Big| \bigcup_{k=0}^{n} \Phi_{0:k}(b_0) \Big| \leq (1 + |\mathbb{X}|)^{|\mathrm{supp}(b_0)|}\ , \tag{45}$$

where $\Phi_{0:k} = \Phi_k \circ \cdots \circ \Phi_0$ is defined in Equation (19).

Proof. For all $k \in \mathbb{N}$, we have

$$\begin{aligned}
\Phi_{0:k}(b_0) &= (\Phi_k \circ \Phi_{k-1} \circ \cdots \circ \Phi_0)(b_0) && \text{(by Equation (19))}\\
&= \big(\mathcal{R} \circ (\mathbb{G}_k)_\star \circ \mathcal{R} \circ (\mathbb{G}_{k-1})_\star \circ \cdots \circ \mathcal{R} \circ (\mathbb{G}_0)_\star\big)(b_0) && \text{(by Equation (44))}\\
&= \big(\mathcal{R} \circ (\mathbb{G}_k)_\star \circ (\mathbb{G}_{k-1})_\star \circ \cdots \circ (\mathbb{G}_0)_\star\big)(b_0)
\end{aligned}$$

by Lemma (19), as the sets $\mathbb{G}_k$ are, by assumption, either all $\big(\overrightarrow{\mathbb{X}}\big)$-mappings sets or all $\big(\overleftarrow{\mathbb{X}}\big)$-mappings sets,

$$\begin{aligned}
&= \big(\mathcal{R} \circ (\mathbb{G}_k \circ \mathbb{G}_{k-1} \circ \cdots \circ \mathbb{G}_0)_\star\big)(b_0) && \text{(as } f_\star \circ h_\star = (f \circ h)_\star\text{)}\\
&= \mathcal{R}\big((\mathbb{G}_{0:k})_\star(b_0)\big)\ .
\end{aligned}$$

Thus we have, for all $n \in \mathbb{N}$, $\Big|\bigcup_{k=0}^{n} \Phi_{0:k}(b_0)\Big| \leq \Big|\big(\bigcup_{k=0}^{n} \mathbb{G}_{0:k}\big)_\star(b_0)\Big|$, and the conclusion follows from the postponed Lemma 21 with $\mathbb{J} = \bigcup_{k=0}^{n} \mathbb{G}_{0:k}$, $\mathbb{Y} = \mathbb{V} = \overline{\mathbb{X}}$, and $\mu = b_0$. □

Note that we can extend the previous Lemma 20 to cases with sequences $\{\mathbb{G}_k\}_{k\in\mathbb{N}}$ of mixes of both $\big(\overrightarrow{\mathbb{X}}\big)$-mappings sets and $\big(\overleftarrow{\mathbb{X}}\big)$-mappings sets. Indeed, forward mappings are also backward mappings by Equation (34a). We can hence write the sequence $\{\mathbb{G}_k\}_{k\in\mathbb{N}}$ as a sequence of only $\big(\overleftarrow{\mathbb{X}}\big)$-mappings sets. In the rest of this paper, we consider sequences of only $\big(\overrightarrow{\mathbb{X}}\big)$-mappings sets or only $\big(\overleftarrow{\mathbb{X}}\big)$-mappings sets, and thus only need Lemma 20.

We can bound the cardinality of the set of pushforward of a given nonnegative measure thanks to the following Lemma 21 (which was previously postponed in the proof of Lemma 20).

Lemma 21. *Let $\mathbb{J} \subset \mathbb{L}(\mathbb{V};\mathbb{Y})$ be a set of mappings from the set $\mathbb{V}$ to the set $\mathbb{Y}$. Assume that the sets $\mathbb{V}$ and $\mathbb{Y}$ are both finite. Then, for any nonnegative measure μ on the set $\mathbb{V}$, we have that*

$$|\mathbb{J}_\star\mu| \leq |\mathbb{Y}|^{|\mathrm{supp}(\mu)|}\ , \tag{46}$$

where we recall that $|\mathbb{J}_\star\mu|$ denotes the cardinal of the set $\big|\{j_\star\mu \,|\, j \in \mathbb{J}\}\big|$ as exposed in Equation (19a).

Proof. Let μ be a given nonnegative measure on $\mathbb{V}$. For any $j \in \mathbb{J}$ we denote by $j_{|\mathrm{supp}(\mu)}$ the restriction of the mapping j to the subset $\mathrm{supp}(\mu) \subset \mathbb{V}$. For all $y \in \mathbb{Y}$, we have that

$$\begin{aligned}
j_\star\mu(y) &= \mu\big(j^{-1}(y)\big) && \text{(by the definition (15) of pushforward measures)}\\
&= \mu\Big(\big(j^{-1}(y) \cap \mathrm{supp}(\mu)\big) \cup \big(j^{-1}(y) \cap (\mathrm{supp}(\mu))^c\big)\Big)\\
&= \mu\big(j^{-1}(y) \cap \mathrm{supp}(\mu)\big) + \underbrace{\mu\big(j^{-1}(y) \cap (\mathrm{supp}(\mu))^c\big)}_{=0}\\
&= \mu\big(j_{|\mathrm{supp}(\mu)}^{-1}(y)\big)\\
&= \Big(\big(j_{|\mathrm{supp}(\mu)}\big)_\star\mu\Big)(y)\ . && \text{(by (15))}
\end{aligned}$$

Thus, defining $\mathbb{J}_{|\mathrm{supp}(\mu)} = \{j_{|\mathrm{supp}(\mu)} \,|\, j \in \mathbb{J}\}$, we get that

$$|\{j_\star\mu \,|\, j \in \mathbb{J}\}| = |\{(j_{|\mathrm{supp}(\mu)})_\star\mu \,|\, j \in \mathbb{J}\}| \leq |\mathbb{J}_{|\mathrm{supp}(\mu)}| \leq |\mathbb{Y}^{\mathrm{supp}(\mu)}| = |\mathbb{Y}|^{|\mathrm{supp}(\mu)|}\ .$$

This ends the proof. □

We now present a lemma on the conservation of the cardinality of the support of a measure through a composition of sets of mappings, if we have conservation of the cardinality for each individual set.

Lemma 22. *Let $\{\Phi_k\}_{k\in\mathbb{N}}$ be a sequence of self-mappings on the set $\mathbb{B}$ and assume that, for all $k \in \mathbb{N}$, we have that*

$$\forall b \in \mathbb{B}\ , \qquad \sum_{h\in\Phi_k} |\mathrm{supp}\big(h(b)_{|\mathbb{X}}\big)| \leq |\mathrm{supp}(b_{|\mathbb{X}})|\ . \tag{47}$$

Then, for any $b_0 \in \Delta(\mathbb{X})$, we have the following bound

$$\forall k \in \mathbb{N}\ , \quad \big|\Phi_{0:k}(b_0) \setminus \{\delta_\partial\}\big| \leq |\mathrm{supp}(b_0)|\ , \tag{48}$$

where $\Phi_{0:k}(b_0) = \Phi_k \circ \cdots \circ \Phi_0(b_0)$ is defined in Equation (19c).

Proof. Let a belief $b_0 \in \Delta(\mathbb{X})$ be given. As a preliminary result we prove, by forward induction on $k \in \mathbb{N}$, that

$$\forall k \in \mathbb{N}\ , \qquad \sum_{b\in\Phi_{0:k}(b_0)} \big|\mathrm{supp}(b_{|\mathbb{X}})\big| \leq |\mathrm{supp}(b_0)|\ . \tag{49}$$

First, we consider the case $k = 0$. As $\Phi_{0:0} = \Phi_0$ the result follows from Equation (47) used for $k = 0$ and $b = b_0$. Second, we consider $0 < k$, and, assuming that Equation (49) is satisfied for k, we prove that it is also satisfied for $k+1$ as follows:

$$\begin{aligned}
\sum_{b\in\Phi_{0:k+1}(b_0)} \big|\mathrm{supp}(b_{|\mathbb{X}})\big| &= \sum_{h\in\Phi_{0:k+1}} \big|\mathrm{supp}\big((h(b_0))_{|\mathbb{X}}\big)\big| && \text{(by (19))}\\
&= \sum_{h'\in\Phi_{k+1}, h''\in\Phi_{0:k}} \Big|\mathrm{supp}\Big(\big(h'(h''(b_0))\big)_{|\mathbb{X}}\Big)\Big| && (\text{as } \Phi_{0:k+1} = \Phi_{k+1}\circ\Phi_{0:k})\\
&= \sum_{h''\in\Phi_{0:k}} \Big(\sum_{h'\in\Phi_{k+1}} \Big|\mathrm{supp}\Big(\big(h'(h''(b_0))\big)_{|\mathbb{X}}\Big)\Big|\Big)\\
&\leq \sum_{h''\in\Phi_{0:k}} \Big|\mathrm{supp}\Big(\big(h''(b_0)\big)_{|\mathbb{X}}\Big)\Big| && \text{(using Equation (47) for } k \text{ and } b = h''(b_0))\\
&= \sum_{b\in\Phi_{0:k}(b_0)} \big|\mathrm{supp}\big(b_{|\mathbb{X}}\big)\big| && \text{(by (19))}\\
&\leq |\mathrm{supp}(b_0)|\ . && \text{(by induction assumption (49) on } k)
\end{aligned}$$

We conclude that Equation (49) is satisfied for all $k \in \mathbb{N}$.

Now, we turn to the proof of Equation (48). We make the following observation: if $b \in \Delta(\mathbb{X})$, then we have that $|\mathrm{supp}(b_{|\mathbb{X}})| \geq 1$ and if $b = \delta_\partial$ then $|\mathrm{supp}(b_{|\mathbb{X}})| = 0$. Thus, we have that

$$\begin{aligned}
|\Phi_{0:k}(b_0) \setminus \{\delta_\partial\}| &= \sum_{b\in\Phi_{0:k}(b_0)\setminus\{\delta_\partial\}} 1 && (50)\\
&\leq \sum_{b\in\Phi_{0:k}(b_0)\setminus\{\delta_\partial\}} |\mathrm{supp}(b_{|\mathbb{X}})| && (\text{as } |\mathrm{supp}(b_{|\mathbb{X}})| \geq 1 \text{ for } b \in \Phi_{0:k}(b_0) \setminus \{\delta_\partial\})\\
&= \sum_{b\in\Phi_{0:k}(b_0)} |\mathrm{supp}(b_{|\mathbb{X}})| && (\text{as } |\mathrm{supp}(\delta_{\partial|\mathbb{X}})| = 0)\\
&\leq |\mathrm{supp}(b_0)|\ , && \text{(by (49))}
\end{aligned}$$

which gives Equation (48). That concludes the proof. □

Lemma 23. *Let $\{h^k\}_{k\in\mathbb{N}}$ be a sequence of self-mappings on the set $\overline{\mathbb{X}}$ and, for all $k \in \mathbb{N}$, let $\{X_i^k\}_{i\in I_k}$ be a finite family of two by two disjoints subsets of $\mathbb{X}$. Let $\{\mathbb{G}_k\}_{k\in\mathbb{N}}$ be the sequence of self-mappings on the set $\overline{\mathbb{X}}$, of the following form*

$$\forall k \in \mathbb{N}\ , \ \ \mathbb{G}_k = \big\{ h^k{}_{\overrightarrow{X_i^k}} \,\big|\, i \in I_k \big\} \subset \overline{\mathbb{X}}^{\overline{\mathbb{X}}}\ , \tag{51}$$

where $h^k{}_{\overrightarrow{X_i^k}} : \overline{\mathbb{X}} \to \overline{\mathbb{X}}$ are built following Equation (32). *Consider the sequence $\{\Phi_k\}_{k\in\mathbb{N}}$ of self-mappings on the set $\mathbb{B}$, given, for all $k \in \mathbb{N}$, by $\Phi_k = \mathcal{R} \circ (\mathbb{G}_k)_\star$ and the associated sequence $(\Phi_{0:k})_{k\in\mathbb{N}}$ as defined in Equation* (19). *Then, given $b_0 \in \Delta(\mathbb{X})$, we have*

$$\forall k \in \mathbb{N}\ , \ \ \big|\Phi_{0:k}(b_0) \setminus \{\delta_\partial\}\big| \leq |\mathrm{supp}(b_0)|\ . \tag{52}$$

Proof. The proof relies on postponed Lemma 24 from which we obtain that the mappings Φ_k satisfy Equation (47) for all $k \in \mathbb{N}$, and on Lemma 22.

First, as a preliminary fact, we have that, for all $\mu \in \Delta(\overline{\mathbb{X}})$, $\mathrm{supp}\Big(\big(\mathcal{R}(\mu)\big)_{|\mathbb{X}}\Big) = \mathrm{supp}(\mu_{|\mathbb{X}})$. Indeed, by (17), if $\mu(\mathbb{X}) = 0$, then $\mathrm{supp}\Big(\big(\mathcal{R}(\mu)\big)_{|\mathbb{X}}\Big) = \mathrm{supp}\big((\delta_\partial)_{|\mathbb{X}}\big) = \emptyset = \mathrm{supp}\big(\mu_{|\mathbb{X}}\big)$; whereas if $\mu(\mathbb{X}) \neq 0$, then we have $\mathrm{supp}\Big(\big(\mathcal{R}(\mu)\big)_{|\mathbb{X}}\Big) = \mathrm{supp}\big((\frac{\mu_{|\mathbb{X}}}{\mu(\mathbb{X})}, 0)_{|\mathbb{X}}\big) = \mathrm{supp}\big(\frac{\mu_{|\mathbb{X}}}{\mu(\mathbb{X})}\big) = \mathrm{supp}(\mu_{|\mathbb{X}})$.

Second, we show that the mappings Φ_k satisfy Equation (47) for all $k \in \mathbb{N}$. For that purpose, we fix $k \in \mathbb{N}$, and $b \in \mathbb{B}$ and we successively have

$$\begin{aligned}
\sum_{h \in \Phi_k} \big|\mathrm{supp}\big(h(b)_{|\mathbb{X}}\big)\big| &= \sum_{i \in I_k} \Big|\mathrm{supp}\Big(\big((\mathcal{R} \circ (h^k{}_{\overrightarrow{X_i^k}})_\star)(b)\big)_{|\mathbb{X}}\Big)\Big| \\
&\qquad \text{(by definition of } \Phi_k = \mathcal{R} \circ (\mathbb{G}_k)_\star \text{ and } \mathbb{G}_k \text{ in (51))} \\
&= \sum_{i \in I_k} \Big|\mathrm{supp}\Big(\big((h^k{}_{\overrightarrow{X_i^k}})_\star(b)\big)_{|\mathbb{X}}\Big)\Big| \\
&\qquad \text{(as, by the preliminary fact, } \forall \mu \in \Delta(\overline{\mathbb{X}}),\ \mathrm{supp}\Big(\big(\mathcal{R}(\mu)\big)_{|\mathbb{X}}\Big) = \mathrm{supp}(\mu_{|\mathbb{X}})) \\
&\leq \big|\mathrm{supp}\big(b_{|_{h^{-1}(\sqcup_{i \in I_k} X_i^k)}}\big)\big|
\end{aligned}$$

(by (55) in Lemma 24, applied with $\mathbb{Y} = \mathbb{V} = \mathbb{X}$ and $V = X$, $V_i = X_i^k$ for $i \in I = I_k$)

$$\leq \big|\mathrm{supp}\big(b_{|\mathbb{X}}\big)\big| \,. \qquad \text{(as } h^{-1}(\sqcup_{i \in I_k} X_i^k) \subset \mathbb{X})$$

Third, as the assumptions given in Equation (47) are satisfied, the result follows by Lemma 22. □

We now present the postponed technical Lemma 24.

Lemma 24. *Let $h \in \mathbb{L}(\mathbb{Y}; \mathbb{V})$ be a mapping from the set $\mathbb{Y}$ to the set $\mathbb{V}$ and assume that the sets $\mathbb{Y}$ and $\mathbb{V}$ are both finite. Let $V \subset \mathbb{V}$ be a subset of $\mathbb{V}$. We define the mapping*[5] *$h_V : \mathbb{Y} \to \mathbb{V} \cup \{\partial_\mathbb{V}\}$ taking values in the extended set $\overline{\mathbb{V}} = \mathbb{V} \cup \{\partial_\mathbb{V}\}$ as follows*

$$h_V : y \in \mathbb{Y} \mapsto \begin{cases} h(y) & \text{if} \quad h(y) \in V \,, \\ \partial_\mathbb{V} & \text{elsewhere} \,. \end{cases} \tag{53}$$

Then, for any nonnegative measure μ on the set $\mathbb{Y}$, we have that

$$\Big|\mathrm{supp}\Big(\big((h_V)_\star \mu\big)_{|\mathbb{V}}\Big)\Big| \leq \big|\mathrm{supp}\big(\mu_{|_{h^{-1}(V)}}\big)\big| \,. \tag{54}$$

Moreover, for any finite family $\{V_i\}_{i \in I}$ of pairwise disjoints subsets of $\mathbb{V}$, we have that

$$\sum_{i \in I} \Big|\mathrm{supp}\Big(\big((h_{V_i})_\star \mu\big)_{|\mathbb{V}}\Big)\Big| \leq \big|\mathrm{supp}\big(\mu_{|_{h^{-1}(\sqcup_{i \in I} V_i)}}\big)\big| \,. \tag{55}$$

Proof. We prove Equation (54). Let $\mu \in \Delta(\mathbb{Y})$ be given. First, we note that, if the set $\mathrm{supp}\big(\big((h_V)_\star \mu\big)_{|\mathbb{V}}\big)$ is empty, the result is obvious. Second, we assume that $\mathrm{supp}\big(\big((h_V)_\star \mu\big)_{|\mathbb{V}}\big) \neq \emptyset$ and consider $v \in \mathrm{supp}\big(\big((h_V)_\star \mu\big)_{|\mathbb{V}}\big)$. Thus, v is restricted to belong to $\mathbb{V}$ and, by definition of a pushforward measure, it must satisfy $\mu\big(h_V^{-1}(v)\big) \neq 0$. This implies that $h_V^{-1}(v) \neq \emptyset$ and, using the definition of h_V (in Equation (53)), we obtain that v must belong to V. We conclude that there must exist $y \in h_V^{-1}(v)$ such that $\mu(y) \neq 0$ which, combined with the fact that the mapping h_V^{-1} coincides with the mapping h^{-1} on V, gives that $y \in h^{-1}(v) \cap \mathrm{supp}(\mu)$. Now, consider the set-valued mapping $\Gamma : \mathrm{supp}\big(\big((h_V)_\star \mu\big)_{|\mathbb{V}}\big) \rightrightarrows \mathbb{Y}\,,\ y \mapsto h^{-1}(v) \cap \mathrm{supp}(\mu)$.

By construction, the set-valued mapping Γ takes values in the subsets of $\mathrm{supp}(\mu_{|_{h^{-1}(V)}})$, and we have just proved that it takes values in the nonempty subsets of $\mu_{|_{h^{-1}(V)}}$. Moreover, the set-valued

[5] Note that the mapping h_V is slightly different from $h_{\overrightarrow{V}}$. Indeed $h_{\overrightarrow{V}}$ are defined for self-mappings, whereas h_V is defined for an extended codomain (set of destinations).

mapping Γ is injective as, for all pairs $(v', v'') \in V^2$ of distinct elements, $v' \neq v''$, we must have that $h^{-1}(v') \cap h^{-1}(v'') = \emptyset$, as otherwise there would exist an element $y \in \mathbb{Y}$ such that $h(y) = v'$ and $h(y) = v''$, which is not possible. Thus, the image of Γ is a partition of a subset of $\mathrm{supp}(\mu_{|_{h^{-1}(V)}})$ and we conclude that

$$\left|\mathrm{supp}\big(\big((h_V)_\star\mu\big)_{|_{\mathbb{V}}}\big)\right| = \left|\Gamma\big(\mathrm{supp}\big(\big((h_V)_\star\mu\big)_{|_{\mathbb{V}}}\big)\big)\right| \leq \left|\mathrm{supp}(\mu_{|_{h^{-1}(V)}})\right| ,$$

which gives Equation (54).

Now, we turn to the proof of Inequality (55). We successively have

$$\begin{aligned}\sum_{i\in I}\left|\mathrm{supp}\Big(\big((h_{V_i})_\star\mu\big)_{|_{\mathbb{V}}}\Big)\right| &\leq \sum_{i\in I}\left|\mathrm{supp}\big(\mu_{|_{h^{-1}(V_i)}}\big)\right| && \text{(by (54) for each } i \in I)\\ &= \left|\mathrm{supp}\big(\mu_{|_{\sqcup_{i\in I}h^{-1}(V_i)}}\big)\right|\end{aligned}$$

(as the family of subsets $\{h^{-1}(V_i)\}_{i\in I}$ is composed of pairwise disjoints subsets as it was the case for the family $\{V_i\}_{i\in I}$)

$$= \left|\mathrm{supp}\big(\mu_{|_{h^{-1}(\sqcup_{i\in I}V_i)}}\big)\right| , \qquad \text{(as } h^{-1}(\sqcup_{i\in I}V_i) = \sqcup_{i\in I}h^{-1}(V_i))$$

which concludes the proof. □

This technical Lemma 24 shows that the cardinality of the support of a measure decreases when the measure is transported by a pushforward measure induced by a mapping of the form given by Equation (53). A similar result

$$\forall t \in \mathcal{T} \ , \ \forall b \in \mathbb{B} \ , \ \forall u \in \mathbb{U} \ , \ \sum_{o\in\mathbb{O}}\left|\mathrm{supp}\big(\tau_t(b,u,o)\big)\right| \leq \left|\mathrm{supp}(b)\right| ,$$

is given in (Littman, 1996, Lemma 6.2) but only for the mappings $(\tau_t)_{t\in\mathcal{T}}$ defined in Equation (9), and with a proof not explicitly connected to pushforward measures.

We now present the postponed proof of Lemma 7, presented in page 8.

Proof of Lemma 7. Fix $(u, o) \in \mathbb{U} \times \mathbb{O}$, $t \in \mathcal{T} \setminus \{T\}$, and $b \in \mathbb{B}$ and denote by $X \subset \mathbb{X}$ the subset $X = \big(h^u_{t+1}\big)^{-1}(o)$.

We need to prove that we have

$$\tau_t(b, u, o) = \mathcal{R} \circ (F^{u,o}_t)_\star(b) \ . \tag{56}$$

Using Equation (7), we have that

$$Q_{t+1}(b, u, o) = b\big((h^u_{t+1} \circ f^u_t)^{-1}(o)\big) = b\big((f^u_t)^{-1}(X)\big) \ . \tag{57}$$

Now, using the expression of τ_t in Equation (9) combined with Equation (57) and the definition of X, we obtain, for all $\overline{x} \in \overline{\mathbb{X}}$, that

$$\tau_t(b, u, o)(\overline{x}) = \begin{cases} \dfrac{b\big((f^u_t)^{-1}(\overline{x})\big)\mathbf{1}_X(\overline{x})}{b\big((f^u_t)^{-1}(X)\big)} & \text{if } b\big((f^u_t)^{-1}(X)\big) \neq 0 \ , \\ 0 & \text{otherwise} \ . \end{cases} \tag{58}$$

Then, Equation (56) follows from Lemma 18 applied with the mapping $h = f^u_t$ and with the subset $X = \big(h^u_{t+1}\big)^{-1}(o)$, as we have

$$F^{u,o}_t = f^u_t\xrightarrow[(h^u_{t+1})^{-1}(o)]{} \ , \tag{59}$$

where $f^u_t\xrightarrow[(h^u_{t+1})^{-1}(o)]{}$ is defined in Equation (32).

This ends the proof. □

We now present the postponed proof of Lemma 8.

Proof of Lemma 8. We first prove Equation (24). As a preliminary fact, by applying Lemma 7, Notation (19a) and the definitions of sets $\mathbb{T}_t^{\mathcal{D}}$ and $\mathbb{F}_t^{\mathcal{D}}$ (Equations (20)-(21)), we obtain that, for all time $t \in \mathcal{T} \setminus \{T\}$,

$$\mathbb{T}_t^{\mathcal{D}} = \mathcal{R} \circ (\mathbb{F}_t^{\mathcal{D}})_\star \ . \tag{60}$$

Second, for all times $(t, t') \in \big(\mathcal{T} \setminus \{T\}\big)^2$ and for all pairs of controls and observations $(u, u') \in \mathbb{U}^2$ and $(o, o') \in \mathbb{O}^2$, we can apply Lemma 19 on mappings $F_t^{u,o}$ and $F_{t'}^{u',o'}$. Indeed, by Equation (59), the mappings $F_t^{u,o}$ and $F_{t'}^{u',o'}$ are X-forward mappings. We hence have by Equation (43) that $\mathcal{R} \circ F_t^{u,o} \circ \mathcal{R} \circ F_{t'}^{u',o'} = \mathcal{R} \circ F_t^{u,o} \circ F_{t'}^{u',o'}$. Combined with Equation (60), this leads to

$$\mathbb{T}_{0:t}^{\mathcal{D}} = \mathcal{R} \circ (\mathbb{F}_{0:t}^{\mathcal{D}})_\star \ , \tag{61}$$

i.e. it leads to Equation (24).

Now, let $b_0 \in \Delta(\mathbb{X})$. We prove Equation (23) by induction on $t > 0$. By Definition 2 of the set of reachable beliefs, we have

$$\mathbb{B}_1^{\mathrm{R},\mathcal{D}}(b_0) \overset{(11)}{=} \tau_0\big(\{b_0\}, \mathbb{U}, \mathbb{O}\big) \overset{(20)}{=} \mathbb{T}_0^{\mathcal{D}}(b_0) \overset{(60)}{=} \mathcal{R} \circ \big(\mathbb{F}_0^{\mathcal{D}}\big)_\star(b_0) \ ,$$

i.e. Equation (23) stands at time 1. Now, assuming Equation (23) is true up to time $t \in \mathcal{T} \setminus \{T\}$, $t > 0$, we have

$$\mathbb{B}_{t+1}^{\mathrm{R},\mathcal{D}}(b_0) \overset{(11)}{=} \tau_t\big(\mathbb{B}_t^{\mathrm{R},\mathcal{D}}(b_0), \mathbb{U}, \mathbb{O}\big) \overset{(20)}{=} \mathbb{T}_t^{\mathcal{D}}\big(\mathbb{B}_t^{\mathrm{R},\mathcal{D}}(b_0)\big) \overset{(23)}{=} \mathbb{T}_t^{\mathcal{D}} \circ \mathbb{T}_{0:t-1}^{\mathcal{D}}(b_0) \overset{(19)}{=} \mathbb{T}_{0:t}^{\mathcal{D}}(b_0) \overset{(61)}{=} \mathcal{R} \circ \big(\mathbb{F}_{0:t}^{\mathcal{D}}\big)_\star(b_0) \ .$$

By induction on time t, we hence have Equation (23).

Meanwhile, Equation (25) comes from the definition of $\mathbb{T}_{[\![1,T]\!]}^{\mathcal{D}}$ (see Equation (12)), the definition of $\mathbb{F}^{\mathcal{D}}$ (Equation (22)) and Equation (23). □

We can now present the detailed proof of Theorem 4.

Proof of Theorem 4. Let $b_0 \in \Delta(\mathbb{X})$ be given.

We first prove the inequality $|\mathbb{B}_{[\![1,T]\!]}^{\mathrm{R},\mathcal{D}}(b_0)| \leq (1 + |\mathbb{X}|)^{|\mathrm{supp}(b_0)|}$, before proving the inequality $\big|\mathbb{B}_{[\![1,T]\!]}^{\mathrm{R},\mathcal{D}}(b_0)\big| \leq 1 + |\mathrm{supp}(b_0)||\mathbb{U}|^{|\mathcal{T}|}$.

First, by Lemma (8), we have $\mathbb{B}_{[\![1,T]\!]}^{\mathrm{R},\mathcal{D}}(b_0) = \mathbb{T}^{\mathcal{D}}(b_0)$. We hence have

$$|\mathbb{B}_{[\![1,t]\!]}^{\mathrm{R},\mathcal{D}}(b_0)| \overset{(23)}{=} |\mathbb{T}^{\mathcal{D}}(b_0)| \overset{(24)}{=} \Big|\bigcup_{i=0}^{T-1} \mathbb{T}_{0:i}^{\mathcal{D}}(b_0)\Big| \overset{(45)}{\leq} (1 + |\mathbb{X}|)^{|\mathrm{supp}(b_0)|} \ .$$

The last inequality is given by Equation (45), obtained by applying Lemma 20. As all the elements of $\mathbb{F}_t^{\mathcal{D}}$ are of the form given in Equation (16), the two sequences $\{\mathbb{F}_t^{\mathcal{D}}\}_{t \in [\![0,T-1]\!]}$ and $\{\mathbb{T}_t^{\mathcal{D}}\}_{t \in [\![0,T-1]\!]}$ satisfy the assumptions of Lemma 20 where the role of $\{\Phi_k\}_{k \in \mathbb{N}}$ is taken by $\{\mathbb{T}_t^{\mathcal{D}}\}_{t \in [\![0,T-1]\!]}$ and the role of $\{\mathbb{G}_k\}_{k \in \mathbb{N}}$ is taken by $\{\mathbb{F}_t^{\mathcal{D}}\}_{t \in [\![0,T-1]\!]}$ (the proof of Lemma 7 states that set $\mathbb{F}_t^{\mathcal{D}}$ is an $\big(\overrightarrow{\mathbb{X}}\big)$-mappings set).

We now prove that we have

$$\big|\mathbb{B}_{[\![1,T]\!]}^{\mathrm{R},\mathcal{D}}(b_0)\big| \leq 1 + |\mathrm{supp}(b_0)||\mathbb{U}|^{|\mathcal{T}|} \ , \tag{62}$$

in order to obtain Inequality (13). With the help of the representation of the beliefs evolution mappings given by Lemma 7, Inequality (62) is obtained as an application of Lemma 23 that we detail now.

For each $t \in \mathcal{T} \setminus \{T\}$ and each $u_t \in \mathbb{U}$ we introduce the sets $\mathbb{T}_t^{\mathcal{D},u_t} = \big\{\tau_t(\cdot, u_t, o) \,\big|\, o \in \mathbb{O}\big\}$ and $\mathbb{F}_t^{\mathcal{D},u_t} = \big\{F_t^{u_t,o} \,\big|\, o \in \mathbb{O}\big\}$. Using set notations described in Equations (19), we obtain that $\mathbb{T}_t^{\mathcal{D},u_t} = \mathcal{R} \circ (\mathbb{F}_t^{\mathcal{D},u_t})_\star$. Then, using the definition of $\mathbb{B}_t^{\mathrm{R},\mathcal{D}}(b_0)$ in Equation (11), we have that, for all time $t \in \mathcal{T}$, $t > 0$,

$$\mathbb{B}_t^{\mathrm{R},\mathcal{D}}(b_0) = \bigcup_{u_{0:t-1} \in \mathbb{U}_{0:t-1}} \mathbb{T}_{t-1}^{\mathcal{D},u_{t-1}} \circ \mathbb{T}_{t-2}^{\mathcal{D},u_{t-2}} \circ \cdots \circ \mathbb{T}_0^{\mathcal{D},u_0}(b_0) = \bigcup_{u_{0:t-1} \in \mathbb{U}_{0:t-1}} \mathbb{T}_{0:t-1}^{\mathcal{D},u_{0:t-1}}(b_0) \ . \tag{63}$$

For a fixed sequence of controls $u_{0:t} \in \mathbb{U}_{0:t}$, the associated sequences of mappings $\{\mathbb{T}_t^{\mathcal{D},u_t}\}_{t \in \mathcal{T}}$ and $\{\mathbb{F}_t^{\mathcal{D},u_t}\}_{t \in \mathcal{T}}$ satisfy the assumptions of Lemma 23, where the role of $\{\Phi_k\}_{k \in \mathbb{N}}$ is taken by $\{\mathbb{T}_t^{\mathcal{D},u_t}\}_{t \in [\![-1,T]\!]}$,

the role of $\{\mathbb{G}_k\}_{k\in\mathbb{N}}$ is taken by $\{\mathbb{F}_t^{\mathcal{D},u_t}\}_{t\in[\![-1,T]\!]}$ and the role of the family of disjoint sets $\{X_i^k\}_{i\in I_k}$ is taken by the family $\{(h_t^u)^{-1}(o)\}_{o\in\mathbb{O},t\in[\![-1,T]\!]}$ (the proof of Lemma 7 states that the set $\mathbb{F}_t^{\mathcal{D}}$ is an $(\overrightarrow{\mathbb{X}})$-mappings set). We hence get that

$$\forall t \in \mathcal{T} \setminus \{T\}\ ,\ \left|\mathbb{T}^{\mathcal{D},u_{0:t}}{}_{0:t}(b_0) \setminus \{\delta_\partial\}\right| \leq |\mathrm{supp}(b_0)|\ . \tag{64}$$

Finally, we obtain

$$\begin{aligned}
\left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)\right| &= \Big|\bigcup_{t=1}^{T}\big(\mathbb{B}^{\mathrm{R},\mathcal{D}}_t(b_0)\big)\Big| && \text{(using Equation (12))}\\
&\leq 1 + \Big|\bigcup_{t=1}^{T}\big(\mathbb{B}^{\mathrm{R},\mathcal{D}}_t(b_0)\setminus\{\delta_\partial\}\big)\Big| && \text{(by removing } \delta_\partial \text{ from } \mathbb{B}^{\mathrm{R},\mathcal{D}}_t(b_0) \text{ for all } t)\\
&= 1 + \Big|\bigcup_{t=0}^{T-1}\bigcup_{u_{0:t}\in\mathbb{U}_{0:t}}\big(\mathbb{T}^{\mathcal{D},u_{0:t}}_{0:t}(b_0)\setminus\{\delta_\partial\}\big)\Big| && \text{(using Equation (63))}\\
&\leq 1 + \sum_{t=0}^{T-1}\sum_{u_{0:t}\in\mathbb{U}_{0:t}}\big|\big(\mathbb{T}^{\mathcal{D},u_{0:t}}_{0:t}(b_0)\setminus\{\delta_\partial\}\big)\big| && \text{(as } |A\cup B| \leq |A| + |B|)\\
&\leq 1 + \sum_{t=0}^{T-1}\sum_{u_{0:t}\in\mathbb{U}_{0:t}}|\mathrm{supp}(b_0)| && \text{(using Equation (64))}\\
&\leq 1 + \sum_{t=0}^{T-1}|\mathbb{U}|^{t+1}|\mathrm{supp}(b_0)| && \text{(as } \mathbb{U}_{0:t} = \mathbb{U}^{t+1})\\
&\leq 1 + |\mathbb{U}|\Big(\frac{|\mathbb{U}|^T - 1}{|\mathbb{U}| - 1}\Big)|\mathrm{supp}(b_0)| && \text{(as } \textstyle\sum_{i=0}^{N} x^i = \frac{x^{N+1}-1}{x-1} \text{ for } x \neq 1)\\
&\leq 1 + |\mathbb{U}|^{|\mathcal{T}|}|\mathrm{supp}(b_0)|\ . && \text{(as } |\mathcal{T}| = T+1 \text{ and } |\mathbb{U}| > 1)
\end{aligned}$$

We have established the Inequality (62) and this concludes the proof. □

A.2 Complementary result on (∂)-separated mapping sets

In this subsection, we present complementary results on (∂)-separated mapping sets by applying the framework presented in Appendix A.1. We notably apply the notion of forward and backward mappings, presented in Equations (32) and (33), and the notion of pushforward measures, defined in Equation (15) in §3.3.

First, in §A.2.1, we present and prove the lemmata used in the proofs of §4. Second, in §A.2.2, we present a few examples of Separated Det-Pomdps.

A.2.1 Properties of (∂)-separated mapping sets

Lemma 25. *Let $\mathbb{G}$ be an $(\mathbb{M}, \overleftarrow{\mathbb{X}})$-mappings set as defined in Definition 17. If $\mathbb{M}$ is a separated mapping set, then $\mathbb{G}$ is a (∂)-separated mapping set.*

Proof. Let g_1 and g_2 be two mappings in $\mathbb{G}$. In order to prove that $\mathbb{G}$ is a (∂)-separated mapping set, using Definition 10, we need to prove that the restrictions of the two mappings g_1 and g_2 on the subset $A = g_1^{-1}(\mathbb{X}) \cap g_2^{-1}(\mathbb{X})$ are separated. Using the property of the set $\mathbb{G}$, there exist $m^1 \in \mathbb{M}$ (resp. $m^2 \in \mathbb{M}$) and $X_1 \subset \mathbb{X}$ (resp. $X_2 \subset \mathbb{X}$) such that $g_1 = m^1{}_{\overleftarrow{X_1}}$ (resp. $g_2 = m^2{}_{\overleftarrow{X_2}}$). Combined with the definition of $m^1{}_{\overleftarrow{X_1}}$ in Equation (33), this gives that $g_1^{-1}(\mathbb{X}) = (m^1)^{-1}(X_1)$ (resp. $g_2^{-1}(\mathbb{X}) = (m^2)^{-1}(X_2)$). We therefore obtain the equality $A = (m^1)^{-1}(X_1) \cap (m^2)^{-1}(X_2)$.

First, if the set A is empty, it is immediate to prove that g_1 and g_2 are (∂)-separated. Second, assuming that A is not empty and using again the fact that $g_1 = m^1{}_{\overleftarrow{X_1}}$, we obtain that g_1 coincides with m^1 on the set A, and in the same way we obtain that g_2 coincides with m^2 on the set A.

Now, as m^1 and m^2 belong to a separated mapping set, they are separated mappings, and therefore their restrictions to A are also separated. We conclude that the restrictions of g_1 and g_2 on the subset $A = {g_1}^{-1}(\mathbb{X}) \cap {g_2}^{-1}(\mathbb{X})$ are separated. This ends the proof. $\square$

A direct consequence of Lemma 25 is the following Corollary 26.

Corollary 26. *Let $\{\mathbb{M}_k\}_{k\in\mathbb{N}}$ be a sequence of sets of self-mappings on the set $\overline{\mathbb{X}}$. Let $\{\mathbb{G}_k\}_{k\in\mathbb{N}}$ be a sequence of sets of self-mappings on the set $\overline{\mathbb{X}}$, such that, for all $k \in \mathbb{N}$, $\mathbb{G}_k$ is an $\big(\mathbb{M}_k, \overleftarrow{\mathbb{X}}\big)$-mappings set. If the set $\cup_{k\in\mathbb{N}}\big(\mathbb{M}_k \circ \mathbb{M}_{k-1} \circ \cdots \circ \mathbb{M}_0\big)$ of mappings is a separated mapping set, then the set $\cup_{k\in\mathbb{N}}\big(\mathbb{G}_k \circ \mathbb{G}_{k-1} \circ \cdots \circ \mathbb{G}_0\big)$ is a (∂)-separated mapping set.*

Proof. Let $\mathbb{G}_1$ and $\mathbb{G}_2$ be respectively an $\big(\mathbb{M}_1, \overleftarrow{\mathbb{X}}\big)$-mappings set and an $\big(\mathbb{M}_2, \overleftarrow{\mathbb{X}}\big)$-mappings set. Then, we have that

$$\begin{aligned}
\mathbb{G}_1 \circ \mathbb{G}_2 &= \big\{g_1 \circ g_2 \,\big|\, g_1 \in \mathbb{G}_1 \text{ and } g_2 \in \mathbb{G}_2\big\} && \text{(by Notation (19b))}\\
&\subset \big\{{m^1}_{\overleftarrow{X_1}} \circ {m^2}_{\overleftarrow{X_2}} \,\big|\, m^1 \in \mathbb{M}_1\ ,\ \ m^2 \in \mathbb{M}_2\ ,\ \ X_1 \subset \mathbb{X}\ ,\ \ X_2 \subset \mathbb{X}\big\} && \text{(by (35))}\\
&\subset \big\{(m^1 \circ m^2)_{\overleftarrow{X_2 \cap (m^2)^{-1}(X_1)}} \,\big|\, m^1 \in \mathbb{M}_1\ ,\ \ m^2 \in \mathbb{M}_2\ ,\ \ X_1 \subset \mathbb{X}\ ,\ \ X_2 \subset \mathbb{X}\big\} && \text{(by (36))}\\
&\subset \big\{m_X \,\big|\, m \in \mathbb{M}_1 \circ \mathbb{M}_2 \text{ and } X \subset \mathbb{X}\big\}\ .
\end{aligned}$$

We have obtained that $\mathbb{G}_1 \circ \mathbb{G}_2$ is a $\big(\mathbb{M}_1 \circ \mathbb{M}_2, \overleftarrow{\mathbb{X}}\big)$-mappings set. Thus, if $\mathbb{M}_1 \circ \mathbb{M}_2$ is a separated mapping set, then the set $\mathbb{G}_1 \circ \mathbb{G}_2$ is a (∂)-separated mapping set by using Lemma 25. The end of the proof follows by induction on the number of compositions of sets, and by straightforward arguments when considering unions of $\big(\overleftarrow{\mathbb{X}}\big)$-mappings sets. $\square$

Before presenting bounds on the cardinality of a (∂)-separated mapping set, we present Lemma 27.

Lemma 27. *Let $\mathbb{J} \subset \mathbb{L}(\mathbb{X};\mathbb{Y})$ be a set of mappings from the finite set $\mathbb{X}$ to the finite set $\mathbb{Y}$. Assume that for all pairs of mappings $(j, j') \in \mathbb{J}^2$, if there exists $x \in \mathbb{X}$ such that $j(x) = j'(x)$, then $j = j'$. Then, we have that*

$$|\mathbb{J}| \leq |\mathbb{Y}|\ . \tag{65}$$

Proof. Fix $\overline{x} \in \mathbb{X}$ and consider the evaluation mapping $\gamma_{\overline{x}} : \mathbb{J} \to \mathbb{Y}$ defined by $\gamma_{\overline{x}}(j) = j(\overline{x})$ for all $j \in \mathbb{J}$. The image $\gamma_{\overline{x}}(\mathbb{J})$ of the set $\mathbb{J}$ by the mapping $\gamma_{\overline{x}}$ is indeed the subset $\{j(\overline{x}) \,|\, j \in \mathbb{J}\}$ of $\mathbb{Y}$. First, the codomain of the mapping $\gamma_{\overline{x}}$ being the finite set $\mathbb{Y}$, we immediately have that

$$\big|\gamma_{\overline{x}}(\mathbb{J})\big| \leq |\mathbb{Y}|\ . \tag{66}$$

Second, the mapping $\gamma_{\overline{x}}$ is injective. Indeed, using the assumption on the set $\mathbb{J}$, two distinct mappings j and j' in the set $\mathbb{J}$ must satisfy $\gamma_{\overline{x}}(j) = j(\overline{x}) \neq j'(\overline{x}) = \gamma_{\overline{x}}(j')$. Thus, we must have the equality $|\mathbb{J}| = \big|\gamma_{\overline{x}}(\mathbb{J})\big|$ which, combined with Equation (66), gives Inequality (65), and concludes the proof. $\square$

We now use the previous Lemma 27 to bound the cardinality of a (∂)-separated mapping set.

Lemma 28. *Let $\overline{\mathbb{X}} = \mathbb{X} \cup \{\partial\}$, and a (∂)-separated mapping set $\mathbb{G}$ of self-mappings on the set $\overline{\mathbb{X}}$. Moreover, assume that, for all $g \in \mathbb{G}$, $g(\partial) = \partial$. For any subsets X and X' of the set $\overline{\mathbb{X}}$, we define $\mathbb{G}_{X\to X'}$ as follows*

$$\mathbb{G}_{X\to X'} = \big\{g \in \mathbb{G} \,\big|\, g^{-1}(\mathbb{X}) = X, g(X) \subset X'\big\}\ . \tag{67}$$

Then, we have

$$\big|\mathbb{G}_{X\to X'}\big| \begin{cases} \leq |X'| & \text{if } X \subset \mathbb{X}\ ,\\ = 0 & \text{if } X \cap \{\partial\} \neq \emptyset\ . \end{cases} \tag{68}$$

Proof. Fix $X \subset \overline{\mathbb{X}}$ and $X' \subset \overline{\mathbb{X}}$. First, we consider the case where $X \cap \{\partial\} \neq \emptyset$. As we have assumed that $g(\partial) = \partial$, for all $g \in \mathbb{G}$, we obtain that $g^{-1}(\mathbb{X}) \cap \{\partial\} = \emptyset$. Thus, we conclude that $|\mathbb{G}_{X\to X'}| = |\emptyset| = 0$. Second, we consider the case where $X \subset \mathbb{X}$ and consider the mapping

$$\Gamma : \mathbb{G}_{X\to X'} \to {X'}^X\ ,\ \ g \mapsto g_{|X}\ . \tag{69}$$

The mapping Γ is injective. Indeed, if two mappings in $\mathbb{G}_{X\to X'}$ have the same restriction on X, they coincide on $\overline{\mathbb{X}}$ as they are both constant on the set $\overline{\mathbb{X}} \setminus X$ with value ∂. We therefore obtain that

$$\left|\mathbb{G}_{X\to X'}\right| = \left|\Gamma(\mathbb{G}_{X\to X'})\right| . \tag{70}$$

Now, the set $\mathbb{G}' = \Gamma(\mathbb{G}_{X\to X'})$ is a subset of mappings from X to X'. As $\mathbb{G}$ is a (∂)-separated mapping set, we obtain that $\mathbb{G}'$ is a separated set of mappings from X to X'. Indeed, consider a pair of mappings $(g'_1, g'_2) \in \mathbb{G}'^2$ and assume that there exists $x \in X$ such that $g'_1(x) = g'_2(x)$. Using the definition of $\mathbb{G}'$, we have that $g'_1(x)$ and $g'_2(x)$ are both non equal to ∂. Moreover, there exists g_1 and g_2 in $\mathbb{G}_{X\to X'}$ such that $g'_1 = \Gamma(g_1)$ and $g'_2 = \Gamma(g_2)$. Using again the definition of $\mathbb{G}' = \Gamma(\mathbb{G}_{X\to X'})$ we obtain that $g_1(x) = g_2(x) \neq \partial$. Now, as $\mathbb{G}$ is a (∂)-separated mapping set, we obtain that the two mappings g_1 and g_2 coincide on X since they both do not take the value ∂ on X. We conclude that their restrictions on X, the mappings g'_1 and g'_2, coincide. Using Lemma 27 in Subsection A.2 we obtain that

$$\left|\Gamma(\mathbb{G}_{X\to X'})\right| \leq |X'| , \tag{71}$$

which, combined with Equation (70), gives Equation (68). This concludes the proof. □

We now present the postponed proof of Proposition 12, presented in page 11.

Proof of Proposition 12. The proof of Proposition 12 is a direct consequence of Corollary 26.

We assume that the set $\bigcup_{t\in\mathcal{T}} f_{0:t}^{\mathbb{U}^{t+1}} = \{f_{0:t}^{u_{0:t}} \,|\, \forall t \in \mathcal{T} \setminus \{T\}, \forall u_{0:t} \in \mathbb{U}^{t+1}\}$ of the composition of the evolution functions of Problem (2) is a separated mapping set. We then prove that Problem (2) is a Separated Det-Pomdp.

First, for all time t and for all pair $(u, o) \in \mathbb{U} \times \mathbb{O}$, we have $F_t^{u,o} = f_t^u \xrightarrow[(h_{t+1}^u)^{-1}(o)]{}$ (see Equation (59)). Thus, by Equation (34a), there exists $X \subset \mathbb{X}$ such that $F_t^{u,o} = f_t^u \underset{X}{\leftarrow}$. Hence, $\mathbb{F}_t^{\mathcal{D}}$ is of the same form as in Equation (51), with the role of set Φ_k taken by $\{f_t^{\mathbb{U}}\}$.

We hence have that $\mathbb{F}^{\mathcal{D}} = \bigcup_{t\in\mathcal{T}} \mathbb{F}_{0:t}^{\mathcal{D}}$ is a (∂)-separated mapping set by Corollary 26, where the role of $\{\mathbb{G}_k\}_{k\in\mathbb{N}}$ is taken by $\{\mathbb{F}_t^{\mathcal{D}}\}_{t\in\mathcal{T}\setminus\{T\}}$ and the role of $\{\Phi_k\}_{k\in\mathbb{N}}$ is taken by $\{f_t^{\mathbb{U}}\}_{t\in\mathcal{T}\setminus\{T\}}$.

Therefore, as $\mathbb{F}^{\mathcal{D}}$ is a (∂)-separated mapping set, Problem (2) is a Separated Det-Pomdp. □

We now present the postponed proof of Theorem 13, presented in page 11.

Proof of Theorem 13. We start by giving preliminary bounds on $\left|\left(\mathcal{R} \circ (\mathbb{F}_{X\to\mathbb{X}}^{\mathcal{D}})_\star\right)(b_0) \setminus \{\delta_\partial\}\right|$, where $\mathbb{F}_{X\to\mathbb{X}}^{\mathcal{D}}$ is defined by Equation (67), i.e.

$$\mathbb{F}_{X\to\mathbb{X}}^{\mathcal{D}} = \left\{F \in \mathbb{F}^{\mathcal{D}} \,\middle|\, F^{-1}(\mathbb{X}) = X, F(X) \subset \mathbb{X}\right\} ,$$

where $\mathbb{F}^{\mathcal{D}}$ is defined in Equation (21). We consider three cases depending on the cardinality of the subset X:

1. When $|X| = 0$, we have that $X = \emptyset$ and $\left(\mathcal{R} \circ (\mathbb{F}_{\emptyset\to\mathbb{X}}^{\mathcal{D}})_\star\right)(b_0) \setminus \{\delta_\partial\} = \emptyset$, and thus

$$\left|\left(\mathcal{R} \circ (\mathbb{F}_{X\to\mathbb{X}}^{\mathcal{D}})_\star\right)(b_0) \setminus \{\delta_\partial\}\right| = 0 . \tag{72a}$$

2. When $|X| = 1$, we have that $\left(\mathcal{R} \circ (\mathbb{F}_{X\to\mathbb{X}}^{\mathcal{D}})_\star\right)(b_0) \setminus \{\delta_\partial\} \subset \left\{\delta_x \,\middle|\, x \in \mathbb{X}\right\}$, as the only probability distributions of $\Delta(\mathbb{X})$ with support of cardinality at most 1 are the vertices $\left\{\delta_x \,\middle|\, x \in \mathbb{X}\right\}$ of the simplex $\Delta(\mathbb{X})$, and thus

$$\left|\left(\mathcal{R} \circ (\mathbb{F}_{X\to\mathbb{X}}^{\mathcal{D}})_\star\right)(b_0) \setminus \{\delta_\partial\}\right| \leq \left|\left\{\delta_x \,\middle|\, x \in \mathbb{X}\right\}\right| = |\mathbb{X}| . \tag{72b}$$

3. For $|X| \geq 2$, we have by Lemma 28 in Appendix A.1, applied with $\mathbb{G} = \mathbb{F}$ (as $\mathbb{F}$ is a (∂)-separated mapping set) that

$$\left|\left(\mathcal{R} \circ (\mathbb{F}_{X\to\mathbb{X}}^{\mathcal{D}})_\star\right)(b_0) \setminus \{\delta_\partial\}\right| \leq \left|(\mathbb{F}_{X\to\mathbb{X}}^{\mathcal{D}})_\star\right| \leq |\mathbb{X}| . \tag{72c}$$

We have by Equation (25) that $\left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)\right| = |\mathbb{T}^{\mathcal{D}}(b_0)|$. We now detail the cardinality of $\mathbb{T}^{\mathcal{D}}(b_0)$:

$$\begin{aligned}
\left|\mathbb{T}^{\mathcal{D}}(b_0) \setminus \{\delta_\partial\}\right| &= \left|\big(\mathcal{R} \circ (\mathbb{F}^{\mathcal{D}})_\star\big)(b_0) \setminus \{\delta_\partial\}\right| \\
&= \left|\left(\mathcal{R} \circ \Big(\bigcup_{X \subset \mathbb{X}} \mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}}\Big)_\star\right)(b_0) \setminus \{\delta_\partial\}\right| && \left(\text{as } \textstyle\bigcup_{X \subset \mathbb{X}} \mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}} = \mathbb{F}^{\mathcal{D}}\right) \\
&= \left|\bigcup_{X \subset \mathbb{X}} \big(\mathcal{R} \circ (\mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}})_\star\big)(b_0) \setminus \{\delta_\partial\}\right|
\end{aligned}$$

as $\forall (F, F') \in \left(\mathbb{F}^{\mathcal{D}}\right)^2$, $\mathcal{R} \circ \big(F \cup F'\big) = \mathcal{R} \circ F \cup \mathcal{R} \circ F'$,

$$= \left|\bigcup_{X \subset \mathrm{supp}(b_0)} \big(\mathcal{R} \circ (\mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}})_\star\big)(b_0) \setminus \{\delta_\partial\}\right|$$

as $\big(\mathcal{R} \circ (\mathbb{F}^{\mathcal{D}}_{X \cap \mathrm{supp}(b_0) \to \mathbb{X}})_\star\big)(b_0) = \big(\mathcal{R} \circ (\mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}})_\star\big)(b_0)$ by Equation (38) in Lemma 18,

$$\begin{aligned}
&\leq \sum_{X \subset \mathrm{supp}(b_0)} \left|\big(\mathcal{R} \circ (\mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}})_\star\big)(b_0) \setminus \{\delta_\partial\}\right| \\
&= \sum_{k \geq 0} \sum_{\substack{X \subset \mathrm{supp}(b_0) \\ |X| = k}} \left|\big(\mathcal{R} \circ (\mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}})_\star\big)(b_0) \setminus \{\delta_\partial\}\right| \\
&\leq |\mathbb{X}| + \sum_{\substack{X \subset \mathrm{supp}(b_0) \\ |X| \geq 2}} |\mathbb{X}| && \text{(by Equations (72))} \\
&= |\mathbb{X}| + \big(2^{|\mathrm{supp}(b_0)|} - |\mathrm{supp}(b_0)| - 1\big)|\mathbb{X}| \ , && (73)
\end{aligned}$$

where the last equality comes from the fact that $\left|\{X \subset \mathrm{supp}(b_0) \,|\, |X| \geq 2\}\right|$ is given by

$$\left|\{X \subset \mathrm{supp}(b_0) \,|\, |X| \geq 2\}\right| = \underbrace{\left|\{X \subset \mathbb{X} \,\big|\, X \subset \mathrm{supp}(b_0)\}\right|}_{2^{|\mathrm{supp}(b_0)|}} - \underbrace{\left|\{X \subset \mathrm{supp}(b_0) \,\big|\, |X| = 1\}\right|}_{=|\mathrm{supp}(b_0)|} - \underbrace{\left|\{X \subset \mathrm{supp}(b_0) \,\big|\, |X| = 0\}\right|}_{=1} \ .$$

We hence obtain that

$$\left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)\right| \overset{(24)}{=} |\mathbb{T}^{\mathcal{D}}(b_0)| \overset{(73)}{\leq} 1 + \big(2^{|\mathrm{supp}(b_0)|} - |\mathrm{supp}(b_0)|\big)|\mathbb{X}| \ .$$

This ends the proof. □

We now present examples of Separated Det-Pomdps.

A.2.2 Examples of Separated Det-Pomdps

In this subsection, we present examples of Separated Det-Pomdps. Indeed, a direct consequence of Proposition 12 is that, if the evolution mappings of a Det-Pomdp belong to a separated mapping set, then the Det-Pomdp is a Separated Det-Pomdp. We now present examples of such evolution mappings.

Corollary 29. *Consider a* Det-Pomdp *optimization problem given by Problem* (2) *which satisfies the finite sets Assumption 1. The notations are those of Problem* (2)*. Assuming that, for all time $t \in \mathcal{T} \setminus \{T\}$, there exist mappings g_t such that, for all states $x \in \mathbb{X} \subset \mathbb{R}^n$,*

$$f_t(x, u) = x + g_t(u) \ , \tag{74}$$

then Problem (2) *is a Separated* Det-Pomdp*.*

Proof. This corollary is a direct result of Proposition 12. Indeed, we only need to prove that $\cup_{t\in\mathcal{T}}\big(f_{0:t}^{\mathbb{U}^{t+1}}\big)$ is a separated mapping set.

Let $t_1 \leq t'_1$ and $t_2 \leq t'_2$ be such that $[\![t_1, t'_1]\!] \subset \mathcal{T}$ and $[\![t_2, t'_2]\!] \subset \mathcal{T}$. Let $u_{t_1:t'_1} \in \mathbb{U}^{t'_1-t_1+1}$ and $u'_{t_2:t'_2} \in \mathbb{U}^{t'_2-t_2+1}$ be two sequences of controls. We have, by using Equation (74), that $f_{t_1:t'_1}^{u_{t_1:t'_1}} : \mathbb{X} \to \mathbb{X}, x \mapsto x + \sum_{t\in[\![t_1,t'_1]\!]} g_t(u_t)$, and $f_{t_2:t'_2}^{u'_{t_2:t'_2}} : \mathbb{X} \to \mathbb{X}, x \mapsto x + \sum_{t\in[\![t_2,t'_2]\!]} g_t(u'_t)$.

If there exists a state $x \in \mathbb{X}$ such that $f_{t_1:t'_1}^{u_{t_1:t'_1}}(x) = f_{t_2:t'_2}^{u'_{t_2:t'_2}}(x)$, we hence have $\sum_{t\in[\![t_1,t'_1]\!]} g_t(u_t) = \sum_{t\in[\![t_2,t'_2]\!]} g_t(u'_t)$.

Thus $f_{t_1:t'_1}^{u_{t_1:t'_1}}(x) = f_{t_2:t'_2}^{u'_{t_2:t'_2}}(x) \Rightarrow f_{t_1:t'_1}^{u_{t_1:t'_1}} = f_{t_2:t'_2}^{u'_{t_2:t'_2}}$. Therefore, the set $\cup_{t\in\mathcal{T}}\big(f_{0:t}^{\mathbb{U}^{t+1}}\big) = \{f_{0:t}^{u_{0:t}} \,|\, \forall t \in \mathcal{T} \setminus \{T\}, \forall u_{0:t} \in \mathbb{U}^{t+1}\}$ of composition of the evolution mappings is a separated mapping set. We conclude by Proposition 12 that Problem (2) is a Separated Det-Pomdp. □

Corollary 30. *Consider a* Det-Pomdp *optimization problem given by Problem* (2) *which satisfies the finite sets Assumption 1. The notations are those of Problem* (2)*. Assuming that, for all time $t \in \mathcal{T}\setminus\{T\}$, there exist mappings g_t such that for all states $x \in \mathbb{X} \subset \mathbb{R}^n$,*

$$f_t(x, u) = x \times g_t(u) \ , \tag{75}$$

and assuming that $0 \notin \mathbb{X}$, then Problem (2) *is a Separated* Det-Pomdp.

Proof. Let $t_1 \leq t'_1$ and $t_2 \leq t'_2$ such that $[\![t_1, t'_1]\!] \subset \mathcal{T}$ and $[\![t_2, t'_2]\!] \subset \mathcal{T}$. Let $u_{t_1:t'_1} \in \mathbb{U}^{t'_1-t_1+1}$ and $u'_{t_2:t'_2} \in \mathbb{U}^{t'_2-t_2+1}$ be two sequences of controls . We have, by using Equation (75), $f_{t_1:t'_1}^{u_{t_1:t'_1}} : \mathbb{X} \to \mathbb{X}, x \mapsto x \times \prod_{t\in[\![t_1,t'_1]\!]} g_t(u_t)$, and $f_{t_2:t'_2}^{u'_{t_2:t'_2}} : \mathbb{X} \to \mathbb{X}, x \mapsto x \times \prod_{t\in[\![t_2,t'_2]\!]} g_t(u'_t)$.

If there exists a state $x \in \mathbb{X}$ such that $f_{t_1:t'_1}^{u_{t_1:t'_1}}(x) = f_{t_2:t'_2}^{u'_{t_2:t'_2}}(x)$, we hence have, as $x \neq 0$, $\prod_{t\in[\![t_1,t'_1]\!]} g_t(u_t) = \prod_{t\in[\![t_2,t'_2]\!]} g_t(u'_t)$.

Thus $f_{t_1:t'_1}^{u_{t_1:t'_1}}(x) = f_{t_2:t'_2}^{u'_{t_2:t'_2}}(x) \Rightarrow f_{t_1:t'_1}^{u_{t_1:t'_1}} = f_{t_2:t'_2}^{u'_{t_2:t'_2}}$. Therefore, the set of compositions of the evolution functions $\cup_{t\in\mathcal{T}}\big(f_{0:t}^{\mathbb{U}^{t+1}}\big) = \{f_{0:t}^{u_{0:t}} \,|\, \forall t \in \mathcal{T} \setminus \{T\}, \forall u_{0:t} \in \mathbb{U}^{t+1}\}$ is a separated mapping set. □

References

K. J. Åström. Optimal control of Markov processes with incomplete state information. *Journal of Mathematical Analysis and Applications*, 10(1):174–205, Feb. 1965. doi: 10.1016/0022-247X(65)90154-X.

R. Bellman. *Dynamic programming.* Princeton Univ. Pr, Princeton, NJ, 1957.

D. P. Bertsekas. *Dynamic Programming and Optimal Control.* Athena Scientific, Belmont, Massachusetts, second edition, 2000. Volumes 1 and 2.

D. P. Bertsekas and S. E. Shreve. *Stochastic optimal control: the discrete time case.* Number v. 139 in Mathematics in science and engineering. Academic Press, New York, 1978.

B. Bonet. Deterministic pomdps revisited. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence*, UAI '09, page 59–66, Arlington, Virginia, USA, 2009. AUAI Press.

H. Geffner and B. Bonet. Solving Large POMDPs by Real Time Dynamic Programming. In *Proc. Fall AAAI Symposium on POMDPS*, Orlando, FL, 1998.

H. Kurniawati, D. Hsu, and W. Sun Lee. SARSOP: Efficient Point-Based POMDP Planning by Approximating Optimally Reachable Belief Spaces. In *Robotics: Science and Systems IV*. Robotics: Science and Systems Foundation, June 2008. doi: 10.15607/RSS.2008.IV.009.

M. L. Littman. *Algorithms for Sequential Decision Making.* PhD thesis, Brown University, 1996.

J. Pajarinen and V. Kyrki. Robotic manipulation of multiple objects as a POMDP. *Artificial Intelligence*, 247:213–228, June 2017. doi: 10.1016/j.artint.2015.04.001.

M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley Series in Probability and Statistics. Wiley, 1 edition, Apr. 1994. doi: 10.1002/9780470316887.

R. D. Smallwood and E. J. Sondik. The Optimal Control of Partially Observable Markov Processes over a Finite Horizon. *Operations Research*, 21(5):1071–1088, Oct. 1973. doi: 10.1287/opre.21.5.1071.

L. N. Steimle, D. L. Kaufman, and B. T. Denton. Multi-model markov decision processes. *IISE Transactions*, 53(10):1124–1139, 2021. doi: 10.1080/24725854.2021.1895454.

R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*. The MIT Press, second edition, 2018.

D. J. White. A Survey of Applications of Markov Decision Processes. *The Journal of the Operational Research Society*, 44(11):1073, Nov. 1993. doi: 10.2307/2583870.