Design and operation management of oil-fields taking into account partially observed uncertainties

Cyrille Vessaire

THÈSE DE DOCTORAT
de l'École des Ponts ParisTech

Design and operation management of oil-fields taking into account partially observed uncertainties

École doctorale MSTIC, Mathématiques, Sciences et Technologies de l'Information et de la Communication

Mathématiques Appliquées

Thèse préparée au laboratoire CERMICS,
au sein de l'équipe Optimisation et Systèmes

Thèse soutenue le 16 décembre 2022, par
Cyrille VESSAIRE

Composition du jury

Ignacio GROSSMANN Pr, Carnegie Mellon University	*Rapporteur*
Bernardo PAGNONCELLI Pr, SKEMA Business School	*Rapporteur*
Jean-Philippe CHANCELIER Pr, CERMICS	*Directeur de Thèse*
Marianne AKIAN Pr, INRIA and École Polytechnique	*Examinateur*
Michel DE LARA Pr, CERMICS	*Examinateur*
Anna ROBERT Dr, TotalEnergies	*Examinateur*
Alejandro RODRÌGUEZ-MARTÌNEZ Dr, TotalEnergies	*Examinateur*
Pierre CARPENTIER Dr, CERMICS	*Invité*

Résumé

La gestion de l'exploitation des champs pétroliers consiste à piloter un réseau de production pétrolière - composé d'un réservoir (une formation géologique contenant des hydrocarbures), de puits et de oléoducs reliant ces puits - afin de produire des hydrocarbures. Une particularité de la gestion des champs pétroliers est qu'au début de l'exploitation du réservoir, nous n'avons qu'une connaissance partielle du contenu du réservoir, à savoir une distribution de probabilité de son état initial. Ensuite, nous obtenons des informations sur le contenu du réservoir au fur et à mesure qu'il se vide, c'est-à-dire au fur et à mesure que nous appliquons des contrôles. L'objectif de cette thèse est de gérer un champ pétrolier en tenant compte de cette particularité.

Dans un premier temps, nous proposons une nouvelle formulation du problème de gestion d'un réseau de production pétrolier pour le cas totalement observé. Le modèle proposé diffère des pratiques courantes où le réservoir est approximé par des courbes de déclin ou par des simulateurs de type boîte-noire. Nous modélisons le réservoir comme un système dynamique contrôlé (non linéaire) en utilisant des équations de bilan de matière, en supposant que les fluides suivent un *modèle de black-oil* et que le réservoir a un comportement de type "tank". L'état du système dynamique a cinq dimensions : le volume total de pétrole, de gaz et d'eau dans le réservoir, le volume total des pores et la pression du réservoir. Nous utilisons un algorithme de programmation dynamique pour résoudre numériquement le problème d'optimisation multi-étapes sur deux instances spécifiques où la dimension de l'état peut être réduite. Plus précisément, la première application consiste à optimiser la production d'un réservoir de gaz qui est subdivisé en deux parties, ce qui conduit à un état bidimensionnel (une dimension par partie), tandis que la seconde application est l'optimisation d'un réservoir de pétrole avec injection d'eau, ce qui conduit à un état bidimensionnel. Le principal avantage de notre approche est qu'elle permet de traiter des cas avec des réservoirs interconnectés et de la réinjection de fluides.

Deuxièmement, nous étudions une classe de problèmes intéressante pour la gestion d'un problème de réseau de production pétrolière sous observation partielle, à savoir les problèmes d'optimisation par processus de décision de Markov partiellement observés (POMDP). Plus précisément, nous améliorons les résultats de Littman concernant une sous-classe de POMDPs appelée DET-POMDP. Les DET-POMDPs sont intéressants car la taille des ensembles d'états atteignables utilisés dans les algorithmes de programmation dynamique est bornée, que ce soit dans le cas d'un horizon fini ou infini. Nous donnons des améliorations des limites présentées par Littman [1996]. Ensuite, en ajoutant des conditions supplémentaires sur la dynamique et les observations, nous définissons une sous-classe de DET-POMDPs dont les limites sont encore améliorées : Separated DET-POMDP. Grâce à cette sous-classe, nous sommes en mesure de repousser davantage la malédiction de la dimensionnalité. Nous obtenons ainsi des problèmes partiellement observés qui sont traitables par des algorithmes de programmation dynamique, alors qu'ils sont généralement résolus par approximation des fonctions de valeur de Bellman lorsque l'on considère les algorithmes généraux de POMDPs.

Troisièmement, nous élargissons la formulation totalement observée du problème de gestion d'un réseau de production de pétrole afin de prendre en compte l'observation partielle du contenu du réservoir. Ceci conduit à la classe de Separated DET-POMDP précédemment discutée. Nous sommes donc en mesure d'utiliser la Programmation Dynamique pour trouver la planification optimale de la production. Nous résolvons à nouveau numériquement le problème d'optimisation partiellement observé sur les deux cas présentés dans le cas totalement observé : un réservoir de gaz sec à deux réservoirs, et un réservoir de pétrole avec injection d'eau.

Abstract

The operation management of oil-fields consists in piloting an oil production network – composed of a reservoir (geological formation containing hydrocarbons), wells and pipes coupling those wells – in order to produce some hydrocarbons. A key particularity of the management of oil-fields is that, at the beginning of the reservoir exploitation, we have only partial knowledge of the content of the reservoir, namely a probability distribution of its initial state. Then, we obtain more and more information on the content of the reservoir as it depletes, i.e as we apply controls. The aim of this thesis is to manage an oil field while taking into account this key particularity.

First, we propose a new formulation of the management of an oil production network problem for the fully observed case. The proposed model differs from common practices where the reservoir of the oil production network is approximated by decline curves or by black-box simulators. We model the reservoir as a controlled (non-linear) dynamical system by using material balance equations, under the assumption that the fluids follow a black-oil model and that the reservoir has a tank-like behavior. The state of the dynamical system has five dimensions: the total volume of respectively oil, gas, and water in the reservoir; the total pore volume; and the reservoir pressure. We use a dynamic programming algorithm to numerically solve the multistage optimization problem on two specific instances of the general optimization problem where the state dimension can be reduced. More precisely, the first numerical application consists in optimizing the production of a dry gas reservoir which is subdivided into two tanks and which leads to a two-dimensional state (one dimension per tank), whereas the second numerical application tackles an oil reservoir with water injection which leads to a two-dimensional state. The key advantage of our approach is that it handles interconnected tanks and allows for optimization beyond first recovery of oil.

Second, we study a class of problems which is of interest for the management of an oil production network problem under partial observation, namely Partially Observed Markov Decision Process (POMDP) optimization problems. More specifically, we improve on results by Littman [1996] concerning a subclass of POMDP called DET-POMDP. DET-POMDPs are of interest as the size of the sets of reachable states used in dynamic programming algorithms are bounded, be it in the finite or infinite horizon cases. We give improvements of the bounds presented by Littman [1996]. Then, by adding further conditions on the dynamics and observations, we define a sub-class of DET-POMDP whose bounds are again improved: Separated DET-POMDP. Through this subclass, we are able to further push back the curse of dimensionality. We hence obtain partially observed problems that are tractable by Dynamic Programming algorithms, whereas they are usually solved by approximating the Bellman value functions when considering general POMDPs algorithms.

Third, we expand on the fully observed formulation of the management of an oil production network problem in order to take into account the partial observation of the content of the reservoir. This leads to the previously discussed Separated DET-POMDP class. We are therefore able to use Dynamic Programming to find the optimal production planning. We once again numerically solve the partially observed optimization problem on the two cases

presented in the fully observed case: a two tanks dry gas reservoir, and an oil reservoir with water injection.

Remerciements

Comme le veut la tradition, mes premiers remerciements vont à mon directeur de thèse, Jean-Philippe Chancelier. J'ai beaucoup appris grâce à toi. Tout au long de ma thèse, ton bureau a toujours été ouvert, et tu as toujours été disponible pour répondre à mes questions. Merci beaucoup pour tous tes conseils, pour toutes tes relectures et pour la rigueur que tu m'as inculqué. Je suis honoré que tu ait été mon directeur, et je te remercie chaleureusement pour ces trois années riches en enseignement.

Je tiens également à remercier Alejandro Rodríguez-Martínez pour son soutien tout au long de la thèse. Directeur officieux, sans tes apports et conseils, la partie industrielle de cette thèse n'aurait pu prendre forme. Tu as toujours su te rendre disponible malgré ton emploi du temps souvent chargé, et bien entendu pris le temps de trouver les données et de suggérer le dépôt de brevet.

Bien entendu, je remercie également Michel De Lara et Pierre Carpentier. Nos nombreuses réunions, remarques et critiques pertinentes m'ont permis, j'espère, de gagner en clarté et d'améliorer mes présentations et capacités de rédactions. De même, je souhaite remercier Anna Robert pour avoir été présente et d'avoir fait le pont entre TotalEnergies et le CERMICS en début de thèse.

I would like to sincerely thank the reviewers of my thesis, Professor Ignacio Grossmann and Professor Bernardo Pagnoncelli, for taking the time to thoroughly read the full manuscript. I am deeply honored to have both of you on my jury. Dans le même esprit, je remercie Marianne Akian d'avoir accepté de faire partie de mon jury de thèse.

Faire une thèse au CERMICS impliquant nécessairement une certaine logistique, je souhaite chaleureusement remercier Isabelle Simunic et Stéphanie Bonnel pour toutes leurs (nombreuses) aides au cours de ces trois années. De même, il me faut nécessairement remercier les doctorants du deuxième et troisième étages. Tous les ans, des nouveaux visages arrivent, et c'est toujours un plaisir de les découvrir. Malgré les confinements successifs, on a toujours essayé de maintenir un contact et de se soucier des autres, que ce soit les plus anciens ou les plus jeunes. Ce cadre bienveillant rend à mon sens le CERMICS comme l'un des meilleurs environnements de travail (ou de non-travail). Ce fut pour moi un privilège d'avoir pu en faire partie. Je remercie donc, sans ordre particulier (en fait l'ordre alphabétique), autant les anciens et que les nouveaux: Albéric, Adrien, Adèle, Alfred, Benoît, Camila, Clément, Coco, Dylan, Edoardo, Eloïse, Emanuele, Ezechiel, Fabian, Gaspard, Guillaume, Hélène, Henri, Hervé, Jean, Julien, Kacem, Laurent, Léo, Louis, Maël, Michel, Nerea, Noé, Oumaïma, Raian, Régis, Roberta, Robert, Rutger, Solal, Sébastien, Shiva, Thomas, Vitor, William, Zoé, et tous ceux que j'oublie. Bien sur, je remercie également tous les postdocs. Enfin, je remercie également l'École des Ponts en général, autant pour la formation et ses enseignants que pour les différentes rencontres qu'elle a rendu possible, ses élèves, ses clubs.

Bien entendu, je remercie mes amis, notamment Jeanne, Rémi et Romain, ainsi que toute ma famille, et ce pour leur soutien et pour avoir réussi à me supporter toutes ces années.

Pour clore ces remerciements, je tiens bien entendu à remercier Google, Wikipedia, StackOverflow, TeX StackExchange et le raccourci CTRL+Z d'exister et de m'avoir simplifier la vie pendant ces trois années, et certainement permis de conserver ma santé mentale au cours de cette thèse.

Contents

Chapter 1

Introduction en français

Contents

Nous donnons d'abord dans le Paragraphe 1.1 des informations contextuelles sur cette thèse. Ensuite, dans le Paragraphe 1.2, nous présentons le sujet de ce manuscrit : l'optimisation des systèmes de production pétroliers en tenant compte des incertitudes. Troisièmement, dans le Paragraphe 1.3, nous exposons la structure de ce document. Enfin, dans le Paragraphe 1.4, nous détaillons nos principales contributions.

1.1 Contexte de cette thèse

Le sujet de cette thèse est la gestion des systèmes de production pétroliers sous incertitudes, c'est-à-dire comment piloter la production au cours du temps d'un champ pétrolier ou gazier afin de maximiser un critère économique en tenant compte d'incertitudes, comme par exemple l'observation partielle des ressources présentes dans le sous-sol. Cette thèse a été réalisée en collaboration avec l'équipe Integrated Asset Modeling (IAM) de TotalEnergies, qui est chargée de l'intégration des modèles de simulation, du réservoir à la production, afin d'apporter un support aussi bien sur l'évaluation que sur la conception de plans de développement de plans de développement des champs pétrolifères. La coopération a débuté grâce à Rémy Marmier, alors responsable de l'équipe IAM de TotalEnergies, tandis qu'Alejandro Rodríguez-Martínez est devenu l'interlocuteur principal, notamment pour la validation des modèles physiques et les applications industrielles. Anna Robert, qui fait partie de la branche R&D de TotalEnergies, faisait également partie des interlocuteurs. Les travaux de thèse ont abouti à un dépôt de brevet impliquant l'auteur de ce manuscrit, le laboratoire CERMICS de l'ENPC et TotalEnergies. Nous détaillons maintenant le sujet de cette thèse, à savoir l'optimisation de systèmes de production pétroliers.

1.2 Optimisation du système de production pétrolière

Comme il devient de plus en plus complexe et coûteux d'extraire du pétrole, il devient de plus en plus important pour les entreprises productrices d'optimiser la production les champs pétroliers. En effet, les projets pétroliers et gaziers s'étalent sur plusieurs décennies et impliquent des investissements massifs et une planification complexe. De plus, la gestion des champs pétroliers est confrontée à de nombreuses incertitudes, que ce soit en raison de la volatilité des prix du pétrole ou de l'impossibilité d'observer directement les ressources souterraines. L'optimisation stochastique multi-étapes (en observation partielle) est donc un outil pertinent pour optimiser la performance globale de tels projets.

La littérature sur l'optimisation des champs pétroliers tend à se concentrer sur la manière d'améliorer la performance de ses différents composants pris séparément (par exemple l'amélioration de la production d'un puits ou des performances de pompes). Une partie beaucoup moins importante de la littérature concerne une approche holistique de l'optimisation des champs pétroliers, et, au sein de cette approche, seule une fraction limitée couvre une approche d'optimisation mathématique, qui est l'approche adoptée dans ce document.

Dans cette thèse, nous présentons un cadre mathématique qui, premièrement, est utilisé pour représenter la dynamique d'un système de production de pétrole et de gaz et, deuxièmement, permet de calculer une politique de contrôle qui maximise un critère intertemporel. Une politique de contrôle est un ensemble de fonctions qui prennent comme entrées des quantités disponibles à chaque instant pour un décideur non anticipatif, telles que l'historique des observations et de la production du système de production de pétrole et de gaz, et qui renvoient une décision (telle que la façon d'opérer le système de production ou la liste des puits à considérer pour le prochain forage) à appliquer à cet instant donné.

Tout d'abord, dans le Paragraphe 1.2.1, nous présentons quelques éléments de contexte sur la gestion des champs pétroliers. Deuxièmement, dans le Paragraphe 1.2.2, nous présentons les sources d'incertitudes présentes dans les systèmes de production pétroliers. Troisièmement, dans le Paragraphe 2.2.3, nous présentons quelques problèmes d'optimisation associés à la gestion des champs pétroliers. Enfin, dans le Paragraphe 2.2.4, nous spécifions le problème qui est étudié dans cette thèse.

1.2.1 Contexte de la conception et de la gestion de l'exploitation des champs pétroliers

Considérons une entreprise qui dispose, grâce à un contrat avec les autorités compétentes, des droits d'exploitation et d'extraction des hydrocarbures d'un champ pétrolier. Cette entreprise utilise le terme "produire" pour l'extraction du pétrole, et appelle "réseau de production" ou "système de production" l'infrastructure utilisée dans un champ pétrolier pour extraire les hydrocarbures. Nous utilisons donc cette terminologie dans la suite du document. Le but de l'entreprise est d'optimiser une fonction objectif (telle que la Valeur Actualisée Nette) sur toute la durée de vie de la concession, et d'y parvenir en appliquant certaines décisions. Nous utilisons donc le terme "décideur" pour désigner l'entreprise.

Dans un champ pétrolier donné, il peut y avoir plusieurs formations géologiques qui contiennent certaines ressources (différents mélanges de pétrole, de gaz et d'eau). Nous appelons ces formations des "réservoirs". Idéalement, les ressources sont stockées dans les pores d'un réservoir (de la même manière que l'eau peut être retenue dans une éponge) à une pression très élevée. C'est cette pression qui permet au pétrole de s'écouler lorsque un puits est foré. Généralement, il y a aussi une grande formation géologique qui contient de l'eau près d'un réservoir. Cette formation est appelée "aquifère". La présence et les caractéristiques de l'aquifère peuvent avoir un impact important sur la gestion d'un réseau de production, car elles peuvent conduire l'eau à pénétrer dans le réservoir et à remplacer le pétrole produit. Cela peut être à la fois un atout et un inconvénient majeur pour le système de production : elle peut soit améliorer le taux d'extraction en amortissant la chute de pression dans le réservoir, soit inonder une partie des infrastructures et donc empêcher l'utilisation future de ces dernières.

Le réseau de production est quand à lui composé d'un grand nombre d'infrastructures différentes, telles que des puits (forages dans le sol permettant la production de pétrole et/ou de gaz), des oléoducs (utilisés pour transporter le pétrole d'un point à un autre du champ), des FPSO (Floating Production Storage and Offloading, navires utilisés dans les champs offshore pour traiter et stocker le pétrole avant qu'il ne soit exporté par un pétrolier), d'autres installations aux points de sortie, ... Chacun de ces composants a ses propres spécifications.

Nous détaillons maintenant les incertitudes à prendre en compte dans la gestion d'un système de production pétrolier.

1.2.2 Incertitudes dans les systèmes de production pétrolière

La première et la plus évidente des incertitudes est la volatilité des prix du pétrole. Personne ne peut prédire avec précision le prix futur du pétrole. De plus, les prix du pétrole dépendent fortement de la situation géopolitique: les tendances à long terme peuvent radicalement changer en raison de décisions politiques. La prise en compte de cette incertitude dans un cadre d'optimisation ne pose pas, dans une certaine mesure, de difficultés particulières lorsque les distributions des prix futurs sont données. Cependant, bien que la recherche des distributions de probabilité pertinentes pour les prix soit un élément important du processus, elle dépasse le cadre des problèmes d'optimisation et de leur résolution et ne sera pas couverte par cette thèse.

La deuxième source d'incertitude provient du fait que nous ne pouvons pas observer directement un réservoir, ni son contenu. Nous ne savons pas combien il reste d'hydrocarbure en son sein, ni sa répartition. Cependant, nous pouvons parfois avoir une bonne idée de ce qui se trouve dans la croûte terrestre grâce à différentes techniques, comme les études sismiques. Au final, cependant, nous avons tout au plus une vue partielle d'un réservoir. A partir de cette observation partielle, nous ne pouvons qu'inférer des trajectoires incertaines du contenu du réservoir car nous n'avons aucune connaissance précise de l'état initial du réservoir. Nous considérons cependant dans cette thèse que nous avons accès à une certaine distribution de probabilité de l'état initial.

Enfin, l'incertitude peut également prendre la forme de risques sur l'infrastructure elle-même. La gestion d'un champ comporte certains aléas, comme l'ont malheureusement montré les marées noires et autres accidents relatifs à la production de pétrole. De plus, les infrastructures peuvent tomber en panne. Il est possible de faire une analyse des risques sur le réseau de production, comme par exemple la quantité de pétrole libérée si un oléoduc est endommagé. Cependant, nous ne couvrirons pas ce genre d'incertitudes dans cette thèse. En effet, nous nous intéressons principalement à la manière de prendre en compte le fait que le réservoir est partiellement observé.

Nous présentons maintenant un aperçu des problèmes d'optimisation de la gestion des champs pétroliers.

1.2.3 Aperçu de quelques problèmes d'optimisation des systèmes de production pétrolière

Les systèmes de production pétroliers étant des projets complexes, de multiples problèmes d'optimisation se posent au cours des différentes phases de ces projets. En effet, la durée de vie d'un champ est généralement décomposée en cinq phases : l'exploration, où le but est de trouver des réservoirs contenant des hydrocarbures; l'évaluation, où l'on estime les valeurs des champs pétroliers ; le développement, où les infrastructures sont planifiées et installées ; la production, où les hydrocarbures sont finalement produits ; l'abandon, où les champs cessent de produire et où les infrastructures sont mises hors service et retirées. Des problèmes d'optimisation peuvent être formalisés pour toutes ces phases. Cependant, selon la revue de littérature Khor et al. [2017], l'optimisation mathématique se concentre sur les phases de développement et de production : soit pour concevoir un système de production, soit pour le gérer.

Nous pouvons identifier trois principaux problèmes d'optimisation industrielle du système de production pétrolière:

1 la conception d'un système de production pétrolier, c'est-à-dire l'optimisation des décisions prises lors de la phase de développement ;

2 la gestion d'un système de production pétrolier déjà donné, c'est-à-dire l'optimisation de la phase de production ;

1+2 la conception et la gestion d'un système de production pétrolier, c'est-à-dire l'optimisation des deux phases de développement et de production.

Nous nous concentrons dans la suite sur le deuxième problème industriel : l'optimisation de la phase de production, que nous décrivons maintenant.

1.2.4 Gestion d'un système de production pétrolière donné

Le problème d'optimisation dont il est question dans cette thèse est la gestion d'un système de production pétrolier donné. Étant donné un réseau de production composé de puits, de conduites et de points de sortie, et étant donné un ensemble fini de pas de temps $\mathcal{T} = \{0, \dots, T\}$, nous cherchons une politique de production qui maximise un indicateur économique tel que la Valeur Actuelle Nette tout en respectant des contraintes physiques sur le réseau de production (telles que maintenir la pression dans une plage donnée en certains points du réseau, le respect d'un planning de maintenance, etc.).

Lorsque l'on considère un réseau de production donné, le décideur a à sa disposition un ensemble de contrôles u_t qu'il peut appliquer sur ce système à l'instant t. Ces contrôles sont des actions sur un réseau de production fixe : on ne peut pas changer la sélection des puits ou modifier le réseau de canalisations, on agit uniquement sur des actionneurs placés à l'intérieur (comme les vannes présentes sur les oléoducs qui peuvent être ouvertes ou fermées). De plus, nous modélisons les réservoirs et le réseau de production comme des systèmes dynamiques à l'aide d'équations physiques. Nous désignons par x_t l'état de ce système dynamique et par u_t les commandes au temps $t \in \mathcal{T}$. Nous formulons l'optimisation de la gestion d'un système de production pétrolière donné comme un problème de contrôle optimal qui prend la forme suivante dans le cas déterministe.

$$\max_{u_t, x_t} \sum_{t=0}^{T-1} \mathcal{L}_t(x_t, u_t) + \mathcal{K}(x_T) \tag{1.1a}$$

$$s.c. \quad x_{t+1} = f(x_t, u_t) \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , \tag{1.1b}$$

$$u_t \in \mathcal{U}_t^{ad}(x_t) \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ . \tag{1.1c}$$

L'objectif de cette thèse est dans un premier temps de détailler la formulation du Problème (1.1), avant d'étendre la formulation précédente pour prendre en compte l'observation partielle de l'état.

Nous présentons maintenant la structure de ce document.

1.3 Plan de la thèse

Le manuscrit est composé des quatre chapitres suivants :

- Dans le Chapitre 3, nous présentons les différentes lois de comportements utilisée dans la construction de la formulation de la gestion d'un système de production pétrolière. Ces équations portent soit sur les fluides dans le réservoir et sur le comportement du réservoir, soit sur le réseau de production lui-même.

- Dans le Chapitre 4, nous détaillons la formulation de l'optimisation d'un système de production pétrolier donné en considérant que l'on a une observation complète du contenu du réservoir, et que nous connaissons les prix futurs. Il s'agit donc d'une formulation déterministe du problème d'optimisation. Nous détaillons les dérivations nécessaires pour représenter un réservoir par un système dynamique contrôlé, avant de présenter quelques applications numériques. Il s'agit de la transcription d'un article accepté dans *Computers & Chemical Engineering*, Vessaire et al. [2022].

- Dans le Chapitre 5, nous présentons un cadre mathématique permettant de prendre en compte l'observation partielle lorsqu'on considère un système dynamique avec des fonctions d'évolution et d'observation déterministes, les Deterministic Partially Observed Markov Decision Processes (Det-Pomdps, processus de décision markovien déterministe partiellement observés). Nous approfondissons les travaux de Littman [1996] et présentons un sous-ensemble des Det-Pomdps, les Separated Det-Pomdps, qui possède des propriétés permettant de repousser la malédiction de la dimensionnalité.

- Dans le Chapitre 6, nous présentons une formulation du problème de gestion pétrolier qui prend en compte l'observation partielle du contenu du réservoir. Nous démontrons ensuite que la formulation obtenue est équivalente à un problème d'optimisation de type Separated Det-Pomdp, avant de présenter quelques applications numériques.

Nous détaillons maintenant nos principales contributions.

1.4 Contributions principales

- Le Chapitre 3 synthétise la littérature sur les lois de comportement physiques utilisées dans le reste de la thèse.

- Dans le Chapitre 4, nous présentons tout d'abord un modèle physique décrivant le comportement du réservoir au cours du temps. Il est constitué d'un système dynamique contrôlé qui donne l'évolution dans le temps des quantités physiques caractérisant le champ pétrolier exploité. Ces équations sont dérivées d'équations de *bilan de matière* sur le réservoir et sous l'hypothèse que les fluides contenus dans le réservoir suivent un modèle connu sous le nom de *modèle black-oil*. Deuxièmement, nous donnons une formulation déterministe d'un problème d'optimisation multi-étapes pour un système de production de pétrole et de gaz, régi par le système dynamique contrôlé introduit dans la première partie. Enfin, des solutions numériques du problème d'optimisation sont présentées et comparées à d'autres formulations possibles.

Décrire l'évolution dans le temps d'un réservoir comme un système dynamique contrôlé et dériver les équations d'évolution à partir d'équations de *bilan matière* et de *modèle de black-oil* n'est pas courant dans la littérature sur la gestion des champs pétroliers. De plus, l'utilisation de la programmation dynamique pour résoudre un problème d'optimisation multi-étape sur un système dynamique contrôlé est classique mais soumise à la malédiction de la dimensionnalité. C'est-à-dire que lorsque l'état du système est grand (ce qui est le cas ici puisque l'état est de dimension 5), il devient difficile de résoudre numériquement le problème. C'est pourquoi nous présentons des applications numériques sur des cas où la dimension de l'état se réduit à la dimension 1 (réservoir de gaz et réservoir de pétrole où la pression est maintenue constante par l'injection d'eau).

Plus précisément, le contenu du Chapitre 4 est le suivant:

- Dans le Paragraphe 4.3, nous décrivons l'évolution du réservoir dans le temps comme un système dynamique contrôlé. C'est-à-dire que, en désignant par x_t l'état du système et par u_t les contrôles appliqués à un instant t, l'évolution de l'état est donnée par une application f détaillé en (4.3), tel que:

$$x_{t+1} = f(x_t, u_t) \ .$$

  Dans cette description, l'état x du système est de dimension cinq et est composé de la quantité des trois fluides du *modèle de black-oil* (pétrole, gaz et eau), de la pression du réservoir et du volume total des pores (le volume où les fluides sont stockés), tandis que les contrôles u sont les actions applicables au réseau de production : ouverture ou fermeture d'oléoducs, choix de la pression en tête de puits, fonctionnement des pompes. L'expression mathématique de l'application f est donnée dans la Proposition 4.2, p.45, et la manière de dériver cette expression à l'aide d'équations de *bilan matière* et de *modèle de black-oil* est donnée dans l'Annexe 4.A.

- Dans le Paragraphe 4.2, nous formulons la gestion optimale système de production pétrolier comme un problème de contrôle optimal sur un horizon fini (voir le Problème (4.1)) en utilisant la dynamique d'état f et en supposant que l'état du réservoir est observé dans le temps. La solution optimale du Problème (4.1) est obtenue à l'aide d'algorithmes classiques de Programmation Dynamique (voir Proposition 4.1 et Algorithme 1).

- Dans le Paragraphe 4.4, nous présentons des applications numériques sur deux cas: un réservoir de gaz et un réservoir de pétrole où la pression est maintenue constante par injection d'eau. De plus, nous comparons notre méthode avec une approche d'optimisation standard de l'industrie pétrolière et gazière sur les mêmes problèmes.

• Dans le Chapitre 5, nous présentons notre principale contribution théorique, consacrée à l'étude des Deterministic Partially Observed Markov Decision Processes

(Det-Pomdp). Il s'agit d'un travail préparatoire pour le Chapitre 6, car nous montrons au Chapitre 6 que le problème d'optimisation d'un système de production pétrolier en tenant compte de l'observation partielle du réservoir appartient à la classe des Det-Pomdps. Plus généralement, la formulation d'un problème d'optimisation dans le temps sous observation partielle de l'état avec une dynamique d'état contrôlée et perturbée par des bruits exogènes tombe dans la classe des problèmes Pomdps. Les Pomdps ont été largement étudiés, et des algorithmes numériques basés sur la Programmation Dynamique ont été spécifiquement développés pour les Pomdps. Cependant, il est bien connu qu'ils sont numériquement difficiles à résoudre car sujets à la malédiction de la dimension. Nous présentons au Paragraphe 5.1.1 une sous-classe connue de Pomdps appelée Det-Pomdp. Ensuite, au Paragraphe 5.2, nous présentons la programmation dynamique sur les "croyances" (distributions de probabilité sur les états) pour les Det-Pomdps avec contraintes sur les contrôles dépendant de l'état non observé. Deuxièmement, dans le Paragraphe 5.3, nous présentons une nouvelle représentation pour la dynamique des croyances en utilisant des *mesure-image*. Troisièmement, dans le Paragraphe 5.4, nous présentons des résultats de complexité de la Programmation Dynamique sur les croyances pour les Det-Pomdps. Quatrièmement, dans le Paragraphe 5.5, nous introduisons une sous-classe encore plus simple de Det-Pomdps, les Separated Det-Pomdps.

Nous détaillons maintenant les principales contributions présentées dans le Chapitre 5.

- Dans le Paragraphe 5.2, nous étendons les équations de Programmation Dynamique sur les croyances de [Bertsekas and Shreve, 1978] pour les Pomdps sans contraintes aux Det-Pomdps avec contraintes sur les contrôles dépendant de l'état non observé. En effet, une hypothèse clé dans [Bertsekas and Shreve, 1978] pour écrire les équations de Programmation Dynamique sur les croyances est qu'il n'y a pas de contraintes d'admissibilité sur les contrôles. Comme les applications imposent la présence de telles contraintes, nous présentons cette extension dans la Proposition 5.2, qui nous donne les équations de Programmation Dynamique (5.16).
- Dans le Paragraphe 5.3, nous exprimons la dynamique des croyances dans Det-Pomdps en utilisant la notion de *mesure-image*, telle que présentée dans le Lemme 5.4. Cette nouvelle représentation est la base de toutes les améliorations des bornes sur la cardinalité de l'ensemble des croyances atteignables.
- Dans le Paragraphe 5.4, nous améliorons la limite de Littman sur la cardinalité de l'ensemble des croyances atteignables, $\left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}\right|$, pour les Det-Pomdps, de $(1+|\mathbb{X}|)^{|\mathbb{X}|}$, à $(1+|\mathbb{X}|)^{|\mathrm{supp}(b_0)|}$ (voir Théorème 5.9). Notons que cette borne ne dépend pas du nombre de pas de temps. De plus, dans le Théorème 5.10, nous donnons une borne dépendant du temps sur la cardinalité de l'ensemble des croyances atteignables, $1+|\mathrm{supp}(b_0)||\mathbb{U}|^{|\mathcal{T}|}$. Ces bornes sont intéressantes, car la complexité bien connue de la Programmation Dynamique sur les croyances est de $O(|\mathcal{T}||\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)||\mathbb{U}||\mathbb{O}|)$ (voir la Proposition 5.7).

   – Dans le Paragraphe 5.5, nous introduisons une sous-classe de DET-POMDPs, les Separated DET-POMDPs. Comme le montre le Corollaire 5.16, l'intérêt des Separated DET-POMDPs est qu'elle repousse la malédiction de la dimension pour la Programmation Dynamique avec croyances. En effet, elle améliore la borne de $(1 + |\mathbb{X}|)^{|\mathrm{supp}(b_0)|}$ à $1 + \left(2^{|\mathrm{supp}(b_0)|} - |\mathrm{supp}(b_0)|\right)|\mathbb{X}|$ (voir Corollaire 5.16). De plus, cette borne est serrée (voir la Proposition 5.18).

- Dans le Chapitre 6, nous reformulons l'optimisation d'un système de production de pétrole et de gaz comme un problème d'observation partielle. Nous prenons en compte le fait que le système de production de pétrole est partiellement observé. La dynamique contrôlée du champ de production pétrolière donnée au Chapitre 4 est la même, ainsi que le critère que nous cherchons à optimiser. Cependant, nous ne considérons plus dans ce chapitre qu'à chaque instant t, l'état du système x_t est connu par le décideur. La stochasticité du problème est introduite par le fait que l'état initial du système n'est pas connu, mais on suppose que l'on connaît sa distribution de probabilité. Nous prouvons dans le Paragraphe 6.2.3 que le modèle du réservoir présenté au Chapitre 4 combiné à une observation partielle du contenu du réservoir conduit est équivalent à un problème d'optimisation de type Separated DET-POMDP. Il entre donc dans la classe étudiée au Chapitre 5.

  Nous détaillons maintenant les principales contributions présentées dans le Chapitre 6:

   – Dans le Paragraphe 6.2.1, nous formulons la gestion optimale d'un système de production pétrolière partiellement observé comme un DET-POMDP (voir le Problème (6.1)). Cette formulation est dérivée du Problème (4.1) de gestion optimale d'un système de production pétrolière. Dans cette formulation, l'état et les contrôles sont les mêmes que dans le Problème (4.1). Nous introduisons également les observations du système (décrites au Paragraphe 6.2.2) qui sont de dimension trois et se composent de la pression du réservoir, de la *water-cut* (proportion d'eau produite par volume de liquide produit) et du *gaz-oil ratio* (proportion de gaz produit par volume de pétrole produit). Nous supposons que l'observation est donnée par une application h qui ne dépend que de l'état du système. Plus précisément, en désignant par x_t l'état au temps t et o_t l'observation au temps t, on a $o_t = h(x_t)$.

   – Au Paragraphe 6.2.3, nous prouvons dans le Lemme 6.1 et la Proposition 6.2 que, sous une hypothèse clé qui est satisfaite en pratique (à savoir que la croyance initiale est compatible avec l'observation initiale), le Problème (6.1) est équivalent à un problème d'optimisation de la classe Separated DET-POMDP qui partage les mêmes ensembles d'états, de contrôles, d'observations et de pas de temps. Par conséquent, nous pouvons appliquer les résultats du Chapitre 5 au Problème (6.1). Nous pouvons donc résoudre le Problème (6.1) grâce à l'Algorithme 4, et sa complexité est donnée par le Corollaire 5.16.

– Dans le Paragraphe 6.3, nous présentons des applications numériques sur les deux cas présentés dans le Chapitre 4, c'est-à-dire un réservoir de gaz et un réservoir de pétrole où la pression est maintenue constante par l'injection d'eau.

Chapter 2

Introduction

Contents

We first give in §2.1 contextual information for this thesis. Second, in §2.2, we present the subject of this manuscript: the optimization of petroleum production systems taking into account uncertainties. Third, in §2.3, we lay out the structure of this document. Finally, in §2.4, we detail our main contributions.

2.1 Context of this thesis

This thesis started on 2019, September 1 and ended on 2022, November 30. It was the second thesis conducted as part of a partnership between TotalEnergies and the École Nationale des Ponts et Chaussées (ENPC). This partnership was signed in 2018, at the initiative of Philippe Ricoux, and contains two Ph.D. subjects: stochastic optimization for the procurement of crude oil in refineries, and stochastic optimization for petroleum production systems.

The subject of this thesis is the management of petroleum production systems under uncertainties, that is, how to pilot the production over time of an oil or gas field in order to maximize profit. This thesis was made in cooperation with the Integrated Asset Modeling (IAM) team at TotalEnergies, which is tasked with the integration of simulation models from reservoir to process in order to support the evaluation and the design of oil-field development plans. The cooperation first began thanks to Remy Marmier, then head of the IAM team, while Alejandro Rodríguez-Martínez became the main interlocutor, notably for the validation of the physical models and the industrial applications. Anna Robert, part of TotalEnergies' R&D branch, was also part of the interlocutors. The cooperation between the CERMICS laboratory at ENPC and TotalEnergies have led to a patent submission.

We now detail the subject of this thesis, namely the optimization of petroleum production system.

2.2 Optimization of petroleum production system

Given that oil is becoming more and more expensive to extract, it becomes more and more pressing to optimize oil-fields. Indeed, oil and gas projects span over several decades and involve massive investments and complex planning. Moreover, the management of oil-fields faces many uncertainties, be it due to the volatility of oil prices, or due to the inability to directly fully observe what is trapped in the ground. Stochastic multistage optimization is therefore a relevant tool to optimize the whole performance of such projects.

Literature on the optimization of oil-field tends to focus on how to improve the performance of its different components (e.g. improving the production of one well). A far smaller part of the literature concerns a holistic approach to the optimization of oil-fields, and only a limited fraction covers a mathematical optimization approach, which is the approach adopted in this document.

In this thesis, we present a mathematical framework that, first, is used to represent the dynamics of an oil and gas production system and, second, returns a control policy which maximizes an inter-temporal criterion. A control policy is a set of functions that take as inputs quantities available to a non-anticipative decision maker such as the history of observations and production of the oil gas production system, and return a decision (such as how to operate the production system or the list of wells to consider drilling next) to be applied at a given time.

First, in §2.2.1, we present some backgrounds on the management of oil-fields. Second, in §2.2.2, we present the sources of uncertainties present in petroleum production systems. Third, in §2.2.3, we present some optimization problems associated with oil-fields management. Finally, in §2.2.4, we specify the problem which is studied in this thesis.

2.2.1 Context of oil-fields design and operation management

Let's consider a company that has, thanks to a contract with the relevant authorities, rights to exploit and extract oil from an oil-field. Such company uses the term "produce" for the extraction of oil, and calls "production network" or "production system" the infrastructure used in an oil-field to extract hydrocarbons. We therefore use that terminology in the rest of the document. The goal of the company is to optimize an objective function (such as the Net Present Value) over the full life of the concession, and do so by applying some decisions. We hence use the term "decision maker" to refer to the company.

In a given oil-field, there can be multiple geological formations that contain some resources (different mixes of oil, gas and water). We call those formations “reservoirs”. Ideally, the resources are stored in the pores of a reservoir (in the same way water can be held in a sponge) at a very high pressure. That pressure is what allows the oil to flow out when a well is drilled. Commonly, there is also a large geological formation that contains water near a reservoir. That formation is called an “aquifer”. The presence and characteristics of the aquifer may have large impacts on the management of a production network, as it can lead to water entering the reservoir and replacing the produced oil. It can be both a boon or a bane to the production system: it can either improve the rate of extraction by cushioning the pressure drop in the reservoir, or it can flood part of the infrastructure and thus prevent the use of said infrastructure afterward.

The production network is made of a lot of different infrastructures, such as wells (bores in the earth that allows the production of oil and/or gas), pipes (used to transport oil from one point of the field to another), FPSO (Floating Production Storage and Offloading, ships used in offshore fields to process and store oil before it is exported through a tanker), other exit-points facilities, etc. Each of those components has their own specifications.

We now detail the uncertainties one need to consider when managing a petroleum production system.

2.2.2 Uncertainties in petroleum production systems

There are multiple sources of uncertainties encountered over the lifetime of a field. The first and most obvious one is the volatility of oil prices. No one can accurately predict the future price of oil. Moreover, oil prices are highly dependent on the geopolitical status, hence long-term trends can drastically change due to political decisions. Taking into account that uncertainty in an optimization framework is to a certain extent straightforward when distributions of future prices are given. However, while finding the relevant probability distributions for the prices is an important component of the process, it is beyond the scope of optimization problems.

The second source of uncertainty comes from the fact that we cannot directly observe a reservoir. We do not know how much stock is left, nor do we know where it is. However, we can sometimes have a good picture of what is in Earth’s crust through different techniques, such as seismic studies. In the end, however, we have at most a partial view of a reservoir. From that partial observation, we can only infer uncertain trajectories of the content of the reservoir as we have no accurate knowledge of the initial state of the reservoir. We however consider that we have access to some probability distribution of the initial state.

Finally, uncertainty can also take the form of risks on the infrastructure itself. There are some hazards in the management of a field, as has been unfortunately shown by accidental oil spills. Indeed, infrastructure can fail. It is possible to do a risk analysis on the production network, such as how much oil is released if a pipe is damaged. However, we will not cover such these kinds of uncertainties in this thesis.

In this thesis, we are mainly interested on how to take into account the fact that the physical system is partially observed. We will thus not consider stochastic prices or risks on the infrastructures.

We now present an overview of the optimization problems for the management of oil-fields.

2.2.3 Overview of some optimization of petroleum production system problems

As petroleum production systems are complex projects, multiple optimization problems arise during the different phases of the projects. Indeed, the lifetime of a field is usually decomposed into five phases: exploration, where the goals are to find reservoirs containing hydrocarbons; appraisal, where the values of oil-fields are evaluated; development, where the infrastructure is planned and installed; production, where hydrocarbons are finally produced; abandonment, where the fields stop producing and infrastructure are decommissioned and removed. Optimization problems can be formalized for all those phases. However, according to the survey Khor et al. [2017], mathematical optimization focuses on the development and production phases: either to design a production system, or to manage it.

We may identify three main industrial optimization problems of petroleum production system:

1 the design of a petroleum production system, i.e. the optimization of the decisions made at the development phase;

2 the management of an already given petroleum production system, i.e. the optimization of the production phase;

1+2 the design and management of a petroleum production system, i.e. the optimization of both the development and production phase.

We focus in the sequel on the second industrial problem: the optimization of the production phase, which we now describe.

2.2.4 Management of a given petroleum production system

The optimization problem which is the focus of this thesis is the management of a given petroleum production system. Given a production network composed of wells, pipe and exit points, and given a finite set of time-steps $\mathcal{T} = \{0, \dots, T\}$, we look for a production policy that maximizes an economic indicator such as the Net Present Value while respecting physical constraints, on the production network (such as having the pressure in a given range at certain points in the network, respecting a maintenance planning, etc.).

When considering a given production network, we have a set of controls $\mathbb{U}$ we can apply on that system. Those controls are actions on a fixed production network: we cannot change the selection of wells or modify the network of pipe, we only act on in-place actuators (such as valves present on pipes that can be opened or closed). Moreover, we model the reservoirs and production network as dynamical systems with the help of physical equations. We denote by $x_t \in \mathbb{X}$ the state of that dynamical system (with $\mathbb{X}$ the set of states) and by $u_t \in \mathbb{U}$ the controls at time $t \in \mathcal{T}$. Its evolution is given by the mapping $f : \mathbb{X} \times \mathbb{U} \to \mathbb{X}$. We formulate the optimization of the management of a given petroleum production system as an optimal control problem which takes the following form in the deterministic case.

$$\max_{u_t, x_t} \sum_{t=0}^{T-1} \mathcal{L}_t(x_t, u_t) + \mathcal{K}(x_T) \tag{2.1a}$$

$$s.c. \quad x_{t+1} = f(x_t, u_t) \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , \tag{2.1b}$$

$$u_t \in \mathcal{U}_t^{ad}(x_t) \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ . \tag{2.1c}$$

At each time t, the controls must belong to an admissibility set that depends on the state x_t, $\mathcal{U}_t^{ad}(x_t)$. Moreover, we optimize an additive cost function, with the instantaneous cost at time $t \in \mathcal{T} \setminus \{T\}$ given by the mapping $\mathcal{L}_t : \mathbb{X} \times \mathbb{U} \to \mathbb{R}$, while the final cost is given by the mapping $\mathcal{K} : \mathbb{X} \to \mathbb{R}$.

The objective of this thesis is to first detail the formulation of Problem (2.1), before extending the previous formulation in order to take into account partial observation of the state.

We now present the structure of this document.

2.3 Outline of the thesis

The manuscript is composed of the four following chapters:

- In Chapter 3, we present an overview of the constitutive equations necessary to build the formulation of the management of a petroleum production system. Those equations either focus on the fluids in the reservoir and on the reservoir's behavior, or they focus on the production network itself.

- In Chapter 4, we detail the formulation of the optimization of a given petroleum production system when considering that we have perfect observation of the content of the reservoir, and we know the future prices. It is therefore a deterministic formulation of the optimization problem. We detail how we can represent the reservoir as a controlled dynamical system, before presenting some numerical applications. It is the transcript of an article accepted in *Computers & Chemical Engineering*, Vessaire et al. [2022].

- In Chapter 5, we present a mathematical framework to take into account partial observation when considering a dynamical system with deterministic evolution and observation functions, Deterministic Partially Observed Markov Decision Process(Det-Pomdp). We further expand works by Littman [1996], and present a subset of Det-Pomdp, Separated Det-Pomdp, which possesses properties which further push back the curse of dimensionality.

- In Chapter 6, we present a formulation of the management problem that takes into account the partial observation of the content of the reservoir. We then demonstrate that the resulting formulation is equivalent to a Separated Det-Pomdp optimization problem, before presenting some numerical applications.

We now detail our main contributions.

2.4 Main contributions

- Chapter 3 gives some background on the physical constitutive equations used in the rest of this thesis

- In Chapter 4, first a physical model that describes the reservoir's behavior over time is given. It consists of a controlled dynamical system which gives the evolution over time of physical quantities which characterize the oil field under exploitation. These equations are derived from *material balance equations* on the reservoir and under the assumption that the fluids contained in the reservoir follow a model known under the name of *black-oil models*. Second, a deterministic formulation of an optimization problem over time for an oil and gas production system, governed by the controlled dynamical system introduced in the first part, is given. Numerical solution of the optimization problem are given and compared to other possible formulations.

  Describing the evolution over time of a reservoir as a controlled dynamical system and deriving the evolution equations from *material balance equations* and *black-oil models* is not common in the oil field management community. Then, using dynamic programming to solve an optimization problem over time for a controlled dynamical system is classical but subject to the curse of dimensionality. That is, when the state of the system is large (which is the case here since the state is of dimension 5) it becomes difficult to numerically solve the problem. That's why we have made numerical applications on cases where the state dimension boils down to dimension 1 (gas reservoir and an oil reservoir where pressure is kept constant through water injection).

  More precisely, the content of Chapter 4 is the following:

  - In §4.3, we describe the reservoir evolution over time as a controlled dynamical system. That is, denoting by x_t the state of the system and by u_t the controls applied a time t, the state evolution is given by a mapping f:

$$x_{t+1} = f(x_t, u_t) \ .$$

In this description, the state x of the system is of dimension five and is composed of the amount of the three fluids of the black oil model (oil, gas and water), the reservoir pressure and the total pore volume (the volume where the fluids are stored) and the controls u are the possible actions on the production network: opening or closing pipes, choosing the well-head pressure, operating pumps. The mathematical expression for the mapping f is given in Proposition 4.2, and the way to derive this expression using *material balance equations* and *black-oil models* is given in Appendix 4.A.

- In §4.2, we formulate, given the state dynamics f and assuming that the state of the reservoir is observed over time, the optimal management of a petroleum production system as an optimal control problem over a finite horizon (see Problem (4.1)). The optimal solution of Problem (4.1) is obtained with classical Dynamic Programming algorithms (see Proposition 4.1 and Algorithm 1).
- In §4.4, numerical applications on two distinct cases, a gas reservoir and an oil reservoir where pressure is kept constant through water injection, are given together with comparisons to a standard optimization approach of the oil and gas industry on the same problems.

- In Chapter 5, we present our main theoretical contribution, devoted to the study of Deterministic Partially Observed Markov Decision Process (Det-Pomdp), and is a preparation work for Chapter 6 as the oil and gas optimization problem under partial observation is shown to belong to the Det-Pomdp class in Chapter 6. More generally, the formulation of an optimization problem over time under partial observation of the state with controlled state dynamics perturbed by exogenous noises falls into the class of Pomdp problems. Pomdps have been widely studied, and numerical algorithms based on Dynamic Programming have been specifically developed for Pomdps. However, it is well-known that they are numerically difficult and subject to the curse of dimensionality. We present in §5.1.1 a known subclass of Pomdp called Det-Pomdp. Then §5.2, we present Dynamic Programming on beliefs for Det-Pomdp with constraints. Second, in §5.3, we present a new representation for the belief dynamic using pushforward measures. Third, in §5.4, we present complexity bounds for Dynamic Programming on beliefs for Det-Pomdp. Fourth, in §5.5, we introduce an even simpler subclass of Det-Pomdp, Separated Det-Pomdp.

  We now detail the main contributions presented in Chapter 5.

  - In §5.2, we extend [Bertsekas and Shreve, 1978] Dynamic Programming equations with beliefs for unconstrained Pomdps to Det-Pomdps with constraints. Indeed, a key assumption in [Bertsekas and Shreve, 1978] to write Dynamic Programming equations with beliefs is that there are no admissibility constraints on the controls. As the applications dictate the presence of such constraints, we present this extension in Proposition 5.2, which gives us Dynamic Programming Equations (5.16).

- In §5.3, we express the belief dynamics in Det-Pomdps using the notion of *pushforward* (or *image-measure*), as presented in Lemma 5.4. This new representation is the basis for all the improvements of the bounds on the cardinality of the set of reachable beliefs.
- In §5.4, we improve Littman's bound on the cardinality of the set of reachable beliefs, $\left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}\right|$, for Det-Pomdp, from $(1+|\mathbb{X}|)^{|\mathbb{X}|}$, to $(1+|\mathbb{X}|)^{|\mathrm{supp}(b_0)|}$ (see Theorem 5.9). Note that this bound does not depend on the number of time steps. Moreover, in Theorem 5.10, we give a time-dependent bound $1+|\mathrm{supp}(b_0)||\mathbb{U}|^{|\mathcal{T}|}$ on the cardinality of the set of reachable beliefs. Those bounds are of interest, as the well-known complexity of Dynamic Programming on beliefs is $O(|\mathcal{T}||\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)||\mathbb{U}||\mathbb{O}|)$ (see Proposition 5.7).
- In §5.5, we introduce a subclass of Det-Pomdps, Separated Det-Pomdps. As shown in Corollary 5.16, the interest of Separated Det-Pomdps is that it pushes back the curse of dimensionality for Dynamic Programming with beliefs. Indeed, it improves the bound from $(1+|\mathbb{X}|)^{|\mathrm{supp}(b_0)|}$ to $1+\left(2^{|\mathrm{supp}(b_0)|}-|\mathrm{supp}(b_0)|\right)|\mathbb{X}|$ (see Corollary 5.16). Moreover, this bound is tight (see Proposition 5.18).

- In Chapter 6, the optimization of an oil and gas production system under partial observation problem is addressed. We take into account the fact that the oil production system is partially observed. The controlled dynamics of the oil production field given in Chapter 4 are the same as is the criterion that we try to optimize. The new fact is that we do not consider that at each time t, the state of the system x_t is given to the decision maker. Stochasticity is introduced as we suppose that the initial state of the system is not known, but we assume that we know its probability distribution. We prove, in §6.2.3, that the model of the reservoir presented in Chapter 4 combined with a partial observation of the content of the reservoir leads is equivalent to a Separated Det-Pomdp optimization problem. It thus falls into the class studied in Chapter 5.

  More precisely, the content of Chapter 6 is the following:

  - In §6.2.1, we formulate the optimal management of a partially observed petroleum production system as a Det-Pomdp (see Problem (6.1)). This formulation is derived from the optimal management of a petroleum production system Problem (4.1). In this formulation, the state and controls are the same as in Problem (4.1). Meanwhile, the observations of the system (described in §6.2.2) are of dimension three, and are composed of the reservoir pressure, the water-cut (proportion of water produced per volume of liquid produced) and the gas-oil ratio (proportion of gas produced per volume of oil produced). We assume that the observation is given by the mapping h which only depends on the state of the system. That is, denoting by x_t the state at time t and o_t the observation at time t,

$$o_t = h(x_t) \ .$$

– In §6.2.3, we prove in Lemma 6.1 and Proposition 6.2 that, under a key assumption which is satisfied in practice, Problem (6.1) is equivalent to a Separated Det-Pomdp optimization problem that shares the same sets of states, controls, observations and time steps. Consequently, we can apply the results of Chapter 5 to Problem (6.1). We can hence solve Problem (6.1) thanks to Algorithm 4, and its complexity is given by Corollary 5.16.

– In §6.3, we present numerical applications on the two cases presented in Chapter 4, i.e. a gas reservoir and an oil reservoir where pressure is kept constant through water injection.

Chapter 3

Physics for oil-field problems

Contents

3.1 Introduction

The goal of this chapter is to present the different components needed for the formulations of optimization problems on oil-fields. We will notably describe oil-fields. An oil-field is a set of infrastructure made to extract a given resource (oil and/or gas). It can be naturally decomposed in two parts: the production network (the infrastructure itself), and the reservoir where the resource is kept. We will properly define the different sets and variables associated with those two components, starting with the production network.

Note that, as a production network is an assembly of multiple parts, an accurate description of each one of them would be too long and too complex. In this chapter, we will only focus on the relevant constitutive equations needed for the formulation of the problems previously presented.

3.2 Representing the production network as a graph

The production network is made of multiple components such as pipes, intersections, valves or wells. Its role is to transport the resources from where they are extracted (the wells) to an exit-point such as an FPSO (Floating Production Storage and Offloading). We represent that network thanks to a planar graph (a graph whose vertex are in the $\mathbb{R}^2$ plane), as shown in Figure 3.1. In this section, we first define how the components of the network are modeled in the graph. Then, we specify the variables necessary to model them. Finally, we outline the equations that link them together.

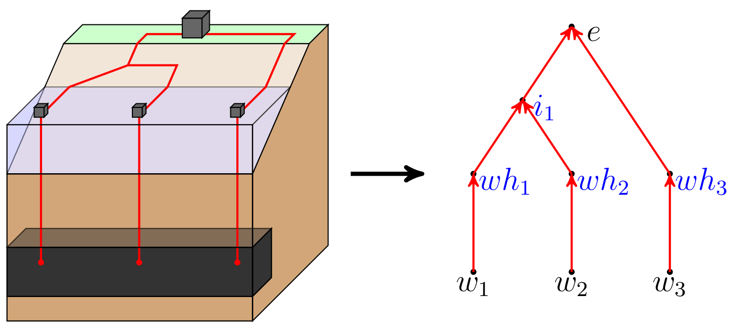


FIGURE 3.1: Representing a production network as a graph

Note that many variables can depend on time. However, we have yet to precise the form that time takes (discrete, continuous, . . .) in our model to accurately define them. Therefore, time will be implicit in this section.

3.2.1 Definition of the production network graph

There are two components in a graph: nodes and arcs.

3.2.1.1 Nodes

The nodes represent pertinent positions in the network. There can be multiple components on those positions (such as wells, intersections, exit-points, ...). We note $\mathbb{V}$ the set of nodes of the graph, and we partition $\mathbb{V}$ as:

$$\mathbb{V} = \mathbb{V}_{in} \sqcup \mathbb{V}_{mid} \sqcup \mathbb{V}_{out} \ . \tag{3.1}$$

The different subsets are defined in Table 3.1. They detail what components are present on each node (each component type is given a subset). If a node is part of a subset, that means that a corresponding component is placed on that node.

For example, in the graph presented in Figure 3.1, we have $\mathbb{V}_{in} = \{w_1, w_2, w_3\}$, $\mathbb{V}_{mid} = \{wh_1, wh_2, wh_3, i_1\}$, $\mathbb{V}_{out} = \{e\}$ and $\mathbb{V}_c = \{wh_1, wh_2, wh_3\}$. This means that well-heads are a part of the intersection set (well-heads are particular intersections), i.e. that $\mathbb{V}_c \subset \mathbb{V}_{mid}$.

Set	Definition
$\mathbb{V}_{in}$	Set of wells (the bottom of the wells)
$\mathbb{V}_{mid}$	Set of intersections
$\mathbb{V}_{out}$	Set of exit-points
$\mathbb{V}_{c}$	Set of the well-heads

Table 3.1: Definition of the subsets of $\mathbb{V}$

3.2.1.2 Arcs

In a production network, we have pipes, wells and valves. We use arcs between nodes to represent those pipes. We note $\mathbb{A}$ the set of arcs: $\mathbb{A} \subset \mathbb{V}^2$. We also define some relevant subsets of $\mathbb{A}$ in Table 3.2.

Set	Definition
$\mathbb{A}_{well}$	Set of well pipes
$\mathbb{A}_{c}$	Set of pipes with valves

Table 3.2: Definition of subsets of $\mathbb{A}$

We notably need to differentiate wells from other pipes. Indeed, a well is a special kind of pipe, as the physics inside it are a bit different from the rest of the pipes. We thus need to specify them with the set $\mathbb{A}_{well} \subset \mathbb{A}$. Valves, on the other hand, allow us to pilot the production network. We can open or close them, and thus close some pipes. We thus create $\mathbb{A}_c \subset \mathbb{A}$ to tell us which pipes we can control.

3.2.2 Variables in the graph

We have defined different sets of the graph, and what they represent. We will now present the variables needed for the formulation of optimization problems.

- Position of a node.
  When we design a network, we need variables representing the position of a node. We note $y_v \in \mathbb{R}^2$ the position of the node $v \in \mathbb{V}$. Note that y_v is a vector, each component being a coordinate of the position of v.

- Construction of a pipe.
  We also need a variable for the construction of a pipe. We note $p_{(v,v')}$ the Boolean that is at 1 if we build the edge $(v, v') \in \mathbb{A}$, and at 0 otherwise.

- Pressure.
  At each point of the network, there is pressure. Pressure is thus defined on all $\mathbb{V}$.

We note P_v the pressure at the vertex v. We also consider that pressure is positive, hence $P_v \geq 0 \;\; \forall v \in \mathbb{V}$

- Flow.
In each arc of the network, there can be a flow. Moreover, there can be multiple kinds of fluids. We thus define by $q_{(v,v')}$ the flow in the arc (v, v'). $q_{(v,v')}$ is a vector, each component being the flow of a given fluid. Usually, we will consider that q has three components: oil, gas and water. We note them respectively q^{O}, q^{G} and q^{W}. Therefore, we have:
$$q_a = \begin{pmatrix} q_a^{\text{O}} \\ q_a^{\text{G}} \\ q_a^{\text{W}} \end{pmatrix} \;, \;\; \forall a \in \mathbb{A} \;. \tag{3.2}$$
If we need to have more components, we explicitly mention it.

- State of a pipe.
Some pipes in the network have valves that allow the oilfield management entity to open or close a given pipe. We thus need to have a Boolean o_a that is at 1 if the arc a is open, and at 0 if it is closed. That variable is only defined on the arcs $a \in \mathbb{A}_c$.

3.2.3 Constitutive equations and constraints for the network

We will now present two constitutive equations linking the different variables in the production network, and two constraints.

3.2.3.1 Flow conservation at intersection

The first constitutive equation is the conservation of the flow at an intersection:

$$\sum_{(v,v'),v \in \mathbb{V}_{in} \cup \mathbb{V}_{mid}} q_{(v,v')} = \sum_{(v',v''),v'' \in \mathbb{V}_{mid}} q_{(v',v'')} \;, \;\; \forall v' \in \mathbb{V}_{mid} \;. \tag{3.3}$$

Thus, at every node $v' \in \mathbb{V}_{mid}$, we have conservation of the flow. This constraint does not exist at the wells (i.e. for $v \in \mathbb{V}_{in}$), as they are sources from which the flow comes from. It does not exist at the exit points either (i.e. for $v \in \mathbb{V}_{out}$), as they are the places where the flow exits the network.

3.2.3.2 Pressure drop in pipes

The second constitutive equation is the pressure drop in a pipe. We consider that there is a polynomial relation for the pressure drop:

$$P_v - P_{v'} = \Delta^p_{(v,v')}(q_{(v,v')}) \;, \;\; \forall (v, v') \in \mathbb{A} \;, \tag{3.4}$$

where, $\Delta^p_{(v,v')}$ is a polynomial function that depends on the pipe $(v, v') \in \mathbb{A}$. The functions Δ^p depends on the length of the pipe, but also on the material used, its section, etc. We assume that $\Delta^p_{(v,v')}$ is stationary, as, for now, we do not model the degradation of the pipes.

3.2.3.3 Flow capacity in pipes

The first constraint is on the flow rate in the pipes. Indeed, the flow that can pass through a pipe is limited:

$$q_a^{min} \leq q_a \leq q_a^{max} \ , \ \ \forall a \in \mathbb{A} \ . \tag{3.5}$$

We have $q_a^{min} = -q_a^{max}$, $\forall (v, v') \in \mathbb{A}$, and q^{min} is a stationary vector. A positive flow value in arc (v, v') means that the flow goes from v to v'. A negative flow value means that the flow goes from v' to v.

3.2.3.4 Pressure range

The second constraint is on the range of pressures at a given point:

$$P_v^{min} \leq P_v \leq P_v^{max} \ , \ \ \forall v \in \mathbb{V} \ , \tag{3.6}$$

where P^{min} and P_v^{max} are stationary parameter.

3.2.4 Wells and production functions

A well is an infrastructure which allows the extraction of oil and/or gas from underground. We consider that it is made of two parts: the well perforations, where the fluids are flowing in the well, and a pipe that allows the fluids to flow from the bottom of the well to the well-head and the rest of the production network.

We consider that the production of the well $w \in \mathbb{V}_{in}$ is characterized by a production mapping called "Inflow Performance Relationship", which we denote by $\textsc{Ipr}_w$. It is a mapping of the "bottom-hole pressure" (i.e. the pressure at the well perforations) and the content of the reservoir near the well.

Let P_w^{R} be the reservoir pressure near well $w \in \mathbb{V}_{in}$, S_w^{W} be the saturation of water near well w (i.e. the proportion of water), S_w^{G} be the saturation of gas near well w and P_w^{BH} be the bottom-hole pressure of well w. The flow $q_{(w,v)}$ in the outgoing pipe of well w, $(w, v) \in \mathbb{A}_{well}$, is given by the mapping $\textsc{Ipr}_w : \mathbb{R}_+^2 \times [0, 1]^2 \to \mathbb{R}^3$

$$\textsc{Ipr}_w : (P_w^{\text{BH}}, P_w^{\text{R}}, S_w^{\text{W}}, S_w^{\text{G}}) \mapsto \begin{pmatrix} q^{\text{O}}_{(w,v)} \\ q^{\text{G}}_{(w,v)} \\ q^{\text{W}}_{(w,v)} \end{pmatrix} = q_{(w,v)} \ , \ \ (w, v) \in \mathbb{A}_{well} \ . \tag{3.7}$$

There is an extensive literature on possible mathematical expressions of the mappings $\textsc{Ipr}$ (Archer et al. [2003], Al-Rbeawi [2019]), as the possible expressions are highly dependent on the type of the considered well (its geometry, position, structure) and on the reservoir's characteristics (porosity of the rock). Note that, most of the time, the mathematical expressions of the mappings $\textsc{Ipr}$ are not "invertibles". For example, we may have $q^{\mathrm{W}}_{(w,v)} = 0$ over a range of water saturation S^{W}_w, even though we produce some oil $q^{\mathrm{O}}_{(w,v)}$, or we may also have $q^{\mathrm{G}}_{(w,v)} = 0$ over a range of gas saturation S^{G}_w.

We give below two $\textsc{Ipr}$ mappings used in numerical studies in Chapter 4 (Equations (4.18) and (4.21)), but first we make a detour devoted to another representation of the Inflow Performance Relationship, denoted by $\widetilde{\textsc{Ipr}}_w : \mathbb{R}^2_+ \times [0,1] \times \mathbb{R}_+ \to \mathbb{R}^3$, which is a function of the bottom-hole pressure P^{BH}, the reservoir pressure P^{R}, and of two ratios, the water-cut w^{CT} and the gas-oil ratio g^{OR}, which we now define.

The water-cut w^{CT}_w is the proportion of water produced when we extract a volume of liquid. If the production of liquid is not zero (i.e. if $q^{\mathrm{O}}_{(w,v)} + q^{\mathrm{W}}_{(w,v)} \neq 0$), we have

$$w^{\mathrm{CT}}_w = \frac{q^{\mathrm{W}}_{(w,v)}}{q^{\mathrm{W}}_{(w,v)} + q^{\mathrm{O}}_{(w,v)}} \; . \tag{3.8}$$

We define the water-cut function $\mathrm{W}^{\mathrm{CT}}_w : \mathbb{R}^2_+ \times [0,1] \times \mathbb{R}_+ \to \mathbb{R}^3$ as the function which returns the water-cut value given the variables P^{BH}_w , P^{R}_w, S^{W}_w, S^{G}_w. Hence, it is given as the composition of the function $\Psi^{\mathrm{wct}} : \mathbb{R}^3 \to \mathbb{R}, x \mapsto \frac{x^{(3)}}{x^{(1)}+x^{(3)}}$ with the mapping $\textsc{Ipr}_w$:

$$\mathrm{W}^{\mathrm{CT}}_w = \Psi^{\mathrm{wct}} \circ \textsc{Ipr}_w \; . \tag{3.9}$$

Moreover, as shown in Li and Li [2014], the water-cut function only depends on the oil saturation $S^{\mathrm{O}}_w = 1 - S^{\mathrm{W}}_w - S^{\mathrm{G}}_w$ and on the water saturation S^{W}_w, i.e. $\mathrm{W}^{\mathrm{CT}}_w : [0,1]^2 \to [0,1]$.

The gas-oil ratio g^{OR} is the proportion of gas produced when we extract a volume of oil. If the production of oil is not zero, we have

$$g^{\mathrm{OR}}_w = \frac{q^{\mathrm{G}}_{(w,v)}}{q^{\mathrm{O}}_{(w,v)}} \; . \tag{3.10}$$

We define the gas-oil ratio function $\mathrm{G}^{\mathrm{OR}}_w : \mathbb{R}^2_+ \times [0,1] \times \mathbb{R}_+ \to \mathbb{R}^3$ as the function which returns the gas-oil ratio given the variables P^{BH}_w , P^{R}_w, S^{W}_w, S^{G}_w. Hence, it is given as the composition of the function $\Psi^{\mathrm{gor}} : \mathbb{R}^3 \to \mathbb{R}, x \mapsto \frac{x^{(2)}}{x^{(1)}}$ with the mapping $\textsc{Ipr}_w$:

$$\mathrm{G}^{\mathrm{OR}}_w = \Psi^{\mathrm{gor}} \circ \textsc{Ipr}_w \; . \tag{3.11}$$

Moreover, as shown in Archer et al. [2003], the gas-oil function only depends on the gas saturation S^{G} and on the water saturation S^{W}, i.e. $\mathrm{G}^{\mathrm{OR}}_w : [0,1]^2 \to \mathbb{R}_+$.

Using the water-cut function and the gas-oil ratio function, it is proved in Al-Rbeawi [2019] that there exists a mapping $\widetilde{\textsc{Ipr}}_w$ such that for all P^{BH}_w, P^{R}_w, S^{W}_w, S^{G}_w we have

$$\widetilde{\textsc{Ipr}}_w\big(P^{\mathrm{BH}}_w, P^{\mathrm{R}}_w, \mathrm{W}^{\mathrm{CT}}_w(1 - S^{\mathrm{W}}_w - S^{\mathrm{G}}_w, S^{\mathrm{W}}_w), \mathrm{G}^{\mathrm{OR}}_w(S^{\mathrm{W}}_w, S^{\mathrm{G}}_w)\big) = \textsc{Ipr}_w(P^{\mathrm{BH}}_w, P^{\mathrm{R}}_w, S^{\mathrm{W}}_w, S^{\mathrm{G}}_w) \; . \tag{3.12}$$

As W^{CT} and G^{OR} are obtained as the compositions of functions Ψ^{wct} and Ψ^{gor} with mapping $\textsc{Ipr}$, there is a mapping $h : \mathbb{R}^3 \to \mathbb{R} \times [0,1] \times \mathbb{R}$ such that

$$h \circ \textsc{Ipr}_w(P^{\mathrm{BH}}_w, P^{\mathrm{R}}_w, S^{\mathrm{W}}_w, S^{\mathrm{G}}_w) = (P^{\mathrm{R}}_w, \mathrm{W}^{\mathrm{CT}}_w(1 - S^{\mathrm{W}}_w - S^{\mathrm{G}}_w, S^{\mathrm{W}}_w), \mathrm{G}^{\mathrm{OR}}_w(S^{\mathrm{W}}_w, S^{\mathrm{G}}_w)) \ .$$

We hence rewrite Equation (3.12) as:

$$\widetilde{\textsc{Ipr}}_w\big(P^{\mathrm{BH}}_w, h \circ \textsc{Ipr}_w(P^{\mathrm{BH}}_w, P^{\mathrm{R}}_w, S^{\mathrm{W}}_w, S^{\mathrm{G}}_w)\big) = \textsc{Ipr}_w(P^{\mathrm{BH}}_w, P^{\mathrm{R}}_w, S^{\mathrm{W}}_w, S^{\mathrm{G}}_w) \ .$$

The construction of the mapping $\widetilde{\textsc{Ipr}}$ is detailed in Al-Rbeawi [2019], where the author expresses the multi-flow of fluids produced $q_{(w,v)}$ as a function of the relative permeability of each phase depending on their saturation. The construction of the mapping $\widetilde{\textsc{Ipr}}$ is possible thanks to some key properties of the Inflow Performance Relationship mappings that we detail now.

($\textsc{Ipr}_1$) For a fixed ordered pair of water saturation and gas saturation $(S^{\mathrm{W}}_w, S^{\mathrm{G}}_w) \in [0,1]^2$, the flow of oil $q^{\mathrm{O}}_{(w,v)}$ is an increasing function of the difference between the reservoir pressure P^{R}_w and the bottom-hole pressure P^{BH}_w (see [Dake, 1983, Chap. 6, Table 1]). For the rest of the fluids (i.e. for $q^{\mathrm{G}}_{(w,v)}$ and $q^{\mathrm{W}}_{(w,v)}$), they are either the null functions, or they are increasing functions.

($\textsc{Ipr}_2$) The water-cut function $\mathrm{W}^{\mathrm{CT}}_w$ as defined in Equation (3.9) is a nondecreasing function of the water saturation S^{W}_w. This point is illustrated in Li and Li [2014], where the authors notably studied the impact of the water saturation on the recovery of oil.

($\textsc{Ipr}_3$) The gas-oil $\mathrm{G}^{\mathrm{OR}}_w$ function as defined in Equation (3.11) is a nondecreasing function of the gas S^{G}_w, as illustrated by the relative permeability of gas in Archer et al. [2003].

($\textsc{Ipr}_4$) For a given pair of water-cut and the gas-oil ratio $(w^{\mathrm{CT}}, g^{\mathrm{OR}}) \in [0,1] \times \mathbb{R}_+$, there is a unique difference of pressure between the reservoir pressure and the bottom-hole pressure leading to a given flow of oil $q^{\mathrm{O}}_{(w,v)}$. Otherwise stated, for a given tuple $(w^{\mathrm{CT}}, g^{\mathrm{OR}}, q^{\mathrm{O}}_{(w,v)}) \in [0,1] \times \mathbb{R}_+ \times \mathbb{R}$, there exists a unique difference of pressure $\Delta P \in \mathbb{R}$ such that the following system of equations has at least one solution $(P^{\mathrm{BH}}_w, P^{\mathrm{R}}_w, S^{\mathrm{W}}_w, S^{\mathrm{G}}_w)$:

$$\begin{aligned}
P^{\mathrm{R}}_w - P^{\mathrm{BH}}_w &= \Delta P \ , \\
\Psi^{\mathrm{wct}} \circ \textsc{Ipr}_w(P^{\mathrm{BH}}_w, P^{\mathrm{R}}_w, S^{\mathrm{W}}_w, S^{\mathrm{G}}_w) &= w^{\mathrm{CT}} \ , \\
\Psi^{\mathrm{gor}} \circ \textsc{Ipr}_w(P^{\mathrm{BH}}_w, P^{\mathrm{R}}_w, S^{\mathrm{W}}_w, S^{\mathrm{G}}_w) &= g^{\mathrm{OR}} \ , \\
\textsc{Ipr}_w(P^{\mathrm{BH}}_w, P^{\mathrm{R}}_w, S^{\mathrm{W}}_w, S^{\mathrm{G}}_w) &= q^{\mathrm{O}}_{(w,v)} \ .
\end{aligned}$$

Indeed, the function that, for a fixed pair of water-cut and the gas-oil ratio $(w^{\mathrm{CT}}, g^{\mathrm{OR}}) \in [0,1] \times \mathbb{R}_+$, takes as input a difference of pressure and returns the oil production is increasing with the difference of pressure, as illustrated in Daoud et al. [2017].

The existence of the mapping $\widetilde{\textsc{Ipr}}_w$ is derived by combining the four Points ($\textsc{Ipr}_1$)-($\textsc{Ipr}_4$). Moreover, it is established that the mapping $\widetilde{\textsc{Ipr}}_w(P^{\mathrm{BH}}_w, \cdot)$ is invertible.

We now present the two simplifications $\textsc{Ipr}_w$ that we will use in the applications presented in this thesis:

- When considering a well which produces only oil and water, the function $\textsc{Ipr}$ can be simplified to a simplified Darcy's law. Indeed, when there is no gas near the well (i.e. $S_w^{\text{G}} = 0$), we obtain the following simplification for the production of oil and water. We can assume that the *total flow* $q_{(w,v)}^{\text{R}}$ is given by

  $$q_{(w,v)}^{\text{R}} = \alpha_w (P_w^{\text{R}} - P_w^{\text{BH}}) \ ,$$

  where $q_{(w,v)}^{\text{R}}$ is given by

  $$q_{(w,v)}^{\text{R}} = q_{(w,v)}^{\text{O}} + \underbrace{q_{(w,v)}^{\text{G}}}_{=0} + q_{(w,v)}^{\text{W}} \ ,$$

  with α_w the productivity index of the well w and $q_{(w,v)}^{\text{R}}$ the total flow which consists of a mix of oil and water.

  The partition of the total flow between the flow of oil and water is given by the water cut function W_w^{CT}

- When considering a well that only produce gas, the Inflow Performance Relationship $\textsc{Ipr}$ is simplified. We hence obtain

  $$q_{(w,v)}^{\text{G}} = \textsc{Ipr}_w^{\text{G}} (P_w^{\text{R}} - P_w^{\text{BH}}) \ .$$

3.3 Representing the exploitation of a reservoir

A reservoir is a geological formation that contains some resources (oil and/or gas). It is a key element of an oil-field. Thus, how we model it can have a significant impact on the optimization of oil-fields. There are many possible models for the reservoir. In this section, we will present two of such models: the decline curves, and a black-oil tank model. In §3.3.1 we present the decline curves model. In §3.3.2, we will present the black-oil model.

All reservoir models are time dependent, and are usually written with discrete time. We will therefore assume that we have discrete time steps which belongs to the set $\mathcal{T}$, defined as:

$$\mathcal{T} = \{1, \ldots, T\}$$

Each $t \in \mathcal{T}$ represents the time step $[t, t+1)$. and the horizon T is the final step considered.

3.3.1 Description of the exploitation of a reservoir through decline curves

The simplest representation of the reservoir are decline curves or oil-deliverability curves. Usually, decline curves describe the bottom-hole pressure or gives the maximal production possible at a well depending on the amount of oil that has been produced. They thus return the reservoir pressure at one well depending on what has already been produced from it. They also return the water cut and the gas oil-ratio. Decline curves can be good approximation of the exploitation of a reservoir when considering primary extraction (i.e. producing hydrocarbon without re-injection in the reservoir) of a one tank reservoir. We also note $Q_{v,t}$ the cumulated oil extracted at v and t, i.e.:

$$Q_{v,t} = \sum_{\tau \leq t} q_{v,\tau} \ , \ \ \forall (v,t) \in \mathbb{V}_{in} \times \mathcal{T} \ .$$

With those notations, the decline curves looks like in Figure 3.2. Usually, decline-curves are functions g that return the maximal production:

$$q_{v,t} \leq g(Q_{v,t}) \ , \ \ \forall t \in \mathcal{T} \ .$$

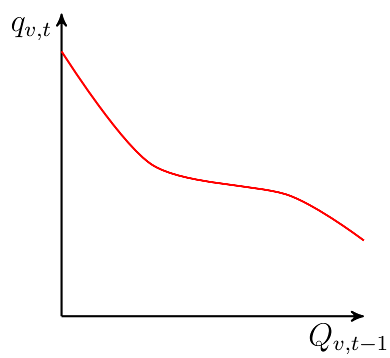


FIGURE 3.2: Illustration of a decline curve for a given well v. The X-axis represents the total amount of oil that has been produced on that well. The Y-axis represents the resulting maximal possible production.

With this representation, we can easily describe each well as a dynamical system. The main issue is that we naturally get one state per well. This means that we won't be able to use the decline curves when there are multiple wells, as we would have a state too large for any computation.

3.3.2 Black-oil reservoir model

The second representation of a reservoir is the black-oil tank model. This model allows us to have a simplistic, but global description of the reservoir, with interaction between wells. We will first describe the reservoir and the assumptions for the black-oil tank-like model. We will then write the material balance equations that result from that description.

Note that, usually, such equations are written with parameters in mass units. However, we will use parameters in standard volume instead of mass. Since the standard volumes are defined for a given pressure and temperature, standard volume is proportional to mass through density (which is a constant at the standard conditions of pressure and temperature). Thus, all the following mass conservation equations will be written in standard volume units, and not in mass units.

3.3.2.1 Reservoir description

We consider a tank-like reservoir, like in Figure 3.3. We also consider that the fluids in the reservoir follow the black-oil model. This means that the reservoir can contain three fluids: oil, gas and water. We consider that their physical properties are stationary (constant in time) and are uniform in the reservoir (constant in space). We also consider that those fluids can be divided in up to two phases in the reservoir: a liquid phase, and a gaseous phase. In the liquid phase, we can have a mix of oil, dissolved gas and water. In the gaseous phase, we can only have free gas.

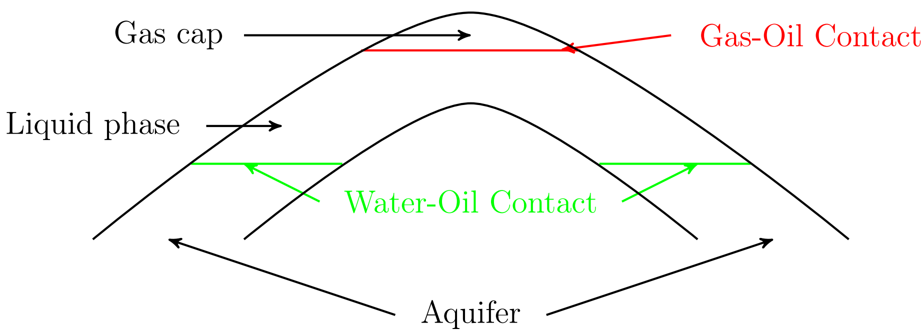


FIGURE 3.3: Representation of a tank-like reservoir

The standard volume taken by those components are defined in Table 3.3. These components are stored in the rocks' pores, and the total pore volume is written V^{P}.

Notation	Definition
V^{O}	Standard oil volume (liquid phase)
V^{DG}	Standard dissolved gas volume (liquid phase)
V^{W}	Standard water volume (liquid phase)
V^{G}	Standard free gas volume (gaseous phase)

TABLE 3.3: Definitions of the volume in the reservoir

We also consider that the reservoir pressure is uniform in the reservoir, so that both phases and all fluids will be at the same pressure, written P^{R}. The reservoir temperature is uniform and stationary. Temperature will therefore not appear in any constitutive equations, as it is a constant. Since we assume that we have a tank-like reservoir, we assume that the free gas travels instantly to the gas cap. We also assume that the reservoir's cap-rock cannot fracture. It means that there is no “leak”, and that we can write mass conservation equations.

The last component of the model of the reservoir is the aquifer model. An aquifer is a geological formation around the reservoir which contains water and not hydrocarbons. When there is an aquifer, we consider that it is sufficiently large to be considered infinite compared to the reservoir. We therefore won't need to explicitly account for water in the aquifer or for changes in it.

All the variables, $V^{\mathrm{O}}, V^{\mathrm{G}}, V^{\mathrm{DG}}, V^{\mathrm{W}}, V^{\mathrm{P}}$ and P^{R} can change over time. They will therefore be indexed by a $t \in \mathcal{T}$, where $\mathcal{T}$ is the set of time-steps.

Now that we have presented the different components of a reservoir, we can write material balance equations to describe their evolution.

3.3.2.2 Oil material balance

The oil material conservation gives a dynamical equation for the amount of oil V_t^{O}:

$$V_{t+1}^{\mathrm{O}} = V_t^{\mathrm{O}} - F_t^{\mathrm{O}} \ , \ \ \forall t \in \mathcal{T} \ , \tag{3.13}$$

where F_t^{O} the standard volume of oil extracted from the reservoir during the interval $[t, t+1)$. The oil material conservation equation states that the variation of the amount of oil in the reservoir between two time-steps is equal to the oil production between the two time-steps.

3.3.2.3 Gas material balance

We express here the mass conservation for the gas. We proceed as follows. First, during interval $[t, t+1)$ a quantity F_t^{G} of gas is extracted from the reservoir. Second, during the same interval, the quantity of oil has evolved from V_t^{O} to V_{t+1}^{O} (using (3.13)) and the pressure has evolved from P_t^{R} to P_{t+1}^{R}. At any time, we assume that the quantity of dissolved gas in the oil V^{DG} is given by a function of the reservoir pressure P^{R} and the amount of oil in the reservoir V^{O} as

$$V^{\mathrm{DG}} = \delta(V^{\mathrm{O}}, P^{\mathrm{R}}) = V^{\mathrm{O}} \cdot R_s(P^{\mathrm{R}}) \ , \tag{3.14}$$

where R_s is a piece-wise linear function (represented in Figure 3.4) which returns the solution gas dissolved in the oil.

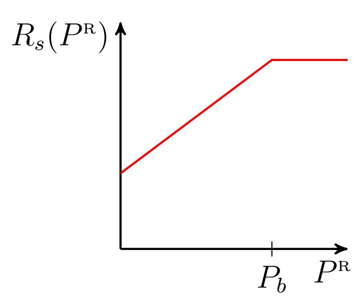


FIGURE 3.4: representation of the function R_s, that returns the proportion of gas that is dissolved in the oil for a given reservoir pressure

Thus, the quantity of dissolved gas has evolved from $V_t^{\mathrm{DG}} = \delta(V_t^{\mathrm{O}}, P_t^{\mathrm{R}})$ to $V_{t+1}^{\mathrm{DG}} = \delta(V_{t+1}^{\mathrm{O}}, P_{t+1}^{\mathrm{R}})$. Therefore, the quantity of liberated gas $(V_t^{\mathrm{DG}} - V_{t+1}^{\mathrm{DG}})$ must be added to the gas mass conservation equation. Thus, we have a mass conservation equation for the free gas that can be written:

$$\begin{aligned}
V_{t+1}^{\mathrm{G}} &= V_t^{\mathrm{G}} - F_t^{\mathrm{G}} + \underbrace{(V_t^{\mathrm{DG}} - V_{t+1}^{\mathrm{DG}})}_{liberated\ gas} \\
&= V_t^{\mathrm{G}} - F_t^{\mathrm{G}} + \big(V_t^{\mathrm{O}} \cdot R_s(P_t^{\mathrm{R}}) - V_{t+1}^{\mathrm{O}} \cdot R_s(P_{t+1}^{\mathrm{R}})\big) \qquad \text{(using Equation (3.14))} \\
&= V_t^{\mathrm{G}} - F_t^{\mathrm{G}} + \Big(V_t^{\mathrm{O}} \cdot R_s(P_t^{\mathrm{R}}) - \big(V_t^{\mathrm{O}} - F_t^{\mathrm{O}}\big) \cdot R_s(P_{t+1}^{\mathrm{R}})\Big)
\end{aligned}$$

(using Equation (3.13) to transform V_{t+1}^{O} as an expression with terms depending only on t)

Hence, we have

$$V_{t+1}^{\mathrm{G}} = V_t^{\mathrm{G}} - F_t^{\mathrm{G}} + \Big(V_t^{\mathrm{O}} \cdot \big(R_s(P_t^{\mathrm{R}}) - R_s(P_{t+1}^{\mathrm{R}})\big) + F_t^{\mathrm{O}} \cdot R_s(P_{t+1}^{\mathrm{R}})\Big) \,. \tag{3.15}$$

3.3.2.4 Water material balance

The water material balance gives us the dynamics equation

$$V_{t+1}^{\mathrm{W}} = V_t^{\mathrm{W}} - F_t^{\mathrm{W}} + F_t^{\mathrm{WE}} \,, \tag{3.16}$$

with F_t^{W} the amount of water produced during interval $[t, t+1)$, and F_t^{WE} the water encroachment (i.e. the inflow of water from the aquifer to the reservoir) between t and $t+1$.

The mass conservation equation states that the variation of the amount of water in the reservoir between two time-steps is equal to the difference between the water production and the water encroachment between the two time-steps.

3.3.2.5 Total pore volume change

In this section, we express the change in the total pore volume between two time steps. We present two methods, which we use in different cases:

- a finite difference method, presented in Dake [1983], which can always be used but leads to approximations;
- a solution of a differential equation, more precise, but which can only be used in certain conditions as it often leads to transcendental equations.

We first present the finite difference method. According to Dake [1983], the change in the total pore volume satisfies:

$$\frac{V_{t+1}^{\mathrm{P}} - V_t^{\mathrm{P}}}{V_t^{\mathrm{P}}} = c_f(P_{t+1}^{\mathrm{R}} - P_t^{\mathrm{R}}) \,, \tag{3.17}$$

where c_f gives the compressibility of the rocks in the reservoir. It is assumed to be a positive constant in the pressure range considered. Equation (3.17) express the fact that the relative pore volume change between two time-steps is proportional to the difference of pressure between the two time-steps.

However, Equation (3.17) is obtained by a first order Taylor approximation of the changes of the total pore volume depending on the pressure. It may raise some problems when used for example with a gas storage reservoir. Indeed, if we start at a volume V_0^{P}, then we first raise the pressure by ΔP at time 1, and then change it again by $-\Delta P$, then we obtain that $V_2^{\mathrm{P}} \neq V_0^{\mathrm{P}}$ as we have a second order error:

$$\begin{aligned} V_2^{\mathrm{P}} &= V_1^{\mathrm{P}}(1 - c_f \Delta P) \\ &= V_0^{\mathrm{P}}(1 + c_f \Delta P)(1 - c_f \Delta P) \\ &= V_0^{\mathrm{P}}\left(1 - (c_f \Delta P)^2\right) . \end{aligned}$$

This contradicts the physics which states that we should have returned to the initial volume, i.e. $V_2^{\mathrm{P}} = V_0^{\mathrm{P}}$.

We now present the second method, based on the solution of a differential equation. In order to reduce the number of variables necessary to describe the reservoir, we would like a function giving the reservoir volume V_t^{P} depending on the reservoir pressure P_t^{R}. Indeed, the pore volume is a function of the pressure, and it should not change depending on the history of the reservoir. To get that function, we need to get back to the definition of c_f:

$$c_f = \frac{1}{V^{\mathrm{P}}} \left(\frac{\partial V^{\mathrm{P}}}{\partial P} \right)_{\mathrm{T}} ,$$

where T is the temperature in the reservoir. Since we consider that the temperature T is stationary (tank-like assumption), the pore volume only depends on the reservoir pressure, and we end up with the following differential equation:

$$(V^{\mathrm{P}})' - c_f V^{\mathrm{P}} = 0 .$$

Since c_f is assumed to be a positive constant, we obtain the following expression for the pore volume:

$$V^{\mathrm{P}}(P^{\mathrm{R}}) = a e^{c_f P^{\mathrm{R}}} , \tag{3.18}$$

where a is a given positive constant.

The main issue with that expression is that, when combined with the other material balance equations (i.e. with Equations (3.13)-(3.16)), it leads to a transcendental equation. We thus cannot get an expression of the pressure in the reservoir in the general case when considering Equation (3.18). In most cases, we will therefore use Equation (3.17) despite the second order error.

There are three main cases where we can use Equation (3.18):

- A reservoir which only contains gas.

- A reservoir which contains gas and water.
- A reservoir which contains oil and water but where pressure stays below the bubble point.

3.3.2.6 Saturation conservation

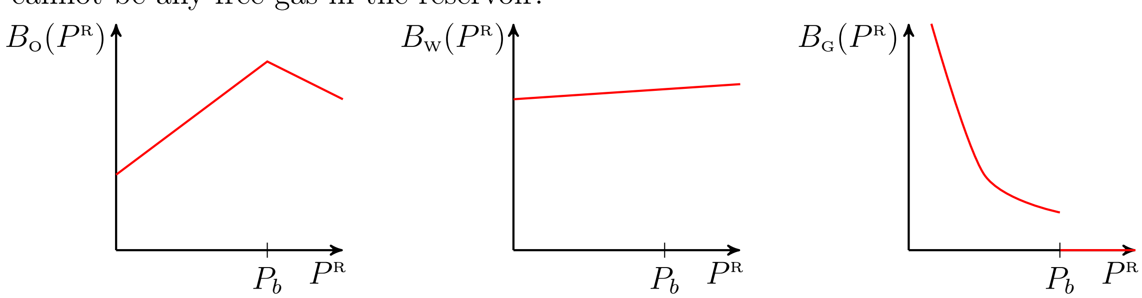
We now present a conservation equation for the reservoir. First, since the reservoir is defined by the total pore volume, we can define the saturations of each fluid contained in the reservoir. The saturation of a given fluid correspond to the proportion of the pore volume taken by that fluid. We must have a conservation of the saturations in the reservoir, as there cannot be any void. Hence:

$$S_t^{\text{O}} + S_t^{\text{G}} + S_t^{\text{W}} = 1 \ , \ \ \forall t \in \mathcal{T} \ . \tag{3.19}$$

To get the saturation, we need to transform the standard volumes into volumes at reservoir conditions. We need the functions (called factor in reservoir engineering):

Notation	Definition
B_{O}	Oil formation factor
B_{G}	Gas formation factor
B_{W}	Water formation factor

Those functions take the pressure of the reservoir as input and return the volume taken by one standard cubic meter of a fluid when at that reservoir pressure. The functions are derived from the physical properties of the fluids as they are part of the PVT (Pressure Volume Temperature) functions, and are stationary. In the black-oil model, the functions B_{O} and B_{W} can be approximated by piece-wise linear functions. The function B_{G} is inversely proportional to the pressure until the so-called bubble point pressure P_b, beyond which there cannot be any free gas in the reservoir.



Note that in our cases, we extend B_{G} beyond the bubble point with a constant value 0. We thus obtain a function which is defined for all possible pressures and with finite values. Note that the value of the constant used to extend the function B_{G} doesn't matter as long as it is finite, as every time this function is used, it is multiplied with a term whose value is 0 when the pressure is higher than the bubble point P_b.

For each phase, the saturation can be computed at each time t from the pore volume (i.e. the available volume in the reservoir), the quantity of the phase in the reservoir and the reservoir pressure. We have, for each fluid:

Notation	Definition	Formula
S_t^{O}	Oil saturation	$\frac{V_t^{\mathrm{O}} \cdot B_{\mathrm{O}}(P_t^{\mathrm{R}})}{V_t^{\mathrm{P}}}$
S^{G}	Free gas saturation	$\frac{V_t^{\mathrm{G}} \cdot B_{\mathrm{G}}(P_t^{\mathrm{R}})}{V_t^{\mathrm{P}}}$
S^{W}	Water saturation	$\frac{V_t^{\mathrm{W}} \cdot B_{\mathrm{W}}(P_t^{\mathrm{R}})}{V_t^{\mathrm{P}}}$

With the saturations, we can rewrite the saturation Equation (3.19) as:

$$V_t^{\mathrm{O}} \cdot B_{\mathrm{O}}(P_t^{\mathrm{R}}) + V_t^{\mathrm{G}} \cdot B_{\mathrm{G}}(P_t^{\mathrm{R}}) + V_t^{\mathrm{W}} \cdot B_{\mathrm{W}}(P_t^{\mathrm{R}}) = V_t^{\mathrm{P}} \ , \ \ \forall t \in \mathcal{T} \ . \tag{3.20}$$

3.3.2.7 Model for the water encroachment

We now detail how we model the water encroachment (the inflow of water in the reservoir that comes from the aquifer). In Dake [1983], it is considered that the water encroachment F^{WE} (used in Equation (3.16)) is not known. There are two main methods to take the water encroachment into account:

- We either assume that the water encroachment is governed by constitutive equations. However, the parameters of those equations are uncertain.
- Or we assume that the water encroachment is a stochastic process governed by known probability distributions.

The second approach is not used by oil companies. Indeed, a stochastic process for the water encroachment might be difficult to define. Therefore, it seems difficult to get an accurate description of the reservoir by modeling the water encroachment by a stochastic process. We will thus only focus on the first approach.

We will describe the following models for the water encroachment:

- Pot aquifer model,
- Schilthuis' steady-state,
- Hurst's modified steady-state,
- The Van Everdingen-Hurst unsteady-state,
- The Carter-Tracy unsteady-state.

Pot aquifer model. In this model, we consider that the water encroachment F^{WE} comes from the expansion of the aquifer. That means that, when pressure drops in the reservoir, the aquifer's volume changes and some water comes into the reservoir.

This model can accurately describe a "small" aquifer, and we can then consider it to follow the same equations as the reservoir. We then get that the total water encroachment follows

$$F^{\text{WE}} = (c_{\text{W}}^{\text{AQU}} + c_f^{\text{AQU}})V^{\text{AQU,ini}}(P^{\text{AQU,ini}} - P) \ ,$$

with $V^{\text{AQU,ini}}$ the initial volume of the aquifer, $c_{\text{W}}^{\text{AQU}}$ the water compressibility for the aquifer water, c_f^{AQU} the aquifer pore compressibility and F^{WE} the total water encroachment.

We therefore get that the water encroachment between t and $t+1$ is:

$$F_t^{\text{WE}} = (c_{\text{W}}^{\text{AQU}} + c_f^{\text{AQU}})V^{\text{AQU,ini}}(P_t^{\text{R}} - P_{t+1}^{\text{R}}) \ . \tag{3.21}$$

Schilthuis' steady-state. In this model, we consider that the water encroachment F^{WE} is governed by a Darcy's equation, and that the flow is in a steady-state. Thus, we consider that the aquifer pressure P^{AQU} is stationary. We then have

$$F_t^{\text{WE}} = C(P^{\text{AQU,ini}} - P_t^{\text{R}}) \ , \tag{3.22}$$

where C is a given constant that depends on the geometry and on the type of rocks of the aquifer-reservoir limit.

Hurst's modified steady-state. The issue with the previous model is that the geometry of the aquifer-reservoir contact surface may change due to the water encroachment. Indeed, as the total water encroachment rises, the oil volume is reduced as its pressure decreases, and the aquifer-reservoir contact surface may get smaller. The Hurst's model's goal is to take that change into account. It states that

$$F_t^{\text{WE}} = \frac{C}{\ln(a \cdot t)}(P^{\text{AQU,ini}} - P_t^{\text{R}}) \ , \tag{3.23}$$

where a is a constant depending on the aquifer.

Van Everdingen-Hurst unsteady-state. Previously, we considered that the pressure of the aquifer was uniform. Van Everdingen and Hurst proposed to take into account the diffusion of the pressure in the aquifer. Van Everdingen and Hurst state that the pressure of the aquifer verifies:

$$\frac{\partial P}{\partial r_D^2} + \frac{1}{r_{\text{AQU}}}\frac{\partial P}{\partial r_D} = \frac{\partial P}{\partial t} \ .$$

The pressure in the aquifer also needs to verify boundary conditions.

Van Everdingen and Hurst proposed that the water encroachment be of the form:

$$\frac{\partial F_t^{\text{WE}}}{\partial t} = \frac{B}{h\big(\tau_D(t)\big)}\big(P^{\text{AQU}}(t) - P^{\text{R}}(t)\big) , \tag{3.24}$$

where B is a parameter that is defined by the aquifer-reservoir border, and the function τ_D depends on the physical properties of the water and rocks in the aquifer. Finally, h is a function that depends on the ratio of the size of the reservoir and the size of the aquifer.

Van Everdingen and Hurst further detail those parameters depending on the geometry of the aquifer, and more specifically if we are in an "edge" or "bottom" water-drive (two specific form of geology of the reservoir and the aquifer, where parameters can computed more efficiently).

The issue is that Equation (3.24) needs to be integrated between each time-step.

Carter-Tracy equations. The Carter-Tracy model was made to make the previous Van Everdingen-Hurst model easier to compute. It states that

$$F_t^{\text{WE}} = \big(\tau_D(t) - \tau_D(t-1)\big)\frac{B(P_{t-1}^{\text{R}} - P_t^{\text{R}}) - F_{t-1}^{\text{WE}} h_1\big(\tau_D(t)\big)}{h_2\big(\tau_D(t)\big) - \tau_D(t-1)h_1\big(\tau_D(t)\big)} , \tag{3.25}$$

where h_1 and h_2 are functions that depend on the ratio of the size of the aquifer and the size of the reservoir.

3.4 Conclusion

In this chapter, we have presented all the relevant constitutive equations used in the following chapters. The reservoir and fluids description are used to represent the reservoir as a controlled dynamical system, as detailed in Chapter 4. Meanwhile, the constitutive equations for the production network are used to define admissible control set of the mathematical formulations of the management of a petroleum production system.

Chapter 4

Multistage Optimization of a Petroleum Production System with Material Balance Model

Contents

This chapter is a transcript of an article accepted in *Computers & Chemical Engineering* (see Vessaire et al. [2022]).

4.1 Introduction

Oil and gas projects usually span over several decades and involve complex planning and decision-making. Therefore, multistage optimization is a relevant tool to address the long-term performance of such projects. This is the focus of this paper.

The lifetime of a field usually consists of five phases: exploration, where reservoirs containing hydrocarbon are found; appraisal, to give a value to a field; development, where infrastructures are planned and installed; production, where hydrocarbon is finally produced; abandonment, where the field stops producing and infrastructures are decommissioned and removed. In this paper, we focus on the production phase. We consider that the infrastructure has already been installed in the development phase, and we thus focus on finding a production schedule that maximizes the profit over the full production phase.

Now, we position our contribution with respect to the currently available literature. According to the survey [Khor et al., 2017], there is extensive research on how to optimize the production phase, with multiple approaches. The authors present three main methods for the optimization of petroleum production systems: sensitivity analysis by employing simulation tools, heuristic rules and mathematical optimization, the approach of this paper. Most of the literature resorts to the first two approaches.

Regarding mathematical optimization, most works on the topic have considered black-box simulators to describe the reservoir dynamics: Hepguler et al. [1997] consider integrating both a network model and a proprietary reservoir model (a commercial simulation software for reservoir modeling); Gerogiorgis et al. [2006] combine a proprietary reservoir simulator with a general optimization formulation. In Sarma et al. [2006] a closed-loop multistage optimal control approach with a simulator that can be updated with new data from sensors is considered. It is also a standard practice to add some optimization layer over a commercial reservoir simulator to locally improve a production planning, such as modifying the pressure on different points of the petroleum production system to locally improve an operational solution (see ECLIPSE by Schlumberger, or GAP and MBAL by Petroleum Experts). In theory, such approach could be amenable to dynamic programming. However, this is not done in practice due to the the computation time of a single simulation run.

A limited fraction of the literature addresses the problem as a multistage optimization problem, such as in Iyer et al. [1998], Gupta and Grossmann [2012], Marmier et al. [2019]. In those papers, the formulation relies on dynamical models based on decline curves (or type curves). In short, decline curves are functions that take as input the cumulative production and return the maximal well rate. In the context of mathematical optimization, decline curves were first assumed to be linear, such as in Bohannon [1970], before being assumed to be piecewise linear in Frair and Devine [1975] or polynomial in Marmier et al. [2019], or even being assumed to be given by a set of logical relationships for shale gas in Hong et al. [2020], when algorithms could treat those refinements. The decline curves are generally constructed by using a foresight of the optimal solution that is looked after, as they are usually generated by assuming a production schedule. In Satter and Iqbal [2016], the authors write that, usually, decline curves analysis is performed under one key assumption: the wells produce at "constant bottom-hole pressure". They also state that "in reality, such a condition may not be observed". Note that decline curves can, in some cases, provide an accurate representation of the reservoir if the wells that constitute the oil field are independent of each other, and when we are only considering first recovery of oil and gas (i.e. when we are only producing fluids in the reservoir and without any injection of gas and water in the reservoir). Despite those shortcomings, mathematical formulations using decline curves are commonly used in oilfield development studies. For example, two case studies, one in Brazil [Silva and Guedes Soares, 2021] and one in New Mexico [Davis, 2021] use decline curves to solve a multistage optimization problem.

Part of the literature also tries to develop a middle ground between using a black-box reservoir simulator and using decline curves. For example, some papers use parametrized surrogate models (also called proxy models). Parameters of the surrogate model are to be adjusted to fit simulators output or real data (see [Caballero and Grossmann, 2008]). Numerous applications following the methodology developed in Caballero and Grossmann [2008] have been done, each one being characterized by a specific surrogate model: in Lei et al. [2022], a proxy model (presented in Lei et al. [2021]) that takes into account the decommission timing and costs in the development planning is used; whereas in Camponogara et al. [2017], the authors use MILP as a proxy model and apply it to a case in the Santos Basin; finally, in Moolya et al. [2022], the authors also use a MILP surrogate model combined with aggregation and disaggregation methods in well placement problems. In Epelle and Gerogiorgis [2020], the authors compare the performances between MILP and MINLP formulations of the surrogate model.

In this paper, we represent the reservoir as a controlled dynamical system based on black-oil model and conservation laws (mass balance equations) for a tank-like reservoir instead of using decline curves or surrogate models based on a reservoir simulator. Mass balance equations belong nowadays to the folklore of petroleum engineering and have been described many times in the reservoir modeling literature (see Dake [1983]). We formulate the management problem as a multistage optimization problem, and we use the dynamic programming algorithm to solve it (see Bertsekas [2000]). To the best of our knowledge, this approach is new in the oil and gas literature. This formulation is well adapted to first and secondary recovery of oil and gas cases. Moreover, multistage optimization and dynamic programming are well adapted to tackle more complex formulations with uncertain parameters and partial observations.

4.2 Formulation of the management of a petroleum production system as a multistage optimization problem

We consider a production system composed of a reservoir and production assets (pipes, wells, chokes). We represent the topology of the production assets as a simple graph $\mathcal{G} = (\mathbb{V}, \mathbb{A})$, where $\mathbb{V}$ is the set of vertices and $\mathbb{A} \subset \mathbb{V}^2$ is the set of edges. Controls are variables indexed by either vertices or edges. We place the different production assets on the graph, with the pipes as the edges of the graph, and the rest of the assets, such as the well-heads, positioned on the vertices of the graph. This is illustrated in Figure 4.1. The wells' perforations are represented as vertices (w_i in Figure 4.1) where the fluids produced enter the graph. On the other vertices, we have assets such as the well-head chokes (wh_i in Figure 4.1), or joints between different pipes (noted i_1). We can also have valves to open or close pipes. Finally, we have an export point (on the vertex e).

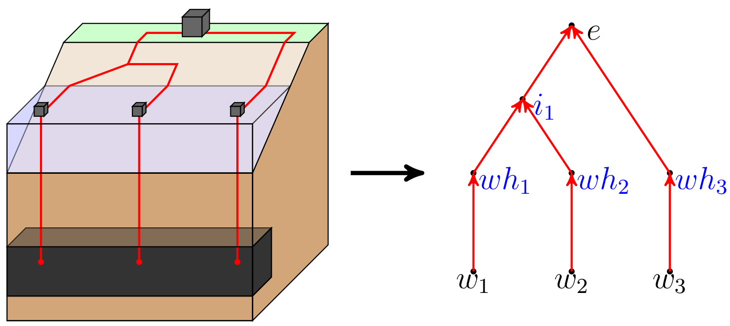


FIGURE 4.1: Representing a production network as a graph

All the relevant operational constraints and features - such as pressure loss on the pipes, mass balance of the fluids at each vertex, allowed pressures, and flow rate ranges in the different assets or unavailability due to maintenance - are modeled as constraints using variables defined on the edges and vertices of the graph. Indeed, the graph allows us to define the different controls we can apply to the system, such as opening or closing valves or changing the well-head pressures. Detailed formulations on the production network can be seen in [Gupta and Grossmann, 2012]. We will not explicit it in the general case as this is not our main focus, and we only present numerical applications without taking into account the production network.

As we aim to optimize the system over the whole production phase (i.e. over multiple years), we consider multiple time steps belonging to a finite set $\mathcal{T} = \{0, 1, 2, \ldots, T\}$ where the parameter T is a natural number. Those time steps are usually monthly[1], but under certain conditions other time steps may be considered.

We propose (and are going to detail) a general formulation of the petroleum production system optimization problem as follows

$$
\begin{aligned}
\mathcal{J}^\star(x_0) = \max_{x,u} \quad & \sum_{t=0}^{T-1} \rho^t \mathcal{L}_t(x_t, u_t) + \rho^T \mathcal{K}(x_T) && \text{(4.1a)} \\
s.t. \quad & x_0 \ given \ , && \text{(4.1b)} \\
& x_{t+1} = f(x_t, u_t) \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , && \text{(4.1c)} \\
& u_t \in \mathcal{U}_t^{ad}(x_t) \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ . && \text{(4.1d)}
\end{aligned}
$$

[1] Numerical applications will be done with monthly time steps and a horizon T of 15 or 20 years

The variables in Problem (4.1) are: *i*) the state of the reservoir $x_t \in \mathbb{X} \subset \mathbb{R}^n$ (with $\mathbb{X}$ the state space); *ii*) the controls $u_t \in \mathbb{U} \subset \mathbb{R}^p$ (with $\mathbb{U}$ the control space), which are the decisions that can be taken at time step t (for example, the pressure $P_{v,t}$ at the different vertices $v \in \mathbb{V}$ of the graph, and the Boolean $o_{a,t}$ stating if a pipe $a \in \mathbb{A}$ of the graph is opened or closed). The reservoir is defined as a controlled dynamical system, with state x_t, control u_t and an evolution function f of the controlled dynamical system, whose construction is the focus of Section 4.3. At every time step t, when the decision maker takes decision u_t, an instantaneous gain denoted by $\mathcal{L}_t(x_t, u_t)$ occurs. In the last stage, the final state x_T is valued as $\mathcal{K}(x_T)$. We denote by ρ the discount factor. We finally obtain the objective function seen to the right of the max in Equation (4.1a) by adding all terms. The known initial state of the reservoir is defined in Equation (4.1b). The controlled dynamics of the reservoir is given in Equation (4.1c). Equation (4.1d) states that, at each time step t, the allowed controls belong to an admissibility set that depends on x_t. The dependence is noted by $\mathcal{U}_t^{ad}(x_t)$, which is for each time step t a set-valued mapping that takes a given state x_t of the reservoir and returns the set of allowed controls. As far as the petroleum application is concerned, the admissibility set notably depends on the reservoir pressure, which constrains the different pressures in the petroleum production system. It also depends on the production network itself: some pipes can be controlled, while others cannot; facilities have planned or unplanned downtimes, etc. Extensive formulations of the admissibility set of the production depending on the reservoir pressure can be seen in Iyer et al. [1998].

The petroleum production system optimization problem, as formulated in (4.1), is a classical deterministic discrete time optimal control problem. It is known that this problem can be solved by dynamic programming and that the resulting optimal control at time t is a function of the current state at time t.

In order to solve Problem (4.1), we use a family of value functions $\mathcal{J}_t : \mathbb{X} \mapsto \mathbb{R}$, where we recall that $\mathbb{X}$ is the state space. We call *policy* $\mu = \{\mu_0, \ldots, \mu_{T-1}\}$ a set of mappings $\mu_t : \mathbb{X} \to \mathbb{U}$ from states x into admissible controls u. We have the following proposition (see [Bertsekas, 2016, Chap. 1]).

Proposition 4.1. *For every initial state $x_0 \in \mathbb{X}$, the optimal cost $\mathcal{J}^*(x_0)$ of Problem (4.1) is equal to $\mathcal{J}_0(x_0)$, given by the last step of the following algorithm, which proceeds backward in time from final time step T to initial time step 0:*

$$\mathcal{J}_T(x) = \rho^T \mathcal{K}(x) \ , \ \ \forall x \in \mathbb{X} \ , \tag{4.2a}$$

$$\begin{aligned} \mathcal{J}_t(x) = \max_{u \in \mathcal{U}_t^{ad}(x)} \Big(& \rho^t \mathcal{L}_t(x, u) \\ & + \mathcal{J}_{t+1}\left(f(x, u)\right) \Big) \ , \ \ \forall x \in \mathbb{X}, \forall t \in \mathcal{T} \setminus \{T\} \ . \end{aligned} \tag{4.2b}$$

Furthermore, if $u^\star = \mu_t^\star(x)$ maximizes the right-hand side of (4.2b) for each x and t, then the policy $\mu^\star = \left\{\mu_0^\star, \ldots, \mu_{T-1}^\star\right\}$ is optimal.

To solve Problem (4.1), we compute $\mathcal{J}_0$. To do so, we use a dynamic programming algorithm (see Algorithm 1). For that purpose, we discretize the controls, that now belong to a finite set denoted by $\mathbb{U}_d$, and the states that belong to a finite set $\mathbb{X}_d$. Numerically, we also use a multi-linear interpolation for the value functions between the states.

Algorithm 1: dynamic programming algorithm used to solve Problem (4.1)

```
for x ∈ X_d do
    J_T(x) = ρ^T K(x);
for t = T − 1, ..., 1 do
    for x ∈ X_d do
        best_value = - ∞;
        best_controls = 0;
        for u ∈ U_d do
            current_value = ρ^t L_t(x, u) + J_{t+1}( f(x, u) );
            if current_value ≥ best_value then
                best_value = current_value;
                best_controls = u;
        J_t(x) = best_value;
        μ_t(x) = best_controls;
return (J_t, μ_t)_{t∈T}
```

4.3 Formulation of the reservoir extraction as a controlled dynamical system

In this section, we show how to represent the time evolution of the reservoir as a dynamical system, that is, involving a state x, a control u and an evolution function f such that, for each time step t, we have $x_{t+1} = f(x_t, u_t)$. It is shown in Appendix 4.A that a possible state - which is the one we henceforth consider, for modeling the reservoir when using the black-oil model and conservation laws for a tank-like reservoir - is the 5-dimensional vector $x_t = (V_t^{\mathrm{O}}, V_t^{\mathrm{G}}, V_t^{\mathrm{W}}, V_t^{\mathrm{P}}, P_t^{\mathrm{R}})$. Its components are defined in Table 4.1, where Sm^3 stands for standard cubic meter (the volume taken by a fluid at standard pressure and temperature condition: 1.01325 Bara and 15°C), and Bara stands for absolute pressure in Bar.

More precisely, to obtain the evolution function f of the content of the reservoir between time t and $t+1$, we compute the amounts of fluids (oil, gas, water) produced during the period $[t, t+1[$. We denote them by $(F_t^{\mathrm{O}}, F_t^{\mathrm{G}}, F_t^{\mathrm{W}})$ and they are described in Table 4.2. We obtain the production values with a mapping $\Phi = (\Phi^{(1)}, \Phi^{(2)}, \Phi^{(3)}) : \mathbb{X} \times \mathbb{U} \to \mathbb{R}^3$ such that $(F_t^{\mathrm{O}}, F_t^{\mathrm{G}}, F_t^{\mathrm{W}}) = \Phi(x, u)$. The production mapping Φ depends on the form and specifications

Symbol	Definition
V^{O}_t	Amount of oil in the reservoir (Sm^3) at time t
V^{G}_t	Amount of free gas in the reservoir (Sm^3) at time t
V^{W}_t	Amount of water in the reservoir (Sm^3) at time t
V^{P}_t	Total pore volume of the reservoir (m^3) at time t
P^{R}_t	Reservoir pressure (Bara) at time t

TABLE 4.1: Definition of the components of the state

of the production network. We present two examples of such Φ in the numerical applications of Section 4.4, with details in Appendix 4.A.

Symbol	Definition
F^{O}_t	Volume of oil produced (Sm^3) during $[t, t+1[$
F^{G}_t	Volume of gas produced (Sm^3) during $[t, t+1[$
F^{W}_t	Volume of water produced (Sm^3) during $[t, t+1[$

TABLE 4.2: Definition of the productions

We make the following assumptions on the reservoir (as formulated in Dake [1983]): first, the fluids contained in the reservoir follow a *black-oil* model; second, we consider that we have a tank-like reservoir. Thanks to those two standards assumptions, we can formulate the reservoir and the production system as a controlled dynamical system.

Proposition 4.2. *There exists a function $\Xi : \mathbb{X} \times \mathbb{U} \to \mathbb{R}$ such that the following function $f : \mathbb{X} \times \mathbb{U} \to \mathbb{R}^5$*

$$
f : (x, u) \mapsto \begin{pmatrix} x^{(1)} - \Phi^{(1)}(x, u) \\ x^{(2)} - \Phi^{(2)}(x, u) + \Big[x^{(1)} R_s(x^{(5)}) \\ \quad - \big(x^{(1)} - \Phi^{(1)}(x, u)\big) R_s\big(\Xi(x, u)\big) \Big] \\ x^{(3)} - \Phi^{(3)}(x, u) \\ x^{(4)} \left(1 + c_f(\Xi(x, u) - x^{(5)})\right) \\ \Xi(x, u) \end{pmatrix} \tag{4.3}
$$

is the dynamics of the reservoir in (4.1c) (with $x = (x^{(1)}, \ldots, x^{(5)})$, R_s a given function of the reservoir pressure called the solution gas *function, and c_f a given parameter called the* pore compressibility *of the reservoir).*

Proof. See Appendix 4.A. □

4.4 Two numerical applications

We now present two numerical applications that illustrate how the material balance formulation can be used. The numerical applications are done on simple reservoirs. In §4.4.1, the first application is a gas reservoir that can be modeled with two tanks and with a connection, of known transmissivity, linking them together. It illustrates how the formulation can be applied to complex cases with multiple tanks. In §4.4.2, the second application we consider is an oil reservoir where pressure is kept constant through water injection. This shows how we can take into account injection to go beyond the first recovery of oil and gas. All numerical applications were performed on a computer equipped with a Core i7-4700K and 16 GB of memory.

4.4.1 A gas reservoir with one well

In the first application, we consider a real gas reservoir, for which production data are available. The recorded data come from a field approaching abandonment. We only considered a sub-field of a much larger field, the sub-field being constituted of an isolated reservoir with one well.

Our goal here is to show how simple cases can be tackled with the material balance formulation, and that the formulation can also be applied to cases with multiple tanks. We first present a state reduction of this case. We then present a model with one tank, and then a model with two tanks, mimicking an evolutive construction of the reservoir model. Indeed, when optimizing a real petroleum production system, the models are improved as data are analyzed. Hence, reservoir models will get more complex to fit the gathered exploitation data, such as going from a one tank model to a two tanks model. We therefore present the models following such timetable, going from a cruder to a more refined reservoir model.

Characteristics of the case. The geology of this particular sub-field makes it perfect for a tank model, as proved by many years of perfectly matched production. Also, the simplicity of the fluids with a high methane purity makes the black-oil model a very realistic assumption. The reservoir can be modeled with either one or two tanks, while the well's perforations are modeled with a known stationary inflow performance relationship, noted IPR^{G}. The two tanks model is illustrated in Figure 4.2. We do not consider the rest of the network, so that we will not have to take into account any vertical lift performance (VLP) necessary to lift oil to the surface. This implies that the only control we consider is the bottom hole flowing pressure (BHFP), P_t, resulting in the problem known as *optimization at the bottom of the well*. We hence assume that there is no "pipe" necessary to move gas from the reservoir to the surface, thus assuming that the network is only constituted of the well-perforations which allow the production of gas. Indeed, optimizing with the bottom hole flowing pressure makes it easier to compare the different reservoir models, as we directly act on the reservoir. Adding the vertical lift performance only adds a layer of complexity to the comparison of the models, while the only benefit would be to get results closer to an actual field production. All in all, adding the vertical lift performance only adds more constraints on the mathematical formulation and may mask the impact of the reservoir model. As the focus of this paper is to present a formulation with a new reservoir model, we decided not to take into account the vertical lift performance. We also did not try to go beyond the two tanks model.

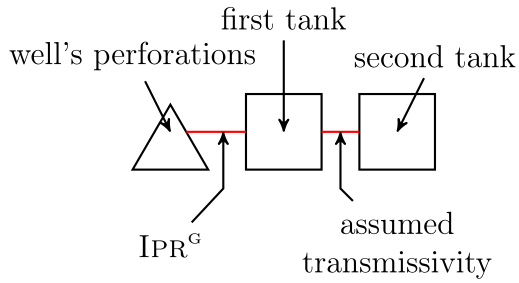


FIGURE 4.2: Representation of the two tanks model

Formulation and state reduction. In this first application, we consider a reservoir that contains only gas and water. We first assume that we only produce some gas, and that no fluids are re-injected in the reservoir. Moreover, we assume that there is no water production, and thus the amount of water remains stationary. Therefore, $V_t^{\mathrm{W}} = V_0^{\mathrm{W}}$ for all $t \in \mathcal{T}$, the initial amount of water V_0^{W} being known. We therefore only need to consider the evolution of the amount of gas, the pressure and the total pore volume as states variables. As shown in Appendix 4.B, we can further reduce the state, and we only need to consider the amount of gas in the reservoir as the reservoir state. Since we do an optimization at the bottom of the well, we only have one control to consider, the bottom-hole flowing pressure, noted P_t. We therefore have state $x_t = V_t^{\mathrm{G}}$ and control $u_t = P_t$.

The optimization problem we consider here is to maximize the revenue of the gas production. At each time t, we sell gas at price r_t, with a discount factor ρ. The general optimization problem (4.1) after state and control reduction when considering the gas reservoir and one tank is given by

$$\max_{(V_t^{\mathrm{G}}, P_t, P_t^{\mathrm{R}}, F_t^{\mathrm{G}})} \sum_{t=0}^{T-1} \rho^t r_t F_t^{\mathrm{G}} \tag{4.4a}$$

$$s.t.\quad V_0^{\mathrm{G}} = x_0 \ , \tag{4.4b}$$

$$P_t^{\mathrm{R}} = \Psi_{1\mathbf{T}}(V_t^{\mathrm{G}}) \ , \ \ \forall t \in \mathcal{T} \ , \tag{4.4c}$$

$$F_t^{\mathrm{G}} = \frac{\mathrm{IPR}^{\mathrm{G}}\left(P_t^{\mathrm{R}} - P_t\right)}{B_{\mathrm{G}}(P_t^{\mathrm{R}})} \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , \tag{4.4d}$$

$$V_{t+1}^{\mathrm{G}} = V_t^{\mathrm{G}} - F_t^{\mathrm{G}} \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , \tag{4.4e}$$

$$F_t^{\mathrm{G}} \geq 0 \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , \tag{4.4f}$$

$$V_t^{\mathrm{G}} \geq 0 \ , \ \ \forall t \in \mathcal{T} \ , \tag{4.4g}$$

$$P_t \geq 0 \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , \tag{4.4h}$$

as detailed in Appendix 4.B.

4.4.1.1 One tank gas reservoir model

Fitting model to real data. We use production data from a sector of a real gas field, to check that the reservoir model described with the Constraints (4.4c) and (4.4e) accurately follows real measurements on the gas field after fitting the model. More precisely, we apply a given real production schedule on a part of the field (only one well), and check that the pressure we simulate in the reservoir is close to the corresponding measured pressure. The historical production spans over 15 years, and we have monthly values, which is why we consider monthly time steps for Problem (4.4).

As can be seen in Figure 4.3, the one tank model fits the observation. However, there is a gap between the simulated and measured pressures whose relative value may exceed 10%. Since the simulated pressure tends to be higher on the first half of the production, we start by underestimating the decline of the production. Then, during the second half of the production, the simulated pressure is lower than the measured pressure, which means we overestimate the decline of the production. This elastic effect is most likely due to the simplification of removing the secondary tank in the model. Indeed, the secondary tank act as a buffer which reacts slowly, explaining the extra pressure at the beginning and then sustaining a better value of the pressure later on.

Optimization of the production on the one tank approximation. We use dynamic programming (see Algorithm 1) to get an optimal production policy. We consider that the revenue per volume of gas is the historical gas spot price of TTF (Netherlands gas market) from 2006 to 2020, and we do not consider any operational cost.

We now present the results of the one tank model. The results are illustrated in Figures 4.4 and 4.5, and summarized in Table 4.3. We notably remark in Figure 4.5 that

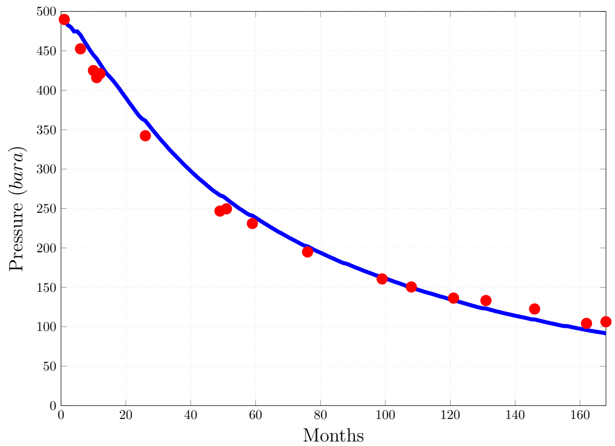


FIGURE 4.3: Comparison of the simulated one tank reservoir pressure to the historical measured pressure when applying the same (historical) production schedule. The blue curve is the simulated pressure in the tank, whereas the red dots are the measured pressures

the optimal production stops when prices are low as we fully take advantage of the perfect knowledge of the future prices.

There is a massive increase in the total gains when using the optimal policy, compared to the real production. We also produce far more over the optimization time period (2,850 MSm^3 instead of 2,250 MSm^3). However, those results are not truly comparable. We do not have access to the criteria used to choose the real production. Optimized and real productions cannot be compared as they do not share the same objective function. Moreover, since the considered case is a small part of a much larger production network, we cannot compare the results to the actual production policy used for fitting the model, which was made with the rest of the network in mind. Furthermore, our optimization is made at the bottom of the well (BHFP). We only take into account the inflow performance of the well, not the vertical lift necessary to bring the gas to the surface. The resulting rates are therefore not fully realistic, reaching values closer to a multi-well development. Finally, the historical production was made without knowing future prices, and could also have been made with other constraints to ensure a minimal production of the field, or having a positive cash-flow (constraints due to the field's exploitation contract). While not directly comparable, this gas reservoir application still illustrates one of the best-case scenarios of the dynamic programming approach, and it shows how much could be gained from using a multistage material balance formulation.

Since the dynamic programming algorithm uses a discretization of the state space $\mathbb{X}_d$ and the control space $\mathbb{U}_d$, we tried different uniform discretizations for the states and controls spaces to prevent any side effects due to the chosen discretization. We do not observe notable changes in the value function past a 10,000 points uniform discretization of the state space and a 20 points discretization of the control space, which are the values we used in this case study. Details on the effect of the discretization can be found in

Appendix 4.C.

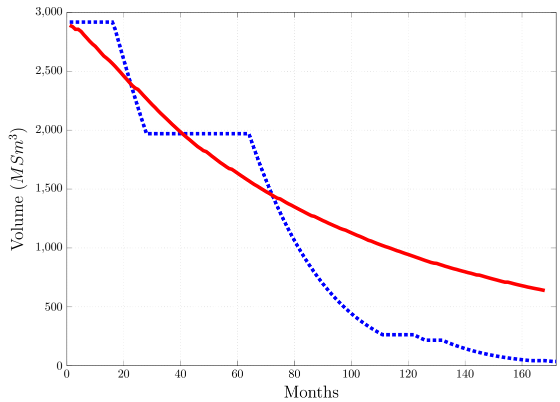


FIGURE 4.4: Evolution of the content of the reservoir in the one tank model. The doted blue curve is the optimal (anticipative) trajectory of the amount of gas, while the red curve is the trajectory with the historical production

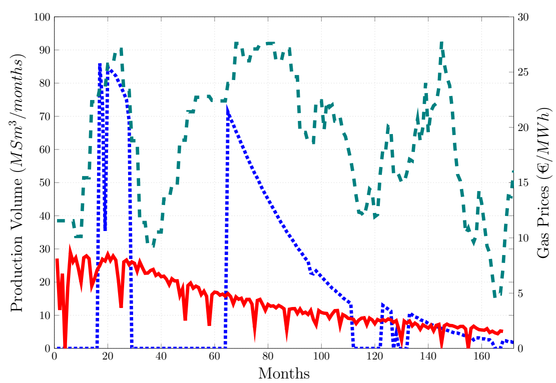


FIGURE 4.5: Trajectories of the production. The dotted blue curve is the optimal (anticipative) production in the one tank model, the red one is the historical production, whereas the dashed green curve is the average monthly gas price

Comparison to policy derived from decline curves. In this paragraph, we compare the material balance formulation to those using decline curves or oil-deliverability curves, such as in Iyer et al. [1998], Gupta and Grossmann [2012, 2014], Marmier et al. [2019]. The decline curves formulation and the way to numerically obtain decline curves are given in Appendix 4.D. The following proposition shows that the decline curves formulation is equivalent to the material balance formulation when considering a one-tank model.

Proposition 4.3. *The formulation using decline curves, written*

$$\max_{u} \quad \sum_{t=0}^{T} \rho^t \mathcal{L}_t(u_t) \tag{4.5a}$$

$$s.t. \quad F_t^{\text{o}} \leq g\left(\sum_{s=0}^{t-1} F_s^{\text{o}}\right) \ , \quad \forall t \in \mathcal{T} \setminus \{0\} \tag{4.5b}$$

$$u_t \in \mathcal{U}_t^{ad}\left(\sum_{s=0}^{t-1} F_s^{\text{o}}\right) \ , \quad \forall t \in \mathcal{T} \ , \tag{4.5c}$$

is equivalent to the material balance formulation when the state of the reservoir is one-dimensional.

Proof. See Appendix 4.D □

We obtain the decline curve g used in Inequality (4.5b) by first computing the maximal production value for the same discrete states as the ones used in the dynamic programming approach. Then, piecewise interpolation between the computed values is used to obtain the value of the decline curve everywhere. It is worth noting that, when using piecewise linear approximation for the decline curves, the maximization problem (4.5) turns out to be a MIP (Mixed Integer Problem) with linear constraints and with more than 170,000 binary variables. We solve that MIP by using the commercial solver Gurobi 9.1. The results are given in Table 4.3. Since the material balance formulation (4.4) uses a one-dimensional state, we obtain similar results between the material balance formulation and the formulation using a decline curve in accordance with Proposition 4.3. The two approaches thus yield similar production policies. Note however that the dynamic programming approach has a lower computation time than a naive implementation of the decline curve formulation. One could decrease the precision on the decline curve formulation, by using fewer points to describe the decline curve. This would improve its computation time. As this is not the focus of this paper, we did not do such refinement of the numerical experiments for the decline curve formulation.

	CPU time (s)	Value (M€)
Material Balance	653	743
Decline Curves	3,882	743

TABLE 4.3: Comparison with regards to CPU time and value between the material balance and decline curve formulation for one tank

4.4.1.2 Two tanks gas reservoir model

Fitting data. We check if the fitted two tanks reservoir model accurately follows real measurement on the gas field. We use the same data as in the one tank case. The two tanks model more accurately fits the observations, as is depicted in Figure 4.6 (we have a gap of less than 5% for each measured point). Since the two tanks model is closer to the observations, we consider that it is the reference of "truth" when comparing results of the one tank approximation and the two tanks model.

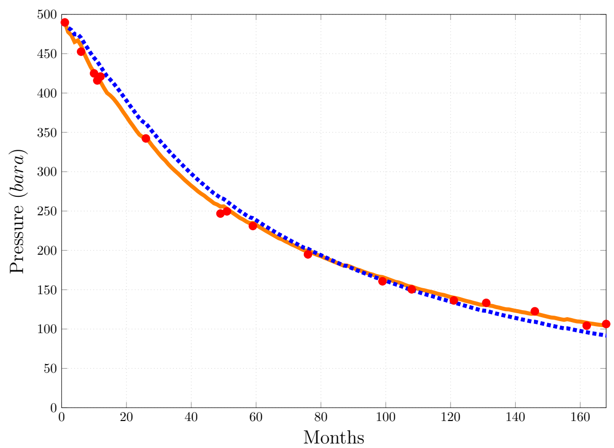


FIGURE 4.6: Comparison of the simulated two tanks reservoir pressure to the measured pressure when applying the same production schedule. The blue dotted curve recalls the pressure obtained using the one tank model. The orange continuous curve is the pressure in the first tank obtained using the two tanks model. The red dots are the measured pressure at the bottom of the well

Optimal production with two tanks. We now present the results of the two tanks model. The only changes compared to the one tank model are on the states and on the dynamics of the reservoir. We use the same prices, and, again, we only do an optimization at the bottom of the well (BHFP). Details on the obtained optimal controls and states trajectory are given in Figure 4.7 and Figure 4.8. Once again, we observe that production stops when prices are low, benefiting fully from anticipating the future prices. We also note that more "pauses" are present in the productions when compared to the one tank model (four instead of three). The "pauses" allow the second tank to replenish the first one (see Figure 4.7). Indeed, production resumes at months 50 to 60, before stopping again for five months. We can then observe that the amount of gas in the first tank is replenished, before we resume production at month 65, at the same date as in the one tank model. We end up producing some more gas than with the one tank model (2,900 Sm^3 instead of 2,850 Sm^3).

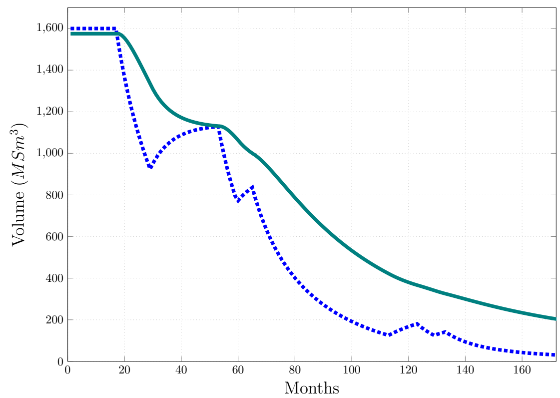


FIGURE 4.7: Evolution of the content of the reservoirs when applying the optimal (anticipative) policy in the two tanks model. The dotted blue curve shows the content of the first tank (linked to the well) while the green curve shows the content of the second tank

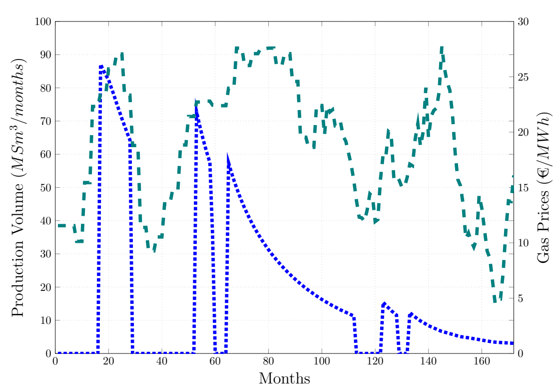


FIGURE 4.8: Trajectory of the optimal production in the two tanks model. The dotted blue curve is the optimal (anticipative) production, whereas the dashed green curve is the monthly gas price

We tried different discretizations for the state space. Notably, using more than 400 possible states per tank and 10 possible controls did not yield any significant improvement in the computed value function. Details on the impact of the discretization are given in Appendix 4.C.

Numerical experiments also reveal that the initial value function $\mathcal{J}_0$ is almost an affine function of the sum of the states. This seems to imply that the one tank and two tanks model should yield similar results. Such a statement does not hold true, as confirmed by the numerical experiments described in the next paragraph.

Comparing the one tank formulation to the two tanks formulation. To compare the results between the two tanks and one tank formulations, we consider that the two tanks material balance model is the reference. A given sequence of controls $(u_t)_{t \in \mathcal{T} \setminus \{T\}}$ admissible for the one tank model is not necessarily admissible for the two tanks model. Indeed, the admissible control set is given by $\mathcal{U}^{ad}(x_t) = [0, \Psi_{\mathbf{1T}}(x_t)]$ for the one tank model and by $[0, \Psi^{(1)}_{\mathbf{2T}}(x^{(1)}_t)]$ for the two tanks model (see Appendix 4.B.1).

Thus, given a sequence of controls $(u_t)_{t \in \mathcal{T} \setminus \{T\}}$ admissible for the one tank model, we produce an admissible sequence of controls for the two tanks model with the use of a *projection* $\Pi_{\mathbf{1T} \to \mathbf{2T}} : \mathbb{U}^T \times \mathbb{X} \to \mathbb{U}^T$ given as follows. The sequence $(\widetilde{u}_t)_{t \in \mathcal{T} \setminus \{T\}} = \Pi_{\mathbf{1T} \to \mathbf{2T}}\big((u_t)_{t \in \mathcal{T} \setminus \{T\}}, x_0\big)$ is computed recursively for all $t \in \mathcal{T} \setminus \{T\}$ by $\widetilde{u}_t = \min\big\{u_t, \Psi^{(1)}_{\mathbf{2T}}(\widetilde{x}^{(1)}_t)\big\}$, where $\widetilde{x}_t$ is defined at time 0 by $\widetilde{x}_0 = x_0$, and for all $t > 0$ by $\widetilde{x}_{t+1} = f_{\mathbf{2T}}(\widetilde{x}_t, \widetilde{u}_t)$. We can get a sequence of admissible controls for the two tanks model by applying the projection $\Pi_{\mathbf{1T} \to \mathbf{2T}}$ on a sequence of admissible controls for the one tank model.

To compare the one tank and two tank models, we project the optimal sequence of controls returned by the dynamic programming algorithm on the one tank formulation thanks to the projection $\Pi_{\mathbf{1T} \to \mathbf{2T}}$. As can be seen in Figure 4.9, the projected sequence of controls differs from the non-projected sequence: the dotted curve, which represents the projected sequence, is below the dashed curve, which represents the optimal sequence for the one tank model.

As depicted in Figures 4.9 and 4.10, the production planning given by the one tank optimization problem differs from the production planning given by the two tanks optimization problem. Moreover, the production planning of the one tank model gives lower gains than anticipated, and is worse than the optimal two tanks model planning. The one tank optimization is thus optimistic on the optimal value of the problem when applied with the reference model. Moreover, there is a 5% difference in value between the one tank and two tanks models (a value of 703 M€ for the translated one tank production planning against 736 M€ for the two tanks production planning). This discrepancy illustrates how having a more accurate model of the reservoir can have a substantial impact on the optimal planning, all other things being equal. It also shows that, contrarily to the assumption presented at the end of the previous paragraph (that the two models could yield similar results if the value function only depended on the sum of the states), the optimal value and control cannot be found with a one tank approximation, and the optimal controls and value functions are not functions of the sum of the states.

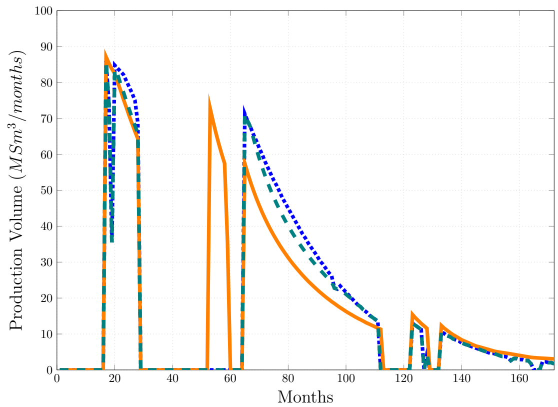


FIGURE 4.9: Comparison of the trajectory of the production with the two tanks model as reference. The dotted blue curve is the production planning in the one tank model, the orange curve is for the two tanks model. The dashed green curve is the production planning of the one tank model projected in the two tanks model

Comparison to decline curves with two tanks. We have numerically compared the decline curve and the material balance formulations in a context where they are known to be equivalent, that is, the one tank formulation. We now produce numerical experiments in a context where the equivalence is not assured: two tanks connected with a known transmissibility. We have generated decline curves for the two tanks formulation by following the procedure described in Appendix 4.D. The results returned by the decline curve formulation provide an admissible production in the two tanks model, as it is constrained by an admissible production schedule. We can therefore directly compare the results obtained by the decline curves approach and the two tanks model. The results of the optimization of the two formulations are compiled in Table 4.4. We end up having close results, with a difference in optimal values of 0.7%, but with a large difference in computing times. However, it appeared that such close results were due to the selected price scenario. Using different prices by randomizing the order in which the different prices appear, the gap between the two approaches widens from 0.5% up to 4%. This implies that the initial price considered was an almost best-case scenario for the decline curves approach. It also shows that the decline curves approach is far less robust to changes in the price data, and that it cannot benefit as efficiently as the material balance formulation of some effects of the two tanks dynamical system, such as waiting for the second tank to empty itself into the first one.

Overall, this application suggests that the material balance approach can work on complex cases, and that dynamic programming is well suited to optimize an oil field. Moreover, there can be differences with results from the decline curves approach, which are likely to grow larger with the complexity of the system.

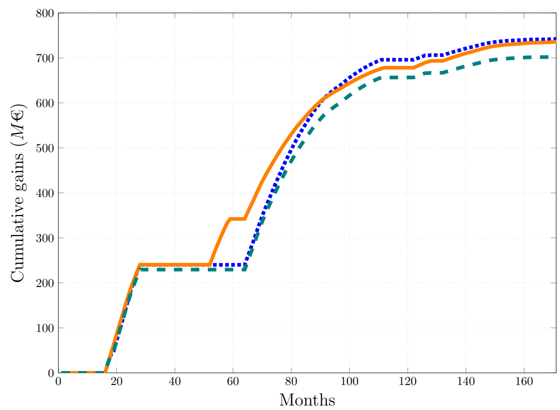


FIGURE 4.10: Cumulated gains with the two tanks model as reference. The dotted blue curve is the cumulated gains of the one tank planning in the one tank model, the orange curve is the cumulated gains of the two tanks planning in the two tanks model, and the dashed green curve is the cumulated of the one tank planning projected in the two tanks model

	CPU time (s)	Value (M€)
Material Balance	706	736
Decline Curves	7,825	731

TABLE 4.4: Comparison with regards to CPU time and value between the material balance and decline curve formulation for two tanks with the initial prices sequence.

4.4.2 An oil reservoir with water injection

The second application is an oil reservoir with water injection. The goal is to demonstrate how the formulation can be used beyond primary recovery cases, on a numerically simple case. We consider that we have one reservoir which contains both oil and water, produced under pressure maintenance by water injection. Moreover, we consider that the initial pressure is above the bubble-point, which eliminates the possibility of having free gas in the reservoir. This allows us to have once again a one-dimensional state: either the water (which we used for the numerical applications), or the oil in the reservoir. We have $x_t = V_t^{\text{W}}$ and $u = P_t$. Here, we want to maximize the revenue of the oil production. The optimization problem (4.1) now becomes

$$
\begin{aligned}
\max_{(V_t^{\text{W}}, P_t, w_t^{\text{CT}})} \quad & \sum_{t=0}^{T-1} \Bigg(\rho^t r_t \alpha \frac{P^{\text{R}} - P_t}{B_{\text{O}}(P^{\text{R}})} (1 - w_t^{\text{CT}}) \\
& \qquad - \rho^t c_t \alpha \frac{P^{\text{R}} - P_t}{B_{\text{W}}(P^{\text{R}})} \Bigg) && \text{(4.6a)} \\
s.t. \quad & w_t^{\text{CT}} = \text{W}^{\text{CT}} \left(\frac{V_t^{\text{W}} B_{\text{W}}(P^{\text{R}})}{V^{\text{P}}} \right) , \ \ \forall t \in \mathcal{T} , && \text{(4.6b)} \\
& V_{t+1}^{\text{W}} = V_t^{\text{W}} - \alpha \frac{P^{\text{R}} - P_t}{B_{\text{W}}(P^{\text{R}})} (w_t^{\text{CT}} - 1) , \ \ \forall t \in \mathcal{T} , && \text{(4.6c)} \\
& F_{min}^{\text{W}} \leq \alpha \frac{P^{\text{R}} - P_t}{B_{\text{W}}(P^{\text{R}})} (w_t^{\text{CT}} - 1) \leq F_{max}^{\text{W}} , \ \ \forall t \in \mathcal{T} , && \text{(4.6d)} \\
& F_{min}^{\text{O}} \leq \alpha \frac{P^{\text{R}} - P_t}{B_{\text{O}}(P^{\text{R}})} (1 - w_t^{\text{CT}}) \leq F_{max}^{\text{O}} , \ \ \forall t \in \mathcal{T} , && \text{(4.6e)} \\
& P_t \geq 0 , \ \ \forall t \in \mathcal{T} . && \text{(4.6f)}
\end{aligned}
$$

The objective function (Equation (4.6a)) is divided in two components. At time t, we consider a discount factor ρ and the price r_t of the oil, whereas injecting water costs c_t per cubic meter. The revenue is hence

$$
\sum_{t=0}^{T-1} \rho^t \left(r_t F_t^{\text{O}} - c_t F_t^{\text{WI}} \right) .
$$

Replacing the produced oil F_t^{O} and the injected water F_t^{WI} by the relevant functions of the controls (see Equations (4.30) and (4.33)) leads to the objective function (4.6a)).

We assume that the water-cut function W^{CT} (the amount of water produced when extracting one cubic meter of liquid at standard conditions) is given by a piecewise linear function. The water-cut depends on the water saturation S^{W} (proportion of water in the reservoir pore volume). Since the reservoir pressure is kept constant, the total pore volume is constant and the water saturation expression is thus $S_t^{\text{W}} = \frac{V_t^{\text{W}} B_{\text{W}}(P^{\text{R}})}{V^{\text{P}}}$. This gives us constraint (4.6b).

Since we want to keep a constant pressure in the reservoir, we need to re-inject enough water to replace the extracted oil. Replacing the oil with water gives a new dynamic for V_t^{W} as in Equation (4.6c). Constraints (4.6d) and (4.6e) details the oil and water produced depending on the control P_t with their respective bounds. The details of the formulation are given in Appendix 4.B.

We do a monthly optimization, with the historical Brent prices for years 2000–2020 as the prices in the objective function (4.6a), and a water injection cost of 1 €/m^3. Details on the resulting trajectory of the content of the reservoir can be found in Figure 4.11, whereas details on the production can be found in Figure 4.12. As previously discussed in §4.4.1, the optimal policy yields more production when prices are high, and stops producing when they are low. The production goes from one bound to the other (zero production, with $P_t = P^{\mathrm{R}}$, and full production, with $P_t = 0$).

The production also does not fully deplete the reservoir, which means that it is not advantageous to completely deplete the reservoir if one wants to maximize the profit over the optimization time frame (there is still 18.2 MSm^3 of oil in the reservoir at time T, as can be seen in Figure 4.11). Indeed, production slowly diminishes with the volume of oil V_t^{O} in the reservoir, as can be seen in Figure 4.12. It is more advantageous to wait for high prices instead of producing, as it would reduce the possible future production. This leads to halting production with some reserves still in the reservoir, as we prefer to wait for a higher price instead of producing when prices are low. As a side effect, numerical experiments reveal that the initial value function $\mathcal{J}_0$ is almost linear with regards to the state x_0. However, we only considered simple constraints on the production. As more constraints will be added to the problem, other behaviors will certainly appear. CPU time was 1,575 s for a 100,000 discretization of the state variable, with a value of 3,376 M€. Impact of the discretization can be found in Appendix 4.C.

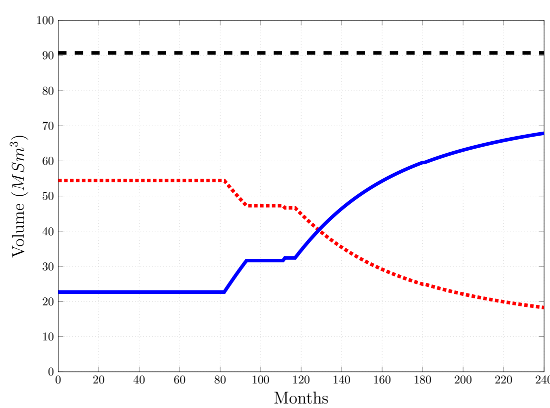


FIGURE 4.11: Evolution of the content of the reservoir when applying the optimal policy in the oil reservoir model. The blue curve shows the volume of water in the reservoir, whereas the dotted red curve is the volume of oil the reservoir. The dashed black curve represents the total pore volume

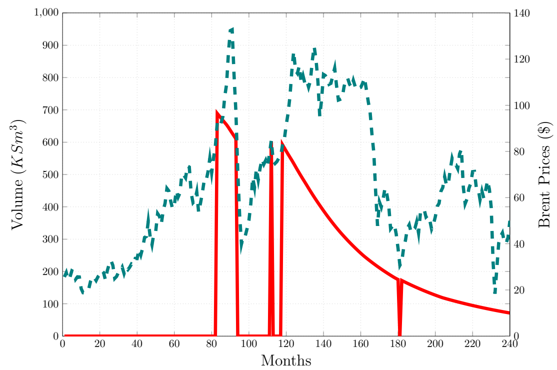


FIGURE 4.12: Trajectory of the optimal production in the oil reservoir model. The red curve is the optimal production, whereas the dashed green curve is the monthly oil price

Overall, this application shows how we can apply the material balance approach beyond first recovery of oil and gas, and that it can be used on different kinds of reservoirs.

4.5 Conclusion

In this paper, we have presented a mathematical formulation for the optimal management over time of an oil production network as a multistage optimization problem. In this formulation, the reservoir is modeled as a controlled (non-linear) dynamical system derived from material balance equations and the black-oil model. The state of the derived dynamic system is of dimension five, which is quite large for numerical resolution via dynamic programming algorithm. However, we were able to use Dynamic Programming to numerically solve the management optimization problem for specific cases of interest with either oil or gas, both presenting a reduced dimensionality of the state. We have also shown that our mathematical formulation is an improvement over decline curves formulation. First, as predicted by the theory, we replicated results from decline curve formulations when considering the first recovery of a one tank system (as seen in §4.4.1.1). Second, in more complex cases with inter-connected tanks, as described in §4.4.1.2, we have shown that we can surpass the NPV returned by the decline curve formulation. Third, we have gone beyond the first recovery of hydrocarbons, as we have shown in §4.4.2, where we took into account water injection.

Finally, it is to be noted that the dynamic programming algorithm can be used in a stochastic framework. As an example, we could add uncertainties to the oil and gas prices, instead of assuming that they are known in advance and thus deterministic. Moreover, an even more realistic formulation with *partial observation* of the content of the reservoir could also be explored. Indeed, in oil production systems, the initial state of the reservoir is not known. Such a formulation is amenable to dynamic programming, as will be explored in future works.

Acknowledgements

We would like to thank TEPNL in general and Erik Hornstra in particular for providing data used in this paper.

4.A Detailed construction of the reservoir as a dynamical system

In this section, we detail the construction of the reservoir as a dynamical system. This serves as the proof of Proposition 4.2.

4.A.1 Constitutive equations assuming the black-oil model for the fluids

The black-oil model relies on the assumption that there are at most three *fluids* in the reservoir: oil, gas and water. Moreover, the fluids can be present in the reservoir in up to two phases: a liquid phase, and a gaseous phase. A black-oil representation of a reservoir can be seen in Figure 4.13. The three fluids, oil, gas and water, can be present in the liquid phase and the gas in the liquid phase is denoted as *dissolved gas*. By contrast, it is assumed that in the gaseous phase, only gas, denoted as *free gas*, can be present.

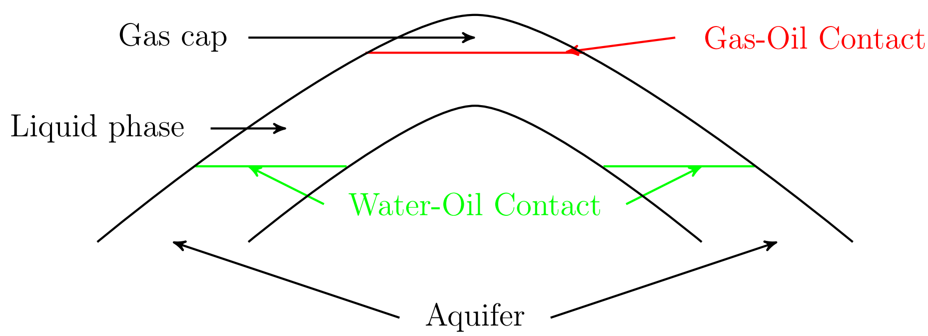


FIGURE 4.13: Black-oil Representation of a reservoir

Therefore, in the black-oil model, we consider the following four components

- V^{O}, the standard volume of oil in the liquid phase,
- V^{G}, the standard volume of free gas in the gaseous phase,
- V^{DG}, the standard volume of dissolved gas in the liquid phase,
- V^{W}, the standard volume of water in the liquid phase,

where *standard volume* is the volume taken by a fluid at standard pressure and temperature condition (1.01325 Bar and 15°C), also known as stock tank conditions. The units of standard volumes are preceded by a capital S, as in Sm^3 for standard cubic meter.

There are functions in the black-oil model to convert standard volumes into in situ volumes in the reservoir under a given pressure and temperature. The set of functions describing the pressure, volume and temperature behavior of the fluids, under the black-oil assumption, is call the PVT (Pressure-Volume-Temperature) model. We consider here a simplified black-oil model, assuming that the temperature in the reservoir is stationary and uniform, which is a common assumption for a geological formation such as a reservoir. There are four PVT functions, one per component, which are given in Table 4.5. The PVT functions only depend on the reservoir pressure under the stationary and uniform temperature assumption. As an example, given the oil standard volume, V^{O}, and the reservoir pressure, P^{R}, the oil volume in the reservoir is given by $V^{\text{O}} \times B_{\text{O}}(P^{\text{R}})$.

Notations	Description
B_{O}	Oil formation volume factor. It is the volume in barrels occupied in the reservoir, at the prevailing pressure and temperature, by one stock tank barrel of oil plus its dissolved gas. (unit: rb/stb)
B_{G}	Gas formation volume factor. It is the volume in barrels that one standard cubic foot of gas will occupy as free gas in the reservoir at the prevailing reservoir pressure and temperature. (unit: rb/scf)
B_{W}	Water formation factor. It is the volume occupied in the reservoir by one stock tank barrel of water. (unit: rb/stb)
R_s	Solution (or dissolved) gas. It is the number of standard cubic feet of gas which will dissolve in one stock tank barrel of oil when both are taken down to the reservoir at the prevailing reservoir pressure and temperature. (unit: scf/stb)

Table 4.5: Definition of the PVT functions

One key characteristic of the black-oil model that we use is due to [Danesh, 1998, chap 2], which states that the sum of the physical volumes in the reservoir associated with the three components V^{O}, V^{G}, V^{W} is a decreasing function

$$P^{\text{R}} \mapsto V^{\text{O}} \times B_{\text{O}}(P^{\text{R}}) + V^{\text{G}} \times B_{\text{G}}(P^{\text{R}}) + V^{\text{W}} \times B_{\text{W}}(P^{\text{R}}) \; , \tag{4.7}$$

of the reservoir pressure.

The last characteristic of the black-oil model concerns the dissolved gas in the oil V^{DG}. It is assumed in Dake [1983] that the standard volume of the dissolved gas V^{DG} is a function of both the standard volume of oil, V^{O}, and the reservoir pressure, P^{R}, as follows

$$V^{\text{DG}} = \delta(V^{\text{O}}, P^{\text{R}}) = V^{\text{O}} \times R_s(P^{\text{R}}) \; . \tag{4.8}$$

4.A.2 Conservation law in the reservoir

We assume that the reservoir structural integrity is guaranteed, so there is no leakage of any fluids at any time. We can therefore write mass conservation equations, which are also named *material balance* equations in the oil literature, for each of the four components introduced in §4.A.1. In order to write the material balance equations of the reservoir, we need to consider the production volumes, F^{O}, F^{G} and F^{W} which are the standard volumes of oil, free gas and water extracted from the reservoir.

Using material balance for the standard volume of oil in the liquid phase, we get

$$V^{\mathrm{O}}_{t+1} = V^{\mathrm{O}}_t - F^{\mathrm{O}}_t \quad \forall t \in \mathcal{T} \setminus \{T\} \ , \tag{4.9}$$

and, for the standard volume of water, we get

$$V^{\mathrm{W}}_{t+1} = V^{\mathrm{W}}_t - F^{\mathrm{W}}_t \quad \forall t \in \mathcal{T} \setminus \{T\} \ . \tag{4.10}$$

The material balance for gas requires some more developments as it mixes the standard volume of free gas and the standard volume of dissolved gas. As given in §4.A.1, at any time, $t \in \mathcal{T}$, the standard volume of dissolved gas in the liquid phase V^{DG}_t is given by Equation (4.8). Therefore, between time t and time $t+1$, the standard volume of dissolved gas evolves from $V^{\mathrm{DG}}_t = \delta(V^{\mathrm{O}}_t, P^{\mathrm{R}}_t)$ to $V^{\mathrm{DG}}_{t+1} = \delta(V^{\mathrm{O}}_{t+1}, P^{\mathrm{R}}_{t+1})$. Hence, the quantity $(V^{\mathrm{DG}}_t - V^{\mathrm{DG}}_{t+1})$ of *liberated* gas must be added to the free gas material balance equation. Thus, for all $t \in \mathcal{T} \setminus \{T\}$, we obtain the following mass conservation equation for the standard volume of free gas

$$\begin{aligned}
V^{\mathrm{G}}_{t+1} &= V^{\mathrm{G}}_t - F^{\mathrm{G}}_t + (V^{\mathrm{DG}}_t - V^{\mathrm{DG}}_{t+1}) \\
&= V^{\mathrm{G}}_t - F^{\mathrm{G}}_t + \big(V^{\mathrm{O}}_t \times R_s(P^{\mathrm{R}}_t) - V^{\mathrm{O}}_{t+1} \times R_s(P^{\mathrm{R}}_{t+1})\big) && \text{(by (4.8))} \\
&= V^{\mathrm{G}}_t - F^{\mathrm{G}}_t + \Big(V^{\mathrm{O}}_t \times R_s(P^{\mathrm{R}}_t) - \big(V^{\mathrm{O}}_t - F^{\mathrm{O}}_t\big) \times R_s(P^{\mathrm{R}}_{t+1})\Big) && \text{(by (4.9))} \\
&= V^{\mathrm{G}}_t - F^{\mathrm{G}}_t + \Big(V^{\mathrm{O}}_t \times \big(R_s(P^{\mathrm{R}}_t) - R_s(P^{\mathrm{R}}_{t+1})\big) \\
&\qquad\qquad + F^{\mathrm{O}}_t \cdot R_s(P^{\mathrm{R}}_{t+1})\Big) \ . && (4.11)
\end{aligned}$$

The last conservation equation is given by a physical volume constraint coming from the fact that all four components of the reservoir are kept in the pores of the reservoir rocks. We note V^{P} the total pore volume of the reservoir. Following Dake [1983] and assuming that the pore compressibility c_f is constant, the total pore volume is a function of the pressure in the reservoir given by

$$V^{\mathrm{P}}_t = V_0 \exp(c_f P^{\mathrm{R}}_t) \ , \ \ \forall t \in \mathcal{T} \ , \tag{4.12}$$

with V_0 the asymptotic reservoir volume when pressure tends to 0.

A linearized version of Equation (4.12) proposed in Dake [1983] is

$$\frac{V^{\mathrm{P}}_{t+1} - V^{\mathrm{P}}_{t}}{V^{\mathrm{P}}_{t}} = c_f \left(P^{\mathrm{R}}_{t+1} - P^{\mathrm{R}}_{t} \right) \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , \tag{4.13}$$

and is used to derive the state dynamics of the reservoir.

Now, we consider the *saturations* of the fluids which are the proportions of the available pore volume taken by each of the three fluids in the reservoir. Denoting by S^{O}, S^{G} and S^{W} the saturations of respectively the oil, free gas and water components, we obtain that the sum of the three saturations must be equal to one over time

$$S^{\mathrm{O}}_t + S^{\mathrm{G}}_t + S^{\mathrm{W}}_t = 1 \ , \ \ \forall t \in \mathcal{T} \ . \tag{4.14}$$

Since, for all $t \in \mathcal{T}$ and $i \in \{\mathrm{O}, \mathrm{G}, \mathrm{W}\}$, we have that

$$S^{i}_t = \frac{V^{i}_t \times B_i(P^{\mathrm{R}}_t)}{V^{\mathrm{P}}_t} \ ,$$

Equation (4.14) gives

$$V^{\mathrm{O}}_t \times B_{\mathrm{O}}(P^{\mathrm{R}}_t) + V^{\mathrm{G}}_t \times B_{\mathrm{G}}(P^{\mathrm{R}}_t) + V^{\mathrm{W}}_t \times B_{\mathrm{W}}(P^{\mathrm{R}}_t) = V^{\mathrm{P}}_t \ , \forall t \in \mathcal{T} \ . \tag{4.15}$$

4.A.3 Construction of a production function

The time evolution of the reservoir is driven by the three production volumes, F^{O}, F^{G} and F^{W} which are the standard volumes of oil, free gas and water extracted from the reservoir.

Thus, the three production volumes may appear as possible controls on the reservoir. However, when adding a production network to the reservoir model, the controls to be considered are no longer production volumes, but decisions made upon the production network, such as opening or closing a pipe, choosing the well-head or bottom hole pressure, etc.

In the general case, we can assume that the physical model of the production network leads to a production function $\Phi : \mathbb{X} \times \mathbb{U} \to \mathbb{R}^3$, which relates the production volumes to the variables of the reservoir $x = (V^{\mathrm{O}}, V^{\mathrm{G}}, V^{\mathrm{W}}, V^{\mathrm{P}}, P^{\mathrm{R}})$ (we will show that x is a possible state of the reservoir) and to the network controls u, giving

$$(F^{\mathrm{O}}, F^{\mathrm{G}}, F^{\mathrm{W}}) = \Phi(x, u) \ . \tag{4.16}$$

When considering only one well, a common assumption is that the production volumes are given by the Inflow Performance Relationship IPR, which is a function of the reservoir pressure P^{R}, the bottom-hole pressure P, the saturation of water S^{W} and the saturation of gas S^{G}. More precisely, we obtain, for a one well model, that

$$F^i = \Phi^i(x,u) = \frac{\mathrm{IPR}^i(u, x^{(5)}, \frac{x^{(3)}B_{\mathrm{W}}(x^{(5)})}{x^{(4)}}, \frac{x^{(2)}B_{\mathrm{G}}(x^{(5)})}{x^{(4)}})}{B_i(x^{(5)})} \ , \ \ \forall i \in \{\mathrm{O}, \mathrm{G}, \mathrm{W}\} \ . \tag{4.17}$$

In the general case, we then need to take into account pressure drop due to the flow in the well itself through the use of a Vertical Lift Performance relationship.

In the two cases presented in Section 4.4, we can further detail the general production function Φ as follows

- For the gas reservoir as exposed in §4.4.1, we assume that the well only produces gas, and we hence obtain the following simplified formulation

$$F_t^{\mathrm{G}} = \Phi^{\mathrm{G}}(x,u) = \frac{\mathrm{IPR}^{\mathrm{G}}(P_t^{\mathrm{R}} - P_t)}{B_{\mathrm{G}}(P_t^{\mathrm{R}})} \ . \tag{4.18}$$

  Indeed, when we only produce gas, there is no need to consider the different saturations. Those saturations are necessary to find the proportion of oil, water and gas produced when applying a difference of pressure $P^{\mathrm{R}} - P$. Having only gas implies that the saturations have no impact on the production.

- When considering that the reservoir does not contain any free gas (i.e. $V^{\mathrm{G}} = 0$ and $S^{\mathrm{G}} = 0$), we obtain the following simplification for the production of oil and water. We assume that the *total production* F_t follows a simplified Darcy's law

$$F_t = \alpha(P_t^{\mathrm{R}} - P_t) \ , \ \ \forall t \in \mathcal{T} \ , \tag{4.19}$$

  where F_t is given by

$$F_t = F_t^{\mathrm{O}} \times B_{\mathrm{O}}(P_t^{\mathrm{R}}) + \underbrace{F_t^{\mathrm{G}} \times B_{\mathrm{G}}(P_t^{\mathrm{R}})}_{=0} + F_t^{\mathrm{W}} \times B_{\mathrm{W}}(P_t^{\mathrm{R}}) \ , \tag{4.20}$$

  with α the productivity index of the well, P_t the bottom-hole pressure of the well and F_t the total production which consists of a mix of oil and water as we have assumed that we have no free gas.

  For the oil reservoir with water injection case presented in §4.4.2, the last assumption we make is that the amount of produced water is given by

$$F_t^{\mathrm{W}} \times B_{\mathrm{W}}(P_t^{\mathrm{R}}) = \alpha(P_t^{\mathrm{R}} - P_t)\mathrm{W}^{\mathrm{CT}}(S_t^{\mathrm{W}}) \ , \tag{4.21}$$

  where W^{CT} is the water-cut function and, as already seen, where the water saturation is

$$S_t^{\mathrm{W}} = \frac{V_t^{\mathrm{W}} B_{\mathrm{W}}(P_t^{\mathrm{R}})}{V_t^{\mathrm{P}}} \ .$$

As we do not use more complex networks, we will not look any deeper into the network controls and their relationship with the general production Φ since those are beyond the scope of this paper.

4.A.4 Reservoir dynamics

We can now write the reservoir time evolution as a controlled dynamical system. The state of the controlled dynamical system is $x = (V^{\text{O}}, V^{\text{G}}, V^{\text{W}}, V^{\text{P}}, P^{\text{R}})$. We also express the production volumes thanks to the general production function, Φ, defined in Equation (4.16).

Now, we show that using Equations (4.9), (4.10), (4.11), (4.13), (4.15) and (4.16) we can build a mapping f such that $x_{t+1} = f(x_t, u_t)$ for all $t \in \mathcal{T}$. We proceed as follows: we consider the conservation Equation (4.15) at time $t+1$, and use Equations (4.9), (4.10), (4.11) and (4.13) to obtain the equation

$$\begin{aligned}(V_t^{\text{O}} - F_t^{\text{O}}) \times B_{\text{O}}(P_{t+1}^{\text{R}}) + (V_t^{\text{W}} - F_t^{\text{W}}) \times B_{\text{W}}(P_{t+1}^{\text{R}}) \\ + \Big[V_t^{\text{G}} - F_t^{\text{G}} + V_t^{\text{O}} \times \big(R_s(P_t^{\text{R}}) - R_s(P_{t+1}^{\text{R}})\big) \\ + F_t^{\text{O}} \times R_s(P_{t+1}^{\text{R}})\Big] \times B_{\text{G}}(P_{t+1}^{\text{R}}) \\ = V_t^{\text{P}} \left(1 + c_f(P_{t+1}^{\text{R}} - P_t^{\text{R}})\right) \ ,\end{aligned} \tag{4.22}$$

which depends on the state and production volumes at time t and of the pressure of the reservoir at time $t+1$. As recalled in §4.A.1, it is established in [Danesh, 1998, chap 2] that the left-hand side of Equation (4.22) is a decreasing function of the reservoir pressure P_{t+1}^{R}. More precisely, the expansion of the oil when gas dissolves into it due to an increase in pressure ΔP is less than the aggregated volume decrease of the free gas and the other fluids due to that same ΔP. To the contrary, the right-hand side of Equation (4.22) is increasing with the reservoir pressure. Hence, Equation (4.22) gives a function $\Xi : \mathbb{X} \times \mathbb{U} \to \mathbb{R}$ such that $\forall t \in \mathcal{T}, P_{t+1}^{\text{R}} = \Xi(x_t, u_t)$.

Moreover, note that when the PVT functions (B_{O}, B_{G}, B_{W} and R_s) are piecewise linear functions, the function Ξ can be computed efficiently. We only need to look at the breaking points of the piecewise linear functions to know on which segment we can invert Equation (4.22), thus getting the reservoir pressure P^{R}.

Combining Equations (4.9), (4.10), (4.11), (4.13) and using function Ξ, we finally obtain the expression of function f given in Equation (4.3).

4.B Material on state reduction

In this section, we detail how the general dynamics can be simplified in specific cases.

4.B.1 Gas reservoir state reduction

We consider a gas reservoir with no gas injection and where there is no water production or extraction, as used in §4.4.1, and we prove that the time evolution of the gas reservoir can be described by a reduced state composed of the standard volume of gas $x_t = V_t^{\mathrm{G}}$.

By assumption, the reservoir contains only gas and a constant volume of water. Thus, the standard volume of water satisfies $V_t^{\mathrm{W}} = V_0^{\mathrm{W}}$ for all $t \in \mathcal{T}$ and the standard volume of oil satisfies $V_t^{\mathrm{O}} = 0$ for all $t \in \mathcal{T}$. Hence, the state dimension can be reduced from dimension 5 to dimension 3.

Now, we show that the state dimension can be reduced to 1. First, we use Equation (4.12) in place of the linearized version (4.13) to obtain that $V_t^{\mathrm{P}} = V_0 \exp(c_f P_t^{\mathrm{R}})$ for all $t \in \mathcal{T}$. Second, we consider Equation (4.15) at time t together with $V_t^{\mathrm{O}} = 0$ and $V_t^{\mathrm{W}} = V_0^{\mathrm{W}}$ and $V_t^{\mathrm{G}} = V_t^{\mathrm{G}} - F_t^{\mathrm{G}}$ to obtain

$$V_t^{\mathrm{G}} \times B_{\mathrm{G}}(P_t^{\mathrm{R}}) + V_0^{\mathrm{W}} \times B_{\mathrm{W}}(P_t^{\mathrm{R}}) = V_0 \exp(c_f P_t^{\mathrm{R}}) \ , \ \ \forall t \in \mathcal{T} \ . \tag{4.23}$$

The left-hand side of Equation (4.23) is a decreasing continuous function of the pressure (the volume of gas and the production being known) which we assume to be piecewise linear (we assume that the PVT functions are piecewise linear), whereas the right-hand side is an increasing and continuous function of the pressure. This implies that there can be at most one reservoir pressure which satisfies Equation (4.23). Moreover, since the left-hand side is piecewise linear, we can compute the reservoir pressure thanks to the $\mathcal{W}$ Lambert function (the inverse relation of $f(w) = we^w$), and since pressure is positive, we use the $\mathcal{W}_0$ branch of the Lambert function. Finally, we obtain a function $\Psi : \mathbb{R} \to \mathbb{R}$ such that the pressure

$$P_t^{\mathrm{R}} = \Psi(V_t^{\mathrm{G}}) \ , \ \ \forall t \in \mathcal{T} \ , \tag{4.24}$$

is the solution of Equation (4.23).

As the pressure, P_t^{R}, is given as a function of V_t^{G} and the pore volume, V_t^{P}, is given as a function of the pressure, P_t^{R}, we obtain a reduced state of dimension 1 given by the standard volume of gas V_t^{G}.

The only thing missing in order to get formulation (4.4) is to explicit the production function. The production of gas is given by Equation (4.18). As the reservoir pressure is given by the function Ψ, the production of gas when considering a one tank reservoir is given by

$$F^{\mathrm{G}} = \frac{\mathrm{IPR}^{\mathrm{G}}(\Psi(V^{\mathrm{G}}) - P)}{B_{\mathrm{G}}(\Psi(V^{\mathrm{G}}))} \ .$$

In the numerics, it is assumed that $\mathrm{IPR}^{\mathrm{G}}$, the inflow performance relationship of the well, is a piecewise linear function.

We consider two different models in §4.4.1: a one tank reservoir and a two tanks reservoir, as illustrated by Figure 4.2. In both cases, we have only one well and, as the optimization is done at the bottom of the well, the unique control is given by $u_t = P_t$. The state in the one tank case is $x_t = V_t^{\mathrm{G}}$, whereas it is $x_t = \big((V_t^{\mathrm{G}})^{(1)}, (V_t^{\mathrm{G}})^{(2)}\big)$ for the two tanks case.

We denote by $\Psi_{\mathbf{1T}}$ the function which returns the reservoir pressure of the one tank case given a volume of gas in the reservoir (as defined in Equation (4.24)), and $\Psi_{\mathbf{2T}}$ the function for the producing tank pressure in the two tanks case.

The general production function $\Phi_{\mathbf{1T}}$ of the one tank case is hence given by

$$\Phi^{\mathrm{G}}_{\mathbf{1T}}(x_t, u_t) = \frac{\mathrm{IPR}^{\mathrm{G}}\left(\Psi_{\mathbf{1T}}(x_t) - u_t\right)}{B_{\mathrm{G}}(\Psi_{\mathbf{1T}}(x_t))} = F^{\mathrm{G}}_t \ . \tag{4.25}$$

For the two tanks case, we consider that the well only produces gas from the first tank. The general production function $\Phi_{\mathbf{2T}}$ of the two tanks case is thus given by

$$\Phi^{\mathrm{G}}_{\mathbf{2T}}(x_t, u_t) = \frac{\mathrm{IPR}^{\mathrm{G}}\left(\Psi^{(1)}_{\mathbf{2T}}(x_t) - u_t\right)}{B_{\mathrm{G}}(\Psi^{(1)}_{\mathbf{2T}}(x_t))} = F^{\mathrm{G}}_t \ . \tag{4.26}$$

In the Formulation (4.4) (for the one tank case), we split $\Phi^{\mathrm{G}}_{\mathbf{1T}}$ in Constraints (4.4c) and (4.4d) to explicit the reservoir pressure and to mirror Equation (4.18).

Moreover, since we have only one well and since the IPR function is strictly monotonous, the production function of the well of Equation (4.4d) is injective. In the models considered here (one tank or two tanks), we can thus pass from the controls to the production and from the production to the controls without any ambiguity at a given state: the function $\Phi^{\mathrm{G}}(x, \cdot)$ is a bijection, hence we find the (unique) bottom-hole pressure associated with a given production F^{G} when in state x. Finally, we obtain the admissibility set of the gas reservoir case. As the gas production F^{G}_t must be nonnegative, we obtain that the control must satisfy $P_t \in [0, P^{\mathrm{R}}_t]$ for all time $t \in \mathcal{T}$, which gives the admissible control set

$$\mathcal{U}^{ad}(x_t) = [0, P^{\mathrm{R}}_t] = [0, \Psi_{\mathbf{1T}}(x_t)] \ . \tag{4.27}$$

4.B.2 Oil reservoir with water injection state reduction

Now, we consider an oil reservoir where water injection is used to keep the reservoir pressure constant as in §4.4.2. To eliminate the possibility of having free gas in the reservoir, we assume that the initial pressure in the reservoir is above the bubble-point. Indeed, as we are going to keep the pressure constant, the pressure will always remain above the bubble-point.

We assume that the *produced water*[2] is given by Equation (4.21).

We now prove that the standard volume of water V^{W}_t may be used as a state for describing the reservoir dynamics. To start with, we have that $V^{\mathrm{G}}_t = 0$, $F^{\mathrm{G}}_t = 0$ and $P^{\mathrm{R}}_t = P^{\mathrm{R}}_0$ for all $t \in \mathcal{T}$. Moreover, using Equation (4.12) in place of the linearized version (4.13) we obtain that the pore volume is constant over time and given by $V^{\mathrm{P}}_t = V_0 \exp(c_f P^{\mathrm{R}}_0)$.

[2] Here, the produced water F^{W} is the water that is produced from the well. It should not be confused with the net produced water, which is the difference $F^{\mathrm{W}} - F^{\mathrm{WI}}$ between the water produced and the water injected

Hence, the state dimension can be reduced from dimension 5 to dimension 2 as V_t^{G}, P_t^{R} and V_t^{P} are known over time.

Now, using Equation (4.15) combined with the fact that $V_t^{\mathrm{G}} = 0$, we obtain that

$$V_t^{\mathrm{O}} \times B_{\mathrm{O}}(P_0^{\mathrm{R}}) + V_t^{\mathrm{W}} \times B_{\mathrm{W}}(P_0^{\mathrm{R}}) = V_0^{\mathrm{P}} \ , \forall t \in \mathcal{T} \ . \tag{4.28}$$

Thus, the standard volume of oil in the reservoir is obtained as a function of the standard volume of water as follows

$$V_t^{\mathrm{O}} = \frac{V_0^{\mathrm{P}} - V_t^{\mathrm{W}} \times B_{\mathrm{W}}(P_0^{\mathrm{R}})}{B_{\mathrm{O}}(P_0^{\mathrm{R}})} \ .$$

Moreover, using Equation (4.19) and Equation (4.21), for all time $t \in \mathcal{T}$, we have that

$$F_t^{\mathrm{W}} = \Phi^{\mathrm{W}}(V_t^{\mathrm{W}}, P_t) \tag{4.29a}$$

$$\text{with } \Phi^{\mathrm{W}}(V^{\mathrm{W}}, P) = \frac{\alpha(P_0^{\mathrm{R}} - P)\mathrm{W}^{\mathrm{CT}}\Big(\frac{V^{\mathrm{W}}}{V_0^{\mathrm{P}}} B_{\mathrm{W}}(P_0^{\mathrm{R}})\Big)}{B_{\mathrm{W}}(P_0^{\mathrm{R}})} \ , \tag{4.29b}$$

and

$$F_t^{\mathrm{O}} = \Phi^{\mathrm{O}}(V_t^{\mathrm{W}}, P_t) \tag{4.30a}$$

$$\text{with } \Phi^{\mathrm{O}}(V^{\mathrm{W}}, P) = \frac{\alpha(P_0^{\mathrm{R}} - P)\Bigg(1 - \mathrm{W}^{\mathrm{CT}}\Big(\frac{V^{\mathrm{W}}}{V_0^{\mathrm{P}}} B_{\mathrm{W}}(P_0^{\mathrm{R}})\Big)\Bigg)}{B_{\mathrm{O}}(P_0^{\mathrm{R}})} \ . \tag{4.30b}$$

Now, we turn to the time evolution of the standard volume of water. Equation (4.10) must be changed as we need to introduce the injected water F_t^{WI} at time t to obtain

$$V_{t+1}^{\mathrm{W}} = V_t^{\mathrm{W}} - F_t^{\mathrm{W}} + F_t^{\mathrm{WI}} \ , \ \ \forall t \in \mathcal{T} \ . \tag{4.31}$$

It remains to show that the water injection can be deduced from the previous equations. Using Equation (4.15) at time $t+1$ combined with Equation (4.31) and Equation (4.9) gives

$$\big(V_t^{\mathrm{W}} - F_t^{\mathrm{W}} + F_t^{\mathrm{WI}}\big) \times B_{\mathrm{W}}\left(P_0^{\mathrm{R}}\right) + \big(V_t^{\mathrm{O}} - F_t^{\mathrm{O}}\big) \times B_{\mathrm{O}}(P_0^{\mathrm{R}}) = V_0^{\mathrm{P}} \ , \tag{4.32}$$

which, using Equation (4.28), (4.29b) and (4.30b), gives

$$F_t^{\mathrm{WI}} = F_t^{\mathrm{W}} + F_t^{\mathrm{O}} \times \frac{B_{\mathrm{O}}(P_0^{\mathrm{R}})}{B_{\mathrm{W}}\left(P_0^{\mathrm{R}}\right)} = \frac{\alpha(P_0^{\mathrm{R}} - P_t)}{B_{\mathrm{W}}(P_0^{\mathrm{R}})} \ . \tag{4.33}$$

We conclude that we obtain a state dynamics with a one dimensional state $x_t = V_t^{\mathrm{W}}$, a one dimensional control $u_t = P_t$, and state dynamics given by

$$V_{t+1}^{\mathrm{W}} = V_t^{\mathrm{W}} - \frac{\alpha(P_0^{\mathrm{R}} - P_t)\Big(\mathrm{W}^{\mathrm{CT}}\big(V_t^{\mathrm{W}} B_{\mathrm{W}}(P_0^{\mathrm{R}})/V_0^{\mathrm{P}}\big) - 1\Big)}{B_{\mathrm{W}}(P_0^{\mathrm{R}})} \ . \tag{4.34}$$

4.C Details on the impact of the states and controls discretizations

One tank gas reservoir. In the application of §4.4.1.1, we tried different discretization values for the state and control spaces. Results get better each time we increase the number of states or controls used in the loops of Algorithm 1. The optimal values and CPU times are compiled in Table 4.6. Discretization of the control space has less impact than discretization of the state space (there is no significant improvement when using more than 10 possible controls). We used 50 possible controls for the rest of the state discretization analysis to ensure we do not have any issues due to the control space. Moreover, the computation time grows linearly with the number of controls, hence we only got penalized by a factor of 5 for the computation time compared to being at the most efficient level for the discretization of the controls. We can also remark that going beyond 10,000 points for the state discretization yields no discernible improvement (less than 0.2%). However, the computation time grows exponentially with the state discretization. We hence used 10,000 points for the states and 20 controls for the results presented in §4.4.1.1.

State discretization	Value (M€)	CPU time (s)
100	602	1.25
200	689	1.45
500	725	2.50
1,000	736	7.50
2,000	740	25.20
5,000	742	110.00
10,000	743	653.00
20,000	743	2,288.00
50,000	743	8,142.00

Table 4.6: Summary of the impact of the discretization of the state space on the one tank formulation, with 50 possible controls

Two tanks gas reservoir. We tried different discretization values for the two reservoirs problem of §4.4.1.2: 200×200 (i.e. the two reservoirs are discretized with 200 points each), 400×400, 600×600 and $1{,}000 \times 1{,}000$. Results are summarized in Table 4.7, which shows the computation time of the optimization and the optimal value obtained. As can be seen, the computation time grows exponentially with the discretization, as we need to handle more and more values when we get a finer discretization. However, performance remains reasonable for the number of time steps considered. We can also remark that going past a 200×200 discretization of the states of the reservoir does not improve the optimal value. A very small impact is observed from the discretization of the controls. Indeed, almost no improvement is obtained above 10 possible controls (we hence used 50 possible controls in Table 4.7 to ensure the discretization of the controls will not influence the analysis of the discretization of the states). All the results of §4.4.1.2 were therefore computed with the 400×400 discretization for the states, and 20 for the controls.

State discretization	CPU time (s)	Value (M€)
50×50	5.1	730
100×100	28.3	735
200×200	115.3	736
400×400	706.0	736
600×600	3,893.0	736
1000×1000	18,089.0	736

Table 4.7: Impact of the discretization of the state space on the two tanks model, with 50 possible controls

Oil reservoir with water injection. We tried different values for the discretization of the state space of the problem described in §4.4.2. However, the discretization of the controls had no impact, as the controls only took two different values: either no production, or production at the maximal rate. We therefore chose 10 possible controls to ensure we do not missed another behavior during the analysis on the impact of the discretization of the states. Table 4.8 compiles the time to solve and the associated results of the optimization depending on the number of points considered for the discretization of the states space. We note that there is not a lot of gain from going from 10,000 points to 100,000 points in the discretization, whereas computation time grows by more than 100 times.

Discretization	Time steps	CPU time (s)	Value (M€)
1,000	240	0.35	3182
10,000	240	12.05	3358
100,000	240	1575	3376

Table 4.8: Summary of the dynamic programming results for the oil reservoir with water injection

4.D Additional material on the decline curves formulation

Usually, formulations using decline curves, as can be seen in the works of Iyer et al. [1998], are of the form:

$$\max_u \quad \sum_{t=0}^{T} \rho^t \mathcal{L}_t(u_t) \tag{4.35a}$$

$$s.t. \quad F_t^{\mathrm{O}} \leq g\left(\sum_{s=0}^{t-1} F_s^{\mathrm{O}}\right) \ , \quad \forall t \in \mathcal{T} \setminus \{0\} \ , \tag{4.35b}$$

$$u_t \in \mathcal{U}_t^{ad}\left(\sum_{s=0}^{t-1} F_s^{\mathrm{O}}\right) \ , \quad \forall t \in \mathcal{T} \ . \tag{4.35c}$$

Using decline curves, or oil deliverability curves, means using Equation (4.35b) to predict the reservoir's behavior. It states that the maximal rate at time t only depends on the oil cumulated production until time t. In the general case, there is no reason to believe that there is an equivalence between a material balance model for the reservoir and a decline curve represented with a function g.

However, when the state of the material balance formulation can be reduced to a one dimensional state (such as a reservoir which only contains gas), there can be an equivalence between the decline curve and the material balance formulations, as was stated in Proposition 4.3.

Proof of Proposition 4.3. Let us consider the component $\Phi^{\mathrm{G}} : \mathbb{X} \times \mathbb{U} \to \mathbb{R}$ of the production mapping $\Phi : \mathbb{X} \times \mathbb{U} \to \mathbb{R}^3$ such that

$$F_t^{\mathrm{G}} = \Phi^{\mathrm{G}}\left(x_t, u_t\right) \ . \tag{4.36}$$

Therefore, we have

$$F_t^{\mathrm{G}} \leq \max_u \Phi^{\mathrm{G}}\left(x_t, u\right) \ . \tag{4.37}$$

Moreover, having a one-dimensional state greatly simplifies the dynamics, as we only need to consider one fluid. The dynamics thus simplifies to

$$x_{t+1} = f\left(x_t, u_t\right) = x_t - F_t^{\mathrm{G}} \ . \tag{4.38}$$

By propagating the simplified dynamics (4.38) and by re-injecting it in Equation (4.37), we get:

$$F_t^{\mathrm{G}} \leq \underbrace{\max_u \Phi^{\mathrm{G}}\left(x_0 - \sum_{s=0}^{t-1} F_s^{\mathrm{G}}, u\right)}_{g\left(\sum_{s=0}^{t-1} F_s^{\mathrm{G}}\right)} \ . \tag{4.39}$$

Hence, Equation (4.39) defines the function g. The equivalence exists when the state is reduced to one dimension (as similar reasoning can be applied to the other one-dimensional cases). $\square$

However, when considering more complex cases, such as a reservoir with both oil and gas, or when there is water encroachment (influx of water in the reservoir from the aquifer), we cannot have a reduction to a one-dimensional state. Decline curves, or oil deliverability curves, will not be equivalent to the material balance system, as they can only represent a one dimensional dynamical system, where the state is the cumulated production.

Even if we have a state that cannot be reduced to one dimension, we can still propagate the dynamics in Equation (4.36):

$$\begin{aligned} F_t^{\mathrm{G}} &= \Phi^{\mathrm{G}}(x_t, u_t) \\ &= \Phi^{\mathrm{G}}\left(f\left(f\left(\dots f\left(x_0, u_0\right), \dots\right), u_{t-1}\right), u_t\right) \ . \end{aligned}$$

However, there is no reason to believe that there exists a function g depending on the sum of productions in the general case, contrarily to the one-dimensional case. This is why those functions are generated with a given production planning, i.e. a series of controls applied to the reservoir. Given a series of admissible controls $\tilde{\mathcal{U}} = (\tilde{u}_0, \dots, \tilde{u}_T)$, one can create an oil-deliverability curve, that takes as argument the total cumulated production and returns the maximal possible production. It however depends on the underlying production planning $\tilde{\mathcal{U}}$. We can create such function $\tilde{g}_{\tilde{\mathcal{U}}}$ through the Algorithm 2.

Algorithm 2: Finding the points of the piecewise linear function $\tilde{g}_{\tilde{\mathcal{U}}}$

```
control_to_apply = Ũ;
current_state = x_0;
cumulated_production = 0;
max_production = max_u Φ^G(current_state, u);
list_of_points = {(cumulated_production, max_production)};
for t from 1 to T do
    ũ = control_to_apply[t];
    production = Φ^G(current_state, ũ);
    cumulated_production = cumulated_production + production;
    current_state = f(current_state, ũ);
    max_production = max_u Φ^G( current_state, u);
    push(list_of_points, (cumulated_production, max_production));
end
return list_of_points
```

Once we have a list of points of $\tilde{g}_{\tilde{\mathcal{U}}}$, we consider a linear interpolation between those points as the decline curve we use in the optimization problem (4.5).

In [Iyer et al., 1998, Marmier et al., 2019], the authors use decline curves, i.e. oil-deliverability curves with natural depletion at the maximal rate. This means that there is no injection, and the production planning consists of maximal production rates. We can generate those decline curves with a tweaked version of the previous procedure (see Algorithm 3).

Algorithm 3: Finding the points of the piecewise linear function g

```
current_state = x_0;
cumulated_production = 0;
max_production = max_u Φ^G(current_state, u);
list_of_points = {(cumulated_production, max_production)};
for t from 1 to T do
    ũ = arg max_u Φ^G( current_state, u);
    production = Φ^G(current_state, ũ);
    cumulated_production = cumulated_production + production;
    current_state = f(current_state, ũ);
    max_production = max_u Φ^G(current_state, u);
    push(list_of_points, (cumulated_production, max_production));
end
return list_of_points
```

Chapter 5

Deterministic Partially Observed Markov Decision Processes

Contents

5.1 Introduction

In this chapter, we introduce a new subclass of Partially Observed Markov Decision Processes (POMDP) which we call Separated Deterministic Partially Observed Markov Decision Processes (Separated DET-POMDP). The subclass Separated DET-POMDP, which is included in the Deterministic Partially Observed Markov Decision Processes (DET-POMDP) class, presented in [Bonet, 2009], is of interest for the optimization of oil and gas production. In contrast to general POMDP, this subclass of problem can be numerically tractable. The scope of this chapter is on the definitions, the mathematical tools, and the formulations of POMDPs optimization problems. We will present applications of Separated DET-POMDP to the oil and gas production case in Chapter 6.

To optimize controlled dynamical systems under partial observation, one often uses the formalism of Partially Observed Markov Decision Processes. 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 (such as in [Steimle et al., 2021]) to robotics (such as in [Pajarinen and Kyrki, 2017]) to name a few. Algorithms based on Dynamic Programming ([Bellman, 1957]) have been designed to exploit specific structures in POMDPs in order to solve this difficult class of problems. They do so by first reformulating the problem through the use of beliefs (probability distributions over the state space), such 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. Indeed, working with beliefs implies working on the space of distributions over the state space, which is, by nature, an infinite space. Instead of focusing on the general POMDPs, we present a subclass that is relevant for the oil and gas production case, namely, DET-POMDPs. That subclass of problems has been studied by [Littman, 1996] and [Bonet, 2009]. It was first considered as a limit case of POMDPs for 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 some examples in [Bonet, 2009, §2], such as the navigation of a robot in a partially observed terrain. We introduce and study an even simpler class: Separated DET-POMDPs. Indeed, that new class of problems uses a property of the dynamics and observation to push back the curse of dimensionality.

First, in §5.1.1, we define DET-POMDPs. Second, in §5.1.2, we present how the chapter is organized and present our main contributions.

5.1.1 Formulation of Deterministic Partially Observed Markov Decision Processes

A Det-Pomdp is a restricted case of Pomdps, itself an extension of Markov Decision Processes (Mdps). Backgrounds on Mdps can be found in Appendix 5.A.1, whereas backgrounds on Pomdps can be found in Appendix 5.A.2. 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 a function of the state and controls: the *observations*. For Det-Pomdp, the transitions and observations are given by deterministic evolution and observation functions.

First, we present some notations regarding sets. Second, we present the ingredients of a Det-Pomdp. Third, we present the formulation of a Det-Pomdp optimization problem.

Notations for sets.

- Let t and t' be two integers, with $t' \geq t$. The set $\{t, t+1, \dots, t'\}$ is denoted by $[\![t, t']\!]$.
- We denote by $\overline{\mathbb{R}}$ the extended real numbers, that is
$$\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\} \ ,$$
equipped with the extended (upper) addition that extends the usual addition with $(+\infty) + (-\infty) = (-\infty) + (+\infty) = +\infty$ and the extended (lower) multiplication $0 \times (+\infty) = 0 \times (-\infty) = 0$, following conventions as in [Bertsekas and Shreve, 1978]. Note that these conventions are tailored to minimization problems.[1]
- The cardinality of a finite set is the number of its elements and is denoted by $|\cdot|$.

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

$$\mathcal{D} = \big(\mathcal{T}, \mathbb{U}, \mathbb{O}, \mathbb{X}, \{\mathcal{L}_t\}_{t\in\mathcal{T}\setminus\{T\}}, \{f_t\}_{t\in\mathcal{T}\setminus\{T\}}, \{\mathcal{U}_t^{ad}\}_{t\in\mathcal{T}\setminus\{T\}}, \{h_t\}_{t\in\mathcal{T}}\big) \ , \tag{5.1}$$

which we now detail.

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

[1] In the case of maximization problems, we take the opposite convention considering the (lower) addition: $(+\infty) + (-\infty) = (-\infty) + (+\infty) = -\infty$, while we keep the same convention regarding the (lower) multiplication $0 \times (+\infty) = 0 \times (-\infty) = 0$

- $\mathbb{O}$ is the set of observations available to the decision maker.
- $\mathbb{X}$ is the set of states.
- $\{\mathcal{L}_t\}_{t\in\mathcal{T}\setminus\{T\}}$ 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 $\mathcal{K} : \mathbb{X} \to \mathbb{R}\cup\{+\infty\}$.
- $\{f_t\}_{t\in\mathcal{T}\setminus\{T\}}$ 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.
- $\{\mathcal{U}_t^{ad}\}_{t\in\mathcal{T}\setminus\{T\}}$ 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 $x \in \mathbb{X}$, $\mathcal{U}_t^{ad}(x)$ is a subset on $\mathbb{U}$.
- $\{h_t\}_{t\in\mathcal{T}}$ 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 and $\mathbb{P}$ is a probability measure on Ω. We denote by $\mathbb{E}$ the mathematical expectation operator.

In this chapter, we only consider Det-Pomdp which satisfies the following finite sets assumption.

Assumption 5.1 (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 $|\mathbb{Y}|$.

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

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

$$
\begin{aligned}
\mathcal{V}^\star(b_0) = \min_{\boldsymbol{X},\boldsymbol{O},\boldsymbol{U}} \ & \mathbb{E}\Big[\sum_{t=0}^{T-1} \mathcal{L}_t(\boldsymbol{X}_t,\boldsymbol{U}_t) + \mathcal{K}(\boldsymbol{X}_T)\Big] && \text{(5.2a)}\\
s.t. \ & \mathbb{P}_{\boldsymbol{X}_0} = b_0 \ , && \text{(5.2b)}\\
& \boldsymbol{X}_{t+1} = f_t(\boldsymbol{X}_t,\boldsymbol{U}_t) \ , \ \forall t \in \mathcal{T}\setminus\{T\} \ , && \text{(5.2c)}\\
& \boldsymbol{O}_0 = h_0(\boldsymbol{X}_0) \ , && \text{(5.2d)}\\
& \boldsymbol{O}_{t+1} = h_{t+1}(\boldsymbol{X}_{t+1},\boldsymbol{U}_t) \ , \ \forall t \in \mathcal{T}\setminus\{T\} \ , && \text{(5.2e)}\\
& \boldsymbol{U}_t \in \mathcal{U}_t^{ad}(\boldsymbol{X}_t) \ , \ \forall t \in \mathcal{T}\setminus\{T\} \ , && \text{(5.2f)}\\
& \sigma(\boldsymbol{U}_t) \subset \sigma(\boldsymbol{O}_0,\ldots,\boldsymbol{O}_t,\boldsymbol{U}_0,\ldots,\boldsymbol{U}_{t-1}) \ , \ \forall t \in \mathcal{T}\setminus\{T\} \ , && \text{(5.2g)}
\end{aligned}
$$

where we denote by $\mathcal{V}^\star(b_0)$ the optimal value of Problem (5.2), 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 (5.2), there are three processes $\boldsymbol{X} = \{\boldsymbol{X}_t\}_{t\in\mathcal{T}}$, $\boldsymbol{U} = \{\boldsymbol{U}_t\}_{t\in\mathcal{T}\setminus\{T\}}$ and $\boldsymbol{O} = \{\boldsymbol{O}_t\}_{t\in\mathcal{T}}$. 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 (5.2) is given by Equation (5.2a). In this chapter, we only consider the minimization of the expected value in the finite horizon case.

We now detail the constraints of the optimization Problem (5.2).
First, Equation (5.2b) is the *initialization* constraint. As the initial state is not fully known, we instead use the probability distribution b_0 of the initial state of the system for the initialization.
Second, Equation (5.2c) 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 (5.2d) and (5.2e) define the *observations* evolution functions of the system available at each time step.
Fourth, Equation (5.2f) 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 is

$$\boldsymbol{U}_t(\omega) \in \mathcal{U}_t^{ad}(\boldsymbol{X}_t(\omega)) \ , \ \forall \omega \in \Omega, \forall t \in \mathcal{T}\setminus\{T\} \ .$$

However, for the remainder of this chapter, we instead use a more compact notation by omitting the $\omega \in \Omega$:

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

Finally, Equation (5.2g) is the *non-anticipativity* constraint: it defines the information available to the decision maker before choosing a control at each time step. It states that, at time $t \in \mathcal{T}\setminus\{T\}$, the decision maker has access to an *information vector* $\boldsymbol{I}_t$, defined by the recursion

$$\boldsymbol{I}_0 = (\boldsymbol{O}_0) \ , \quad \text{and,} \quad \forall t \in \mathcal{T} \setminus \{T\} \ , \ \ \boldsymbol{I}_{t+1} = (\boldsymbol{I}_t, \boldsymbol{U}_t, \boldsymbol{O}_{t+1}) \ . \tag{5.3}$$

As all sets Ω, $\mathbb{X}$, $\mathbb{U}$ and $\mathbb{O}$ are assumed to be finite by Assumption 5.1, the optimization Problem (5.2) is well defined.

To summarize, a Det-Pomdp is a Pomdp with the following characteristics:

- There are no exogenous uncertainties for the state transition and the observations. The transitions are given by the evolution mappings $f_t : \mathbb{X} \times \mathbb{U} \to \mathbb{X}$ while the observation are given by the observation mappings $h_t : \mathbb{X} \times \mathbb{U} \to \mathbb{O}$. This contrasts with general Pomdp where exogenous uncertainties may enter the evolution mappings and the observation mappings.
- The initial state x_0 is only known through its probability distribution, the initial belief b_0.

5.1.2 Main contributions

The chapter is organized as follows. First, in §5.2, we present Dynamic Programming on beliefs for Det-Pomdp with constraints. Second, in §5.3, we present a new representation for the belief dynamic using pushforward measures. Third, in §5.4, we present complexity bounds for Dynamic Programming on beliefs for Det-Pomdp. Fourth, in §5.5, we introduce a subclass of Det-Pomdp, Separated Det-Pomdp. Finally, in §5.6 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.

We also present backgrounds on Mdp and Pomdp in Appendix 5.A.1 and 5.A.2. Then, in Appendix 5.A.3, we detail the proof of Dynamic Programming on beliefs for Det-Pomdp with constraints. In Appendix 5.A.4, we present technical lemmata and considerations on pushforward measures. Finally, in Appendix 5.A.5, we present some complements on Separated Det-Pomdps.

We now detail our main contributions.

- In §5.2, we extend [Bertsekas and Shreve, 1978] Dynamic Programming equations with beliefs for unconstrained Pomdps to Det-Pomdps with constraints. Indeed, a key assumption in [Bertsekas and Shreve, 1978] to write Dynamic Programming equations with beliefs is that there are no admissibility constraints on the controls. As the applications dictate the presence of such constraints, we present this extension in Proposition 5.2, which gives us Dynamic Programming Equations (5.16).
- In §5.3, we express the belief dynamics in Det-Pomdps using the notion of *pushforward* (or *image-measure*), as presented in Lemma 5.4. This new representation is the basis for all the improvements of the bounds on the cardinality of the set of reachable beliefs.

- In §5.4, we improve Littman [1996] bound on the cardinality of the set of reachable beliefs, $\left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}\right|$, for Det-Pomdp, from $(1+|\mathbb{X}|)^{|\mathbb{X}|}$, to $(1+|\mathbb{X}|)^{|\mathrm{supp}(b_0)|}$ (see Theorem 5.9). Note that this bound does not depend on the number of time steps. Moreover, in Theorem 5.10, we give a time-dependent bound $1+|\mathrm{supp}(b_0)||\mathbb{U}|^{|\mathcal{T}|}$ on the cardinality of the set of reachable beliefs. Those bounds are of interest, as the well-known complexity of Dynamic Programming on beliefs is $O(|\mathcal{T}||\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)||\mathbb{U}||\mathbb{O}|)$ (see Proposition 5.7).
- In §5.5, we introduce a subclass of Det-Pomdps, Separated Det-Pomdps. As shown in Corollary 5.16, the interest of Separated Det-Pomdps is that it pushes back the curse of dimensionality for Dynamic Programming with beliefs. Indeed, it improves the bound from $(1+|\mathbb{X}|)^{|\mathrm{supp}(b_0)|}$ to $1+\big(2^{|\mathrm{supp}(b_0)|}-|\mathrm{supp}(b_0)|\big)|\mathbb{X}|$ (see Corollary 5.16). Moreover, this last bound is tight (see Proposition 5.18).

We now present Dynamic Programming Equations with beliefs.

5.2 Dynamic Programming for Det-Pomdp with constraints

In this section, we present Dynamic Programming Equations with beliefs for Problem (5.2). 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 (see Appendix 5.A.2 for more background on Pomdps). In the general case, those Dynamic Programming equations take as input the *information vector*.

As stated in §5.A.2.3, Dynamic Programming on the information vector I_t, defined in Equation (5.3), is in practice untractable as the dimension of the information vector grows with time. However, under some assumptions detailed in [Bertsekas and Shreve, 1978], we instead create a belief-Mdp where the state is a probability distribution over the state space called beliefs. 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 §5.2.1, we formally define sets and mappings which are necessary for the formulation of the belief-Mdp. Second, in §5.2.2, we present the Dynamic Programming equations on the resulting belief-Mdp.

5.2.1 Beliefs in Det-Pomdp

First, as beliefs are probability distributions, we present some notations for probability distributions on finite sets, and for partial mappings. Second, we present the set of beliefs. Finally, we present the mappings necessary for the formulation of the belief-Mdp, notably the beliefs dynamics.

Notations for probability distributions and measures on finite sets and for partial mappings. In this chapter, we only work with probability distributions on finite sets.

- Let $b \in \Delta(\mathbb{Y})$, and let $y \in \mathbb{Y}$. By convention, we denote by $b(y)$ the measure of the set $\{y\}$ for the probability b:
$$b(y) = b\big(\{y\}\big) \ . \tag{5.4a}$$
Thus, for $Y \subset \mathbb{Y}$, we have, using Equation (5.4a),
$$b(Y) = \sum_{y \in Y} b(y) \ . \tag{5.4b}$$

- For any nonnegative measure μ on the finite set $\mathbb{Y}$, we denote by $\mathrm{supp}(\mu) \subset \mathbb{Y}$ the support of the measure μ
$$\mathrm{supp}(\mu) = \big\{ y \in \mathbb{Y} \,\big|\, \mu(\{y\}) > 0 \big\} \ . \tag{5.5}$$

We also present a notation for partial mappings.

- Let $\mathbb{A}$, $\mathbb{D}$, $\mathbb{F}$ (and $\mathbb{G}$) be sets. Let $g : \mathbb{A} \times \mathbb{D} \to \mathbb{F}, (a,d) \mapsto g(a,d)$. We denote by g^d the mapping
$$g^d : \mathbb{A} \to \mathbb{F}, a \mapsto g(a,d) \ , \tag{5.6}$$
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 $g^{(d,f)} : \mathbb{A} \to \mathbb{G}, a \mapsto g(a,d,f)$.

- We now present some notations for the composition of functions with fixed variables. Let n be a positive integer, let $\big(g_i\big)_{i \in [\![1,n]\!]}$ be n mappings such that, for all $i \in [\![1,n]\!]$, $g_i : \mathbb{A}_i \times \mathbb{D}_i \to \mathbb{A}_{i+1}$, and let $d_{1:n} \in \prod_{i \in [\![1,n]\!]} \mathbb{D}_i$ be an element of the product space of $\mathbb{D}_i$. We denote by $g_{1:n}^{d_{1:n}}$ the composition mapping
$$g_{1:n}^{d_{1:n}} : \mathbb{A}_1 \to \mathbb{A}_{n+1} \ , \ \ g_{1:n}^{d_{1:n}} = g_n^{d_n} \circ g_{n-1}^{d_{n-1}} \circ \cdots \circ g_1^{d_1} \ . \tag{5.7}$$
We illustrate this notation in the case where we have two mappings. Let $g_1 : \mathbb{A}_1 \times \mathbb{D}_1 \to \mathbb{A}_2$ and $g_2 : \mathbb{A}_2 \times \mathbb{D}_2 \to \mathbb{A}_3$, and let $(d_1, d_2) \in \mathbb{D}_1 \times \mathbb{D}_2$:
$$\forall a_1 \in \mathbb{A}_1 \ , \ \ g_{1:2}^{d_{1:2}}(a_1) = g_2^{d_2} \circ g_1^{d_1}(a_1) = g_2\big(g_1(a_1, d_1), d_2\big) \ .$$

We now present the sets necessary for the definition of beliefs.

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 (5.14) 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 ∂:

$$\overline{\mathbb{X}} = \mathbb{X} \cup \{\partial\} \ . \tag{5.8}$$

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

$$\mathbb{B} = \Delta(\mathbb{X}) \cup \{\delta_\partial\} \ , \tag{5.9}$$

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 as defined in Equation (5.5). The probability measure δ_∂ can be called the *cemetery* belief as we will see in Equation (5.14) that the belief dynamics, when reaching the belief state δ_∂, remains in δ_∂ forever. A probability measure $\nu \in \Delta(\overline{\mathbb{X}})$ are represented in some equations by the ordered pair $\big(\nu_{|\mathbb{X}}, \nu(\partial)\big)$, where $\nu_{|\mathbb{X}}$ is a nonnegative measures 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 $\tau_t : \mathbb{B} \times \mathbb{U} \times \mathbb{O} \to \mathbb{B}$.

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

$$\forall t \in \mathcal{T} \setminus \{T\} \ , \ \ Q_{t+1} : (b,u,o) \ni \mathbb{B} \times \mathbb{U} \times \mathbb{O} \mapsto b\big((h^u_{t+1} \circ f^u_t)^{-1}(o)\big) \ , \tag{5.10}$$

where $f^u_t(\cdot)$ and $h^u_t(\cdot)$ are partial mapping that follow the notation defined in Equation (5.6):

$$\forall u \in \mathbb{U}, \quad f^u_t : \mathbb{X} \to \mathbb{X} \ , \ \ x \mapsto f_t(x,u) \ , \tag{5.11}$$

$$\forall u \in \mathbb{U}, \quad h^u_t : \mathbb{X} \to \mathbb{O} \ , \ \ x \mapsto h_t(x,u) \ , \tag{5.12}$$

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, following Notation (5.4b). Note that, we always have that

$$Q_{t+1}(\delta_\partial, u, o) = \delta_\partial\big((h^u_{t+1} \circ f^u_t)^{-1}(o)\big) = 0 \ , \tag{5.13}$$

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

- for all $y \in \mathbb{X}$,

$$\tau_t(b,u,o)(y) = \begin{cases} \dfrac{b\big((f_t^u)^{-1}(y)\big)}{Q_{t+1}(b,u,o)} & \text{if } Q_{t+1}(b,u,o) \neq 0, \text{ and } y \in \left(h_{t+1}^u\right)^{-1}(o) \;, \\ 0 & \text{otherwise,} \end{cases} \tag{5.14}$$

- for $y = \partial$,

$$\tau_t(b,u,o)(\partial) = 1 - \tau_t(b,u,o)(\mathbb{X}) \;. \tag{5.15}$$

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. Under an admissible sequence of controls and observations, it should not be possible to attain such belief δ_∂. Indeed, δ_∂ is used to ensure that the mappings τ_t are well defined for all beliefs, controls and observations.

Lemma 5.1. *We assume that the finite set Assumption 5.1 holds. For all time $t \in \mathcal{T} \setminus \{T\}$, the belief dynamics mapping τ_t given by Equation (5.14) takes its values in the set $\mathbb{B} = \Delta(\mathbb{X}) \cup \{\delta_\partial\}$.*

Proof. Let time $t \in \mathcal{T} \setminus \{T\}$, control $u \in \mathbb{U}$ and observation $o \in \mathbb{O}$ be fixed. First, suppose that $b = \delta_\partial$. Using Equation (5.13), we have that $Q_{t+1}(\delta_\partial, u, o) = 0$ and using Equation (5.14), we obtain that $\tau_t(\delta_\partial, u, o) = \delta_\partial \in \mathbb{B}$.

Second, suppose that $b \in \Delta(\mathbb{X})$. If $Q_{t+1}(b,u,o) = 0$, we obtain by (5.14) that $\tau_t(b,u,o)(y) = 0$ for all $y \in \mathbb{X}$. Therefore, $\tau_t(b,u,o) = \delta_\partial \in \mathbb{B}$. Else, if $Q_{t+1}(b,u,o) \neq 0$, then we have

$$\begin{aligned}
\sum_{y \in \mathbb{X}} \tau_t(b,u,o)(y) &= \sum_{y \in (h_{t+1}^u)^{-1}(o)} \frac{b\big((f_t^u)^{-1}(y)\big)}{Q_{t+1}(b,u,o)} && \text{(by definition of } \tau_t \text{ in Equation (5.14))} \\
&= \frac{b\Big(\sqcup_{y \in (h_{t+1}^u)^{-1}(o)} (f_t^u)^{-1}(y)\Big)}{Q_{t+1}(b,u,o)} \\
& && \text{(probability of a two by two disjoint union } \sqcup) \\
&= \frac{b\Big((f_t^u)^{-1}\big((h_{t+1}^u)^{-1}(o)\big)\Big)}{Q_{t+1}(b,u,o)} && \text{(as } f^{-1}(A) = \cup_{a \in A} f^{-1}\big(\{a\}\big)) \\
&= \frac{b\big((h_{t+1}^u \circ f_t^u)^{-1}(o)\big)}{Q_{t+1}(b,u,o)} \\
&= 1 \;. && \text{(as } b\big((h_{t+1}^u \circ f_t^u)^{-1}(o)\big) = Q_{t+1}(b,u,o) \text{ by (5.10))}
\end{aligned}$$

Hence as $\tau_t(b,u,o)(\mathbb{X}) = 1$, we have by Equation (5.15) that $\tau_t(b,u,o)(\partial) = 0$, and we obtain by Equations (5.13), (5.14) and (5.15) that $\tau_t(b,u,o) \in \Delta(\mathbb{X}) \subset \mathbb{B}$.
This ends the proof. □

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.

5.2.2 Dynamic Programming Equations for DET-POMDPs with constraints

We now show that Dynamic Programming equations on the belief-MDP solve Problem (5.2). 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 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 (5.16a) and (5.16b) together with the expression of the beliefs dynamics $\{\tau_t\}_{t\in\mathcal{T}\setminus\{T\}}$ (see Equation (5.14)) in the case where there are no constraints on the controls in [Littman, 1996]. As stated in §5.A.2.3, 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 (5.2), as it contains Constraint (5.2f). As it is natural to consider such constraints for MDPs, POMDPs and DET-POMDPs, our first contribution is to extend classical results by [Bertsekas and Shreve, 1978] in order to tackle such constraints. This is the purpose of Proposition 5.2, where we identify an admissibility set for beliefs of the form $\mathcal{U}^{\mathrm{b}}(b) = \bigcap_{x\in\mathrm{supp}(b)} \mathcal{U}^{ad}(x)$. Note that, as far as we know, the first Dynamic Programming equations using such sets $\mathcal{U}^{\mathrm{b}}(b)$ were given in [Geffner and Bonet, 1998, §5] with no explicit proof.

Proposition 5.2. *Consider a* DET-POMDP *optimization problem given by Problem* (5.2) *which satisfies the finite sets Assumption 5.1. Let $\mathbb{B} = \Delta(\mathbb{X}) \cup \{\delta_\partial\}$, as defined in Equation* (5.9) *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*

$$V_T : b \in \Delta(\mathbb{X}) \mapsto \sum_{x\in\mathbb{X}} b(x)\mathcal{K}(x) \;, \tag{5.16a}$$

$$V_t : b \in \Delta(\mathbb{X}) \mapsto \min_{u\in\mathcal{U}_t^{\mathrm{b}}(b)} \Big(\mathcal{C}_t(b,u) + \sum_{o\in\mathbb{O}} Q_{t+1}(b,u,o) V_{t+1}\big(\tau_t(b,u,o)\big)\Big) \;, \tag{5.16b}$$

where

$$\mathcal{C}_t(b,u) = \sum_{x\in\mathbb{X}} b(x)\mathcal{L}_t(x,u) \;, \tag{5.16c}$$

and

$$\mathcal{U}_t^{\mathrm{b}}(b) = \bigcap_{x \in \mathrm{supp}(b)} \mathcal{U}_t^{ad}(x) \; . \tag{5.16d}$$

Then, the optimal value of Problem (5.2) and the value of the function V_0 at the initial belief $b = b_0$ are equal, that is, $\mathcal{V}^\star(b_0) = V_0(b_0)$. 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 (5.16b) for each b and t is an optimal policy of Problem (5.2): 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 (5.2).

Proof. The details of the proof are given in Appendix 5.A.3. The proof follows the structure presented in Figure 5.1, where Dynamic Programming is abbreviated to DP.

1. First, we rewrite Problem (5.2) as an equivalent problem, Problem (5.68), without constraint (5.2f) by adding characteristic functions of the constraints to the instantaneous costs. By Lemma 5.19, the two problems are indeed equivalent, and Problem (5.68) follows the framework of [Bertsekas and Shreve, 1978].

2. Second, we apply the results of [Bertsekas and Shreve, 1978] to the reformulated Problem (5.68), and obtain associated Dynamic Programming equations.

3. Third, by Lemma 5.20, the Dynamic Programming equations which solve Problem (5.68) are equivalent to Equations (5.16) presented in Proposition 5.2, thus concluding that Equations (5.16) give the solution of Problem (5.2) as formulated in Proposition 5.2.

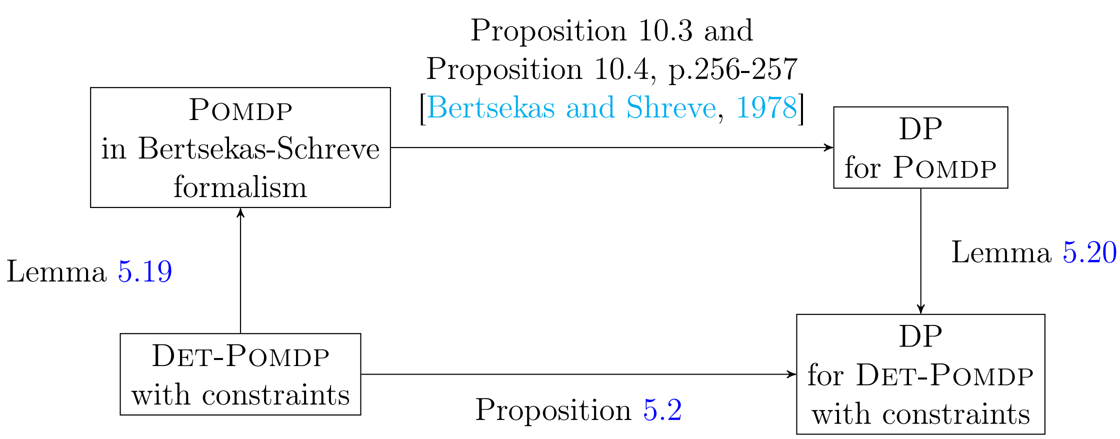


FIGURE 5.1: Illustration of the sketch of proof of Proposition 5.2, where Dynamic Programming is abbreviated to DP.

□

Now that we have presented Dynamic Programming equations on beliefs, we present a new representation of the belief dynamics using pushforward measures.

5.3 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 (5.14). First, we recall the notion of *pushforward measures* when considering finite sets. Second, we introduce the mappings necessary for the new representations. We then present in Lemma 5.4 the representation of the belief dynamics as pushforward measures.

Definition 5.3. *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*

$$\forall d \in \mathbb{D}\ ,\ \ (h_\star\mu)(d) = \mu\big(h^{-1}(d)\big) = \sum_{a\in\mathbb{A},h(a)=d} \mu(a)\ . \tag{5.17}$$

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 5.4, we first introduce two 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 (5.8)), and defined by:

$$F_t^{u,o} : \overline{\mathbb{X}} \to \overline{\mathbb{X}}\ ,\ \ \overline{x} \mapsto \begin{cases} f_t^u(\overline{x}) & \text{if} \quad \overline{x} \neq \partial \text{ and } f_t^u(\overline{x}) \in \big(h_{t+1}^u\big)^{-1}(o)\ , \\ \partial & \text{otherwise.} \end{cases} \tag{5.18}$$

The mapping $F_t^{u,o}$ hence applies the dynamics f_t 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

$$\mathcal{R} : \nu \in \Delta(\overline{\mathbb{X}}) \mapsto \begin{cases} \big(\frac{1}{\nu(\mathbb{X})}\nu_{|\mathbb{X}}, 0\big) & \text{if } \nu(\mathbb{X}) \neq 0\ , \\ \delta_\partial & \text{if } \nu(\mathbb{X}) = 0\ . \end{cases} \tag{5.19}$$

We now express the belief dynamics as pushforward measures.

Lemma 5.4. *Let $(u, o) \in \mathbb{U}\times\mathbb{O}$ be given, and let $t \in \mathcal{T}\setminus\{T\}$. We have*

$$\forall b \in \mathbb{B}\ ,\ \ \tau_t(b, u, o) = \mathcal{R}\circ(F_t^{u,o})_\star(b)\ , \tag{5.20}$$

where the pushforward $(F_t^{u,o})_\star(b)$ follows Notation (5.17).

Proof. The proof is detailed in Appendix 5.A.4. □

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 $(u, o) \in \mathbb{U}\times\mathbb{O}$.

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

FIGURE 5.2: Illustration of the beliefs dynamics as pushforward measures

Corollary 5.5. *Let $(u, u', o, o') \in \mathbb{U}^2 \times \mathbb{O}^2$ be given, and let $t \in \mathcal{T} \setminus \{T-1, T\}$. We have*

$$
\forall b \in \mathbb{B} \ , \ \ \tau_{t+1}^{u',o'} \circ \tau_t^{u,o}(b) = \mathcal{R} \circ (F_{t+1}^{u',o'} \circ F_t^{u,o})_\star(b) \ . \tag{5.21}
$$

Proof. Corollary 5.5 is a direct consequence of Lemma 5.4 and Lemma 5.23 found in Appendix 5.A.4. □

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

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

FIGURE 5.3: Illustration of the composition of belief dynamics as pushforward measures.

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.

5.4 Dynamic Programming complexity for DET-POMDP

In §5.2, we presented Dynamic Programming for DET-POMDP. We now study its *complexity*, i.e. the number of operations necessary to solve Problem (5.2), using the representation of beliefs as pushforward measures presented in §5.3.

According to Proposition 5.2, we can solve Problem (5.2) by computing $V_0(b_0)$ by means of Equations (5.16). Solving Dynamic Programming Equations (5.16) 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. The section is organized as follows. We first define the *set of reachable beliefs* in §5.4.1, and we detail a Dynamic Programming algorithm that solves Problem (5.2). Second, in §5.4.2, we present new bounds on the cardinality of the set of reachable beliefs.

5.4.1 Reachable beliefs and Dynamic Programming complexity

We start by formally defining the set of reachable beliefs, before we present a Dynamic Programming algorithm and give our first complexity result.

5.4.1.1 Set of reachable beliefs

The set of reachable beliefs is defined as follows.

Definition 5.6. *Let $b_0 \in \Delta(\mathbb{X})$ be given and consider the sequence of subsets of the beliefs $\mathbb{B}$ defined by the induction:*

$$\mathbb{B}_0^{\mathrm{R},\mathcal{D}}(b_0) = \{b_0\} \quad and \quad \forall t \in \mathcal{T} \setminus \{T\} \ , \ \mathbb{B}_{t+1}^{\mathrm{R},\mathcal{D}}(b_0) = \tau_t\big(\mathbb{B}_t^{\mathrm{R},\mathcal{D}}(b_0), \mathbb{U}, \mathbb{O}\big) \ , \tag{5.22}$$

where τ_t is defined in Equation (5.14). 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 $\mathbb{B}_{[\![1,t]\!]}^{\mathrm{R},\mathcal{D}}(b_0)$ the union for t' in the time interval $[\![1,t]\!]$ of the reachable beliefs at time t' starting from the initial belief $b_0 \in \Delta(\mathbb{X})$, that is,

$$\forall t \in \mathcal{T} \setminus \{0\} \ , \ \mathbb{B}_{[\![1,t]\!]}^{\mathrm{R},\mathcal{D}}(b_0) = \bigcup_{t'=1}^{t} \mathbb{B}_{t'}^{\mathrm{R},\mathcal{D}}(b_0) \ . \tag{5.23}$$

The set $\mathbb{B}_{[\![1,t]\!]}^{\mathrm{R},\mathcal{D}}$ is called the set of reachable beliefs*.*

We also denote by $\mathbb{B}_{[\![0,t]\!]}^{\mathrm{R},\mathcal{D}}(b_0)$ the union starting from time 0, i.e.

$$\mathbb{B}_{[\![0,t]\!]}^{\mathrm{R},\mathcal{D}}(b_0) = \bigcup_{t'=0}^{t} \mathbb{B}_{t'}^{\mathrm{R},\mathcal{D}}(b) = \{b_0\} \cup \mathbb{B}_{[\![1,t]\!]}^{\mathrm{R},\mathcal{D}}(b_0) \ . \tag{5.24}$$

Note that under Assumption 5.1, the set $\mathbb{B}_{[\![1,T]\!]}^{\mathrm{R},\mathcal{D}}(b_0)$ is finite. Note also 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 $\mathcal{D}$ defined in Equation (5.1), whereas the upper index R stands for reachable.

We now write a Dynamic Programming algorithm over the set of reachable beliefs.

5.4.1.2 Dynamic Programming Algorithm for Det-Pomdps

Here, the initial probability distribution b_0 is fixed. We present a Dynamic Programming algorithm and give our first complexity results.

Proposition 5.7. *Consider a* Det-Pomdp *optimization problem given by Problem* (5.2) *which satisfies the finite sets Assumption 5.1. Let* $b_0 \in \Delta(\mathbb{X})$*. Then, Algorithm 4 (numerically) solves Problem* (5.2) *by Dynamic Programming 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* (5.23).

Proof. First, note that as we consider that Assumption 5.1 holds, $\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)$ is finite and all the loops in Algorithm 4 are finite (all the sets are finite). As there is no recursive calls to functions in the Algorithm, only unitary functions of the problem, Algorithm 4 ends, and its complexity is given by the multiplications of the number of operations in each loop. Moreover, as we have for all time $t \in \mathcal{T}$, $t > 0$, $\mathbb{B}^{\mathrm{R},\mathcal{D}}_t(b_0) \subset \mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)$, we have $|\mathbb{B}^{\mathrm{R},\mathcal{D}}_t(b_0)| \leq |\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)|$. We also have $\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0) \neq \emptyset$[2] and $\mathbb{B}^{\mathrm{R},\mathcal{D}}_0(b_0) = \{b_0\}$, hence $|\mathbb{B}^{\mathrm{R},\mathcal{D}}_0(b_0)| \leq |\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)|$.

Thus, the complexity of Algorithm 4 is $O(|\mathcal{T}||\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)||\mathbb{U}||\mathbb{O}|)$.

We now prove that Algorithm 4 indeed yields the optimal value and an optimal policy of Problem (5.2).

First, as Assumption 5.1 holds, we can apply Proposition 5.2 on Problem (5.2). We can hence solve Problem (5.2) by computing value functions given by Equations (5.16).

Second, by using the definition of the set of reachable beliefs (see Definition 5.6), we obtain that, when we consider the value functions V_t defined in Equations (5.16), at each time $t \in \mathcal{T}$, we only need to compute the value functions at beliefs that belong to $\mathbb{B}^{\mathrm{R},\mathcal{D}}_t(b_0)$ in order to get to $V_0(b_0)$. Indeed, by Definition 5.6, all the beliefs at time $t \in \mathcal{T}$ that can be reached when starting at belief b_0 are, by definition, in $\mathbb{B}^{\mathrm{R},\mathcal{D}}_t(b_0)$, hence we get to $V_0(b_0)$ through the Bellman equation (in Equation (5.16b)).

Third, the loops on the controls u and observations o return, for a given belief, the value functions defined in Equations (5.16). As the first loop proceeds backward in time, Algorithm 4 indeed returns the values of the value functions presented in Proposition 5.2. Thus, we can indeed solve Problem (5.2) thanks to Algorithm 4. □

In order to apply Proposition 5.7 on Problem (5.2) and to get complexity bounds on Algorithm 4, 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.

[2] 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 $o \in \mathbb{O}$, $\tau_0(b_0, u, o) \in \mathbb{B}^{\mathrm{R},\mathcal{D}}_1(b_0) \subset \mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)$

Algorithm 4: Computation of the value functions and policies V_t, π_t

```
for b ∈ 𝔹_T^{R,𝒟}(b_0) do
    V_T(b) = ∑_{x∈𝕏} b(x)𝒦(x) ;
end
for t = T − 1, . . . , 0 do
    for b ∈ 𝔹_t^{R,𝒟}(b_0) do
        best_value = + ∞;
        best_controls = ∅ ;
        for u ∈ 𝒰_t^b(b) do
            current_value = 𝒞_t(b, u) ;
            future_value = 0;
            for o ∈ h_{t+1}(supp(b), u) do
                future_value += Q_{t+1}(b, u, o) * V_{t+1}(τ_t(b, u, o));
            end
            current_value += future_value;
            if current_value < best_value then
                best_value = current_value;
                best_controls = u;
            end
        end
        V_t(b) = best_value;
        π_t(b) = best_controls;
    end
end
return (V_t, π_t)_{t∈𝒯}
```

5.4.2 Bounds on the cardinality of the set of reachable beliefs

Under the finite sets Assumption 5.1, the set of reachable beliefs $\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)$, defined in Equation (5.23), is finite. Here, we present new bounds on its cardinality.

Horizon-independent bounds on the cardinality of the set of reachable beliefs. In Theorem 5.9, we give a bound on the cardinality of the set $\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![0,t]\!]}(b_0) = \mathbb{B}^{\mathrm{R},\mathcal{D}}_{0}(b_0) \cup \mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,t]\!]}(b_0)$ which is independent of time $t \in \mathcal{T}$ (the bound would be the same when considering an infinite horizon) and which improves a previous result recalled in Remark 5.8.

Remark 5.8. *Littman presents in [Littman, 1996, Lemma 6.1] a bound on the set of reachable beliefs starting from belief $b_0 \in \Delta(\mathbb{X})$:*

$$\forall t \in \mathcal{T} \ , \ \left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![0,t]\!]}(b_0)\right| \leq (1 + |\mathbb{X}|)^{|\mathbb{X}|} \ . \tag{5.25}$$

We now present in Equation (5.26) an improvement on the bound given in Equation (5.25) 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 (5.26) is tighter than Equation (5.25).

Theorem 5.9. *Consider a* Det-Pomdp *optimization problem given by Problem (5.2) which satisfies the finite sets Assumption 5.1. For any initial belief $b_0 \in \Delta(\mathbb{X})$, the cardinality of the set of reachable beliefs starting from b_0, defined in Equation (5.23), satisfies the following bound:*

$$\forall t \in \mathcal{T} \ , \ \left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![0,t]\!]}(b_0)\right| \leq (1 + |\mathbb{X}|)^{|\mathrm{supp}(b_0)|} \ . \tag{5.26}$$

Proof. With the help of the representation of the beliefs evolution mappings given by Lemma 5.4, the proof of Theorem 5.9 is a simple application of Lemma 5.24, given in Appendix 5.A.4, that we detail now.

First, we present some notations regarding 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

$$\mathbb{G}(Y) = \left\{ g(y) \,\middle|\, y \in Y \text{ and } g \in \mathbb{G} \right\} \subset \mathbb{V} \ , \tag{5.27a}$$
$$\mathbb{G}_\star(B) = \left\{ g_\star b \,\middle|\, b \in B \text{ and } g \in \mathbb{G} \right\} \subset \Delta(\mathbb{V}) \ , \tag{5.27b}$$
$$\mathbb{G}_\star(b) = \mathbb{G}_\star(\{b\}) \subset \Delta(\mathbb{V}) \ . \tag{5.27c}$$

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

$$\mathbb{G}' \circ \mathbb{G}'' = \left\{ g' \circ g'' \,\middle|\, g' \in \mathbb{G}' \text{ and } g'' \in \mathbb{G}'' \right\} \ . \tag{5.27d}$$

- 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 $\mathbb{G}_{0:k}$ of $\mathbb{L}(\mathbb{Y};\mathbb{Y})$ defined by

$$\mathbb{G}_{0:k} = \mathbb{G}_k \circ \mathbb{G}_{k-1} \circ \cdots \circ \mathbb{G}_0 \ , \ \ \forall k \in \mathbb{N} \ . \tag{5.27e}$$

We now return to the proof. For a fixed value of $u \in \mathbb{U}$, and $o \in \mathbb{O}$, for all $t \in \mathcal{T}\backslash\{T\}$, we have obtained in Lemma 5.4 that $\tau_t(\cdot, u, o) = \mathcal{R} \circ (F_t^{u,o})_\star$. Now, for each $t \in \mathcal{T}$, introducing the sets

$$\mathbb{T}_t^{\mathcal{D}} = \big\{\tau_t(\cdot, u, o) \,\big|\, u \in \mathbb{U}, o \in \mathbb{O}\big\} \subset \mathbb{L}(\mathbb{B};\mathbb{B}) \ , \tag{5.28}$$

$$\mathbb{F}_t^{\mathcal{D}} = \big\{F_t^{u,o} \,\big|\, u \in \mathbb{U}, o \in \mathbb{O}\big\} \subset \mathbb{L}(\overline{\mathbb{X}};\overline{\mathbb{X}}) \ , \tag{5.29}$$

$$\mathbb{F}^{\mathcal{D}} = \bigcup_{t\in\mathcal{T}\backslash\{T\}} \mathbb{F}_{0:t}^{\mathcal{D}} \tag{5.30}$$

where the composition of sets of mapping is given by Notation (5.27d) and (5.27e). Note that $\mathbb{F}_{0:t}^{\mathcal{D}} \neq \mathbb{F}_{[\![0,t]\!]}^{\mathcal{D}}$: $\mathbb{F}_{0:t}^{\mathcal{D}}$ 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.

Using Lemma 5.4 and Notation (5.27b), we obtain that $\mathbb{T}_t^{\mathcal{D}} = \mathcal{R} \circ (\mathbb{F}_t^{\mathcal{D}})_\star$. Moreover, in order to account for $\mathbb{B}_0^{\mathrm{R},\mathcal{D}}(b_0)$, we introduce sets $\mathbb{T}_{-1}^{\mathcal{D}} = \{\tau_{-1}\}$ and $\mathbb{F}_{-1}^{\mathcal{D}} = \{F_{-1}\}$, where τ_{-1} and F_{-1} are given by the identity mappings on $\mathbb{B}$ and $\overline{\mathbb{X}}$:

$$\tau_{-1} : \mathbb{B} \to \mathbb{B} \ , \ \ b \mapsto b \quad \text{and} \qquad F_{-1} : \overline{\mathbb{X}} \to \overline{\mathbb{X}} \ , \ \ x \mapsto x \ . \tag{5.31}$$

Then, using the definition of $\mathbb{B}_t^{\mathrm{R},\mathcal{D}}(b_0)$ in Equation (5.22), we have that, for all time $t \in \mathcal{T}$,

$$\mathbb{B}_t^{\mathrm{R},\mathcal{D}}(b_0) = \mathbb{T}_{t-1}^{\mathcal{D}} \circ \mathbb{T}_{t-2}^{\mathcal{D}} \circ \cdots \circ \mathbb{T}_0^{\mathcal{D}} \circ \mathbb{T}_{-1}^{\mathcal{D}}(b_0) = \mathbb{T}_{(-1):(t-1)}^{\mathcal{D}}(b_0) \ , \tag{5.32}$$

where the composition of sets of mapping is given by Notation (5.27d).

Note that, as $\mathbb{T}_{-1}^{\mathcal{D}}$ is the set that only contains the identity function, we also have for all time $t \in \mathcal{T}$, $t > 0$,

$$\mathbb{B}_t^{\mathrm{R},\mathcal{D}}(b_0) = \mathbb{T}_{t-1}^{\mathcal{D}} \circ \mathbb{T}_{t-2}^{\mathcal{D}} \circ \cdots \circ \mathbb{T}_0^{\mathcal{D}}(b_0) = \mathbb{T}_{0:t-1}^{\mathcal{D}}(b_0) \ . \tag{5.33}$$

Finally, we obtain

$$\forall t \in \mathcal{T} \setminus \{0\} \ , \ \ |\mathbb{B}_{[\![0,t]\!]}^{\mathrm{R},\mathcal{D}}(b_0)| \overset{(5.24)}{=} \Big|\bigcup_{i=0}^{t} \mathbb{B}_i^{\mathrm{R},\mathcal{D}}(b_0)\Big| \overset{(5.32)}{=} \Big|\bigcup_{i=-1}^{t-1} \mathbb{T}_{-1:i}^{\mathcal{D}}(b_0)\Big| \overset{(5.96)}{\leq} (1 + |\mathbb{X}|)^{|\mathrm{supp}(b_0)|} \ .$$

The last inequality is given by Equation (5.96), obtained by applying Lemma 5.24. As all the elements of $\mathbb{F}_t^{\mathcal{D}}$ are of the form given in Equation (5.18), and $\mathbb{F}_{-1}^{\mathcal{D}}$ is a $\big(\overrightarrow{\mathbb{X}}\big)$-mappings set ($F_{-1}$ is the identity function on $\overline{\mathbb{X}}$, hence a $\mathbb{X}$-forward mapping), the two sequences $\{\mathbb{F}_t^{\mathcal{D}}\}_{t\in[\![-1,T]\!]}$ and $\{\mathbb{T}_t^{\mathcal{D}}\}_{t\in[\![-1,T]\!]}$ satisfy the assumptions of Lemma 5.24 where the role of

$\{\Phi_k\}_{k\in\mathbb{N}}$ is taken by $\{\mathbb{T}^{\mathcal{D}}_t\}_{t\in[\![-1,T]\!]}$ and the role of $\{\mathbb{G}_k\}_{k\in\mathbb{N}}$ is taken by $\{\mathbb{F}^{\mathcal{D}}_t\}_{t\in[\![-1,T]\!]}$ (the proof of Lemma 5.4, in Appendix 5.A.4 states that set $\mathbb{F}^{\mathcal{D}}_t$ is a $(\overrightarrow{\mathbb{X}})$-mappings set). □

The number of reachable beliefs of a Det-Pomdp is therefore finite even when considering the case of an infinite horizon. This might look counter-intuitive at first, but since there is no uncertainty beyond the distribution of the initial state, and since we consider finite sets $\mathbb{X}$, $\mathbb{U}$, $\mathbb{O}$ under Assumption 5.1, there is a finite number of possible pushforward measures of the initial belief, hence a finite number of reachable beliefs, time notwithstanding.

We now present another bound that takes into account the time horizon.

Horizon-dependent bounds on the cardinality of the set of reachable beliefs. We previously exposed in Theorem 5.9 a bound on the cardinality of the set of reachable beliefs that does not depend on the horizon of the optimization. We now present a bound that depends on the time span of the Det-Pomdp.

Theorem 5.10. *Consider a* Det-Pomdp *optimization problem given by Problem (5.2) which satisfies the finite sets Assumption 5.1, 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 (5.23), satisfies the following bound*

$$\left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)\right| \leq \min\left((1+|\mathbb{X}|)^{|\mathrm{supp}(b_0)|}, 1+|\mathrm{supp}(b_0)||\mathbb{U}|^{|\mathcal{T}|}\right) . \tag{5.34}$$

Proof. Let $b_0 \in \Delta(\mathbb{X})$ be given. Using Theorem 5.9 and Equation (5.26), we already have the inequality $|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)| \leq (1+|\mathbb{X}|)^{|\mathrm{supp}(b_0)|}$. Thus, it is enough to prove that we have

$$\left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)\right| \leq 1+|\mathrm{supp}(b_0)||\mathbb{U}|^{|\mathcal{T}|} , \tag{5.35}$$

in order to obtain Inequality (5.34). With the help of the representation of the beliefs evolution mappings given by Lemma 5.4, Inequality (5.35) is obtained as an application of Lemma 5.27, given in Appendix 5.A.4, that we detail now.

For a fixed value of $t \in \mathcal{T} \setminus \{T\}$, $u \in \mathbb{U}$, and $o \in \mathbb{O}$, we have obtained in Lemma 5.4 that $\tau_t(\cdot, u, o) = \mathcal{R} \circ (F^{u,o}_t)_\star$, where $\mathcal{R}$ is defined in Equation (5.19), and the pushforward measure $(F^{u,o}_t)_\star$ uses Notation (5.17). Now, for each $t \in \mathcal{T} \setminus \{T\}$ and each $u_t \in \mathbb{U}$ we introduce the sets

$$\mathbb{T}^{\mathcal{D},u_t}_t = \left\{\tau_t(\cdot, u_t, o) \,\middle|\, o \in \mathbb{O}\right\} , \quad \text{and} \quad \mathbb{F}^{\mathcal{D},u_t}_t = \left\{F^{u_t,o}_t \,\middle|\, o \in \mathbb{O}\right\} .$$

Using set notations described in Equations (5.27) we obtain that $\mathbb{T}^{\mathcal{D},u_t}_t = \mathcal{R} \circ (\mathbb{F}^{\mathcal{D},u_t}_t)_\star$. Then, using the definition of $\mathbb{B}^{\mathrm{R},\mathcal{D}}_t(b_0)$ in Equation (5.22) we have that, for all time $t \in \mathcal{T}$, $t > 0$,

$$\mathbb{B}^{\mathrm{R},\mathcal{D}}_t(b_0) = \bigcup_{u_{0:t-1}\in\mathbb{U}_{0:t-1}} \mathbb{T}^{\mathcal{D},u_{t-1}}_{t-1} \circ \mathbb{T}^{\mathcal{D},u_{t-2}}_{t-2} \circ \cdots \circ \mathbb{T}^{\mathcal{D},u_0}_0(b_0) = \bigcup_{u_{0:t-1}\in\mathbb{U}_{0:t-1}} \mathbb{T}^{\mathcal{D},u_{0:t-1}}_{0:t-1}(b_0) . \tag{5.36}$$

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 5.27, 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 5.4, in Appendix 5.A.4 states that set $\mathbb{F}_t^{\mathcal{D}}$ is a $(\overrightarrow{\mathbb{X}})$-mappings set). We hence get

$$\forall t \in \mathcal{T} \setminus \{T\} \ , \ \left|\mathbb{T}^{\mathcal{D},u_{0:t}}{}_{0:t}(b_0) \setminus \{\delta_\partial\}\right| \leq |\mathrm{supp}(b_0)| \ . \tag{5.37}$$

Finally, we obtain

$$\begin{aligned}
\left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)\right| &= \left|\bigcup_{t=1}^{T}\left(\mathbb{B}^{\mathrm{R},\mathcal{D}}_t(b_0)\right)\right| && \text{(using Equation (5.23))}\\
&\leq 1 + \left|\bigcup_{t=1}^{T}\left(\mathbb{B}^{\mathrm{R},\mathcal{D}}_t(b_0)\setminus\{\delta_\partial\}\right)\right| && \text{(by removing } \delta_\partial \text{ from } \mathbb{B}^{\mathrm{R},\mathcal{D}}_t(b_0) \text{ for all } t)\\
&= 1 + \left|\bigcup_{t=0}^{T-1}\bigcup_{u_{0:t}\in\mathbb{U}_{0:t}}\left(\mathbb{T}^{\mathcal{D},u_{0:t}}_{0:t}(b_0)\setminus\{\delta_\partial\}\right)\right| && \text{(using Equation (5.36))}\\
&\leq 1 + \sum_{t=0}^{T-1}\sum_{u_{0:t}\in\mathbb{U}_{0:t}}\left|\left(\mathbb{T}^{\mathcal{D},u_{0:t}}_{0:t}(b_0)\setminus\{\delta_\partial\}\right)\right| && \text{(as } |A\cup B| \leq |A|+|B|)\\
&\leq 1 + \sum_{t=0}^{T-1}\sum_{u_{0:t}\in\mathbb{U}_{0:t}}|\mathrm{supp}(b_0)| && \text{(using Equation (5.37))}\\
&\leq 1 + \sum_{t=0}^{T-1}|\mathbb{U}|^{t+1}|\mathrm{supp}(b_0)| && \text{(as } \mathbb{U}_{0:t} = \mathbb{U}^{t+1})\\
&\leq 1 + |\mathbb{U}|\left(\frac{|\mathbb{U}|^T-1}{|\mathbb{U}|-1}\right)|\mathrm{supp}(b_0)| && \text{(as } \textstyle\sum_{i=0}^N x^i = \frac{x^{N+1}-1}{x-1} \text{ for } x\neq 1)\\
&\leq 1 + |\mathbb{U}|^{|\mathcal{T}|}|\mathrm{supp}(b_0)| \ . && \text{(as } |\mathcal{T}| = T+1 \text{ and } |\mathbb{U}|\geq 2)
\end{aligned}$$

We have established the Inequality (5.35) and this concludes the proof. □

A direct consequence of Proposition 5.7 and Theorem 5.10 is that the complexity of Algorithm 4 is $O\big(|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}||\mathcal{T}||\mathbb{U}||\mathbb{O}|\big)$, i.e.
in $O\left(\min\left((1+|\mathbb{X}|)^{|\mathrm{supp}(b_0)|}, 1+|\mathrm{supp}(b_0)||\mathbb{U}|^{|\mathcal{T}|}\right)|\mathcal{T}||\mathbb{U}||\mathbb{O}|\right)$.

As a side note, we can remark that we could also use Theorem 5.10 to characterize the complexity of 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, i.e., $\mathbb{X}' = \mathbb{X}\times\mathbb{V}^{T+1}\times\mathbb{W}^T$, or $x' = (x, v_0, \dots v_T, w_0, \dots, w_{T-1})$. We model the problem as though the realization of the noises are predetermined, but the decision maker does not know the noises in advance. For all time $t \in \mathcal{T} \setminus \{T\}$, the new evolution functions are then $f'_t : \mathbb{X}' \times \mathbb{U} \to \mathbb{X}', (x', u) \mapsto \big(f_t(x, u, w_t), x'_{|\mathbb{V}^{T+1} \times \mathbb{W}^T}\big)$, while for all time $t \in \mathcal{T}$ the observation functions are $h'_t : \mathbb{X}' \times \mathbb{U} \to \mathbb{O}, (x', u) \mapsto h_t(x, u, v_t)$. We can hence obtain a Det-Pomdp, with states $\mathbb{X}'$, controls $\mathbb{U}$ and observations $\mathbb{O}$, with evolution functions $\{f'_t\}_{t \in \mathcal{T} \setminus \{T\}}$, observation functions $\{h'_t\}_{t \in \mathcal{T}}$ and the same admissibility constraints $\{\mathcal{U}^{ad}_t\}_{t \in \mathcal{T} \setminus \{T\}}$ as the initial Pomdp. However, such reformulation leads to a drastic increase in the dimension of the states set and the cardinality of the initial belief. Indeed, the initial belief contains all possible realizations of the initial state and all the possible noises, i.e. its cardinality is multiplied by a factor $|\mathbb{V}|^{T+1} \times |\mathbb{W}|^{T}$. Hence, we are doubly penalized when considering the bound presented in Theorem 5.10: we both increase $|\mathbb{X}|$ and $|\mathrm{supp}(b_0)|$. This reinforces the point on the difficulty of solving Pomdp as even ones with simple structures are far more difficult than Det-Pomdp.

We now present the subclass Separated Det-Pomdp, which is simpler than Det-Pomdp.

5.5 Separated Deterministic Partially Observed Markov Decision Processes

In this section, we introduce a subclass of Det-Pomdps: Separated Det-Pomdps. First, we define this subclass in §5.5.1. Second, in §5.5.2, we present an improved bound on the cardinality of the set of reachable beliefs for Separated Det-Pomdps compared to Det-Pomdps.

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

Let us first define separated mapping sets.

Definition 5.11. *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, the evaluation mappings on set $\mathbb{G}$ (i.e. the mappings $\mathbb{G} \to \mathbb{Y}_2, g \mapsto g(y)$, for $y \in \mathbb{Y}_1$) are injective. For example, let $\mathbb{Y}_1 = [\![1, n]\!]$ and $\mathbb{Y}_2 = \mathbb{R}$. Then, $\mathbb{G}$ is identified with $G \subset \mathbb{R}^n$. $\mathbb{G}$ 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 (5.8), we want to extend this notion of separated mapping set to tackle the added point ∂ differently.

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 5.12. *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 5.13. *A Separated* Det-Pomdp *is a* Det-Pomdp *such that the set of mappings $\mathbb{F}^{\mathcal{D}}$ defined in Equation (5.30) 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 ∂. This is illustrated in Figure 5.4 which represents a pair F (in blue) and F' (in red and dashed) of (∂)-separated mapping. Indeed, for a given state $x \in \overline{\mathbb{X}}$, there are four possibilities:

- $F(x) = F'(x)$, when $x \in X_1$,
- $F(x) = \partial$ and $F'(x) \neq \partial$, when $x \in X_2$,
- $F(x) \neq \partial$ and $F'(x) = \partial$, when $x \in X_3$,
- $F(x) = F'(x) = \partial$, when $x \in X_4$.

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 5.14. *If the set $\bigcup_{t \in \mathcal{T} \setminus \{T\}} f_{0:t}^{\mathbb{U}^{t+1}} = \{f_{0:t}^{u_{0:t}} \mid \forall t \in \mathcal{T} \setminus \{T\}, \forall u_{0:t} \in \mathbb{U}^{t+1}\}$ of the composition of the evolution functions of Problem (5.2) is a separated mapping set, then Problem (5.2) is a Separated* Det-Pomdp.

Proof. The proof of Proposition 5.14 is a direct consequence of Corollary 5.30. The detailed proof is found in Appendix 5.A.5. □

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

5.5.2 Bound on the cardinality of the set of reachable beliefs for Separated Det-Pomdps

We now present the main interest of Separated Det-Pomdp 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 $1 + \big(2^{|\mathrm{supp}(b_0)|} - |\mathrm{supp}(b_0)|\big)|\mathbb{X}|$

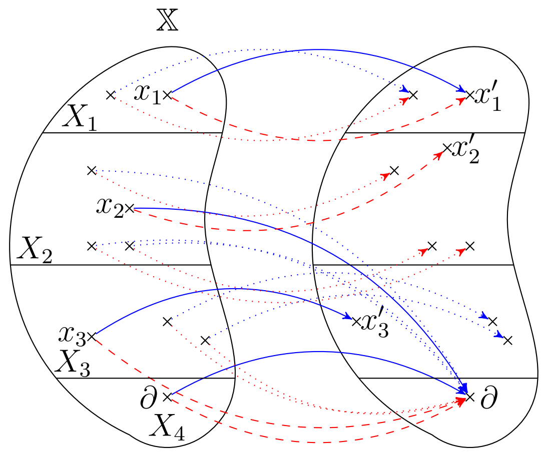


FIGURE 5.4: Illustration of two (∂)-separated mappings. The arrows represent two different mappings F (in blue) and F' (in red and dashed). On the subset $X_1 \subset \mathbb{X}$, F and F' coincide. On the subset X_2, we have $F(X_2) = \{\partial\}$. On the subset X_3, we have $F'(X_3) = \{\partial\}$, and finally, on subset X_4, we have $F(X_4) = F'(X_4) = \{\partial\}$.

Theorem 5.15. *Consider a Separated* DET-POMDP *optimization problem given by Problem (5.2) which satisfies the finite sets Assumption 5.1. 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*

$$\left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)\right| \leq 1 + \left(2^{|\mathrm{supp}(b_0)|} - |\mathrm{supp}(b_0)|\right)|\mathbb{X}| \; . \tag{5.38}$$

Proof. We start by giving preliminary bounds on $\left|\left(\mathcal{R} \circ (\mathbb{F}^{\mathcal{D}}_{X\to\mathbb{X}})_\star\right)(b_0) \setminus \{\delta_\partial\}\right|$, where $\mathbb{F}^{\mathcal{D}}_{X\to\mathbb{X}}$ is defined by Equation (5.113) (in Appendix 5.A.5.1), i.e.

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

where $\mathbb{F}^{\mathcal{D}}$ is defined in Equation (5.29). We consider three cases depending on the cardinality of the subset X:

1. When $|X| = 0$, we have that $X = \emptyset$ and $\left(\mathcal{R} \circ (\mathbb{F}^{\mathcal{D}}_{\emptyset\to\mathbb{X}})_\star\right)(b_0) \setminus \{\delta_\partial\} = \emptyset$, and thus
$$\left|\left(\mathcal{R} \circ (\mathbb{F}^{\mathcal{D}}_{X\to\mathbb{X}})_\star\right)(b_0) \setminus \{\delta_\partial\}\right| = 0 \; . \tag{5.39a}$$

2. When $|X| = 1$, we have that $\left(\mathcal{R} \circ (\mathbb{F}^{\mathcal{D}}_{X\to\mathbb{X}})_\star\right)(b_0) \setminus \{\delta_\partial\} \subset \left\{\delta_x \,\middle|\, x \in \mathbb{X}\right\}$, as the only probability distributions of $\Delta(\mathbb{X})$ which support is of cardinality at most 1 are the vertices of the simplex $\Delta(\mathbb{X})$, and thus
$$\left|\left(\mathcal{R} \circ (\mathbb{F}^{\mathcal{D}}_{X\to\mathbb{X}})_\star\right)(b_0) \setminus \{\delta_\partial\}\right| \leq \left|\left\{\delta_x \,\middle|\, x \in \mathbb{X}\right\}\right| = |\mathbb{X}| \; . \tag{5.39b}$$

3. For $|X| \geq 2$ we have by Lemma 5.32 in Appendix 5.A.4, applied with $\mathbb{G} = \mathbb{F}$ (as $\mathbb{F}$ is a (∂)-separated mapping set) that

$$\left|\big(\mathcal{R} \circ (\mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}})_{\star}\big)(b_0) \setminus \{\delta_\partial\}\right| \leq \left|(\mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}})_{\star}\right| \leq |\mathbb{X}| \; . \tag{5.39c}$$

Now consider the set $\mathbb{T}^{\mathcal{D}}$ defined by

$$\mathbb{T}^{\mathcal{D}} = \bigcup_{t \in \mathcal{T} \setminus \{T\}} \mathbb{T}^{\mathcal{D}}_{0:t} = \mathcal{R} \circ (\mathbb{F}^{\mathcal{D}})_{\star} \; . \tag{5.40}$$

We have

$$\left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)\right| \overset{(5.23)}{=} \left|\bigcup_{t=1}^{T} \mathbb{B}^{\mathrm{R},\mathcal{D}}_{t}(b_0)\right| \overset{(5.33)}{=} \left|\bigcup_{t=0}^{T-1} \mathbb{T}^{\mathcal{D}}_{0:t}(b_0)\right| \overset{(5.40)}{=} |\mathbb{T}^{\mathcal{D}}(b_0)|$$

We now detail the cardinality of $\mathbb{T}^{\mathcal{D}}(b_0)$:

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

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

$$= \left|\bigcup_{X \subset \mathrm{supp}(b_0)} \big(\mathcal{R} \circ (\mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}})_{\star}\big)(b_0) \setminus \{\delta_\partial\}\right| \tag{5.41}$$

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 (5.89) in Lemma 5.22,

$$\begin{aligned}
&\leq \sum_{X \subset \mathrm{supp}(b_0)} \left|\big(\mathcal{R} \circ (\mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}})_{\star}\big)(b_0) \setminus \{\delta_\partial\}\right| \\
&= \sum_{k \geq 0} \sum_{\substack{X \subset \mathrm{supp}(b_0) \\ |X| = k}} \left|\big(\mathcal{R} \circ (\mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}})_{\star}\big)(b_0) \setminus \{\delta_\partial\}\right| && (5.42) \\
&\leq |\mathbb{X}| + \sum_{\substack{X \subset \mathrm{supp}(b_0) \\ |X| \geq 2}} |\mathbb{X}| && (\text{by Equations (5.39)}) \\
&= |\mathbb{X}| + \big(2^{|\mathrm{supp}(b_0)|} - |\mathrm{supp}(b_0)| - 1\big)|\mathbb{X}| \; , && (5.43)
\end{aligned}$$

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

$$\begin{aligned}\big|\{X \subset \mathrm{supp}(b_0) \,|\, |X| \geq 2\}\big| = & \underbrace{\Big|\Big\{X \subset \mathbb{X} \,\Big|\, X \subset \mathrm{supp}(b_0)\Big\}\Big|}_{2^{|\mathrm{supp}(b_0)|}} - \\ & \underbrace{\Big|\Big\{X \subset \mathrm{supp}(b_0) \,\Big|\, |X| = 1\Big\}\Big|}_{=|\mathrm{supp}(b_0)|} - \\ & \underbrace{\Big|\Big\{X \subset \mathrm{supp}(b_0) \,\Big|\, |X| = 0\Big\}\Big|}_{=1} .\end{aligned}$$

We hence obtain that

$$\big|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)\big| \overset{(5.40)}{=} |\mathbb{T}^{\mathcal{D}}(b_0)| \overset{(5.43)}{\leq} 1 + \big(2^{|\mathrm{supp}(b_0)|} - |\mathrm{supp}(b_0)|\big)|\mathbb{X}| .$$

This ends the proof. ☐

We have therefore improved the complexity of Algorithm 4 for Separated Det-Pomdp.

Corollary 5.16. *Consider a Separated* Det-Pomdp *optimization problem given by Problem* (5.2) *which satisfies the finite sets Assumption 5.1. Then Algorithm 4 numerically solves Problem* (5.2) *by Dynamic Programming and its complexity is*

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

Proof. By Proposition 5.7, Algorithm 4 solves Problem (5.2)and its complexity is $O\big(|\mathbb{T}||\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)||\mathbb{U}||\mathbb{O}|\big)$. Then, by Theorem 5.15, we have $\big|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)\big| \leq 1 + \big(2^{|\mathrm{supp}(b_0)|} - |\mathrm{supp}(b_0)|\big)|\mathbb{X}|$, and by Theorem 5.10 we have,

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

☐

As the bound presented in Theorem 5.15 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 $\{f_t\}_{t\in\mathcal{T}}$.

Corollary 5.17. *Assuming that Problem* (5.2) *is a Separated* Det-Pomdp*, that Assumption 5.1 holds, that* $|\mathrm{supp}(b_0)| > 1$*, that the evolution functions* $\{f_t\}_{t\in\mathcal{T}\setminus\{T\}}$ *of Problem* (5.2) *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 5.15 can be improved as*

$$\big|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)\big| \leq 1 + \big(2^{|\mathrm{supp}(b_0)|} - |\mathrm{supp}(b_0)|\big)|A| . \tag{5.44}$$

Proof. Let's assume that there exists $A \subset \mathbb{X}$ such that for all time $t \in \mathcal{T} \setminus \{T\}$, $f_t(A, \mathbb{U}) \subset A$, and that $\text{supp}(b_0) \subset A$.

Then, for all time $t \in \mathcal{T} \setminus \{T\}$, for all sequence of controls $(u_{0:t}) \in \mathbb{U}^{t+1}$, we have that $f_{0:t}^{u_{0:t}}(\text{supp}(b_0)) \subset A$.

Thus,

$$\forall b \in \mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0) \ , \ \ \text{supp}(b) \subset A \ . \tag{5.45}$$

We can therefore consider a new Det-Pomdp, where the set of states is restricted to A, all other things being equal to the initial Det-Pomdp. In other words, we consider a Det-Pomdp defined by the tuple $\mathcal{D}_{|A}$, given by

$$\mathcal{D}_{|A} = \big(\mathcal{T}, \mathbb{U}, \mathbb{O}, A, \{\mathcal{L}_t\}_{t\in\mathcal{T}}, \{f_{t|A}\}_{t\in\mathcal{T}}, \{\mathcal{U}^{ad}_{t|A}\}_{t\in\mathcal{T}}, \{h_{t|A}\}_{t\in\mathcal{T}}\big) \ .$$

We denote by $\mathbb{B}^{\mathrm{R},\mathcal{D}'}(b_{0|A})$ the set of reachable beliefs of this new Det-Pomdp.

By Equation (5.45), there exists a bijection between $\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)$ and $\mathbb{B}^{\mathrm{R},\mathcal{D}_{|A}}_{[\![1,T]\!]}(b_{0|A})$. We hence have

$$|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)| = |\mathbb{B}^{\mathrm{R},\mathcal{D}_{|A}}_{[\![1,T]\!]}(b_{0|A})| \ .$$

Moreover, the new Det-Pomdp (defined by $\mathcal{D}_{|A}$) is also a Separated Det-Pomdp as the dynamics and observation functions stayed the same as in $\mathcal{D}$ (up to a restriction to subset A). As the new Det-Pomdp is also a Separated Det-Pomdp, by applying Theorem 5.15, we have

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

□

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

5.5.3 Existence of Separated Det-Pomdp where the bound is reached

In Theorem 5.15, we have given an improved bound on the cardinality of the set of reachable beliefs for Separated Det-Pomdp. We now prove that the bound can be tight.

Proposition 5.18. *There exist Separated* Det-Pomdp*s such that equality is obtained in Equation* (5.38)*, that is,*

$$\big|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![1,T]\!]}(b_0)\big| = 1 + \big(2^{|\text{supp}(b_0)|} - |\text{supp}(b_0)|\big)|\mathbb{X}| \ . \tag{5.46}$$

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 (5.46). Following the framework of §5.1.1, let:

- $\mathbb{X} = \{x_1, x_2, x_3\}$ three distinct states,
- $\mathbb{O} = \{\bar{o}_1, \bar{o}_2\}$ two distinct observations,

- $\mathbb{U} = \{\bar{u}_1, \bar{u}_2\}$ two distinct controls,
- $\forall x \in \mathbb{X}\ ,\ \ f(x, \bar{u}_1) = x$,
- $\forall i \in \{1, 2, 3\}\ ,\ \ f(x_i, \bar{u}_2) = x_{\mathrm{mod}(i,3)+1}$, where $\mathrm{mod}(i, 3)$ is the remainder of the euclidean division of i by 3,
- $h(x, u) = \begin{cases} \bar{o}_2 & \text{if } x = x_3 \text{ and } u = \bar{u}_1\ , \\ \bar{o}_1 & \text{otherwise}\ . \end{cases}$

We illustrate the mappings $F^{(u,o)}$ defined in Equation (5.18) for this simple case in Figure 5.5, and we illustrate the dynamics and observation functions in Figure 5.6.

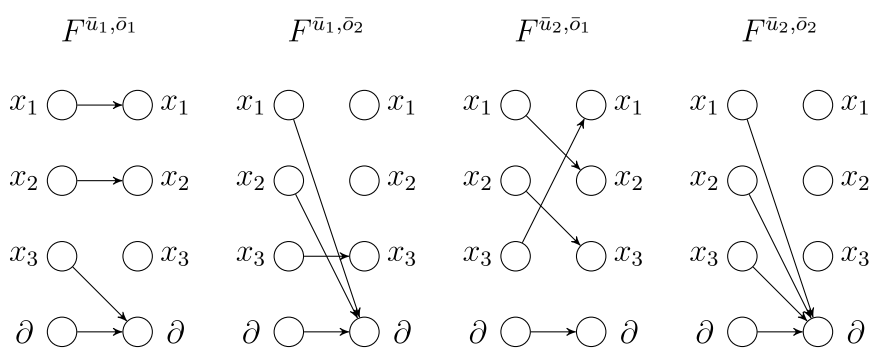


FIGURE 5.5: Representation of the $F^{(u,o)}$ functions in the simple case of §5.5.3

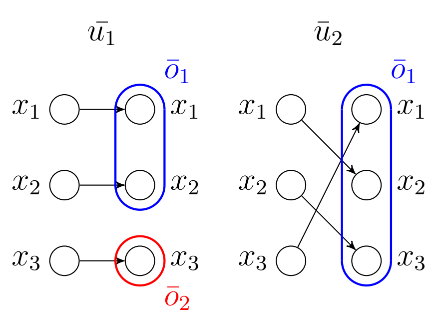


FIGURE 5.6: Representation of the dynamics and the observations depending on the control of the simple case of §5.5.3

By adding a given cost function $\mathcal{L}$, an 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 §5.1.1), where Assumption 5.1 holds.

We now prove that the resulting problem is a Separated DET-POMDP. For that purpose, we enumerate all the possible results of the dynamics before applying Proposition 5.14. For this purpose, let us consider a sequence of controls $(u_1, \dots, u_t) \in \mathbb{U}^t$. By using the composition notation presented in Equation (5.7) on the dynamics (i.e. $f^{u_{1:t}}(x) = f^{u_t} \circ \cdots \circ f^{u_1}(x)$), we have

$$\forall i \in [\![1,3]\!] \ , \ \ f^{u_{1:t}}(x_i) = x_{\mathrm{mod}(i+\gamma(u_{1:t})-1,3)+1} \ ,$$

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

$$\gamma : \mathbb{U}^t \to \mathbb{N} \ , \ \ u_{1:t} \mapsto \Big|\{u_i, i \in [\![1,t]\!] \,|\, u_i = \bar{u}_2\}\Big| \ .$$

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 $\cup_{t \in \mathcal{T}\setminus\{T\}} f_{0:t}^{\mathbb{U}^{t+1}}$ is a separated mapping set. By Proposition 5.14, the optimization problem is hence a Separated Det-Pomdp.

We now present an initial state and the resulting reachable beliefs, and compare its cardinality to the bound of Equation (5.38). Let us consider $\mathrm{supp}(b_0) = \{x_1, x_2\}$. We can apply Theorem 5.15 with such initial belief. Therefore, according to Equation (5.38), there can be at most 7 reachable beliefs (including δ_∂). We now enumerate the possible supports of the reachable beliefs when starting with b_0:

- $\{x_1, x_2\}$, obtained after applying u_1 and observing o_1 on belief b_0:

$$\mathrm{supp}\Big(\big(F^{\bar{u}_1,\bar{o}_1}\big)_\star(b_0)\Big) = \{x_1, x_2\} \ .$$

- $\{x_2, x_3\}$, obtained after applying u_2 and observing o_1 on belief b_0:

$$\mathrm{supp}\Big(\big(F^{\bar{u}_2,\bar{o}_1}\big)_\star(b_0)\Big) = \{x_2, x_3\} \ .$$

- $\{x_3, x_1\}$, obtained after applying u_2 and observing o_1 twice on belief b_0:

$$\mathrm{supp}\Big(\big(F^{\bar{u}_2,\bar{o}_1} \circ F^{\bar{u}_2,\bar{o}_1}\big)_\star(b_0)\Big) = \{x_3, x_1\} \ .$$

- $\{x_3\}$, obtained after applying u_2 and observing o_1, then applying u_1 and observing o_2 on belief b_0:

$$\mathrm{supp}\Big(\big(F^{\bar{u}_1,\bar{o}_2} \circ F^{\bar{u}_2,\bar{o}_1}\big)_\star(b_0)\Big) = \{x_3\} \ .$$

- $\{x_1\}$, obtained after applying u_2 and observing o_1, applying u_1 and observing o_2, then applying u_2 and observing o_1 on belief b_0:

$$\mathrm{supp}\Big(\big(F^{\bar{u}_2,\bar{o}_1} \circ F^{\bar{u}_1,\bar{o}_2} \circ F^{\bar{u}_2,\bar{o}_1}\big)_\star(b_0)\Big) = \{x_1\} \ .$$

- $\{x_2\}$, obtained after applying u_2 and observing o_1, applying u_1 and observing o_2, then applying u_2 and observing o_1 twice on belief b_0:

$$\mathrm{supp}\Big(\big(F^{\bar{u}_2,\bar{o}_1} \circ F^{\bar{u}_2,\bar{o}_1} \circ F^{\bar{u}_1,\bar{o}_2} \circ F^{\bar{u}_2,\bar{o}_1}\big)_\star(b_0)\Big) = \{x_2\} \ .$$

- $\{\partial\}$, obtained after applying u_1 and observing o_2 on belief b_0:

$$\mathrm{supp}\Big(\big(F^{\bar{u}_1,\bar{o}_2}\big)_\star(b_0)\Big) = \{\partial\} \ .$$

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 (5.38) states that there can be at most 7 reachable beliefs, we obtain that we have exactly 7 reachable beliefs and thus Equation (5.46) is obtained. □

Now that we have presented the subclass Separated DET-POMDP, we give a numerical illustration of this subclass.

5.6 An 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 5.7. 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. He/she only knows that the volume of water in the tank is between two quantized levels.

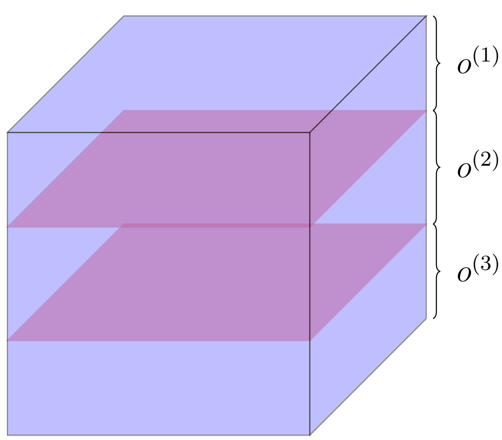


FIGURE 5.7: Illustration of the water tank "quantum" of observation

5.6.1 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.
- 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.
- The controls 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.
- The unitary price of water at each time $t \in \mathcal{T} \setminus \{T\}$ is given by $c_t \in \mathbb{R}$.

Optimization problem. We now adapt the Problem (5.2) to the tank case presented above:

$$\min_{\boldsymbol{X},\boldsymbol{U},\boldsymbol{O}} \mathbb{E}\Big[\sum_{t=0}^{T-1} c_t \boldsymbol{U}_t\Big] \tag{5.47a}$$

$$s.t.\ \mathbb{P}_{\boldsymbol{X}_0} = b_0 \ , \tag{5.47b}$$

$$\boldsymbol{X}_{t+1} = \boldsymbol{X}_t - \boldsymbol{U}_t \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , \tag{5.47c}$$

$$\boldsymbol{U}_t \in \{u^{(i)} \in \mathbb{U} \,|\, u^{(i)} \leq \boldsymbol{X}_t\} \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , \tag{5.47d}$$

$$\boldsymbol{O}_t = \max\{o^{(j)} \in \mathbb{O} \,|\, \boldsymbol{X}_t \geq o^{(j)}\} \ , \ \ \forall t \in \mathcal{T} \ , \tag{5.47e}$$

$$\sigma(\boldsymbol{U}_t) \subset \sigma\left(\boldsymbol{O}_0, \dots, \boldsymbol{O}_t, \boldsymbol{U}_0, \dots, \boldsymbol{U}_{t-1}\right) \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ . \tag{5.47f}$$

Equation (5.47a) represents the objective function of the tank problem, i.e. the implementation of Equation (5.2a) of Problem (5.2). 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 (5.47c).

The observation function h is given by a piecewise constant function which does not depend on the controls u: $h(x) = \max\{o^{(i)} \,|\, x \geq o^{(i)}\}$. We note $[\underline{o}, \overline{o}]$, $(\underline{o}, \overline{o}) \in \mathbb{O}^2$, the interval such that the states are compatible with the observations o_t, i.e.

$$[\underline{o}, \overline{o}] = \{x \in \mathbb{X} \,|\, h(x) = o\} \ . \tag{5.48}$$

This leads to equation (5.47e), which is the implementation of (5.2e).

The admissibility set of the tank problem is given by $\mathcal{U}^{ad}(\boldsymbol{X}_t) = [0, \boldsymbol{X}_t]$ (see Equation (5.47d)). 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 (5.47) has the same form as Problem (5.2). It is therefore a Det-Pomdp and all the relevant results presented in §5.2 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 (5.9). As the evolution functions and observation functions are stationary, the belief dynamics are also stationary.

By Equation (5.47c), we have $(f^u)^{-1}(y) = y + u$. Moreover, we have by the definition of $\underline{o}$ and $\overline{o}$, in Equation (5.48), that $\big(h^u\big)^{-1}(o) = [\underline{o}, \overline{o}]$. Hence, the function Q in (5.10) is here

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

and Equation (5.14) gives

$$\tau(b, u, o)(y) = \begin{cases} \dfrac{b(y+u)}{\displaystyle\sum_{x' \in [\underline{o}-u, \overline{o}-u]} b(x')} & \text{if} \quad y \in [\underline{o}, \overline{o}]\ , \\ 0 & \text{if} \quad y \notin [\underline{o}, \overline{o}] \end{cases} .$$

(where $\underline{o}$ and $\overline{o}$ are defined in Equation (5.48).)

Bellman equations for the partially observed tank problem. As Problem (5.47) is a Det-Pomdp and the finite sets Assumption 5.1 holds, we can apply Proposition 5.2. Equations (5.16a) and (5.16b) are here

$$V_T : \mathbb{B}_T^{\mathrm{R},\mathcal{D}}(b_0) \to \mathbb{R}\ , \quad b \mapsto 0 \tag{5.49a}$$

$$V_t : \mathbb{B}_t^{\mathrm{R},\mathcal{D}}(b_0) \to \mathbb{R}\ , \quad b \mapsto \min_{u \le \min_{x \in \mathrm{supp}(b)} x} \Big(c_t u + \sum_{o \in \mathbb{O}} \sum_{x-u \in [\underline{o}, \overline{o}]} b(x) V_{t+1}\big(\tau(b, u, o)\big) \Big)\ . \tag{5.49b}$$

Indeed, the intersection $\mathcal{U}_t^{\mathrm{b}}(b) = \bigcap_{x \in \mathrm{supp}(b)} \mathcal{U}_t^{ad}(x)$ defined in Equation (5.16d) is

$$\{u^{(i)} \in \mathbb{U} \,|\, u \le \min_{x \in \mathrm{supp}(b)} x\}\ ,$$

as the admissibility set is given by Equation (5.47d), and

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

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 5.33, present in Appendix 5.A.5. Indeed, Corollary 5.33 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.

5.6.2 Numerical applications

We now present some numerical results for the tank problem described by Problem (5.47).

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

- $\mathbb{X} = [\![0, 300]\!]$,
- $\mathbb{U} = [\![0, 9]\!]$,
- $\mathbb{O} = \{0, 1, 20, 40, 60, 80, 100, 120, 140, 160, 180, 200, 220, 240, 260, 280, 300\}$,
- $\mathcal{T} = [\![0, 100]\!]$,
- $\mathrm{supp}(b_0) = [\![260, 300]\!]$, with a randomly generated probability distribution over that support. The distribution used is detailed in Figure 5.8.

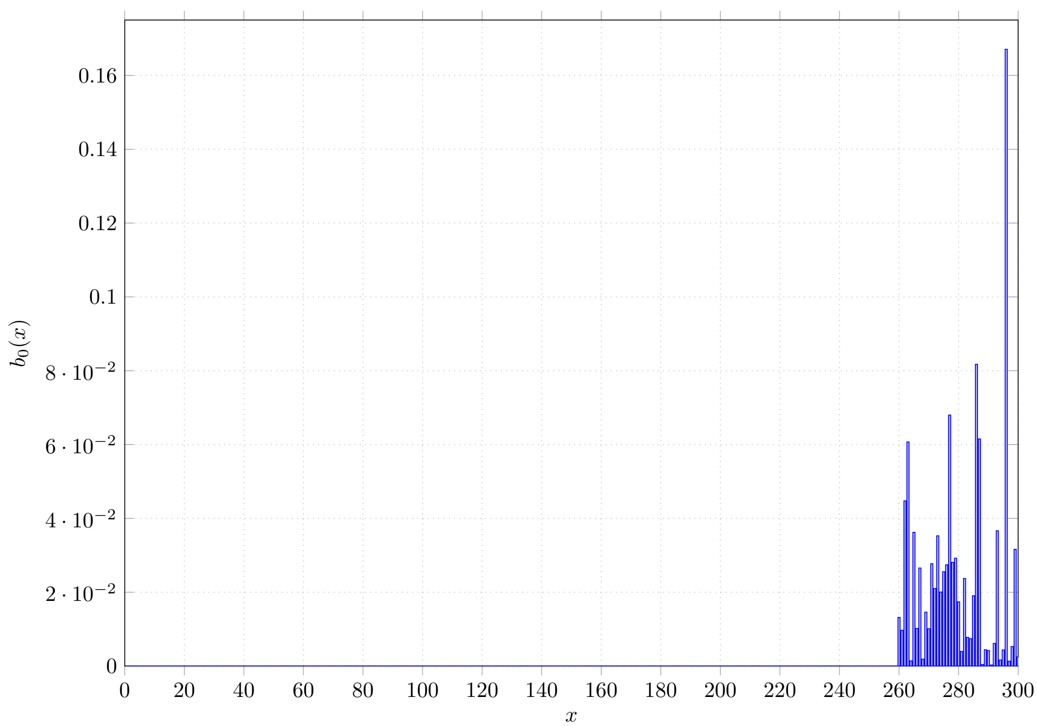


FIGURE 5.8: Probability distribution used as the initial belief b_0 for the numerical applications

When considering the initial belief b_0 presented in Figure 5.8 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 5.9.

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:

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

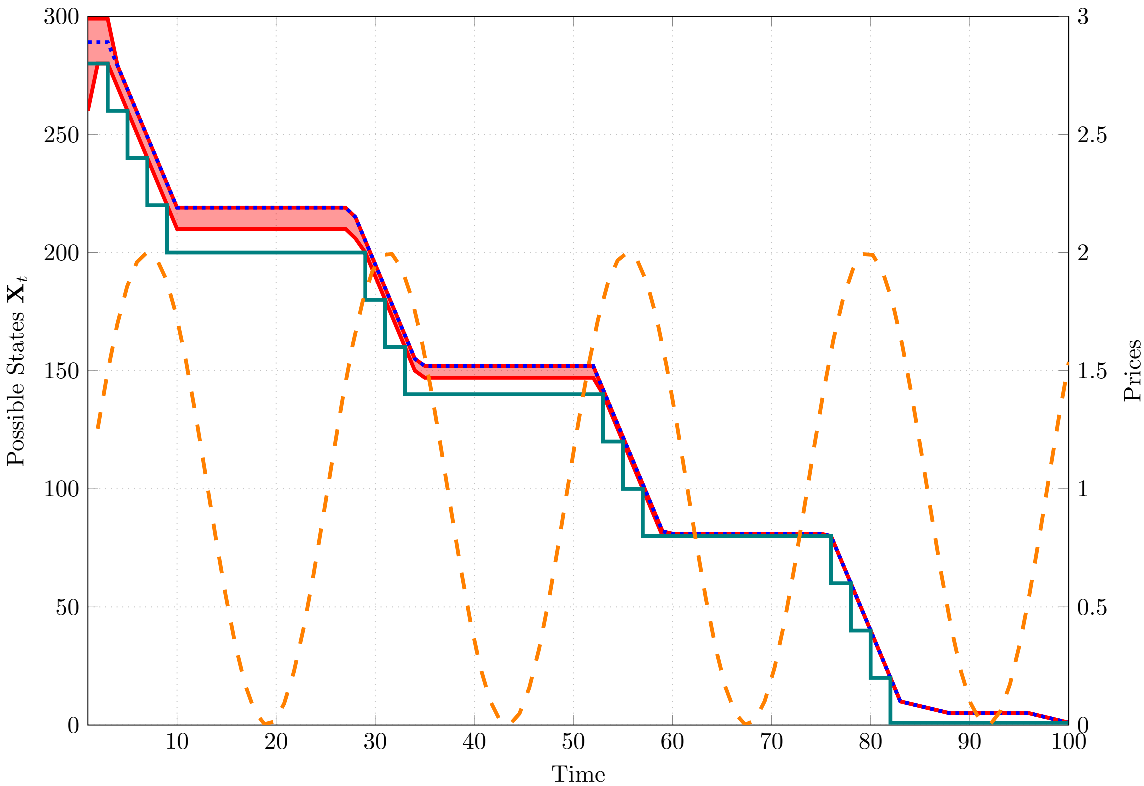


FIGURE 5.9: Representation of a trajectory of the volume of water in the tank when applying the optimal controls when considering the first set of observations. A vertical slice at time t of the red area represents the support of the belief held at time t, the dotted blue curve represents the trajectory 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

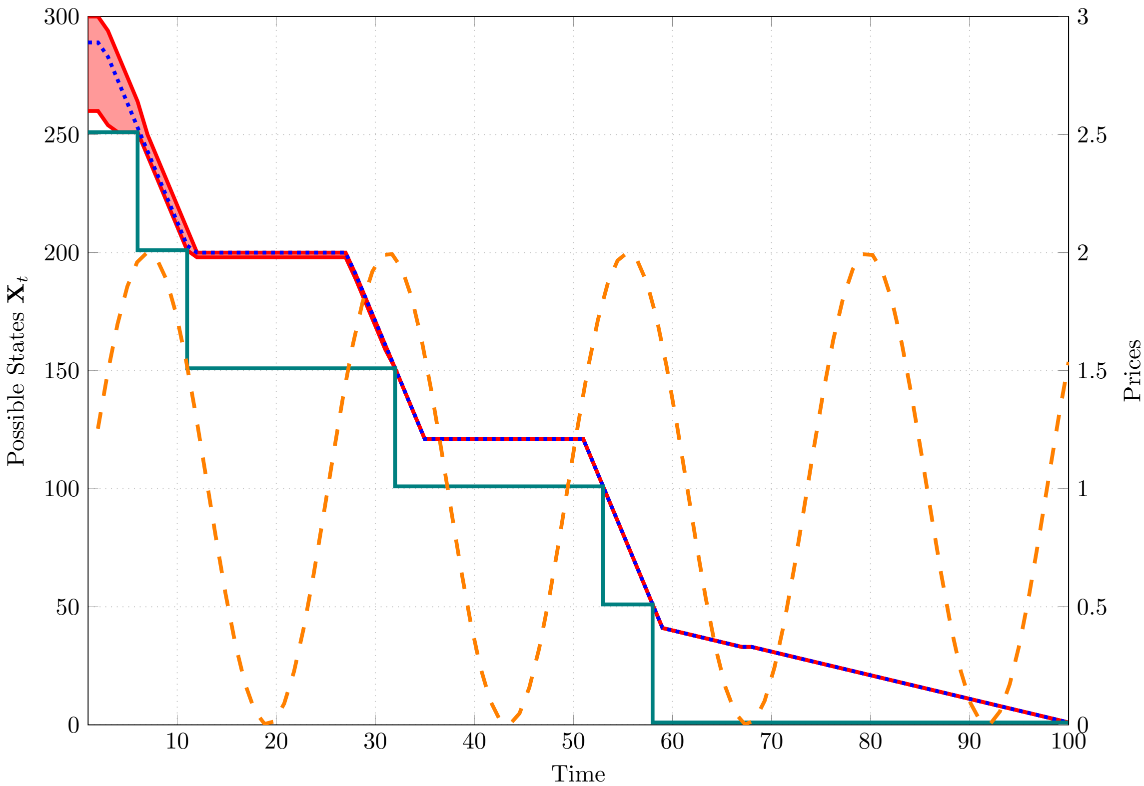


FIGURE 5.10: Representation of a trajectory of the volume of water in the tank when applying the optimal controls when considering the second set of observations. A vertical slice at time t of the red area represents the support of the belief held at time t, the dotted blue curve represents the trajectory 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

When considering the same initial belief and initial state, we obtained a trajectory represented in Figure 5.10.

Figures 5.9 and 5.10 both illustrate some properties of DET-POMDPs. First, in both cases, we see that the support of the beliefs decreases with time.
Second, we remark that such a decrease is due to the observations. Indeed, in Problem (5.47), 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 (5.48)). Thus, the supports of the beliefs are reduced along the limit of those intervals, as is more easily seen in Figure 5.10 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 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 $|\mathrm{supp}(b)| = 1$) much sooner in Figure 5.10 compared to Figure 5.9 (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 Algorithm 4 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 performed 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 5.1.

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 Algorithm 4 found the same value, hence SARSOP indeed converged toward the optimal solution of Problem (5.47).

$\lvert\mathbb{X}\rvert$	$\lvert\mathbb{U}\rvert$	$\lvert\mathbb{O}\rvert$	$\lvert\mathrm{supp}(b_0)\rvert$	T	SARSOP computation time (s)	Algorithm 4 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 5.1: Computation time of different instances of both SARSOP and Algorithm 4

5.7 Conclusion

In this chapter, 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, 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 $\big(1 + \lvert\mathbb{X}\rvert\big)^{\lvert\mathrm{supp}(b_0)\rvert}$ (in the case of Det-Pomdp, see Theorem 5.9) to $2^{\lvert\mathrm{supp}(b_0)\rvert}\lvert\mathbb{X}\rvert$ (Theorem 5.15), as presented in Tables 5.2 and 5.3. 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 5.18).

The Separated Det-Pomdp class of problem 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
Det-Pomdp	$(1+\lvert\mathbb{X}\rvert)^{\lvert\mathbb{X}\rvert}$ [Littman, 1996]
Det-Pomdp improved bounds	$(1+\lvert\mathbb{X}\rvert)^{\lvert\mathrm{supp}(b_0)\rvert}$ Theorem 5.9
Separated Det-Pomdp	$1+\big(2^{\lvert\mathrm{supp}(b_0)\rvert}-\lvert\mathrm{supp}(b_0)\rvert\big)\lvert\mathbb{X}\rvert$ Theorem 5.15

Table 5.2: Summary of the bounds depending on the class of problem for infinite horizon problems

Class	Finite horizon bound
Det-Pomdp	$\min\big((1+\lvert\mathbb{X}\rvert)^{\lvert\mathbb{X}\rvert}, \big(\lvert\mathbb{U}\rvert\lvert\mathbb{O}\rvert\big)^{\lvert\mathcal{T}\rvert}\big)$
Det-Pomdp improved bounds	$\min\big((1+\lvert\mathbb{X}\rvert)^{\lvert\mathrm{supp}(b_0)\rvert}, 1+\lvert\mathrm{supp}(b_0)\rvert\lvert\mathbb{U}\rvert^{\lvert\mathcal{T}\rvert}\big)$ Theorem 5.10
Separated Det-Pomdp	$\min\big(1+\big(2^{\lvert\mathrm{supp}(b_0)\rvert}-\lvert\mathrm{supp}(b_0)\rvert\big)\lvert\mathbb{X}\rvert, 1+\lvert\mathrm{supp}(b_0)\rvert\lvert\mathbb{U}\rvert^{\lvert\mathcal{T}\rvert}\big)$ Corollary 5.16

Table 5.3: Summary of the bounds depending on the class of problem

5.A Appendix

First, in §5.A.1, we present background on Markov Decision Processes. Second, in §5.A.2, we present background on Partially Observed Markov Decision Processes. Third, in §5.A.3, we present lemmata used in the proof of Proposition 5.2 for applying Dynamic Programming for Det-Pomdp with constraints. Fourth, in §5.A.4, we present some technical lemmata used to present bounds on the cardinality of the sets of reachable beliefs. Finally, in §5.A.5, we present complementary results on (∂)-separated mappings sets.

5.A.1 Background on Markov Decision Processes

In this subsection, we present background on Markov Decision Processes (Mdps). First, in §5.A.1.1, we present the general formulation of a Mdp. Second, in §5.A.1.2 we present more specifically the formulation of a discrete time stochastic control problem, i.e. an optimization problem associated with a Mdp. Third, in §5.A.1.3, we present how we can use Dynamic Programming to solve the discrete time stochastic control problem.

5.A.1.1 Formulation of MDPs

MDPs, also commonly referred to as *stochastic dynamic programs*, are models commonly used for numerous applications, some of which are registered in the survey White [1993]. There is an extensive literature on MDPs, notably Puterman [1994] and Bertsekas [2000]. In MDPs, we model how 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.

Here, we focus on *finite horizon finite*-MDPs. As stated in Puterman [1994, chapter 2], a finite horizon finite-MDP can be described by a tuple

$$\left(\mathcal{T} = [\![0,T]\!], \mathbb{X}, \{\mathcal{U}_t^{ad}\}_{t\in\mathcal{T}\setminus\{T\}}, \{p_t\}_{t\in\mathcal{T}\setminus\{T\}}, \{\mathcal{L}_t\}_{t\in\mathcal{T}\setminus\{T\}}, \mathcal{K}\right)^{3},$$

where

- $\mathcal{T} = [\![0,T]\!]$ is the set of time steps for the optimization, with $T \in \mathbb{N}$ the horizon,
- $\mathbb{X}$ is the set of states of the MDP (which has a finite cardinality for finite MDP),
- $\{\mathcal{U}_t^{ad}\}_{t\in\mathcal{T}\setminus\{T\}}$ is the sets of admissible controls: for all time $t \in \mathcal{T}\setminus\{T\}$, the set-valued mappings $\mathcal{U}_t^{ad} : \mathbb{X} \rightrightarrows \mathbb{U}$ (where $\mathbb{U}$ is set of controls, with finite cardinality) provides $\mathcal{U}_t^{ad}(x)$, the set of admissible controls at time t when in state x,
- $\{p_t\}_{t\in\mathcal{T}\setminus\{T\}}$ is the set of discrete transitions kernels: for all times $t \in \mathcal{T} \setminus \{T\}$, $p_t : \mathbb{X} \times \mathbb{U} \times \mathbb{X} \to [0,1], (x,u,x') \mapsto p_t(x'|x,u)$ is the transition kernel that gives the probability of being in state x' at time $t+1$ when applying action u on state x at time t,
- $\{\mathcal{L}_t\}_{t\in\mathcal{T}}$ is the set of reward functions: for all time $t \in \mathcal{T}$, $\mathcal{L}_t : \mathbb{X}\times\mathbb{U} \to \mathbb{R}\cup\{+\infty\}$ is the reward function which defines the reward obtained when applying control u on state x at time t,
- $\mathcal{K} : \mathbb{X} \to \mathbb{R}\cup\{+\infty\}$ is the final reward, i.e. the reward at time T.

Here, we focus on the formulation of discrete time stochastic control problem.

5.A.1.2 Discrete time stochastic control formulation

We now present the formulation of a discrete time stochastic control problem. When considering an MDP, one should recall that decisions are taken sequentially. Thus, when

[3] Here, we have presented the tuple with the notations used in this chapter. However, in Puterman [1994], the tuple is written $\big(T, S, A_s, p_t(\cdot|s,a), r_t(s,a)\big)$, with T the horizon of the MDP, $r_t(s,a)$ the reward functions and $p_t(\cdot|s,a)$ the conditional probability function that determines the probability of transitioning to states s' at time $t+1$ when applying action a on state s at time t.

defining an optimization problem, we need to specify the *non-anticipativity* constraint:

$$\sigma(\boldsymbol{U}_t) \subset \sigma(\boldsymbol{X}_0, \dots, \boldsymbol{X}_t, \boldsymbol{U}_0, \dots, \boldsymbol{U}_{t-1}) \ , \quad \forall t \in \mathcal{T} \setminus \{T\} \ . \tag{5.50}$$

This constraint is used in conjunction with the admissibility constraints defined through the set $\{\mathcal{U}_t^{ad}\}_{t \in \mathcal{T}\setminus\{T\}}$:

$$\forall t \in \mathcal{T} \setminus \{T\} \ , \ \ \boldsymbol{U}_t \in \mathcal{U}^{ad}(\boldsymbol{X}_t) \quad \mathbb{P} - a.s. \tag{5.51}$$

We also need the initialization constraint which fixes the initial state x_0:

$$\boldsymbol{X}_0 = x_0 \ . \tag{5.52}$$

Finally, in order to define an optimization problem, we need to consider an optimization criterion. The most common criterion used is the minimization of the expected value:

$$\min_{\{\boldsymbol{X}_t, \boldsymbol{U}_t\}_{t \in \mathcal{T}}} \mathbb{E}\Big[\sum_{t \in \mathcal{T}\setminus\{T\}} \mathcal{L}_t(\boldsymbol{X}_t, \boldsymbol{U}_t) + \mathcal{K}(\boldsymbol{X}_T) \Big] \ , \tag{5.53}$$

which is measurable as Ω is finite. Other criteria can be used, such as optimizing the cost for the worst case scenario[4] (for example in Bonet and Geffner [2000]), the conditional value at risk (see Chow et al. [2015]) or even a chance constrained optimization of an MDP (see Delage and Mannor [2010]).

Here, we focus on the minimization of the expected value as defined in Equation (5.53). Moreover, although there are formulations of MDPs with continuous time, we focus on the finite horizon finite-MDP.

Note that the above MDP formulation involves state transition kernels. When considering the discrete time stochastic control problem, we instead take the point of view of random processes where the evolutions are given by noisy evolution functions affected by exogenous noises. In that formulation, we consider a set of evolution functions $\{f_t\}_{t \in \mathcal{T}\setminus\{T\}}$ and a set of noises $\mathbb{W}$, with distributions given by the stochastic kernels $\{\mathbb{P}_{w,t}\}_{t \in \mathcal{T}\setminus\{T\}}$. In that second formulation, we have that for all time $t \in \mathcal{T} \setminus \{T\}$, $f_t : \mathbb{X} \times \mathbb{U} \times \mathbb{W} \to \mathbb{X}$ are the evolution functions such that

$$\forall t \in \mathcal{T} \setminus \{T\} \ , \ \ \boldsymbol{X}_{t+1} = f_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_t) \ , \tag{5.54}$$

(i.e. they return the state at time $t+1$ when applying controls u on state x with noise w), while $\mathbb{P}_{w,t} : \mathbb{X} \times \mathbb{U} \times \mathbb{W} \to [0,1], (x, u, w) \mapsto \mathbb{P}_{w,t}(w|x, u)$ are the conditional probability that the realization of the noise is w when applying control u on state x at time t. This formulation still follows the framework of an MDP, as presented in §5.A.1.1.

In the rest of this document, we use the formulation based on random variables and Equation (5.54). The optimization problem of minimizing the expected value for the finite

[4]The optimization of the worst-case scenario is also called robust optimization

horizon and finite-MDP is then:

$$\min_{\{\boldsymbol{X}_t,\boldsymbol{U}_t,\boldsymbol{W}_t\}_{t\in\mathcal{T}}} \mathbb{E}\Big[\sum_{t\in\mathcal{T}\setminus\{T\}} \mathcal{L}_t(\boldsymbol{X}_t,\boldsymbol{U}_t) + \mathcal{K}(\boldsymbol{X}_T)\Big] \tag{5.55a}$$

$$s.t. \quad (5.54), (5.50), (5.51) \text{ and } (5.52) \ .$$

5.A.1.3 Dynamic Programming for MDPs

Once given an optimization criterion (in our case the minimization of the expected value), we can define an optimization problem from a given MDP. The optimization is done over *policies*, i.e. rules for choosing the controls u_t for each time t and each possible sequence of controls and states up to time t. More formally, we define for each time $t \in \mathcal{T}$ a *history vector* h_t associated with a sequence of state and controls:

$$h_0 = x_0 \ , \quad \text{and } \forall t \in \mathcal{T} \setminus \{0\} \ , \ \ h_t = (x_0, u_0, \ldots, x_{t-1}, u_{t-1}, x_t) = (h_{t-1}, u_{t-1}, x_t) \ .$$

We denote by $\{\mathbb{H}_t\}_{t\in\mathcal{T}}$ the sets of history vector at time t, recursively defined as

$$\mathbb{H}_0 = \mathbb{X}_0 \ , \quad \text{and } \forall t \in \mathcal{T} \setminus \{T\} \ , \ \ \mathbb{H}_{t+1} = \mathbb{H}_t \times \mathbb{U}_{t-1} \times \mathbb{X}_t \ , \quad \text{where } \mathbb{X}_t = \mathbb{X} \text{ and } \mathbb{U}_t = \mathbb{U} \ .$$

We hence have, for all time $t \in \mathcal{T}$, $h_t \in \mathbb{H}_t$.

A policy $\pi = \{\pi_t\}_{t\in\mathcal{T}\setminus\{T\}}$ is a set of mappings such that, for all $t \in \mathcal{T}\setminus\{T\}$, $\pi_t : \mathbb{H}_t \to \mathbb{U}$. Moreover, an *admissible policy* π is a policy that satisfies the admissibility constraints, i.e. for all $t \in \mathcal{T} \setminus \{T\}$, for all $h_t \in \mathbb{H}_t$, the control at time t, $u_t = \pi_t(h_t)$ (where $h_t = (h_{t-1}, u_{t-1}, x_t)$), satisfies the admissibility constraint, i.e. $u_t \in \mathcal{U}_t^{ad}(x_t)$. Note that, as policies are set of mappings which for each time $t \in \mathcal{T}\setminus\{T\}$ go from $\mathbb{H}_t$ to $\mathbb{U}$, their output are random variables that satisfy the non-anticipativity constraints defined in Equation (5.50) by construction.

We can solve Problem (5.55) thanks to a Dynamic Programming algorithm which consists in recursively computing mappings called *value functions*, $\{V_t\}_{t\in\mathcal{T}}$, where for all time $t \in \mathcal{T}$, V_t is a function from $\mathbb{H}_t$ to $\overline{\mathbb{R}}$. We compute those value functions thanks to backward induction. Without an additional assumption on the objective function, we need to compute those value functions for each possible realization of a history vector, which is in practice untractable as the dimension of the history vector increases with time, which is colloquially known as one version of the *curse of dimensionality*.

However, it is possible to compute optimal value functions $\{V_t\}_{t\in\mathcal{T}}$ for policies $\{\pi_t\}_{t\in\mathcal{T}\setminus\{T\}}$ from $\mathbb{X}$ to $\mathbb{U}$ under the condition that the optimization criterion is *additive in time*. An optimization criterion is *additive in time* if the cost incurred at time t accumulates over time, and if it only depends on the state at time t. Let $\min g\big((x_t, u_t)_{t\in\mathcal{T}}\big)$ be the optimization criterion (with $g : \mathbb{X}^T \times \mathbb{U}^T \to \overline{\mathbb{R}}$ the objective function). In order to apply Dynamic Programming with value functions from $\mathbb{X}$ to $\overline{\mathbb{R}}$, we suppose that there exists a sequence of functions $\{g_t : \mathbb{X}_t \times \mathbb{U}_t \to \overline{\mathbb{R}}\}_{t\in\mathcal{T}}$ such that

$$g\big((x_t, u_t)_{t\in\mathcal{T}}\big) = \sum_{t\in\mathcal{T}} g_t(x_t, u_t) \ .$$

This condition is satisfied for the minimization of the expected value, i.e. the objective function presented in Equation (5.53).

Under that condition, it is proven in [Puterman, 1994, Theorem 4.4.2, p.89] that there exists an optimal policy such that at time t it only depends on x_t, that is the last element of $h_t = (h_{t_1}, u_{t-1}, x_t)$. Hence, optimal policies contain state feedbacks policies $\pi = \{\pi^\star_t\}_{t\in\mathcal{T}\setminus\{T\}}$ are such that, for all $t \in \mathcal{T} \setminus \{T\}$, $\pi^\star_t : \mathbb{X} \to \mathbb{U}$. Moreover, those policies can be computed thanks to a Dynamic Programming algorithm (see [Puterman, 1994, Theorem 4.5.1, p.92], or [Bertsekas, 2000, Theorem 1]), i.e. by computing the values functions $\{V_t\}_{t\in\mathcal{T}}$:

$$V_T : \mathbb{X} \to \overline{\mathbb{R}}, x \mapsto \mathcal{K}(x) \ , \tag{5.56}$$

$$\forall t \in \mathcal{T} \setminus \{T\}, V_t : \mathbb{X} \to \overline{\mathbb{R}}, x \mapsto \min_{u\in\mathcal{U}^{ad}_t(x)} \mathcal{L}_t(x, u) + \sum_{x'\in\mathbb{X}} p_t(x'|x, u)V_{t+1}(x') \ . \tag{5.57}$$

The optimal state feedbacks are then given by the controls which minimize the right-hand side of Equation (5.57), for each state x and time t.

5.A.2 Background on Partially Observed Markov Decision Processes

In this subsection, we present some on Pomdp. First, in §5.A.2.1, we present the formulation of a Pomdp. Second, in §5.A.2.2, we present more specifically the formulation of a partially observed discrete time stochastic control problem, i.e. an optimization problem associated with a Pomdp. Third, in §5.A.2.3, we present how we can apply Dynamic Programming on a finite horizon optimization problem associated to a Pomdp.

5.A.2.1 Formulation of Pomdps

A Pomdp is an extension of Mdp, where actions are taken without having full knowledge of the state of the dynamic system. The agent does not fully know the state of the dynamical system, but he has access to some information on the state of the system, the observations. In Bertsekas and Shreve [1978], it is assumed that a Pomdp is defined by the following elements

(A_1) Sets of states $\mathbb{X}$, controls $\mathbb{U}$ and observations $\mathbb{O}$, all non-empty Borel spaces[5].

(A_2) A horizon T, which is a positive integer or $+\infty$.

[5] A Borel space is a measurable space which is homeomorphic with a Borel subset of the standard Borel space associated to a Polish space: see [Bertsekas and Shreve, 1978, Definition 7.7, p118]

(A_3) A discount factor $\alpha \in [0,1]$, with the added constraint that the value $\alpha = 1$ is only possible when considering finite horizon.

(A_4) Instantaneous costs $\mathcal{L}_t : \mathbb{X} \times \mathbb{U} \to \overline{\mathbb{R}}$. In the case of a finite horizon, we denote by $\mathcal{K} : \mathbb{X} \to \overline{\mathbb{R}}$ the final reward, i.e. the reward at the time horizon T.

(A_5) $\{p_t\}_{t \in \mathcal{T} \setminus \{T\}}$ is the set of Borel-measurable state transition kernels: for all time $t \in \mathcal{T} \setminus \{T\}$, $p_t : \mathbb{X} \times \mathbb{U} \times \mathcal{B}(\mathbb{X}) \to [0,1], (x, u, X') \mapsto p_t(X'|x, u)$, where $\mathcal{B}(\mathbb{X})$ denotes the Borel σ-field of $\mathbb{X}$. The quantity $p_t(X'|x, u)$ is the probability that the state at time $t+1$ is in the subset $X' \subset \mathbb{X}$ given that the state and control at time t are respectively x and u. By a slight abuse of notation, we will write the transition kernel $p_t(dx'|x, u)$.

(A_6) A set of controls constraints $\mathcal{U}^{\mathrm{I}}_t$, for each time $t \in \mathcal{T} \setminus \{T\}$. They are defined for $t \in \mathcal{T} \setminus \{T\}$, and take as input the *information vector* I_t, defined as:

$$I_0 = o_0 \ , \qquad \text{and } \forall t \in \mathcal{T} \ , \ \ t > 0 \ , \ \ I_t = (o_0, u_0, \dots, u_{t-1}, o_t) \ .$$

We denote by $\mathbb{I}_t = \mathbb{O}_0 \times \mathbb{U}_0 \times \mathbb{O}_1 \cdots \times \mathbb{U}_{t-1} \times \mathbb{O}_t$ the sets of information vectors at time t, where for all time $t \in \mathcal{T}$, $\mathbb{O}_t = \mathbb{O}$ and $\mathbb{U}_t = \mathbb{U}$. Thus, $\mathcal{U}^{\mathrm{I}}_t$ is a set-valued mapping from $\mathbb{I}_t$ to $\mathbb{U}_t$. At time $t \in \mathcal{T} \setminus \{T\}$, the controls u_t must belong to the set $\mathcal{U}^{\mathrm{I}}_t(I_t)$.

(A_7) An initial observation kernel $p_{o,0} : \mathbb{X} \times \mathcal{B}(\mathbb{O}) \to [0,1], (x, O) \mapsto p_{o,0}(O|x)$, which gives the probability that the observation at time 0 is in the subset $O \subset \mathbb{O}_0$ if the initial state is x. Once again, we also write the kernel $p_{o,0}(do|x)$.

(A_8) $\{p_{o,t}\}_{t \in \mathcal{T} \setminus \{0\}}$ is the set of Borel-measurable observation kernels: for all time $t \in \mathcal{T} \setminus \{T\}$, $p_{o,t} : \mathbb{X} \times \mathbb{U} \times \mathcal{B}(\mathbb{O}) \to [0,1], (x, u, O) \mapsto p_{o,t}(O|x, u)$, which is the probability that the observation at time t belongs to the subset $O \subset \mathbb{O}_t$ if the system is in state x at time t after the agent chose control u at time $t-1$. We also write this kernel $p_{o,t}(do|x, u)$.

In the following, we will focus on finite Pomdp, i.e. Pomdp where sets $\mathbb{X}$, $\mathbb{U}$, $\mathbb{O}$ and $\mathcal{T}$ have finite cardinality. We now present an optimization problem associated with a Pomdp.

5.A.2.2 Partially observed discrete time stochastic control formulation

We now present the formulation of a partially observed discrete time stochastic control problem. It is very similar to Problem (5.55). Statement (A_6) defines admissibility constraints of the form

$$\forall t \in \mathcal{T} \setminus \{T\} \ , \ \ \boldsymbol{U}_t \in \mathcal{U}^{ad}_t(\boldsymbol{I}_t) \quad \mathbb{P} - a.s. \tag{5.58}$$

As the agent does not know the initial state, we only suppose that

$$\mathbb{P}_{\boldsymbol{X}_0} = b_0 \ , \tag{5.59}$$

where b_0 is the probability distribution of the initial state of the system.

Finally, as with MDP, we also have a *non-anticipativity* constraint. However, it differs from Equation (5.50) as the agent does not fully know the state of the system. Instead, the agent has access to the *information vector*. Thus, the non-anticipativity constraints for POMDP are:

$$\sigma(\boldsymbol{U}_t) \subset \sigma(\underbrace{\boldsymbol{O}_0, \ldots, \boldsymbol{O}_t, \boldsymbol{U}_0, \ldots, \boldsymbol{U}_{t-1}}_{=\boldsymbol{I}_t}) \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ . \tag{5.60}$$

The last ingredient necessary to properly define an optimization problem is an optimization criterion. Once again, we focus on the minimization of the expected value in the finite horizon case:

$$\min_{\{\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{O}_t\}_{t\in\mathcal{T}}} \mathbb{E}\Big[\sum_{t\in\mathcal{T}\setminus\{T\}} \alpha^t \mathcal{L}_t(\boldsymbol{X}_t, \boldsymbol{U}_t) + \alpha^T \mathcal{K}(\boldsymbol{X}_T) \ .\Big] \tag{5.61}$$

Other criteria also exist, such as minimizing the worst-case scenario (see once again Bonet and Geffner [2000]).

Note that Statement (A_5), which defines the transition kernels is similar to the transition kernels presented in §5.A.1. Thus, the same remark applies, and we can use evolution functions $\{f_t\}_{t\in\mathcal{T}}$ with noises, thus leading to Equation (5.54). We can also apply the same reasoning to the observation kernel: Statements (A_7) and (A_8) are equivalent to another formulation, where we consider a set of observation functions $\{h_t\}_{t\in\mathcal{T}}$ and a set of observation noises $\mathbb{V}$, with distributions given by the stochastic kernels $\{\mathbb{P}_{v,t}\}_{t\in\mathcal{T}\setminus\{T\}}$. In that second formulation, we have that, for all time $t \in \mathcal{T} \setminus \{T\}$, $h_t : \mathbb{X} \times \mathbb{U} \times \mathbb{V} \to \mathbb{O}$ are the evolution functions such that

$$\boldsymbol{O}_0 = h_0(\boldsymbol{X}_0, \boldsymbol{V}_0) \ , \qquad \text{and } \forall t \in \mathcal{T} \setminus \{T\} \ , \ \ \boldsymbol{O}_{t+1} = h_t(\boldsymbol{X}_{t+1}, \boldsymbol{U}_t, \boldsymbol{V}_{t+1}) \ , \tag{5.62}$$

while $\mathbb{P}_{v,t} : \mathbb{X} \times \mathbb{U} \times \mathbb{V} \to [0,1], (x, u, w) \mapsto \mathbb{P}_{v,t}(v|x, u)$ are the conditional probability that the realization of the noise at time $t+1$ is v after obtaining state x at time $t+1$ when applying control u at time t.

We use this second formulation based on random variables and Equations (5.54) (for the dynamics) and (5.62) (for the observations). The optimization problem of minimizing the expected value for the finite horizon and finite-POMDP is then:

$$\min_{\{\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{O}_t\}_{t\in\mathcal{T}}} \mathbb{E}\Big[\sum_{t\in\mathcal{T}\setminus\{T\}} \alpha^t \mathcal{L}_t(\boldsymbol{X}_t, \boldsymbol{U}_t) + \alpha^T \mathcal{K}(\boldsymbol{X}_T)\Big] \tag{5.63a}$$

$$s.t. \quad (5.54)(5.62), (5.58), (5.59) \text{ and } (5.60) \ .$$

5.A.2.3 Dynamic Programming for Pomdps

As with Mdp, solution of Problem (5.63) are *policies*. However, in the case of Pomdp, the policies take as input a probability distribution over the set of states and the *information vector*, and not the *history vector* as the agent does not know the state of the system. Hence, a policy is a set of mappings $\{\pi_t\}_{t\in\mathcal{T}\setminus\{T\}}$, where for all time $t \in \mathcal{T}\setminus\{T\}$, $\pi_t : \Delta(\mathbb{X})\times\mathbb{I}_t \to \mathbb{U}$. Moreover, an *admissible policy* is such that for all time $t \in \mathcal{T}\setminus\{T\}$, for a given initial probability b_0 and for all possible information vector $I_t \in \mathbb{I}_t$, $\pi_t(b_0, I_t) \in \mathcal{U}_t^{\mathrm{I}}(I_t)$.

As with Mdp and the history vector, it is possible to solve an optimization problem associated with a Pomdp which follows Statements (A_1)-(A_8) presented in §5.A.2.1 by applying Dynamic Programming when the optimization criterion is additive in time. However, in order to do so, we need some technical assumptions to guarantee the measurability of the value functions. Those assumptions (called F^+ and F^- in Bertsekas and Shreve [1978]) are satisfied as long as either the positive or negative part of the instantaneous cost functions are finite (i.e. for all $t \in \mathcal{T}\setminus\{T\}$, $\mathcal{L}_t : \mathbb{X}\times\mathbb{U} \to \mathbb{R}\cup\{+\infty\}$ or $\mathcal{L}_t : \mathbb{X}\times\mathbb{U} \to \mathbb{R}\cup\{-\infty\}$). Using [Bertsekas and Shreve, 1978, Proposition 10.3 and Proposition 10.6] (or Equations (4.4)-(4.5) of Bertsekas [2000]), we hence obtain Dynamic Programming equations:

$$
\begin{aligned}
& V_T : \Delta(\mathbb{X})\times\mathbb{I}_T \to \overline{\mathbb{R}}\,, \\
& \qquad (b_0, I_T) \mapsto \mathbb{E}\big[\mathcal{K}(\boldsymbol{X}_T)|b_0, I_T\big] && (5.64)\\
\forall t \in \mathcal{T}\setminus\{T\}\,,\ & V_t : \Delta(\mathbb{X})\times\mathbb{I}_t \to \overline{\mathbb{R}}\,, \\
& \qquad (b_0, I_t) \mapsto \min_{u\in\mathcal{U}_t^{\mathrm{I}}(I_t)} \mathbb{E}\Big[\mathcal{L}_t(\boldsymbol{X}_t, u) + V_{t+1}\big(b_0, (I_t, u, \boldsymbol{O}_t)\big)|b_0, I_t, u\Big]\,. && (5.65)
\end{aligned}
$$

Theoretically, it is possible to solve a Pomdp by computing the value functions presented in Equations (5.64) and (5.65). But, since the information vector grows with each time step, such a method is not tractable in practice. However, it may happen that there exists some *statistics sufficient for controls* in order to compress the information necessary for the decision-maker to act optimally. Such statistics are mappings of the information vectors which should allow the definition of admissible policies. A formal definition of statistics sufficient for controls can be found in [Bertsekas and Shreve, 1978, Definition 10.6]. Statistics sufficient for controls with a lower dimension than the information vector are called a *reduced statistics sufficient for controls*.

Reduced statistics sufficient for controls are described in Åström [1965], where Åström propose a reformulation of a Pomdp as a fully observed Mdp with probability distributions on the set $\mathbb{X}$ as the new states. This new state represents the belief of the decision-maker has of which is the true state of the system and is hence colloquially known as *belief*. Indeed, obtaining a reduced sufficient statistics requires extra assumptions. The first assumption is that the Pomdp not only satisfies Statement (A_6), but it must satisfy the following stronger statement

(A_6') for all time $t \in \mathcal{T} \setminus \{T\}$, for all $I_t \in \mathbb{I}_t$, $\mathcal{U}_t^{\mathrm{I}}(I_t) = \mathbb{U}$.

The POMDP must also satisfy some technical assumptions ($\hat{F}^+$ or $\hat{F}^-$ in Bertsekas and Shreve [1978]). A detailed proof of this reformulation can be found in Bertsekas and Shreve [1978]. We then define a *belief*-MDP, as a MDP, the states of which are beliefs. Once again, under some technical assumptions ($\hat{F}^+$ or $\hat{F}^-$ in Bertsekas and Shreve [1978]), we can apply Dynamic Programming on the optimization problem derived from the belief-MDP, and the resulting optimal policies of the belief-MDP are also optimal for the optimization problem derived from the POMDP.

It should be noted that the main issue with the belief-MDP reformulation is that the set of beliefs is not finite, as it is the set of probability distribution over the set $\mathbb{X}$. Thus, Dynamic Programming suffers from the curse of dimensionality if we cannot restrict the beliefs to a finite subset of $\Delta(\mathbb{X})$. Therefore, heuristics were developed in order to solve POMDP for either the finite or infinite horizon cases. We can mention some algorithms such as SARSOP, presented in Kurniawati et al. [2008], POMCP, presented in Silver and Veness [2010], or DESPOT, presented in Somani et al. [2013]. SARSOP is based on sampling the bellman value functions on a subset of the belief space. To do so, it computes a subset of the reachable belief space, and add cuts in order to converge towards the subset of belief points reachable from the initial belief $b_0 \in \mathbb{B}$ under *optimal sequences* of controls. POMCP is an extension of the *Monte-Carlo tree search* algorithm to POMDPs. It both samples states from a given belief $b \in \mathbb{B}$ and samples histories using a black-box simulator of the POMDP. DESPOT is another extension of the *Monte-Carlo tree search*. However, it samples the reachable belief sets on K scenarios, and then apply Dynamic Programming on the sampled reachable belief sets.

5.A.3 Complementary results on Dynamic Programming for Det-POMDP

In this Appendix, we present the detailed proof of Proposition 5.2. The structure of the proof was presented in Figure 5.1, reproduced here in Figure 5.11.

First, we rewrite Problem (5.2) by removing the Constraints (5.2f). To do so, we modify the instantaneous cost functions (as defined in Statement (A_4)) by adding them characteristic functions.

We denote by $\chi_Y : \mathbb{Y} \to \overline{\mathbb{R}}$ the characteristic function of a subset Y of $\mathbb{Y}$:

$$\chi_Y(y) = \begin{cases} 0 & \text{if } y \in Y \ , \\ +\infty & \text{otherwise} \ . \end{cases} \tag{5.66}$$

For all time $t \in \mathcal{T} \setminus \{T\}$, we introduce the function $\overline{\mathcal{L}}_t$ defined by

$$\overline{\mathcal{L}}_t : \mathbb{X} \times \mathbb{U} \to \overline{\mathbb{R}}, (x, u) \mapsto \mathcal{L}_t(x, u) + \chi_{\mathcal{U}_t^{ad}(x)}(u) \ , \tag{5.67}$$

where $\chi_{\mathcal{U}_t^{ad}(x)}$ is the characteristic function of the admissibility set.

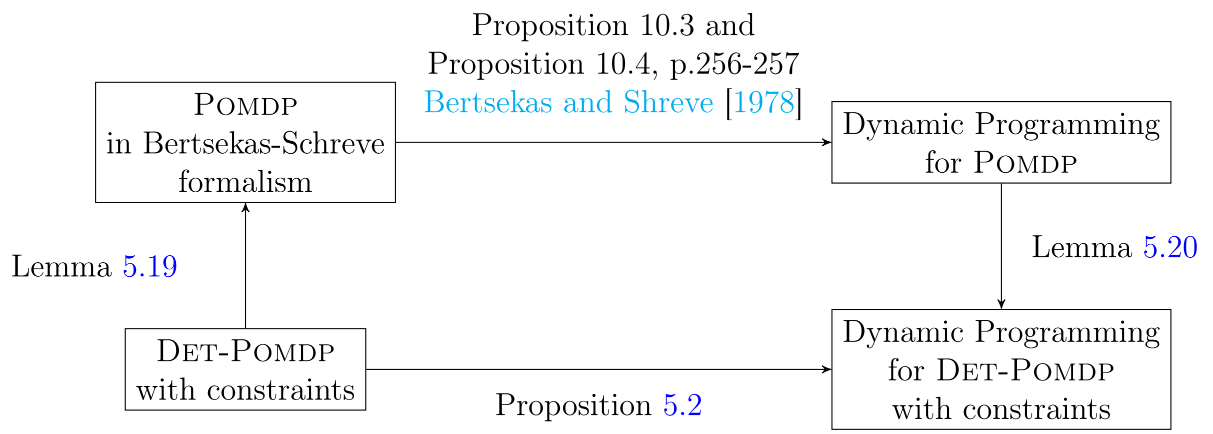


FIGURE 5.11: Illustration of the sketch of proof of Proposition 5.2

We obtain the following optimization Problem:

$$\overline{\mathcal{V}}^{\star}(b_0) = \min_{\boldsymbol{X},\boldsymbol{O},\boldsymbol{U}} \mathbb{E}\Big[\sum_{t=0}^{T-1} \overline{\mathcal{L}}_t(\boldsymbol{X}_t, \boldsymbol{U}_t) + \mathcal{K}(\boldsymbol{X}_T)\Big] \tag{5.68}$$
$$s.t. \quad \text{Equations (5.2b), (5.2c), (5.2d), (5.2e), (5.2g)}$$

Now, we present the Lemma 5.19.

Lemma 5.19. *Under the finite sets Assumption 5.1, Problem* (5.2) *and Problem* (5.68) *are equivalent.*
Moreover, Problem (5.68) *satisfies all the conditions necessary to apply [Bertsekas and Shreve, 1978, Proposition 10.5, p.262] and [Bertsekas and Shreve, 1978, Proposition 10.3, Proposition 10.4, p.256-257]*

Proof. First, we prove that Problem (5.2) and Problem (5.68) are equivalent.

It is standard to reformulate an optimization problem by adding a characteristic function in order to remove a constraint. Here, we need to justify that almost sure constraints can be moved inside the expectation of the sum of the cost functions.

This was already presented in Rockafellar and Wets [1978]. Indeed, in order to prove Theorem 1, Rockafellar and Wets move almost sure constraints in the expectation of the cost functions of the optimization problem. Doing so keeps the cost function as *normal convex intergrand*. Details on *normal convex intergrands* can be seen in [Rockafellar and Wets, 1998, Chapter 14, §D]. In the case of Problem (5.2) and Problem (5.68), Assumption 5.1 ensures that the costs functions are *normal convex intergrand*, and that if the optimal value of Problem (5.2) is finite, optimal policies of Problem (5.2) are also optimal for Problem (5.68), and the reciprocate also holds true.

We now prove that Problem (5.68) follows the framework presented in Bertsekas and Shreve [1978]. In Bertsekas and Shreve [1978], a POMDP satisfies 8 Statements, Statements (A_1)-(A_8).

- Statements (A_1) and (A_2) refer to the definition of sets $\mathbb{X}$, $\mathbb{U}$, $\mathbb{O}$ and $\mathcal{T}$ and are thus satisfied in Problem (5.68).

- Statements (A_3) and (A_4) refer to the definition of the costs functions and a discount factor. The formulation of Problem (5.68) uses a simplified expression equivalent to a discount factor $\alpha = 1$. As $T < +\infty$ by the finite sets Assumption 5.1, Problem (5.68) satisfies the Statement (A_3) regarding the discount factor. Moreover, for all time $t \in \mathcal{T}$, the costs $\overline{\mathcal{L}}_t$ satisfy Statement (A_4).

- Statement (A_5) refers to the definition of state transitions kernels. In Problem (5.68), this role is taken for all time $t \in \mathcal{T}$ by the deterministic evolution functions f_t. We can hence define deterministic transition kernels which satisfy Statement (A_5).

- Statements (A_6) and (A_6') refer to the definition (and absence) of admissibility constraints on the controls, and are both satisfied in Problem (5.68), as we removed any admissibility constraints on the controls.

- Statements (A_7) and (A_8) refer to the definition of the observations kernels. In Problem (5.68), this role is taken for all time $t \in \mathcal{T}$ by the deterministic observation functions h_t. We can hence define deterministic observation kernels which satisfy Statements (A_7) and (A_8).

As Problem (5.68) satisfies Statements (A_1)-(A_8) and Statement (A_6'), it satisfies all the assumptions of [Bertsekas and Shreve, 1978, Proposition 10.5, p.262].

The last assumption necessary to apply [Bertsekas and Shreve, 1978, Proposition 10.3, Proposition 10.4, p.256-257] is a couple of technical assumptions (defined as the couples F^- and $\hat{F}^-$, and F^+ and $\hat{F}^+$). In the case of Problem (5.68), we show that it satisfies the couple $(F^+, \hat{F}^+)$.

Before introducing the couple of assumptions, $(F^+, \hat{F}^+)$, we first introduce the positive and negative parts of real numbers and functions. For any extended real number $y \in \overline{\mathbb{R}}$, we denote by y^+ and y^- the positive and negative part of y, that is,

$$y^+ = \max\left\{0, y\right\} , \qquad \text{and } y^- = \max\left\{0, (-y)\right\} . \tag{5.69}$$

By extension, for any function $g : \mathbb{Y} \to \overline{\mathbb{R}}$, we denote by g^+ and g^- the positive and negative parts of g, i.e.

$$g^+ : \mathbb{Y} \to \overline{\mathbb{R}}, y \mapsto \max\left\{0, g(y)\right\} , \qquad g^- : \mathbb{Y} \to \overline{\mathbb{R}}, y \mapsto \max\left\{0, -g(y)\right\} . \tag{5.70}$$

We now define the couple of assumption $(F^+, \hat{F}^+)$. The condition $(F^+, \hat{F}^+)$ (see [Bertsekas and Shreve, 1978, Definition 10.5, p.249]) is satisfied if, for all admissible policy π

(and $\hat{\pi}$ when using b) and initial state distribution b_0,

$$\mathbb{E}\Big[\sum_{t=0}^{T-1} \overline{\mathcal{L}}_t^- \big(\boldsymbol{X}_t, \pi_t(b_0, \boldsymbol{I}_t)\big) + \mathcal{K}^-(\boldsymbol{X}_T)\Big] < \infty \text{ , and} \tag{5.71}$$

$$\mathbb{E}\Big[\sum_{t=0}^{T-1} \overline{\mathcal{L}}_t^- \big(\hat{\boldsymbol{X}}_t, \hat{\pi}_t(\boldsymbol{B}_t)\big) + \mathcal{K}^-(\boldsymbol{X}_T)\Big] < \infty \text{ ,} \tag{5.72}$$

where for all time $t \in \mathcal{T}$, the random variables $\boldsymbol{X}_t$ and $\hat{\boldsymbol{X}}_t$ are computed thanks to the recursion $\boldsymbol{X}_{t+1} = f_t(\boldsymbol{X}_t, \pi_t(b_0, \boldsymbol{I}_t))$, with $\mathbb{P}_{\boldsymbol{X}_0} = b_0$ for Equation (5.71); and $\hat{\boldsymbol{X}}_{t+1} = f_t(\hat{\boldsymbol{X}}_t, \hat{\pi}_t(\boldsymbol{B}_t))$, with $\boldsymbol{B}_{t+1} = \tau_t(\boldsymbol{B}_t, \hat{\pi}_t(\boldsymbol{B}_t), \boldsymbol{O}_{t+1})$ for Equations (5.72).

Since for all time t, the cost functions $\overline{\mathcal{L}}_t$ take value in $\mathbb{R} \cup \{+\infty\}$ and not $\overline{\mathbb{R}}$, Problem (5.68) indeed satisfies condition $(F^+, \hat{F}^+)$, and hence satisfies all the conditions of [Bertsekas and Shreve, 1978, Proposition 10.3, Proposition 10.4].

This ends the proof. □

Now, by Lemma 5.19, we can apply [Bertsekas and Shreve, 1978, Proposition 10.5, p.262] on Problem (5.68). The beliefs $b \in \mathbb{B}$ are hence *statistics sufficient for controls*. Thus, Dynamic Programming can be applied to the belief-Mdp (by Lemma 5.19, we can apply [Bertsekas and Shreve, 1978, Proposition 10.3, Proposition 10.4, p.256-257], and the Dynamic Programming Equations are given by Equations (46-47), p.259). We thus obtain the following Bellman value functions presented in Bertsekas [2000, Chapter 4] for Problem (5.68):

$$\overline{V}_T : \mathbb{B} \to \overline{\mathbb{R}} \text{ ,} \quad b \mapsto \sum_{x\in\mathbb{X}} b(x)\mathcal{K}(x) \text{ ,} \tag{5.73a}$$

$$\overline{V}_t : \mathbb{B} \to \overline{\mathbb{R}} \text{ ,} \quad b \mapsto \min_{u\in\mathbb{U}} \Big(\overline{\mathcal{C}}_t(b,u) + \sum_{o\in\mathbb{O}, Q_{t+1}(b,u,o)\neq 0} Q_{t+1}(b,u,o)\overline{V}_{t+1}\big(\tau_t(b,u,o)\big)\Big) \text{ ,} \tag{5.73b}$$

where the product of 0 and $+\infty$ is to be computed according to the extended arithmetic convention $0 * (+\infty) = 0$.

We now present Lemma 5.20, which allow us to return to the Dynamic Programming Equations (5.16) presented in Proposition 5.2.

Lemma 5.20. *Under the finite sets Assumption 5.1, the Dynamic Programming Equations of Problem (5.68), Equations (5.73) are equivalent to the Dynamic Programming Equations presented in Proposition 5.2, i.e. Equations (5.16).*

Proof. The proof of Lemma 5.20 is technical, as it involves a number of steps, but is otherwise straightforward and does not present any major difficulty.

We first rewrite Equation (5.73b). We have

$$\begin{aligned}\sum_{o\in\mathbb{O}} Q_{t+1}(b,u,o)\overline{V}_{t+1}\big(\tau_t(b,u,o)\big) &= \sum_{o\in\mathbb{O}, Q_{t+1}(b,u,o)\neq 0} Q_{t+1}(b,u,o)\overline{V}_{t+1}\big(\tau_t(b,u,o)\big) + \\ &\qquad \underbrace{\sum_{o\in\mathbb{O}, Q_{t+1}(b,u,o)=0} 0\times\overline{V}_{t+1}\big(\tau_t(b,u,o)\big)}_{=0 \text{ by the convention } 0\times(+\infty)=0} \\ &= \sum_{o\in\mathbb{O}, Q_{t+1}(b,u,o)\neq 0} Q_{t+1}(b,u,o)\overline{V}_{t+1}\big(\tau_t(b,u,o)\big)\ .\end{aligned}$$

We thus rewrite Equation (5.73b) as

$$\overline{V}_t : \mathbb{B}\to\overline{\mathbb{R}}\ ,\ \ b\mapsto \min_{u\in\mathbb{U}}\Big(\overline{\mathcal{C}}_t(b,u)+\sum_{o\in\mathbb{O}} Q_{t+1}(b,u,o)\overline{V}_{t+1}\big(\tau_t(b,u,o)\big)\Big)\ . \tag{5.74}$$

We now prove that the value functions defined in Equations (5.16) coincide with the value functions $\overline{V}_t$, defined in Equation (5.74) (and used to solve Problem (5.68)).

For the sake of clarity, we note

$$\overline{F}_{t+1} : (b,u,o)\mapsto Q_{t+1}(b,u,o)\overline{V}_{t+1}\big(\tau_t(b,u,o)\big)\ . \tag{5.75}$$

Let $t\in\mathcal{T}\setminus\{T\}$. We start by rewriting the right-hand side of Equation (5.74), with the aim of using $\mathcal{C}$ instead of $\overline{\mathcal{C}}$:

$$\min_{u\in\mathbb{U}}\Big(\overline{\mathcal{C}}_t(b,u)+\sum_{o\in\mathbb{O}}\overline{F}_{t+1}(b,u,o)\Big) = \min_{u\in\mathbb{U}}\Big(\sum_{x\in\mathbb{X}} b(x)\overline{\mathcal{L}}_t(x,u)+\sum_{o\in\mathbb{O}}\overline{F}_{t+1}(b,u,o)\Big)$$

(using Equation (5.16c))

$$= \min_{u\in\mathbb{U}}\Big(\sum_{x\in\mathbb{X}} b(x)\Big[\mathcal{L}_t(x,u)+\chi_{\mathcal{U}_t^{ad}(x)}(u)\Big]+\sum_{o\in\mathbb{O}}\overline{F}_{t+1}(b,u,o)\Big)$$

(using Equation (5.67))

$$= \min_{u\in\mathbb{U}}\Big(\sum_{x\in\mathrm{supp}(b)} b(x)\chi_{\mathcal{U}_t^{ad}(x)}(u)+\sum_{x\in\mathbb{X}} b(x)\mathcal{L}_t(x,u)+\sum_{o\in\mathbb{O}}\overline{F}_{t+1}(b,u,o)\Big)$$

(using the convention $0\times(+\infty)=0$ to only keep the characteristic functions on the support of the belief instead of the whole set $\mathbb{X}$)

$$= \min_{u\in\mathbb{U}}\Big(\sum_{x\in\mathrm{supp}(b)} \chi_{\mathcal{U}_t^{ad}(x)}(u)+\sum_{x\in\mathbb{X}} b(x)\mathcal{L}_t(x,u)+\sum_{o\in\mathbb{O}}\overline{F}_{t+1}(b,u,o)\Big)$$

(as for all $\lambda > 0$ and subset Y, $\lambda\chi_Y = \chi_Y$)

$$= \min_{u\in\mathbb{U}} \Big(\chi_{\bigcap_{x\in\mathrm{supp}(b)} \mathcal{U}_t^{ad}(x)}(u) + \sum_{x\in\mathbb{X}} b(x)\mathcal{L}_t(x,u) + \sum_{o\in\mathbb{O}} \overline{F}_{t+1}(b,u,o) \Big)$$

(as $\chi_{U_1}(u) + \chi_{U_2}(u) = \chi_{U_1\cap U_2}(u)$)

$$= \min_{u\in\mathbb{U}} \Big(\chi_{\mathcal{U}_t^{\mathrm{b}}(b)}(u) + \sum_{x\in\mathbb{X}} b(x)\mathcal{L}_t(x,u) + \sum_{o\in\mathbb{O}} \overline{F}_{t+1}(b,u,o) \Big)$$

(using the definition (5.16d) of $\mathcal{U}_t^{\mathrm{b}}(b)$)

$$= \min_{u\in\mathcal{U}_t^{\mathrm{b}}(b)} \Big(\sum_{x\in\mathbb{X}} b(x)\mathcal{L}_t(x,u) + \sum_{o\in\mathbb{O}} \overline{F}_{t+1}(b,u,o) \Big)$$

(as $\min_{u\in\mathbb{U}} \big(g(u) + \chi_U(u)\big) = \min_{u\in U} g(u)$ for all $U \subset \mathbb{U}$). Hence, we have obtained that

$$\min_{u\in\mathbb{U}} \Big(\overline{\mathcal{C}}_t(b,u) + \sum_{o\in\mathbb{O}} \overline{F}_{t+1}(b,u,o) \Big) = \min_{u\in\mathcal{U}_t^{\mathrm{b}}(b)} \Big(\sum_{x\in\mathbb{X}} b(x)\mathcal{L}_t(x,u) + \sum_{o\in\mathbb{O}} \overline{F}_{t+1}(b,u,o) \Big) \ , \tag{5.76}$$

and, as a consequence, that

$$\begin{aligned} \overline{V}_t(b) &= \min_{u\in\mathbb{U}} \Big(\overline{\mathcal{C}}_t(b,u) + \sum_{o\in\mathbb{O}} Q_{t+1}(b,u,o)\overline{V}_{t+1}\big(\tau_t(b,u,o)\big) \Big) \\ &= \min_{u\in\mathcal{U}_t^{\mathrm{b}}(b)} \Big(\mathcal{C}_t(b,u) + \sum_{o\in\mathbb{O}} Q_{t+1}(b,u,o)\overline{V}_{t+1}\big(\tau_t(b,u,o)\big) \Big) \end{aligned} \tag{5.77}$$

(using Equations (5.74), (5.75) and (5.76)).

Moreover, we obtain, by using Equations (5.16a) and (5.73a), that

$$\forall b \in \mathbb{B} \ , \ \ \overline{V}_T(b) = \sum_{x\in\mathbb{X}} b(x)\mathcal{K}(x) = V_T(b) \ . \tag{5.78}$$

As $\overline{V}_T = V_T$ (Equation (5.78)), and since we have the same expression for the backward induction in Equations (5.16b) and (5.77) for both V_t and $\overline{V}_t$, then by backward induction, we get that

$$\forall t \in \mathcal{T} \ , \ \ \overline{V}_t = V_t \ . \tag{5.79}$$

Moreover, the controls u that minimize the right-hand side of Equation (5.73) also minimize the right-hand side of Equation (5.16). Equations (5.16) and (5.73) hence share the same values and controls.

□

We have now detailed all the steps of the proof of Proposition 5.2, as presented in Figure 5.11.

5.A.4 Technical Lemmata

In this subsection, we present technical lemmata used in the proofs of Theorems 5.9 and 5.10. We first present in §5.A.4.1 notations used in the rest of this subsection and introduce the notions of forward and backward mappings. Second, in §5.A.4.2, we present properties on the composition and pushforward measures by those forward and backward mapping. Third, in §5.A.4.3, we present properties on the cardinality of sets of forward and backward mappings used notably in the proof of Theorems 5.9 and 5.10.

We show the logical relationships between the different lemmata and theorems in Figure 5.12.

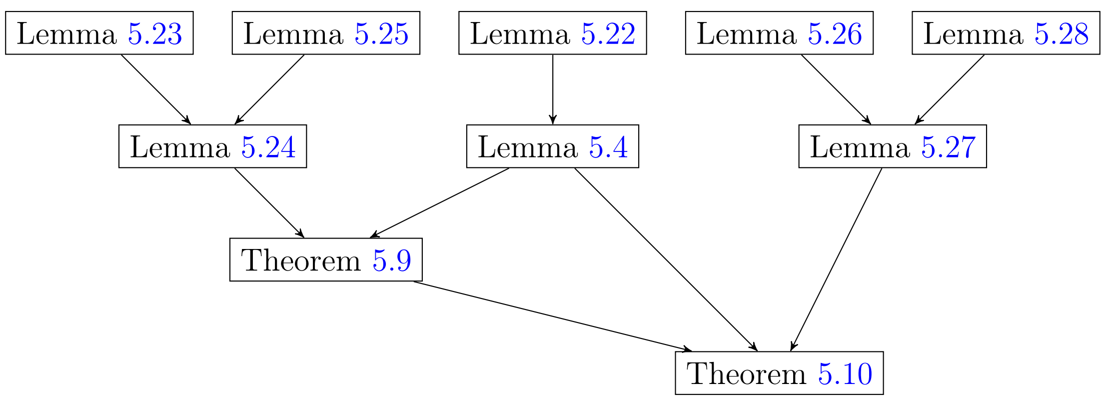


FIGURE 5.12: Diagram of the logical relationships between the different lemmata and theorems of §5.4 and §5.A.4. An arrow between two propositions $A \to B$ states that proposition A is used to prove proposition B.

5.A.4.1 Notations

In §5.A.4, we use the notations presented in §5.2 and §5.3. For the sake of completeness, we first recall the main notations on sets and mapping we use in §5.A.4 and §5.A.5, before defining forward and backward mappings.

Let $\mathbb{X}$ be a finite set. 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 ∂:

$$\overline{\mathbb{X}} = \mathbb{X} \cup \{\partial\} . \tag{5.80}$$

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

$$\mathbb{B} = \Delta(\mathbb{X}) \cup \{\delta_\partial\} , \tag{5.81}$$

where we identify the set of probability distributions $\Delta(\mathbb{X})$ on $\mathbb{X}$ with the subset $\{\mu \in \Delta(\overline{\mathbb{X}}) \,|\, \mathrm{supp}(\mu) \subset \mathbb{X}\} \subset \Delta(\overline{\mathbb{X}})$ and where $\delta_\partial \in \Delta(\overline{\mathbb{X}})$ is the discrete probability measure on $\overline{\mathbb{X}}$ concentrated on the extra point ∂, that is $\delta_\partial(\{\delta\}) = 1$. A probability measure $\nu \in \Delta(\overline{\mathbb{X}})$ will be represented by an ordered pair $(\nu_{|\mathbb{X}}, \nu(\partial))$.

We define a "renormalization" mapping $\mathcal{R} : \Delta(\overline{\mathbb{X}}) \to \Delta(\overline{\mathbb{X}})$ defined by

$$\mathcal{R} : \nu \in \Delta(\overline{\mathbb{X}}) \mapsto \begin{cases} \big(\frac{1}{\nu(\mathbb{X})}\nu_{|\mathbb{X}}, 0\big) & \text{if } \nu(\mathbb{X}) \neq 0 \ , \\ \delta_\partial & \text{if } \nu(\mathbb{X}) = 0 \ . \end{cases} \tag{5.82}$$

It is straightforward to check that the image of the mapping $\mathcal{R}$ is the set $\mathbb{B} = \Delta(\mathbb{X}) \cup \{\delta_\partial\}$ and that we have $\mathcal{R}^{-1}(\delta_\partial) = \delta_\partial$.

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

$$h_{\overrightarrow{X}} : x \in \overline{\mathbb{X}} \mapsto \begin{cases} h(\overline{x}) & \text{if} \quad x \in \mathbb{X} \quad \text{and} \quad h(x) \in X \ , \\ \partial & \text{if} \quad x = \partial \quad \text{or} \quad h(x) \notin X \ . \end{cases} \tag{5.83}$$

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 implies 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

$$h_{\overleftarrow{X}} : x \in \overline{\mathbb{X}} \mapsto \begin{cases} h(x) & \text{if} \quad x \in X \ , \\ \partial & \text{otherwise.} \end{cases} \tag{5.84}$$

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 implies 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

$$\forall X \subset \mathbb{X} \ , \ h_{\overrightarrow{X}} = h_{\overleftarrow{h^{-1}(X)}} \ , \tag{5.85a}$$

$$\forall X \subset \mathbb{X} \ , \ h_{\overrightarrow{X}} = h_{\overrightarrow{X \cap \mathrm{Im}(h)}} \ , \tag{5.85b}$$

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

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

5.A.4.2 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 5.21. *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 a $\big(\mathbb{M}, \overleftarrow{\mathbb{X}}\big)$-mappings set (resp. a $\big(\mathbb{M}, \overrightarrow{\mathbb{X}}\big)$-mappings set) if it satisfies the following property*

$$\mathbb{G} \subset \big\{ h_{\overleftarrow{X}} \,\big|\, h \in \mathbb{M} \text{ and } X \subset \mathbb{X} \big\} , \tag{5.86a}$$

$$\Big(resp. \quad \mathbb{G} \subset \big\{ h_{\overrightarrow{X}} \,\big|\, h \in \mathbb{M} \text{ and } X \subset \mathbb{X} \big\}\Big) , \tag{5.86b}$$

where $h_{\overleftarrow{X}}$ (resp. $h_{\overrightarrow{X}}$) is defined in Equation (5.84) (resp. Equation (5.83)). When $\mathbb{M} = \mathbb{L}(\mathbb{X};\mathbb{X})$, a $\big(\mathbb{M}, \overleftarrow{\mathbb{X}}\big)$-mappings set (resp. a $\big(\mathbb{M}, \overrightarrow{\mathbb{X}}\big)$-mappings set) is just named a $\big(\overleftarrow{\mathbb{X}}\big)$-mappings set (resp. a $\big(\overrightarrow{\mathbb{X}}\big)$-mappings set).

We obtain the following properties.

- If $\mathbb{G}$ is a $\big(\mathbb{M}, \overrightarrow{\mathbb{X}}\big)$-mappings set, then $\mathbb{G}$ is a $\big(\mathbb{M}, \overleftarrow{\mathbb{X}}\big)$-mappings set (using Equality (5.85a)).

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

$$h'_{\overleftarrow{X'}} \circ h_{\overleftarrow{X}} = (h' \circ h)_{\overleftarrow{X \cap h^{-1}(X')}} . \tag{5.87}$$

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

$$\mathbb{G}_{\overleftarrow{X}} = \big\{ g \in \mathbb{G} \,\big|\, \exists h \in \mathbb{L}(\mathbb{X};\mathbb{X}), g = h_{\overleftarrow{X}} \big\} . \tag{5.88}$$

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

$$\big(\mathcal{R} \circ (\mathbb{G}_{\overleftarrow{X \cap \mathrm{supp}(b_0)}})_\star\big)(b_0) = \big(\mathcal{R} \circ (\mathbb{G}_{\overleftarrow{X}})_\star\big)(b_0) . \tag{5.89}$$

  The Equation (5.89) is a consequence of the following Lemma 5.22. Indeed, assuming Lemma 5.22, the expression of $\big(\mathcal{R} \circ (\mathbb{G}_{\overleftarrow{X}})_\star\big)(b_0)$ given by Equation (5.90b) 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 5.22. *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 (5.17), and the mapping*

$\mathcal{R}$ is defined in Equation (5.82), have the following expressions for all $\nu \in \Delta(\overline{\mathbb{X}})$:

$$\big(\mathcal{R} \circ (h_{\overrightarrow{X}})_\star\big)(\nu) = \begin{cases} \left[\overline{x} \in \overline{\mathbb{X}} \mapsto \dfrac{\nu\big(h^{-1}(\overline{x})\big)\mathbf{1}_X(\overline{x})}{\nu\big(h^{-1}(X)\big)}\right] & \textit{if} \quad \nu\big(h^{-1}(X)\big) \neq 0 \ , \\ \delta_\partial & \textit{otherwise,} \end{cases} \tag{5.90a}$$

and

$$\big(\mathcal{R} \circ (h_{\overleftarrow{X}})_\star\big)(\nu) = \begin{cases} \left[\overline{x} \in \overline{\mathbb{X}} \mapsto \dfrac{\nu\big(h^{-1}(\overline{x}) \cap X\big)}{\nu\big(h^{-1}(\mathbb{X}) \cap X\big)}\right] & \textit{if} \quad \nu\big(h^{-1}(\mathbb{X}) \cap X\big) \neq 0 \ , \\ \delta_\partial & \textit{otherwise.} \end{cases} \tag{5.90b}$$

Proof. For any probability measure ν on the finite set $\overline{\mathbb{X}}$, it is straightforward, using the definition of pushforward measure in Equation (5.17), to obtain that the pushforward of the measure ν through the mapping $h_{\overrightarrow{X}}$, as defined in Equation (5.83), is given by

$$\begin{aligned} (h_{\overrightarrow{X}})_\star\nu : \overline{\mathbb{X}} &\to \mathbb{R}_+ \\ \overline{y} \mapsto \nu\big((h_{\overrightarrow{X}})^{-1}(\overline{y})\big) &= \begin{cases} \nu\big(h^{-1}(\overline{y})\big) & \text{if} \quad \overline{y} \in X \ , \\ \Big(1 - \nu\big(h^{-1}(X)\big)\Big) & \text{if} \quad \overline{y} = \partial \ , \\ 0 & \text{if} \quad \overline{y} \neq \partial \text{ and } \overline{y} \notin X \ . \end{cases} \end{aligned} \tag{5.91}$$

Thus, we obtain that

$$\forall x \in \mathbb{X} \ , \quad \big((h_{\overrightarrow{X}})_\star\nu\big)_{|\mathbb{X}}(x) = \nu\big(h^{-1}(x)\big)\mathbf{1}_X(x) \ , \tag{5.92}$$

and that

$$\big((h_{\overrightarrow{X}})_\star\nu\big)(\mathbb{X}) = \sum_{y \in \mathbb{X}} \nu\big(h^{-1}(x)\big)\mathbf{1}_X(x) = \nu\big(h^{-1}(X)\big) \ . \tag{5.93}$$

Hence, using the definition of $\mathcal{R}$ in Equation (5.82), the result follows from Equation (5.90a). The proof of Equation (5.90b) 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.

Lemma 5.23. *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*

$$\mathcal{R} \circ (h_{\overrightarrow{X}})_\star \circ \mathcal{R} \circ (h'_{\overrightarrow{X'}})_\star = \mathcal{R} \circ (h_{\overrightarrow{X}} \circ h'_{\overrightarrow{X'}})_\star \ , \tag{5.94a}$$

$$\mathcal{R} \circ (h_{\overleftarrow{X}})_\star \circ \mathcal{R} \circ (h'_{\overleftarrow{X'}})_\star = \mathcal{R} \circ (h_{\overleftarrow{X}} \circ h'_{\overleftarrow{X'}})_\star \ . \tag{5.94b}$$

Proof. We just prove Equation (5.94a) as the proof follows the same lines for Equation (5.94b). 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 (5.90a) 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

$$\begin{aligned}
&\mathcal{R} \circ (h_{\overrightarrow{X}})_\star \circ \mathcal{R} \circ (h'_{\overrightarrow{X'}})_\star \nu \\
&\quad = \mathcal{R} \circ (h_{\overrightarrow{X}})_\star \circ \mathcal{R}(\nu') && \text{(by replacing } (h'_{\overrightarrow{X'}})_\star \nu \text{ by } \nu')\\
&\quad = \mathcal{R} \circ (h_{\overrightarrow{X}})_\star \big(\frac{1}{\nu'(\mathbb{X})} \nu'_{|\mathbb{X}}, 0\big) && \text{(using } \mathcal{R} \text{ definition in (5.82), with } \nu'(\mathbb{X}) \neq 0)\\
&\quad = \mathcal{R} \circ (h_{\overrightarrow{X}})_\star \big(\frac{1}{\nu'(\mathbb{X})} (\nu'_{|\mathbb{X}}, 0)\big) && \text{(factorizing by } \tfrac{1}{\nu'(\mathbb{X})})\\
&\quad = \mathcal{R}\Big(\frac{1}{\nu'(\mathbb{X})} (h_{\overrightarrow{X}})_\star \big(\nu'_{|\mathbb{X}}, 0\big)\Big) && \text{(as } (h_{\overrightarrow{X}})_\star \text{ is 1-positively homogeneous)}\\
&\quad = \mathcal{R}\big((h_{\overrightarrow{X}})_\star \big(\nu'_{|\mathbb{X}}, 0\big)\big) && \text{(as } \mathcal{R} \text{ is 0-positively homogeneous)}\\
&\quad = \mathcal{R}\big((h_{\overrightarrow{X}})_\star (\nu')\big) && \text{(using the preliminary part)}\\
&\quad = \mathcal{R} \circ (h_{\overrightarrow{X}})_\star \circ (h'_{\overrightarrow{X'}})_\star \nu && \text{(as } \nu' = (h'_{\overrightarrow{X'}})_\star \nu)\\
&\quad = \mathcal{R} \circ (h_{\overrightarrow{X}} \circ h'_{\overrightarrow{X'}})_\star (\nu) \;. && \text{(as } f_\star \circ h_\star = (f \circ h)_\star)
\end{aligned}$$

Second case. We assume that $\nu'(\mathbb{X}) = 0$. Then, we have that $\nu' = \delta_\partial$ 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 (5.94a). □

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.

5.A.4.3 Results on the cardinality of sets of pushforward measures

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

Lemma 5.24. *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* $\mathbb{G}_k$ *sets, for each* $k \in \mathbb{N}$*, 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*

$$\forall k \in \mathbb{N}\ , \quad \Phi_k = \mathcal{R} \circ (\mathbb{G}_k)_\star \ . \tag{5.95}$$

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

$$\forall n \in \mathbb{N}\ , \quad \Big|\bigcup_{k=0}^{n} \Phi_{0:k}(b_0)\Big| \leq (1+|\mathbb{X}|)^{|\mathrm{supp}(b_0)|}\ , \tag{5.96}$$

where $\Phi_{0:k} = \Phi_k \circ \cdots \circ \Phi_0$ *is defined in Equation* (5.27)*.*

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

$$\begin{aligned}
\Phi_{0:k}(b_0) &= (\Phi_k \circ \Phi_{k-1} \circ \cdots \circ \Phi_0)(b_0) && \text{(by Equation (5.27))}\\
&= \big(\mathcal{R} \circ (\mathbb{G}_k)_\star \circ \mathcal{R} \circ (\mathbb{G}_{k-1})_\star \circ \cdots \circ \mathcal{R} \circ (\mathbb{G}_0)_\star\big)(b_0) && \text{(by Equation (5.95))}\\
&= \big(\mathcal{R} \circ (\mathbb{G}_k)_\star \circ (\mathbb{G}_{k-1})_\star \circ \cdots \circ (\mathbb{G}_0)_\star\big)(b_0)
\end{aligned}$$

(by Lemma (5.23), as the sets $\mathbb{G}_k$ are, by assumption, either all $(\overrightarrow{\mathbb{X}})$-mappings sets or all $(\overleftarrow{\mathbb{X}})$-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 5.25 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 could extend previous Lemma 5.24 to cases with sequences $\{\mathbb{G}_k\}_{k\in\mathbb{N}}$ of mixes $(\overrightarrow{\mathbb{X}})$-mappings sets and $(\overleftarrow{\mathbb{X}})$-mappings sets. Indeed, forward mappings are also backward mappings by Equation (5.85a). We can hence write the sequence $\{\mathbb{G}_k\}_{k\in\mathbb{N}}$ as a sequence of only $(\overleftarrow{\mathbb{X}})$-mappings sets. In the rest of this chapter, we consider sequences of only $(\overrightarrow{\mathbb{X}})$-mappings sets or only $(\overleftarrow{\mathbb{X}})$-mappings sets, and thus only need Lemma 5.24.

We can bound the cardinality of the set of pushforward of a given nonnegative measure thanks to the following Lemma 5.25 (which was previously postponed in the proof of Lemma 5.24).

Lemma 5.25. *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*

$$|\mathbb{J}_\star \mu| \leq |\mathbb{Y}|^{|\mathrm{supp}(\mu)|} , \tag{5.97}$$

where we recall that $|\mathbb{J}_\star \mu|$ denote the set $\big|\{j_\star \mu \,|\, j \in \mathbb{J}\}\big|$ as exposed in Equation (5.27b).

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

$$\begin{aligned}
j_\star \mu(y) &= \mu\big(j^{-1}(y)\big) && \text{(by the definition of pushforward measures, in (5.17))} \\
&= \mu\Big(\big(j^{-1}(y) \cap \mathrm{supp}(\mu)\big) \cup \big(j^{-1}(y) \cap (\mathrm{supp}(\mu))^c\big)\Big) \\
&= \mu\big(j^{-1}(y) \cap \mathrm{supp}(\mu)\big) + \underbrace{\mu\big(j^{-1}(y) \cap (\mathrm{supp}(\mu))^c\big)}_{=0} \\
&= \mu\big(j^{-1}_{|\mathrm{supp}(\mu)}(y)\big) \\
&= \big(j_{|\mathrm{supp}(\mu)}\big)_\star \mu(y) . && \text{(by (5.17))}
\end{aligned}$$

Thus, considering $\mathbb{J}_{|\mathrm{supp}(\mu)} = \{j_{|\mathrm{supp}(\mu)} \,|\, j \in \mathbb{J}\}$, 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 5.26. *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*

$$\forall b \in \mathbb{B} , \qquad \sum_{h \in \Phi_k} \big|\mathrm{supp}\big(h(b)_{|\mathbb{X}}\big)\big| \leq |\mathrm{supp}(b_{|\mathbb{X}})| . \tag{5.98}$$

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

$$\forall k \in \mathbb{N} , \ \big|\Phi_{0:k}(b_0) \setminus \{\delta_\partial\}\big| \leq |\mathrm{supp}(b_0)| , \tag{5.99}$$

where $\Phi_{0:k}(b_0) = \Phi_k \circ \cdots \circ \Phi_0(b_0)$ is defined in Equation (5.27e).

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

$$\forall k \in \mathbb{N} \; , \quad \sum_{b \in \Phi_{0:k}(b_0)} \left|\mathrm{supp}(b_{|\mathbb{X}})\right| \le |\mathrm{supp}(b_0)| \; . \tag{5.100}$$

First, we consider the case $k = 0$. As $\Phi_{0:0} = \Phi_0$ the result follows from Equation (5.98) used for $k = 0$ and $b = b_0$. Second, we consider $0 < k$, and, assuming that Equation (5.100) is satisfied for k, we prove that it is also satisfied for $k+1$ as follows:

$$\begin{aligned}
\sum_{b \in \Phi_{0:k+1}(b_0)} \left|\mathrm{supp}(b_{|\mathbb{X}})\right| &= \sum_{h \in \Phi_{0:k+1}} \left|\mathrm{supp}\big((h(b_0))_{|\mathbb{X}}\big)\right| && \text{(by (5.27))} \\
&= \sum_{h' \in \Phi_{k+1}, h'' \in \Phi_{0:k}} \left|\mathrm{supp}\Big(\big(h'(h''(b_0))\big)_{|\mathbb{X}}\Big)\right| && (\text{as } \Phi_{0:k+1} = \Phi_{k+1} \circ \Phi_{0:k}) \\
&= \sum_{h'' \in \Phi_{0:k}} \Big(\sum_{h' \in \Phi_{k+1}} \left|\mathrm{supp}\Big(\big(h'(h''(b_0))\big)_{|\mathbb{X}}\Big)\right| \Big) \\
&\le \sum_{h'' \in \Phi_{0:k}} \left|\mathrm{supp}\Big(\big(h''(b_0)\big)_{|\mathbb{X}}\Big)\right| \\
& && (\text{using Equation (5.98) for } k \text{ and } b = h''(b_0)) \\
&= \sum_{b \in \Phi_{0:k}(b_0)} \left|\mathrm{supp}\big(b_{|\mathbb{X}}\big)\right| && \text{(by (5.27))} \\
&\le |\mathrm{supp}(b_0)| \; . && (\text{by induction assumption on } k)
\end{aligned}$$

We conclude that Equation (5.100) is satisfied for all $k \in \mathbb{N}$.

Now, we turn to the proof of Equation (5.99). We make the following observation: if $b \in \Delta(\mathbb{X})$, then we have that $|\mathrm{supp}(b_{|\mathbb{X}})| \ge 1$ and if $b = \delta_\partial$ then $|\mathrm{supp}(b_{|\mathbb{X}})| = 0$. Thus, we have that

$$\begin{aligned}
|\Phi_{0:k}(b_0) \setminus \{\delta_\partial\}| &= \sum_{b \in \Phi_{0:k}(b_0) \setminus \{\delta_\partial\}} 1 \qquad\qquad (5.101) \\
&\le \sum_{b \in \Phi_{0:k}(b_0) \setminus \{\delta_\partial\}} |\mathrm{supp}(b_{|\mathbb{X}})| && (\text{as } |\mathrm{supp}(b_{|\mathbb{X}})| \ge 1 \text{ for } b \in \Phi_{0:k}(b_0) \setminus \{\delta_\partial\}) \\
&= \sum_{b \in \Phi_{0:k}(b_0)} |\mathrm{supp}(b_{|\mathbb{X}})| && (\text{as } |\mathrm{supp}(\delta_{\partial|\mathbb{X}})| = 0) \\
&\le |\mathrm{supp}(b_0)| \; , && \text{(by (5.100))}
\end{aligned}$$

which gives Equation (5.99). That concludes the proof. $\square$

Lemma 5.27. *Let $\{h^k\}_{k \in \mathbb{N}}$ be a given 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 subset of $\mathbb{X}$. Let $\{\mathbb{G}_k\}_{k \in \mathbb{N}}$ be* *a given sequence of self-mappings on the set $\overline{\mathbb{X}}$, of the following form*

$$\forall k \in \mathbb{N} \ , \quad \mathbb{G}_k = \left\{ h^k{}_{\overrightarrow{X_i^k}} \,\middle|\, i \in I_k \right\} \subset \overline{\mathbb{X}}^{\overline{\mathbb{X}}} \ , \tag{5.102}$$

where $h^k{}_{\overrightarrow{X_i^k}} : \overline{\mathbb{X}} \to \overline{\mathbb{X}}$ are built following Equation (5.83). 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 (5.27). Then, given $b_0 \in \Delta(\mathbb{X})$, we have

$$\forall k \in \mathbb{N} \ , \quad \left| \Phi_{0:k}(b_0) \setminus \{\delta_\partial\} \right| \leq |\mathrm{supp}(b_0)| \ . \tag{5.103}$$

Proof. The proof relies on postponed Lemma 5.28 from which we obtain that the sets Φ_k satisfy Equation (5.98) for all $k \in \mathbb{N}$, and on Lemma 5.26.

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 (5.82), 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 sets Φ_k satisfy Equation (5.98) for all $k \in \mathbb{N}$. For that purpose, we fix $k \in \mathbb{N}$, and $b \in \mathbb{B}$ and we successively have

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

(by (5.106) in Lemma 5.28, applied with $\mathbb{Y} = \mathbb{V} = \mathbb{X}$ and $V = X$, $V_i = X_i^k$ for $i \in I = I_k$)

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

Third, as the assumption given in Equation (5.98) is satisfied, the result follows by Lemma 5.26. □

We now present the postponed technical Lemma 5.28.

Lemma 5.28. *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 given subset of $\mathbb{V}$. We define the mapping $h_V : \mathbb{Y} \to \mathbb{V} \cup \{\partial_{\mathbb{V}}\}$*[6] *taking values in the extended set $\overline{\mathbb{V}} = \mathbb{V} \cup \{\partial_{\mathbb{V}}\}$ as follows*

$$h_V : y \in \mathbb{Y} \mapsto \begin{cases} h(y) & \textit{if} \quad h(y) \in V \ , \\ \partial_{\mathbb{V}} & \textit{elsewhere} \ . \end{cases} \tag{5.104}$$

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

$$\left|\operatorname{supp}\Big(\big((h_V)_\star \mu\big)_{|\mathbb{V}}\Big)\right| \leq \left|\operatorname{supp}\big(\mu_{|_{h^{-1}(V)}}\big)\right| \ . \tag{5.105}$$

Moreover, let $\{V_i\}_{i\in I}$ be a finite family of pairwise disjoints subsets of $\mathbb{V}$. Then, we have that

$$\sum_{i\in I} \left|\operatorname{supp}\Big(\big((h_{V_i})_\star \mu\big)_{|\mathbb{V}}\Big)\right| \leq \left|\operatorname{supp}\big(\mu_{|_{h^{-1}(\sqcup_{i\in I} V_i)}}\big)\right| \ . \tag{5.106}$$

Proof. We prove Equation (5.105). Let $\mu \in \Delta(\mathbb{Y})$ be given. First, we note that, if the set $\operatorname{supp}\big(\big((h_V)_\star \mu\big)_{|\mathbb{V}}\big)$ is empty, the result is obvious. Second, we assume that $\operatorname{supp}\big(\big((h_V)_\star \mu\big)_{|\mathbb{V}}\big) \neq \emptyset$ and consider $v \in \operatorname{supp}\big(\big((h_V)_\star \mu\big)_{|\mathbb{V}}\big)$. 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 (5.104)), 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 $y \in h^{-1}(v) \cap \operatorname{supp}(\mu)$. Now, consider the set-valued mapping

$$\begin{aligned} \Gamma : \operatorname{supp}\big(\big((h_V)_\star \mu\big)_{|\mathbb{V}}\big) &\rightrightarrows \mathbb{Y} \\ y &\mapsto h^{-1}(v) \cap \operatorname{supp}(\mu) \ . \end{aligned}$$

By construction, the set-valued mapping Γ takes values in the subsets of $\operatorname{supp}(\mu_Y)$ with the notation $\mu_Y = \mu_{|_{h^{-1}(V)}}$, and we have just proved that it takes values in the nonempty subsets of $\operatorname{supp}(\mu_Y)$. Moreover, the set-valued mapping Γ is injective, as for all pairs of distinct elements $(v', v'') \in V^2$, $v' \neq v''$, 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 $\operatorname{supp}(\mu_V)$ and we conclude that

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

which gives Equation (5.105).

[6] 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).

Now, we turn to the proof of Inequality (5.106). We successively have

$$\sum_{i\in I} \Big|\mathrm{supp}\Big(\big((h_{V_i})_\star\mu\big)_{|\mathbb{V}}\Big)\Big| \leq \sum_{i\in I} \big|\mathrm{supp}\big(\mu_{|_{h^{-1}(V_i)}}\big)\big| \qquad \text{(by (5.105) for each } i\in I)$$
$$= \big|\mathrm{supp}\big(\mu_{|_{\sqcup_{i\in I} h^{-1}(V_i)}}\big)\big|$$

(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 $\{V_i\}_{i\in I}$)

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

which concludes the proof. □

This technical Lemma 5.28 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 (5.104). A similar result

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

is given in [Littman, 1996, Lemma 6.2] but only for the mappings $(\tau_t)_{t\in\mathcal{T}}$, defined in Equation (5.14) and with a proof not explicitly connected to pushforward measures.

We now present the postponed proof of Lemma 5.4, presented in page 87.

Proof of Lemma 5.4. 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

$$\tau_t(b,u,o) = \mathcal{R} \circ (F^{u,o}_t)_\star(b) \,. \tag{5.107}$$

Using Equation (5.10), we have that

$$Q_{t+1}(b,u,o) = b\big((h^u_{t+1}\circ f^u_t)^{-1}(o)\big) = b\big((f^u_t)^{-1}(X)\big) \,. \tag{5.108}$$

Now, using the expression of τ_t in Equation (5.14) combined with Equation (5.108) and the definition of X we obtain, for all $\overline{x}\in\overline{\mathbb{X}}$, that

$$\tau_t(b,u,o)(\overline{x}) = \begin{cases} \dfrac{b\big((f^u_t)^{-1}(\overline{x})\big)\mathbf{1}_X(\overline{x})}{b\big((f^u_t)^{-1}(X)\big)} & \text{if } b\big((f^u_t)^{-1}(X)\big) \neq 0 \,, \\ 0 & \text{otherwise} \,. \end{cases} \tag{5.109}$$

Then, Equation (5.107) follows from Lemma 5.22 applied with the mapping $h = f_t^u$ and with the subset $X = \left(h_{t+1}^u\right)^{-1}(o)$, as we have

$$F_t^{u,o} = f_t^u{}_{\overrightarrow{(h_{t+1}^u)^{-1}(o)}} \;, \tag{5.110}$$

where $f_t^u{}_{\overrightarrow{(h_{t+1}^u)^{-1}(o)}}$ is defined in Equation (5.83).

This ends the proof. □

5.A.5 Complementary result on (∂)-separated mapping sets

In this subsection, we present complementary results on (∂)-separated mapping sets by applying the framework presented in Appendix 5.A.4. We notably apply the notion of forward and backward mapping, presented in Equations (5.83) and (5.84), and the notion of pushforward measure, defined in §5.3 in Equation (5.17).

First, in §5.A.5.1, we present the lemmata used in the proofs of §5.5. Second, in §5.A.5.2, we present a few examples of Separated Det-Pomdps. Third, in §5.A.5.3, we present a necessary condition in order to have tight bounds in Theorem 5.15.

5.A.5.1 Properties of (∂)-separated mapping set

Lemma 5.29. *Let $\mathbb{G}$ be a $\left(\mathbb{M}, \overleftarrow{\mathbb{X}}\right)$-mappings set as defined in Definition 5.21. 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 5.12, 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 (5.84), 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. □

A direct consequence of Lemma 5.29 is the following Corollary 5.30.

Corollary 5.30. *Let $\{\mathbb{M}_k\}_{k\in\mathbb{N}}$ be a given 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 a $\left(\mathbb{M}_k, \overleftarrow{\mathbb{X}}\right)$-mappings set. If the set $\cup_{k\in\mathbb{N}}\left(\mathbb{M}_k \circ \mathbb{M}_{k-1} \circ \cdots \circ \mathbb{M}_0\right)$ 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 a $\big(\mathbb{M}_1, \overleftarrow{\mathbb{X}}\big)$-mappings set and a $\big(\mathbb{M}_2, \overleftarrow{\mathbb{X}}\big)$-mappings set. Then, we have that

$$\mathbb{G}_1 \circ \mathbb{G}_2 = \big\{ g_1 \circ g_2 \,\big|\, g_1 \in \mathbb{G}_1 \text{ and } g_2 \in \mathbb{G}_2 \big\} \qquad \text{(by Notation (5.27d))}$$
$$\subset \big\{ m^1{}_{\overleftarrow{X_1}} \circ m^2{}_{\overleftarrow{X_2}} \,\big|\, m^1 \in \mathbb{M}_1 \;,\;\; m^2 \in \mathbb{M}_2 \;,\;\; X_1 \subset \mathbb{X} \;,\;\; X_2 \subset \mathbb{X} \big\} \qquad \text{(by (5.86))}$$
$$\subset \big\{ (m^1 \circ m^2)_{\overleftarrow{X_2 \cap (m^2)^{-1}(X_1)}} \,\big|\, m^1 \in \mathbb{M}_1 \;,\;\; m^2 \in \mathbb{M}_2 \;,\;\; X_1 \subset \mathbb{X} \;,\;\; X_2 \subset \mathbb{X} \big\}$$

(by (5.87))

$$\subset \big\{ m_X \,\big|\, m \in \mathbb{M}_1 \circ \mathbb{M}_2 \text{ and } X \subset \mathbb{X} \big\} \; .$$

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 5.29. 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 set. □

Before presenting bounds on the cardinality of a (∂)-separated mapping set, we present Lemma 5.31.

Lemma 5.31. *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*

$$|\mathbb{J}| \leq |\mathbb{Y}| \; . \tag{5.111}$$

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

$$\big|\gamma_{\overline{x}}(\mathbb{J})\big| \leq |\mathbb{Y}| \; . \tag{5.112}$$

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 (5.112), gives Inequality (5.111), and concludes the proof. □

We now use the previous Lemma 5.31 to bound the cardinality of a (∂)-separated mapping set.

Lemma 5.32. *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*

$$\mathbb{G}_{X\to X'} = \left\{ g \in \mathbb{G} \,\middle|\, g^{-1}(\mathbb{X}) = X, g(X) \subset X' \right\} . \tag{5.113}$$

Then, we have

$$\left|\mathbb{G}_{X\to X'}\right| \begin{cases} \leq |X'| & \text{if } X \subset \mathbb{X} , \\ = 0 & \text{if } X \cap \{\partial\} \neq \emptyset . \end{cases} \tag{5.114}$$

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

$$\Gamma : \mathbb{G}_{X\to X'} \to X'^X \ , \ \ g \mapsto g_{|X} \ . \tag{5.115}$$

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

$$\left|\mathbb{G}_{X\to X'}\right| = \left|\Gamma(\mathbb{G}_{X\to X'})\right| . \tag{5.116}$$

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 5.31 in Subsection 5.A.5 we obtain that

$$\left|\Gamma(\mathbb{G}_{X\to X'})\right| \leq |X'| , \tag{5.117}$$

which, combined with Equation (5.116), gives Equation (5.114). This concludes the proof. □

We now present the postponed proof of Proposition 5.14, presented in page 97.

Proof of Proposition 5.14. The proof of Proposition 5.14 is a direct consequence of Corollary 5.30.

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 (5.2) is a separated mapping set. We then prove that Problem (5.2) is a Separated Det-Pomdp.

First, for all time t, for all pair $(u, o) \in \mathbb{U} \times \mathbb{O}$, we have $F_t^{u,o} = f_t^u{}_{\overrightarrow{(h_{t+1}^u)^{-1}(o)}}$ (see Equation (5.110)). Thus, by Equation (5.85a), there exists $X \subset \mathbb{X}$ such that $F_t^{u,o} = f_t^u{}_{\overleftarrow{X}}$. Hence, $\mathbb{F}_t^{\mathcal{D}}$ is of the same form as in Equation (5.102), with the role of set Φ_k taken by $\{f_t^{\mathbb{U}}\}$.

We hence have that $\mathbb{F}^{\mathcal{D}} = \bigcup_{t\in\mathcal{T}} \mathbb{F}^{\mathcal{D}}_{0:t}$ is a (∂)-separated mapping set by Corollary 5.30, where the role of $\{\mathbb{G}_k\}_{k\in\mathbb{N}}$ is taken by $\{\mathbb{F}^{\mathcal{D}}_t\}_{t\in\mathcal{T}\setminus\{T\}}$ and the role of $\{\Phi_k\}_{k\in\mathbb{N}}$ is taken by $\{f^{\mathbb{U}}_t\}_{t\in\mathcal{T}\setminus\{T\}}$.

Therefore, as $\mathbb{F}^{\mathcal{D}}$ is a (∂)-separated mapping set, Problem (5.2) is a Separated DET-POMDP. □

We now present examples of Separated DET-POMDPs.

5.A.5.2 Examples of Separated DET-POMDPs

In this part, we present examples of Separated DET-POMDPs. Indeed, a direct consequence of Proposition 5.14 is that, if the evolution functions 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 functions.

Corollary 5.33. *Consider a* DET-POMDP *optimization problem given by Problem* (5.2) *which satisfies the finite sets Assumption 5.1. The notations are those of Problem* (5.2). *Assuming that, for all time $t \in \mathcal{T} \setminus \{T\}$, there exist mappings g_t such that, for all states $x \in \mathbb{X}$,*

$$f_t(x, u) = x + g_t(u) \ , \tag{5.118}$$

then Problem (5.2) *is a Separated* DET-POMDP.

Proof. This corollary is a direct result of Proposition 5.14. Indeed, we only need to prove that $\cup_{t\in\mathcal{T}}\big(f^{\mathbb{U}^{t+1}}_{0:t}\big)$ 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 (5.118),

$$f^{u_{t_1:t'_1}}_{t_1:t'_1} : \mathbb{X} \to \mathbb{X}, x \mapsto x + \sum_{t\in[\![t_1,t'_1]\!]} g_t(u_t) \ ,$$

and

$$f^{u'_{t_2:t'_2}}_{t_2:t'_2} : \mathbb{X} \to \mathbb{X}, x \mapsto x + \sum_{t\in[\![t_2,t'_2]\!]} g_t(u'_t) \ .$$

If there exists a state $x \in \mathbb{X}$ such that $f^{u_{t_1:t'_1}}_{t_1:t'_1}(x) = f^{u'_{t_2:t'_2}}_{t_2:t'_2}(x)$, 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 $f^{u_{t_1:t'_1}}_{t_1:t'_1}(x) = f^{u'_{t_2:t'_2}}_{t_2:t'_2}(x) \Rightarrow f^{u_{t_1:t'_1}}_{t_1:t'_1} = f^{u'_{t_2:t'_2}}_{t_2:t'_2}$. Therefore, the set $\cup_{t\in\mathcal{T}}\big(f^{\mathbb{U}^{t+1}}_{0:t}\big) = \{f^{u_{0:t}}_{0:t} \mid \forall t \in \mathcal{T} \setminus \{T\}, \forall u_{0:t} \in \mathbb{U}^{t+1}\}$ of composition of the evolution functions is a separated mapping set. We conclude by Proposition 5.14 that Problem (5.2) is a Separated DET-POMDP. □

Corollary 5.34. *Consider a* Det-Pomdp *optimization problem given by Problem* (5.2) *which satisfies the finite sets Assumption 5.1. The notations are those of Problem* (5.2). *Assuming that, for all time* $t \in \mathcal{T} \setminus \{T\}$, *there exist mappings* g_t *such that for all states* $x \in \mathbb{X}$,

$$f_t(x,u) = x \times g_t(u) \ , \tag{5.119}$$

and assuming that $0 \notin \mathbb{X}$, *then Problem* (5.2) *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 (5.119),

$$f_{t_1:t_1'}^{u_{t_1:t_1'}} : \mathbb{X} \to \mathbb{X}, x \mapsto x \times \prod_{t \in [\![t_1, t_1']\!]} g_t(u_t) \ ,$$

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 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. □

5.A.5.3 Necessary condition to attain the bound of the cardinality of the set of reachable beliefs

The problem presented in §5.5.3 achieves the bound on the cardinality of the set of reachable beliefs because of one key property: there is a circulation allowed by the dynamics. That property is a necessary condition in order to achieve the bound.

Proposition 5.35. *Assume that Problem* (5.2) *is a Separated* Det-Pomdp, *that Assumption 5.1 holds, that* $|\mathrm{supp}(b_0)| > 1$ *and that the evolution functions* $\{f_t\}_{t \in \mathcal{T} \setminus \{T\}}$ *of Problem* (5.2) *satisfy the following property: there exists a strict subset* $A \subsetneq \mathbb{X}$ *such that for all time* $t \in \mathcal{T} \setminus \{T\}$, $f_t : A \times \mathbb{U} \to A$. *Then, the bound in Equation* (5.38) *presented in Theorem 5.15 cannot be attained, i.e.*

$$\left| \mathbb{B}_{1:T}^{\mathrm{R},\mathcal{D}}(b_0) \right| < 1 + \big(2^{|\mathrm{supp}(b_0)|} - |\mathrm{supp}(b_0)| \big) |\mathbb{X}| \ . \tag{5.120}$$

Proof. We assume that there exists a subset $A \subsetneq \mathbb{X}$ such that, for all time $t \in \mathcal{T} \setminus \{T\}$, $f_t : A \times \mathbb{U} \to A$.

Note that, due to Corollary 5.17, we now need to prove Proposition 5.35 only for the case where $\mathrm{supp}(b_0) \not\subset A$. Indeed, if $\mathrm{supp}(b_0) \subset A$, by Corollary 5.17, we have Equation (5.44), i.e.

$$\left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{1:T}(b_0)\right| \leq 1 + \left(2^{|\mathrm{supp}(b_0)|} - |\mathrm{supp}(b_0)|\right)|A| \; ,$$

hence we have Equation (5.120) as $|A| < |\mathbb{X}|$.

Let $X \subset A$ a given subset of A. Moreover, recall that $\mathbb{F}^{\mathcal{D}}_{X \to X'}$ is defined as

$$\mathbb{F}^{\mathcal{D}}_{X \to X'} = \left\{ F \in \mathbb{F}^{\mathcal{D}} \,\middle|\, F^{-1}(\mathbb{X}) = X, F(X) \subset X' \right\} . \qquad \text{(by (5.113))}$$

We now prove that we have $\mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}} = \mathbb{F}^{\mathcal{D}}_{X \to A}$.

Indeed, due to the assumption, for all time $t \in \mathcal{T} \backslash \{T\}$, $f_t : A \times \mathbb{U} \to A$. Moreover, recall that due to Equation (5.18), for all time t and pair $(u, o) \in \mathbb{U} \times \mathbb{O}$, $F^{u,o}_t(X) \subset f^u_t(X) \cup \{\partial\}$. As $X \subset A$, we have $f^u_t(X) \subset A$. Hence for all $F \in \mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}}$, $A \cap F(X) = \emptyset$, i.e. $F(X) \subset A$. Thus, for all $F \in \mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}}$, $F \in \mathbb{F}^{\mathcal{D}}_{X \to A}$ as we have $F^{-1}(\mathbb{X}) = X$ (as $F \in \mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}}$), and $F(X) \subset A$. We hence have $\mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}} \subset \mathbb{F}^{\mathcal{D}}_{X \to A}$. Combined with the inclusion $\mathbb{F}^{\mathcal{D}}_{X \to A} \subset \mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}}$ (indeed, $F(X) \subset A$ implies $F(X) \subset \mathbb{X}$, hence for all $F \in \mathbb{F}^{\mathcal{D}}_{X \to A}$, $F \in \mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}}$), we indeed obtain

$$\mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}} = \mathbb{F}^{\mathcal{D}}_{X \to A} \; .$$

Moreover, by Lemma 5.32 applied with $\mathbb{G} = \mathbb{F}^{\mathcal{D}}$, $X = X$ and $X' = A$, we have

$$|\mathbb{F}^{\mathcal{D}}_{X \to A}| \leq |A| \; .$$

Hence, we have

$$\forall X \subset A \;, \;\; |\mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}}| = |\mathbb{F}^{\mathcal{D}}_{X \to A}| \leq |A| \qquad (5.121)$$

Now, we have

$$\left|\mathbb{T}^{\mathcal{D}}(b_0) \setminus \{\delta_\partial\}\right| \leq \sum_{k \geq 0} \sum_{\substack{X \subset \mathrm{supp}(b_0) \\ |X| = k}} \left| \left(\mathcal{R} \circ (\mathbb{F}^{\mathcal{D}}_{X \to \mathbb{X}})_\star\right)(b_0) \setminus \{\delta_\partial\} \right| \qquad \text{(by (5.42))}$$

$$\leq |\mathbb{X}| + \sum_{k \geq 2} \Bigg(\sum_{\substack{X \subset \mathrm{supp}(b_0) \\ |X| = k \text{ and } X \not\subset A}} |\mathbb{X}| + \sum_{\substack{X \subset \mathrm{supp}(b_0) \\ |X| = k \text{ and } X \subset A}} |A| \Bigg)$$

(by Equations (5.39) and (5.121))

$$< |\mathbb{X}| + \sum_{k \geq 2} \sum_{\substack{X \subset \mathrm{supp}(b_0) \\ |X| = k}} |\mathbb{X}|$$

$$= |\mathbb{X}| + \left(2^{|\mathrm{supp}(b_0)|} - |\mathrm{supp}(b_0)| - 1\right)|\mathbb{X}| \; . \qquad (5.122)$$

Finally, we have

$$\left|\mathbb{B}^{\mathrm{R},\mathcal{D}}_{1:T}(b_0)\right| \overset{(5.23)}{=} \left|\bigcup_{t=1}^{T}\mathbb{B}^{\mathrm{R},\mathcal{D}}_{t}(b_0)\right| \overset{(5.33)}{=} \left|\bigcup_{t=0}^{T-1}\mathbb{T}^{\mathcal{D}}_{0:t}(b_0)\right| \overset{(5.40)}{=} |\mathbb{T}^{\mathcal{D}}(b_0)| \overset{(5.122)}{<} 1+\left(2^{|\mathrm{supp}(b_0)|}-|\mathrm{supp}(b_0)|\right)|\mathbb{X}| \ ,$$

which ends the proof □

Chapter 6

Multistage optimization of a partially observed petroleum production system

Contents

6.1 Introduction

Chapter 4 was devoted to the formulation and numerical resolution of a deterministic optimization problem for the management of an oil and gas production system (see Problem (4.1)). In that formulation, we considered that oil prices were known in advance (deterministic oil prices) and that the state of the dynamical system modeling the reservoir dynamics was fully observed (i.e. the optimization problem was formulated under a complete observation assumption).

Relaxing the deterministic assumption for prices and assuming that prices are stochastic could be taken into account by dynamic programming. Indeed, it is possible to apply dynamic programming when prices are assumed to be stagewise independent. More complex dynamical models for prices can also be accounted for, as long as those models do not lead to an extended state which is too large. Indeed, dynamic programming is subject to the curse of dimensionality, and extending the state can render the problem untractable.

However, assuming a full observation of the state is often a too demanding assumption. Indeed, the state variables to consider depend on the model structure of the oil reservoir (which is a geological formation that contains some hydrocarbons): it has 5 dimensions per tank in the reservoir (see Table 4.1), although, as was shown in Chapter 4, it could be reduced to a one dimensional state in specific cases. However, it is not perfectly known when starting to exploit the oil and gas production network: for example, we have only estimates of the initial reserve of hydrocarbons and of the size of each tank. We are therefore led to reformulate the optimization Problem (4.1) in order to take into account *partial observation*. We proceed as follows. To start with, we assume that the initial state of the reservoir is not known, but that we have state partial information given by a probability distribution at the initial time. Therefore, at each time step, the evolution of the state is known through probability distribution evolutions driven by observations and controls (which must be functions of the observations). This leads to Problem (6.1).

This chapter is organized as follows. In §6.2, we reformulate the deterministic optimization Problem (4.1) of an oil and gas production network, studied in Chapter 4, in order to take into account the partial observation of the state. This new formulation leads to an optimization problem Problem (6.1) which is a DET-POMDP optimization problem (see Chapter 5). We show in Lemma 6.1 and Proposition 6.2 that it is equivalent to a Separated DET-POMDP optimization problem. Then, in §6.3, following the same steps as in Chapter 4, we present two numerical applications. First, a gas reservoir with two tanks, and second, an oil reservoir where the pressure is kept constant through water injection.

6.2 Management of a partially observed petroleum production system as a Separated DET-POMDP optimization problem

In this section, we present a problem of optimal management of an oil and gas production network under partial observation. It is derived from the petroleum production system optimization problem, presented in Chapter 4 as Problem (4.1). First, in §6.2.1, we present the formulation of the management of a partially observed petroleum production system as DET-POMDP optimization problem as Problem (6.1). Second, in §6.2.2, we detail the observations we have access to in a petroleum production system. Third, in §6.2.3, we show that there exists a Separated DET-POMDP optimization Problem (6.3) equivalent to optimization Problem (6.1).

6.2.1 General formulation as a Det-Pomdp

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

We reformulate Problem (4.1) to take into account partial observation of the state:

$$\mathcal{V}^\star(b_0) = \max_{\boldsymbol{X},\boldsymbol{O},\boldsymbol{U}} \mathbb{E}\Big[\sum_{t=0}^{T-1} \mathcal{L}_t(\boldsymbol{X}_t, \boldsymbol{U}_t) + \mathcal{K}(\boldsymbol{X}_T)\Big] \tag{6.1a}$$

$$s.t.\ \ \mathbb{P}_{\boldsymbol{X}_0} = b_0 \ , \tag{6.1b}$$

$$\boldsymbol{X}_{t+1} = f(\boldsymbol{X}_t, \boldsymbol{U}_t) \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , \tag{6.1c}$$

$$\boldsymbol{O}_t = h(\boldsymbol{X}_t) \ , \ \ \forall t \in \mathcal{T} \ , \tag{6.1d}$$

$$\boldsymbol{U}_t \in \mathcal{U}_t^{ad}(\boldsymbol{X}_t) \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , \tag{6.1e}$$

$$\sigma(\boldsymbol{U}_t) \subset \sigma(\boldsymbol{O}_0, \dots, \boldsymbol{O}_t, \boldsymbol{U}_0, \dots, \boldsymbol{U}_{t-1}) \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ . \tag{6.1f}$$

The sets of time-steps, $\mathcal{T}$, states, $\mathbb{X}$, and controls, $\mathbb{U}$, are the same in both Problem (6.1) and Problem (4.1). Problem (6.1) requires an additional set, the set $\mathbb{O}$ of observation (which is the focus of §6.2.2).

Now, as we have partial observation of the state of the dynamical system in Problem (6.1), there are now three stochastic processes in the formulation: $\boldsymbol{X} = \big\{\boldsymbol{X}_t\big\}_{t\in\mathcal{T}}$, $\boldsymbol{U} = \big\{\boldsymbol{U}_t\big\}_{t\in\mathcal{T}\setminus\{T\}}$ 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 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.

We now detail the constraints of Problem (6.1).

First, as Problem (6.1) is derived from Problem (4.1), they have equations in common. Indeed, the objective functions (Equations (6.1a) and (4.1a), up to an expectation), the system dynamics (Equations (6.1c) and (4.1c)) and the admissibility constraints (Equations (6.1e) and (4.1d)) are shared by the two problems.

Second, we do not know the initial state of the reservoir, but we assume that we know its distribution, b_0 (a probability distribution over the state space), which we call the *initial belief*. This leads to Equation (6.1b).

Third, the observations of the system are given by Equation (6.1d) as a function of the state. There is an observation function $h : \mathbb{X} \to \mathbb{O}$ such that the observation at time $t \in \mathcal{T}$, when the system is in state x, is given by $h(x)$. Contrarily to general Det-Pomdp optimization Problem (5.2), the observation function does not depends on the controls.

Fourth, we have a measurability constraint on the controls in Equation (6.1f), which takes into account that the controls are constrained to depend only on past observations and past controls (the controls are non-anticipative).

As Problem (6.1) has the similar form as Problem (5.2), it is hence a DET-POMDP optimization problem.

We now detail the observation sets and functions in the case of management of the petroleum production system.

6.2.2 Observations sets and observation functions of the petroleum production system

In §6.2.2.1, we first make some recalls on petroleum production systems. Second, in §6.2.2.2, we derive the associated observation sets and observation functions.

6.2.2.1 Recalls on petroleum production systems

As stated in Chapter 4, a petroleum production system is composed of two components: a reservoir, and production assets. From the physical reservoir models, we obtain a dynamical system that describes the reservoir time evolution, and whose controls are derived from the models of the production assets.

The reservoir is modeled as a set of possibly interconnected tanks. Each tank is modeled as a dynamical system, with a 5 dimensional state, $(V^{\mathrm{O}}, V^{\mathrm{G}}, V^{\mathrm{W}}, V^{\mathrm{P}}, P^{\mathrm{R}})$, whose components are respectively the standard volume of oil, gas, and water in the tank, the tank's total pore volume, and the tank's pressure. In some specific cases, the dimension of the state can be reduced, as was shown in the applications of Chapter 4. Moreover, tanks can be interconnected, and the (constant) transmissivity of the connections can be considered to be only known through a probability distribution. In that case, the transmissivity also becomes a component of the state.

There are three main production assets: wells, chokes, and pipes. Fluids leave a tank through a well, and then flow in a pipe unless a choke closes the pipe. The fluids exit the production system through the exit point, and they are then sold. We represent the production network with a graph $\mathcal{G} = (\mathbb{V}, \mathbb{A})$, where $\mathbb{V}$ is the set of vertices, and $\mathbb{A} \subset \mathbb{V}^2$ is the set of arcs (see Figure 4.1 in Chapter 4). The control u is the opening or closing of pipes $o_a, a \in \mathbb{A}$, and the choice of the well-head pressure $P_w, w \in \mathbb{V}_{in} \subset \mathbb{V}$.

From the wells models, we derive a production function $\Phi : \mathbb{X} \times \mathbb{U} \to \mathbb{R}^3$ (see Equation (4.16), §4.A), which is used in the cost function and the dynamics of the system (see Equation (4.3)). The admissibility set presented in Equation (6.1e) is derived thanks to the models of the pipes and chokes assets.

6.2.2.2 Observations sets and functions of the petroleum production system

The state of the reservoir is not observed. However, we obtain information about the state as we produce hydrocarbons. Indeed, we assume that the production rates are known and that we know the gains at each time step. More specifically, we assume that, at each time step $t \in \mathcal{T} \setminus \{T\}$, the decision maker knows the resulting production values of each well of any admissible control $u \in \mathbb{U}$ at the time of choosing the said control. We assume that this is the only new information the decision maker has access to at time t.

Following this assumption, we only have observations for tanks connected to wells. We have no observation regarding unconnected tanks, nor regarding transmissivity. As an example, we now detail the observation for the case with one well and one tank. More general cases can easily be deduced by increasing the dimensions of the observation and state.

The production mapping for a case with one well and one tank has been given in §4.A. It is given by a mapping $\Phi : \mathbb{X} \times \mathbb{U} \to \mathbb{R}^3$ detailed in Equation (4.17), and given by

$$F^i = \Phi^i(x,u) = \frac{\textsc{Ipr}^i\big(u, x^{(5)}, \frac{x^{(3)} B_{\text{W}}(x^{(5)})}{x^{(4)}}, \frac{x^{(2)} B_{\text{G}}(x^{(5)})}{x^{(4)}}\big)}{B_i(x^{(5)})} \ , \ \ \forall i \in \{\text{O}, \text{G}, \text{W}\} \ ,$$

where IPR stands for *Inflow Performance Relationship* of the well, defined in §3.2.4. The values of the production mapping over time are observed and constitute what we call the *natural observation*. We show now that it exists another equivalent observation function.

For that purpose, we recall that, as described in Equation (3.12), the Inflow Performance Relationship IPR used to define the production mapping can be reformulated as a mapping $\widetilde{\textsc{Ipr}}$. This second mapping takes as input two functions, $\text{W}^{\text{CT}} : \mathbb{X} \to \mathbb{R}$ and $\text{G}^{\text{OR}} : \mathbb{X} \to \mathbb{R}$ which return respectively the water-cut – proportion of water produced when we extract a volume of liquid (see Equation (3.8)) – and the gas-oil ratio – proportion of gas produced when we extract a volume of oil (see Equation (3.10)). Moreover, it is assumed that the two ratios only depend on the state. Thus, there exists a mapping $\widetilde{\Phi} : \mathbb{R}^3 \times \mathbb{U} \to \mathbb{R}^3$, defined thanks to the mappings $\widetilde{\textsc{Ipr}}$ (following Equation (4.17)), such that

$$\Phi(x,u) = \widetilde{\Phi}\big(\text{P}^{\text{Res}}(x), \text{W}^{\text{CT}}(x), \text{G}^{\text{OR}}(x), u\big) \ , \ \ \forall (x,u) \in \mathbb{X} \times \mathbb{U} \ ,$$

with $\text{P}^{\text{Res}} : \mathbb{X} \to \mathbb{R}, x \mapsto x^{(5)}$. Moreover, for a given bottom-hole pressure P^{BH} (i.e. for a given control u), $\widetilde{\textsc{Ipr}}(P^{\text{BH}}, \cdot)$ is invertible (see §3.2.4), hence $\widetilde{\Phi}(\cdot, u)$ (using the partial mapping notation (5.6)) is invertible.

Thus, while the natural observations are given by the production function, we instead use as observations the vector $(P^{\text{R}}, w^{\text{CT}}, g^{\text{OR}})$. The observation function h is given by

$$h : \mathbb{X} \to \mathbb{O}, x \mapsto \begin{pmatrix} \text{P}^{\text{Res}}(x) \\ \text{W}^{\text{CT}}(x) \\ \text{G}^{\text{OR}}(x) \end{pmatrix} \ . \tag{6.2}$$

Now that we have specified the observations, we study whether Problem (6.1) is a Separated DET-POMDP optimization problem or not.

6.2.3 Management of a partially observed petroleum production system and Separated Det-Pomdp

Problem (6.1) is a Det-Pomdp optimization problem, but is not necessarily a Separated Det-Pomdp optimization problem. Fortunately, by adding a key assumption that is satisfied in practice, there exists a Separated Det-Pomdp optimization problem which is equivalent to Problem (6.1).

Assumption 6.1 (Initial consistence assumption)**.** *We assume that the initial belief $b_0 \in \mathbb{B}$ is consistent: all the states $x \in \mathbb{X}$ which belong to the support of b_0 share the same observation, i.e.*

$$\exists o \in \mathbb{O}\ , \quad \textit{such that} \quad h\big(\mathrm{supp}(b_0)\big) = \{o\}\ ,$$

where supp *is the support function, defined in Equation* (5.5)*.*

Under the initial consistence assumption, the observation $\boldsymbol{O}_0 = h(\boldsymbol{X}_0)$ at initial time is a constant random variable, as, using Constraints (6.1b) and (6.1d), we have that $\boldsymbol{O}_0 \in h\big(\mathrm{supp}(\mathbb{P}_{\boldsymbol{X}_0})\big) = h\big(\mathrm{supp}(b_0)\big)$ and the set $h\big(\mathrm{supp}(b_0)\big)$ is a singleton. Moreover, the decision maker has at his disposal at the initial time an initial belief b_0 and a first observation o_0 resulting from geo-physicists measures. The observation o_0 implies that the unknown (deterministic) initial state must belong to the set $h^{-1}(o_0)$ and it is thus "reasonable" for the decision maker to replace his initial belief b_0 with the new belief $(b_0)_{|h^{-1}(o_0)}/b_0(h^{-1}(o_0))$ that is the conditional probability of $\boldsymbol{X}_0$ knowing that $\boldsymbol{X}_0 \in h^{-1}(o_0)$. Indeed, this operation is not possible if $b_0(h^{-1}(o_0)) \neq 0$, but this would mean that the decision maker's initial belief and initial observation are incompatible. Then, the new belief satisfies the initial consistence assumption.

First, we present Problem (6.3). It has the same form as Problem (6.8), found in Appendix 6.A:

$$\mathcal{J}^\star(b_0, o_0) = \min_{\boldsymbol{X},\boldsymbol{O},\boldsymbol{U}} \mathbb{E}\Big[\sum_{t=0}^{T-1} \mathcal{L}_t(\boldsymbol{X}_t, \boldsymbol{U}_t) + \mathcal{K}(\boldsymbol{X}_T)\Big] \tag{6.3a}$$

$$s.t.\ \ \mathbb{P}_{\boldsymbol{X}_0} = b_0\ , \tag{6.3b}$$

$$\boldsymbol{X}_{t+1} = f(\boldsymbol{X}_t, \boldsymbol{U}_t)\ ,\ \ \forall t \in \mathcal{T} \setminus \{T\}\ , \tag{6.3c}$$

$$\boldsymbol{O}_0 = o_0 \tag{6.3d}$$

$$\mathrm{supp}(b_0) \subset h^{-1}(o_0) \tag{6.3e}$$

$$\boldsymbol{O}_t = h(\boldsymbol{X}_t)\ ,\ \ \forall t \in \mathcal{T} \setminus \{0, T\}\ , \tag{6.3f}$$

$$\boldsymbol{U}_t \in \mathcal{U}_t^{ad}(\boldsymbol{X}_t)\ ,\ \ \forall t \in \mathcal{T} \setminus \{T\}\ , \tag{6.3g}$$

$$\sigma(\boldsymbol{U}_t) \subset \sigma(\boldsymbol{O}_0, \dots, \boldsymbol{O}_t, \boldsymbol{U}_0, \dots, \boldsymbol{U}_{t-1})\ ,\ \ \forall t \in \mathcal{T} \setminus \{T\}\ . \tag{6.3h}$$

Note that when Constraint (6.3e) is not satisfied, then Problem (6.3) takes the value $+\infty$. Moreover, Problem (6.3) shares most of its constraints and its objective function with Problem (6.1). Indeed, the only change is that Constraint (6.1d) at time $t = 0$ in Problem (6.1) is replaced with Constraints (6.3d) and (6.3e) in Problem (6.3).

Lemma 6.1. *Assume that the initial belief b_0 satisfies the initial consistence Assumption 6.1. Then, the value of Problem (6.3), $\mathcal{J}^\star(b_0, o_0)$, equals $+\infty$ if the set $h\big(\mathrm{supp}(b_0)\big)$ – which is a singleton by Assumption 6.1 – is not the singleton $\{o_0\}$. If the observation o_0 in Constraint (6.3d) satisfies $\{o_0\} = h\big(\mathrm{supp}(b_0)\big)$, then Problem (6.1) and Problem (6.3) are equivalent. That is, Problem (6.1) and Problem (6.3) share the same sets of reachable beliefs, and $\mathcal{V}^\star(b_0) = \mathcal{J}^\star(b_0, o_0)$.*

Proof. First, if the observation o_0 and initial belief b_0 are such that $h\big(\mathrm{supp}(b_0)\big) \neq \{o_0\}$, then Constraint (6.3e) of Problem (6.3) is not satisfied. Problem (6.3) is therefore not feasible, and its value is hence equal by convention to $+\infty$.

Second, we prove that if $h\big(\mathrm{supp}(b_0)\big) = \{o_0\}$, Problem (6.1) and Problem (6.3) are equivalent. As Problem (6.1) and Problem (6.3) share the same evolution functions and, for all time $t > 0$, the same observation functions, they share the same belief dynamics $\{f_t\}_{t\in\mathcal{T}\setminus\{T\}}$. Indeed the sequence of mappings $\{\tau_t\}_{t\in\mathcal{T}\setminus\{T\}}$, defined in Equation (5.14), does not depend on the observation function at time 0 but only depends on the sequences of mappings $\{f_t\}_{t\in\mathcal{T}\setminus\{T\}}$ and $\{h_t\}_{t\in\mathcal{T}\setminus\{0\}}$. As the sets of reachable beliefs are constructed by applying the belief dynamics $\{f_t\}_{t\in\mathcal{T}\setminus\{T\}}$ on the initial belief b_0 (see the definition of $\mathbb{B}_t^{\mathrm{R},\mathcal{D}}$ in Equation (5.22)), Problem (6.1) and Problem (6.3) thus share the same sets of reachable beliefs as they apply the same belief dynamics $\{f_t\}_{t\in\mathcal{T}\setminus\{T\}}$ to the same initial belief b_0.

Under the initial consistence Assumption 6.1, we have that $\{o_0\} = h(\mathrm{supp}(b_0))$. Hence, for that observation o_0 and initial belief b_0, Constraints (6.3d) and (6.3e) are satisfied in Problem (6.3), and we have $\boldsymbol{O}_0 = o_0$ in Problem (6.3). Furthermore, by Constraint (6.1b), which gives the probability distribution of the initial state, and Constraint (6.1d) at time $t = 0$, which gives the observation as a function of the state, we also have $\boldsymbol{O}_0 = o_0$ in Problem (6.1). As the rest of the constraints and the objective functions are shared by Problem (6.1) and Problem (6.3), and since the stochastic processes $\boldsymbol{X}$ and $\boldsymbol{O}$ have the same initialization in both problems, we have $\mathcal{V}^\star(b_0) = \mathcal{J}^\star(b_0, o_0)$.

This ends the proof. □

Moreover, Problem (6.3) is a Separated DET-POMDP optimization problem, as we show below.

Proposition 6.2. *For all ordered pairs $(b_0, o_0) \in \mathbb{B} \times \mathbb{O}$ in the effective domain of $\mathcal{J}^\star$, Problem (6.3) is a Separated* DET-POMDP *optimization problem.*

Proof. The proof is organized as follows. First, we detail the form of the mappings F for the oil and gas case. Second, we prove that Problem (6.3) is a Separated DET-POMDP optimization problem thanks to Lemma 6.8 proved in Appendix 6.A.

We detail the mappings F as defined in the proof of Theorem 5.9. We consider an extended state space $\overline{\mathbb{X}} = \mathbb{X} \cup \{\partial\}$, and define, for all pair of controls and observations $(u, o) \in \mathbb{U} \times \mathbb{O}$, a new mapping $F^{u,o}$ by

$$F^{u,o} : \overline{\mathbb{X}} \to \overline{\mathbb{X}}, x \mapsto \begin{cases} f(x,u) \text{ if } h(f(x,u)) = o, \\ \partial \text{ otherwise.} \end{cases}$$

We recall that the expression of the mapping f was given in Equation (4.3), using the mapping Φ. We instead use the mapping $\widetilde{\Phi} : \mathbb{O} \times \mathbb{U} \to \mathbb{R}^3$, presented in §6.2.2.2. Moreover, we have $x^{(5)} = h^{(1)}(x)$ for all $x \in \mathbb{X}$. Hence, for all pairs of states and controls $(x, u) \in \mathbb{X} \times \mathbb{U}$, we have that $f^{(5)}(x, u) = h^{(1)}\big(f(x, u)\big)$.

For all pairs of controls and observations $(u, o) \in \mathbb{U} \times \mathbb{O}$, the mapping $F^{u,o}$ is hence given by:

$$F^{u,o} : x \mapsto \begin{cases} \begin{pmatrix} x^{(1)} - \widetilde{\Phi}^{(1)}(h(x), u) \\ x^{(2)} - \widetilde{\Phi}^{(2)}(h(x), u) + x^{(1)} R_s(x^{(5)}) - \\ \qquad \big(x^{(1)} - \widetilde{\Phi}^{(1)}(h(x), u)\big) R_s\big(o^{(1)}\big) \\ x^{(3)} - \widetilde{\Phi}^{(3)}(h(x), u) \\ x^{(4)}\big(1 + c_f(o^{(1)} - x^{(5)})\big) \\ o^{(1)} \end{pmatrix} & \text{if } h(f(x, u)) = o, \\ \partial \text{ otherwise.} \end{cases} \tag{6.4}$$

Hence, by Lemma 6.8 (where the roles of g_1, g_2 and g_3 are taken by $-\tilde{\Phi}^{(1)}$, $-\tilde{\Phi}^{(2)}$ and $-\tilde{\Phi}^{(3)}$ respectively, m by R_s, and a by c_f), Problem (6.3) is indeed a Separated Det-Pomdp optimization problem. □

6.3 Numerical applications

We now present two numerical applications, derived from the applications presented in Chapter 4. The first application, in §6.3.1, is a gas reservoir that can be modeled with two interconnected tanks. The second application, in §6.3.2, is an oil reservoir where pressure is kept constant through water injection.

In the following, we assume that the finite set Assumption 5.1 and the initial consistence Assumption 6.1 hold true for all the numerical applications. By Proposition 5.7, we use Algorithm 4 to solve the numerical applications. The following results were performed on a computer equipped with a Core i7-12700KF and 64 GB of memory, using Julia v1.7.3.

6.3.1 First application: a gas reservoir with one well

We consider the case of a gas reservoir with one well and two tanks, as illustrated in Figure 6.1. The deterministic version of that problem was treated in §4.4.1. We now consider that we do not know the initial content of the reservoir, nor do we know the transmissivity between the two tanks.

Formulation First, we recall that the control of the gas reservoir case is the bottom-hole pressure P of the well. Second, as we only produce gas, the observation is the pressure in the tank connected to the well, i.e. the pressure $P^{\text{R},1\mathbf{T}}$ in the first tank.

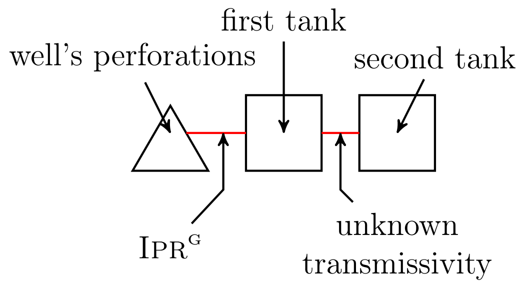


Figure 6.1: Representation of the gas reservoir

Third, the state has 5 dimensions: $x = (V^{\text{G},1\mathbf{T}}, V^{\text{G},2\mathbf{T}}, V_0^{\text{P},1\mathbf{T}}, V_0^{\text{P},2\mathbf{T}}, Ts)$, with $V^{\text{G},1\mathbf{T}}$ and $V^{\text{G},2\mathbf{T}}$ the total volume of gas in each of the two tanks, $V_0^{\text{P},1\mathbf{T}}$ and $V_0^{\text{P},2\mathbf{T}}$ the initial pore volume in each of the two tanks, and Ts the transmissivity parameter. Indeed, we recall that, when the content of the reservoir is known, the state can be reduced to two dimensions: the total volume of gas in each of the two tanks, $(V^{\text{G},1\mathbf{T}}, V^{\text{G},2\mathbf{T}})$, and, as presented in §4.B, there exist two mappings $\Psi_{1\mathbf{T}}$ and $\Psi_{2\mathbf{T}}$ such that the pressure in the two tanks is given by $P^{\text{R},1\mathbf{T}} = \Psi_{1\mathbf{T}}(V^{\text{G},1\mathbf{T}})$ and $P^{\text{R},2\mathbf{T}} = \Psi_{2\mathbf{T}}(V^{\text{G},2\mathbf{T}})$ (see Equation (4.24)). However, those mappings depend on the unknown initial pore volume of each tank, $V_0^{\text{P},1\mathbf{T}}$ and $V_0^{\text{P},2\mathbf{T}}$, which is why we instead need to use mappings $\tilde{\Psi}_{1\mathbf{T}} : \mathbb{R}^2 \to \mathbb{R}$ and $\tilde{\Psi}_{2\mathbf{T}} : \mathbb{R}^2 \to \mathbb{R}$, such that $P^{\text{R},1\mathbf{T}} = \Psi_{1\mathbf{T}}(V^{\text{G},1\mathbf{T}}, V_0^{\text{P},1\mathbf{T}})$ and $P^{\text{R},2\mathbf{T}} = \Psi_{2\mathbf{T}}(V^{\text{G},2\mathbf{T}}, V_0^{\text{P},2\mathbf{T}})$. Finally, the evolution functions take into account exchanges between the two tanks using mapping $\Theta : \mathbb{R}^3 \to \mathbb{R}$, which takes as input the reservoir pressure in the two tanks $P^{\text{R},1\mathbf{T}}$ and $P^{\text{R},2\mathbf{T}}$, and the unknown transmissivity parameter Ts. We hence extend the state with three stationary values in the partially observed case: $V^{\text{P},1\mathbf{T}}$, $V^{\text{P},2\mathbf{T}}$ and Ts.

Fourth, the observation function $h : \mathbb{X} \to \mathbb{O}$ is the projection on the first component, i.e. $h : (P^{\text{R},1\mathbf{T}}, V^{\text{G},2\mathbf{T}}, V^{\text{P},1\mathbf{T}}, V^{\text{P},2\mathbf{T}}, Ts) \mapsto P^{\text{R},1\mathbf{T}}$.

The full formulation follows

$$
\begin{aligned}
\max \quad & \mathbb{E}\left[\sum_{t=0}^{T-1} \rho^t r_t \boldsymbol{F}_t^{\mathrm{G}}\right] && (6.5a)\\
s.t. \quad & \mathbb{P}_{\left(\boldsymbol{V}_0^{\mathrm{G},1\mathbf{T}}, \boldsymbol{V}_0^{\mathrm{G},2\mathbf{T}}, \boldsymbol{V}_0^{\mathrm{P},1\mathbf{T}}, \boldsymbol{V}_0^{\mathrm{P},2\mathbf{T}}, \boldsymbol{Ts}_0\right)} = b_0 \ , && (6.5b)\\
& \boldsymbol{P}_t^{\mathrm{R},1\mathbf{T}} = \tilde{\Psi}_{1\mathbf{T}}(\boldsymbol{V}_t^{\mathrm{G},1\mathbf{T}}, \boldsymbol{V}_0^{\mathrm{P},1\mathbf{T}}) \ , \ \ \forall t \in \mathcal{T} \ , && (6.5c)\\
& \boldsymbol{P}_t^{\mathrm{R},2\mathbf{T}} = \tilde{\Psi}_{2\mathbf{T}}(\boldsymbol{V}_t^{\mathrm{G},2\mathbf{T}}, \boldsymbol{V}_0^{\mathrm{P},2\mathbf{T}}) \ , \ \ \forall t \in \mathcal{T} \ , && (6.5d)\\
& \boldsymbol{O}_t = h(\boldsymbol{P}_t^{\mathrm{R},1\mathbf{T}}, \boldsymbol{V}_t^{\mathrm{G},2\mathbf{T}}, \boldsymbol{V}_0^{\mathrm{P},1\mathbf{T}}, \boldsymbol{V}_0^{\mathrm{P},2\mathbf{T}}, \boldsymbol{Ts}_0) \ , \ \ \forall t \in \mathcal{T} \ , && (6.5e)\\
& \boldsymbol{F}_t^{\mathrm{G}} = \frac{\mathrm{I{\scriptsize PR}}^{\mathrm{G}}\left(\boldsymbol{O}_t - \boldsymbol{P}_t\right)}{B_{\mathrm{G}}(\boldsymbol{O}_t)} \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , && (6.5f)\\
& \boldsymbol{F}_t^{\mathrm{R}} = \Theta(\boldsymbol{P}_t^{\mathrm{R},1\mathbf{T}}, \boldsymbol{P}_t^{\mathrm{R},2\mathbf{T}}, \boldsymbol{Ts}_0) \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , && (6.5g)\\
& \boldsymbol{V}_{t+1}^{\mathrm{G},1\mathbf{T}} = \boldsymbol{V}_t^{\mathrm{G},1\mathbf{T}} - \boldsymbol{F}_t^{\mathrm{G}} + \boldsymbol{F}_t^{\mathrm{R}} \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , && (6.5h)\\
& \boldsymbol{V}_{t+1}^{\mathrm{G},2\mathbf{T}} = \boldsymbol{V}_t^{\mathrm{G},2\mathbf{T}} - \boldsymbol{F}_t^{\mathrm{R}} \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , && (6.5i)\\
& \boldsymbol{F}_t^{\mathrm{G}} \geq 0 \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , && (6.5j)\\
& \boldsymbol{P}_t \geq 0 \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , && (6.5k)\\
& \sigma(\boldsymbol{P}_t) \subset \sigma(\boldsymbol{O}_0, \dots, \boldsymbol{O}_t, \boldsymbol{P}_0, \dots, \boldsymbol{P}_{t-1}) \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ . && (6.5l)
\end{aligned}
$$

Characteristic of the gas reservoir application. Once again, we consider that, as in §4.4.1, the revenue per volume of gas is the historical gas spot price of TTF (Netherlands gas market) from 2006 to 2020, and we do not consider any operational costs.

We now detail the rest of the data.

- The possible controls are $u \in \{0, 100, 200, 300, 400, 460\}$ (in Bar).
- The possible observations are $o \in \{o^{(i)} \mid o^{(i)} = 10 \times i, i \in [\![0, 46]\!]\}$ (in Bar).
- There are 173 monthly time steps.
- We consider that there are 27 possible initial states, with three possible values of the initial pore volume for each tank and three possible values for the transmissivity, while the volume of gas in each tank is such that the reservoir pressure in the two tanks is of 480 Bar, i.e. $\boldsymbol{V}_0^{\mathrm{G},1\mathbf{T}} = \big(\tilde{\Psi}_{1\mathbf{T}}^{\boldsymbol{V}_0^{\mathrm{P},1\mathbf{T}}}\big)^{-1}(480)$ and $\boldsymbol{V}_0^{\mathrm{G},2\mathbf{T}} = \big(\tilde{\Psi}_{2\mathbf{T}}^{\boldsymbol{V}_0^{\mathrm{P},2\mathbf{T}}}\big)^{-1}(480)$ (using the partial mapping notation (5.6), the mappings $\tilde{\Psi}_{i\mathbf{T}}^{\boldsymbol{V}^{\mathrm{P},i\mathbf{T}}}$ are strictly increasing on $\mathbb{R}_+$, hence invertible).
- The values of the transmissivity are $Ts \in \{1.87, 18.71, 112.26\}$ (in m^3/Bar/months), with a finite probability distribution $p = [0.1, 0.8, 0.1]$. The distribution of the transmissivity is $\mathcal{L}_{\boldsymbol{Ts}} = \sum_{i \in [\![1,3]\!]} p_i \delta_{\boldsymbol{Ts}_i}$

- The values of the initial pore volume of the first tank are $V^{\mathrm{P},1\mathbf{T}} \in \{1200, 1600, 2000\}$ (in Mm^3), with a finite probability distribution $p = [0.1, 0.8, 0.1]$. The distribution of the initial pore volume of the first tank is $\mathcal{L}_{\boldsymbol{V}^{\mathrm{P},1\mathbf{T}}} = \sum_{i \in [\![1,3]\!]} p_i \delta_{\boldsymbol{V}_i^{\mathrm{P},1\mathbf{T}}}$

- The values of the initial pore volume of the second tank are $V^{\mathrm{P},2\mathbf{T}} \in \{1181.25, 1575, 1968.75\}$ (in Mm^3), with a finite probability distribution $p = [0.1, 0.8, 0.1]$. The distribution of the initial pore volume of the second tank is $\mathcal{L}_{\boldsymbol{V}^{\mathrm{P},2\mathbf{T}}} = \sum_{i \in [\![1,3]\!]} p_i \delta_{\boldsymbol{V}_i^{\mathrm{P},2\mathbf{T}}}$

- We consider that there are 200 possible values of the volume of gas in each tank for each value of the initial pore volume, corresponding to pressure ranging from 0 Bar to 480 Bar.

The distribution of the initial state is the product of the law of each of its components, as we assume they are independent, i.e. $\mathcal{L}_{\boldsymbol{X}_0} = \mathcal{L}_{\boldsymbol{V}^{\mathrm{P},1\mathbf{T}}} \otimes \mathcal{L}_{\boldsymbol{V}^{\mathrm{P},2\mathbf{T}}} \otimes \mathcal{L}_{\boldsymbol{Ts}}$.

Optimization of the production We use Algorithm 4 to solve Problem (6.5). We hence obtain an optimal policy $\{\pi_t^\star : \mathbb{B} \to \mathbb{U}\}_{t \in \mathcal{T} \setminus \{T\}}$. As it is not possible to have a descriptive drawing of the policy on the set of reachable beliefs, we instead computed the trajectories for each of the possible initial states x_0 consistent with the initial belief (i.e. such that $x_0 \in \mathrm{supp}(b_0)$). Indeed, the evolution of the beliefs depends on the sequence of observations $\{o_t\}_{\mathcal{T}}$. In order to simulate the observation process $\{o_t\}_{t \in \mathcal{T} \setminus \{T\}}$, we compute the evolution of the state process when applying the optimal policy. A trajectory is thus computed by the following induction. First, we take a pair (b_0, x_0), with $x_0 \in \mathrm{supp}(b_0)$. Then, for all time $t \in \mathcal{T} \setminus \{T\}$, we get

$$
\begin{aligned}
b_{t+1} &= \tau\Big(b_t, \pi_t^\star(b_t), h\Big(f\big(x_t, \pi_t^\star(b_t)\big)\Big)\Big) \; , \\
x_{t+1} &= f\big(x_t, \pi_t^\star(b_t)\big) \; .
\end{aligned}
$$

Note that this method could also be applied with an initial state $x_0 \notin \mathrm{supp}(b_0)$. However, there is no guarantee that the evolution of the observation would be consistent with the belief: we would then attain the cemetery state δ_∂, and be unable to compute a new control. In this section, we only study cases where the initial state is consistent with the initial belief.

We represent the evolution over time of relevant parameters along the trajectories of each of the possible initial states as listed in Equation (6.6) in Figures 6.2, 6.3, 6.4, 6.5, 6.6 and 6.7, where the states $x_{0,n}$ correspond to

$$
x_{0,i+3\times(j-1)+9\times(k-1)} = \Big(\big(\tilde{\Psi}_{1\mathbf{T}}^{V_k^{\mathrm{P},1\mathbf{T}}}\big)^{-1}(480), \big(\tilde{\Psi}_{2\mathbf{T}}^{V_j^{\mathrm{P},2\mathbf{T}}}\big)^{-1}(480), V_k^{\mathrm{P},1\mathbf{T}}, V_j^{\mathrm{P},2\mathbf{T}}, Ts_i\Big) \; . \tag{6.6}
$$

First, in Figure 6.2, we draw the evolution over time of the observation process (that is, the pressure in the first tank) for each of the possible initial states as listed in Equation (6.6). Note that no possible initial state is distinguishable from the others thanks to

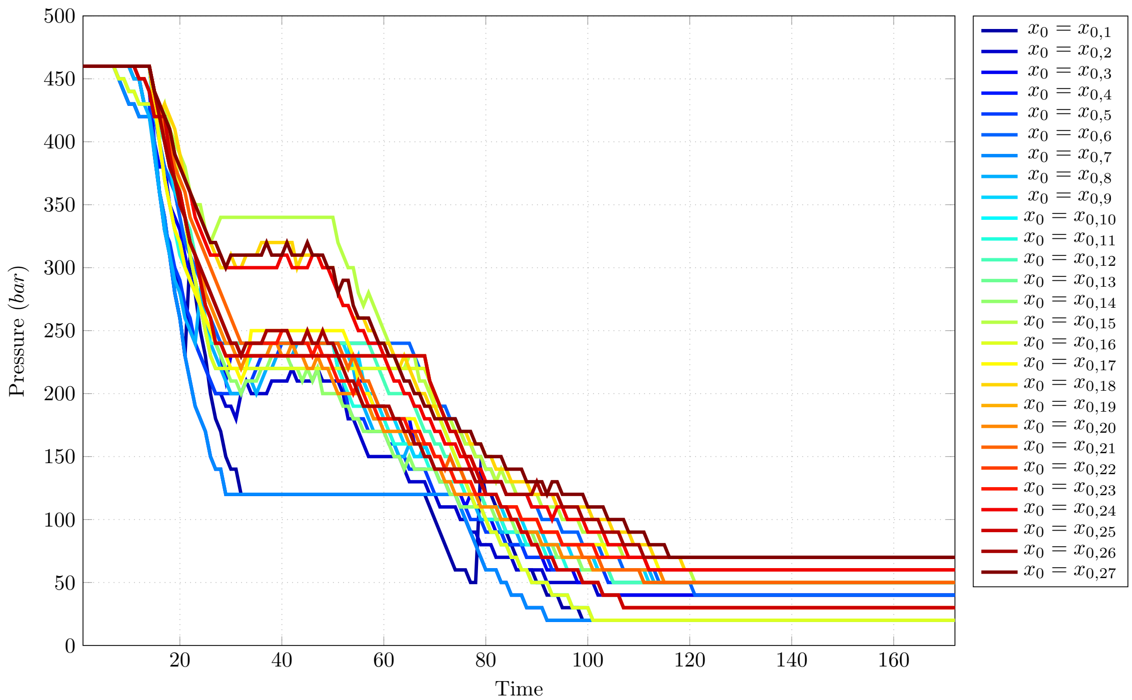


FIGURE 6.2: Evolution over time of the observation (i.e. the pressure in the first tank) when applying the optimal policy for all the possible initial states as listed in Equation (6.6). Each color represents a given initial state

the observation until the 9th month, as all observations curves coincide. Moreover, all pairs of observation curves have separated at least once by the 50th month, thus all initial states have been distinguished thanks to the observation at that date. Hence, by the middle of the optimization horizon, we have obtained enough information to identify the initial state. This is consistent with experience feedback from real production networks.

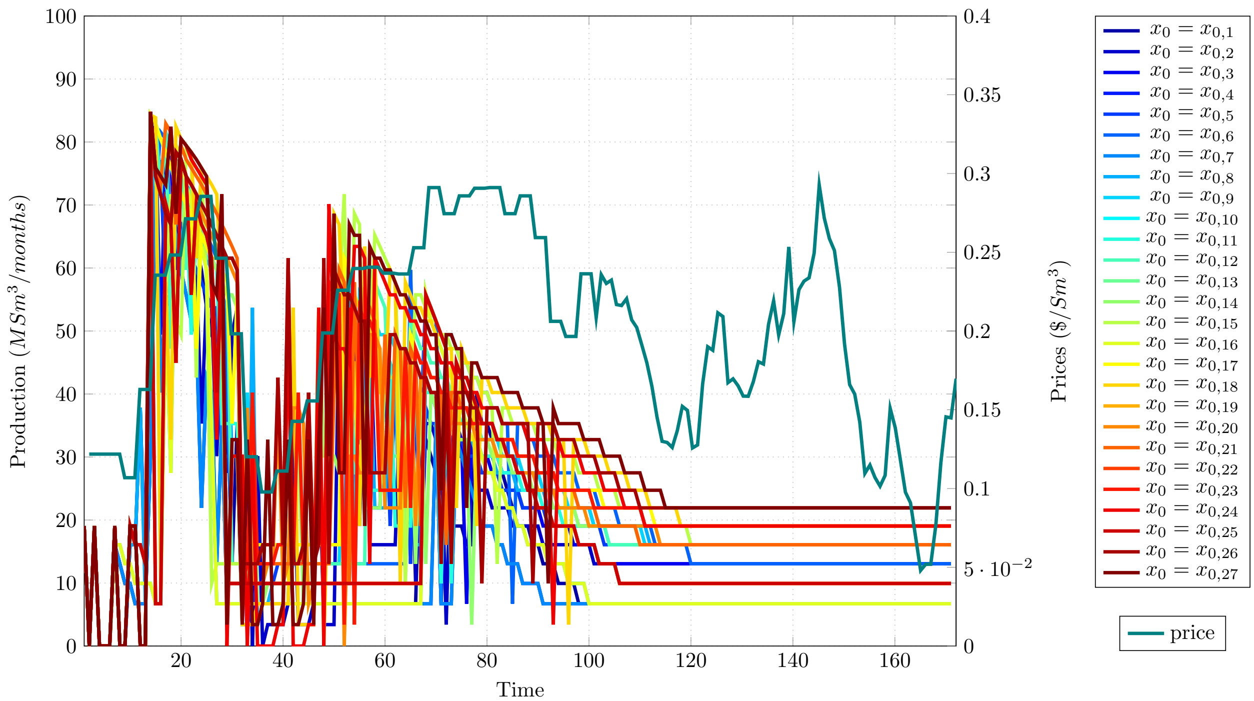


FIGURE 6.3: Evolution over time of the production (i.e. the control) when applying the optimal policy for all the possible initial states as listed in Equation (6.6), with the price curve in dark green. Each other color represents a given initial state

Second, in Figure 6.3, we draw the evolution over time of the production (that is, an image of the controls) for each of the possible initial states as listed in Equation (6.6). Note that there is a small production at the beginning in order to obtain information (before the 8th month) even though prices are less interesting than at a later date. The rest of the policy is more similar to the deterministic case (see §4.4.1.2), as we tend to produce more when prices are high.

Third, in Figure 6.4, we draw the evolution over time of the pressure in the two tanks (i.e. an image of the states which is more comparable between the initial states) for each of the possible initial states as listed in Equation (6.6). It illustrates how the second tank replenishes the first tank.

Fourth, in Figure 6.5, we draw the evolution over time of the cumulated production and the total volume of gas present in the reservoir for each of the possible initial states as listed in Equation (6.6). It illustrates how the trajectory of the cumulated production may

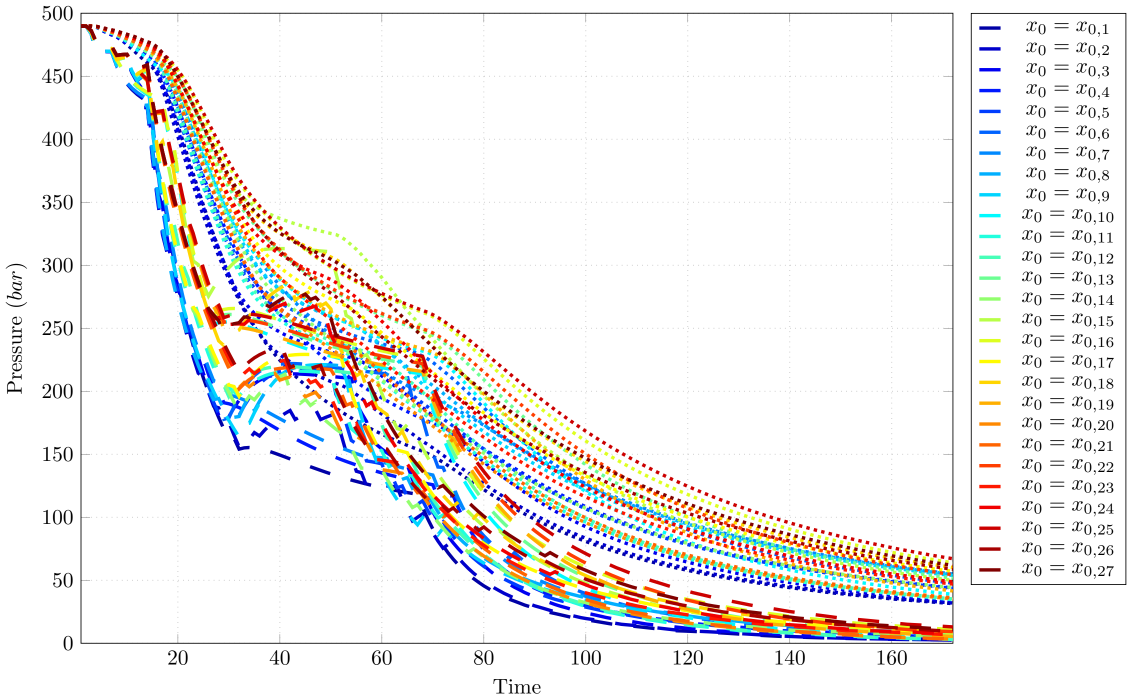


FIGURE 6.4: Evolution over time of the pressure when applying the optimal policy for all the possible initial states as listed in Equation (6.6). The pressures in the first tank are represented by the colored dashed curves, while the pressures in the second tank are represented by the colored dotted curves. Each color represents a given initial state

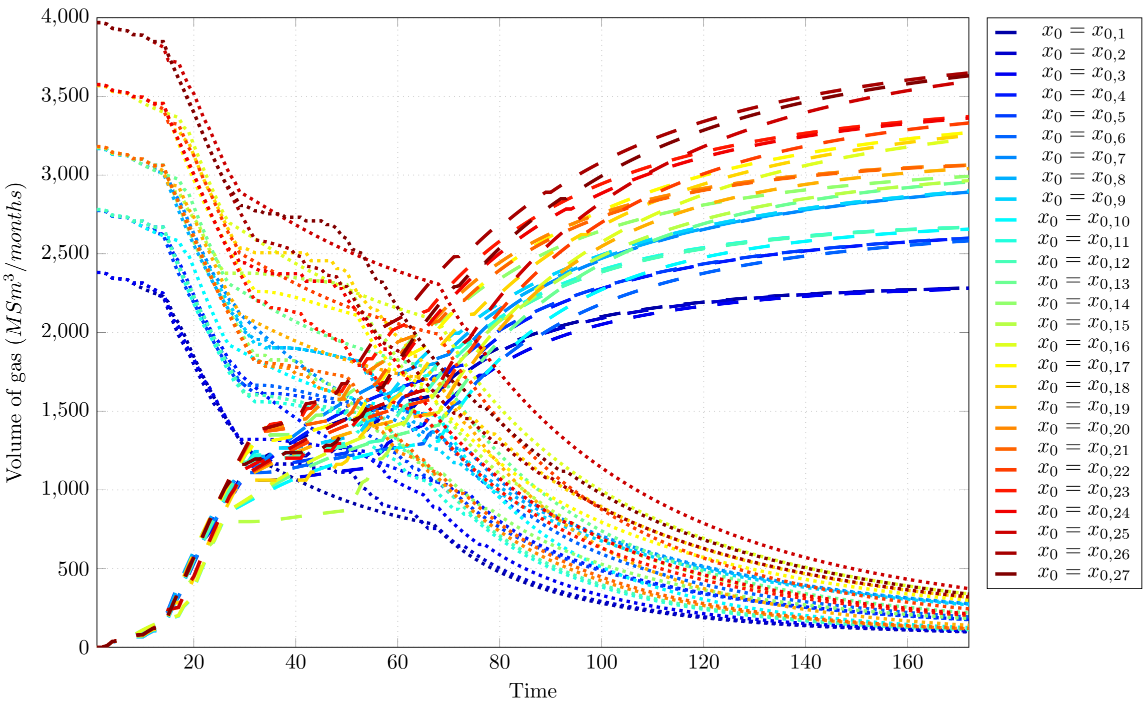


FIGURE 6.5: Evolution over time of the cumulative production and the total volume of in place gas when applying the optimal policy for all the possible initial states as listed in Equation (6.6). The cumulative productions are represented by the colored dashed curves, while the total gas in place are represented by the colored dotted curves. Each color represents a given initial state

vary depending on the belief, notably that, in certain cases, we slow down the production between the 20th and 50th months, whereas we do not in other cases.

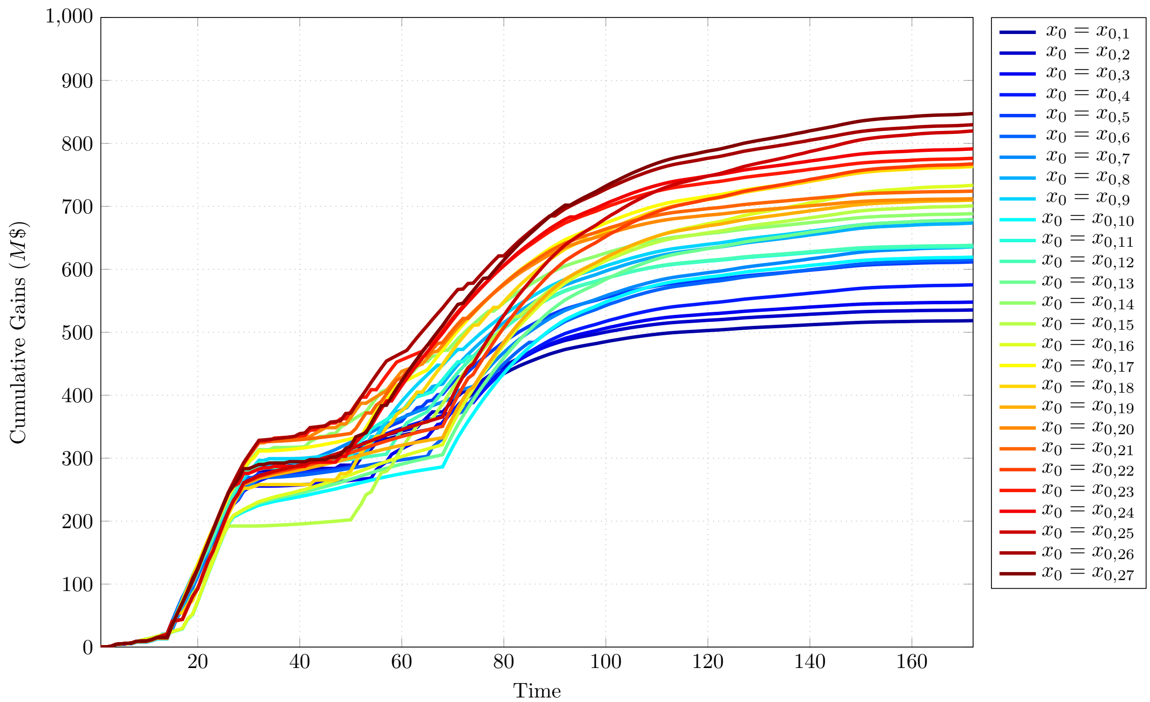


FIGURE 6.6: Evolution over time of the cumulative gains when applying the optimal policy for all the possible initial states. Each color represents a given initial state

Finally, in Figures 6.6 and 6.7, we draw the evolution over time of the cumulated gain and the recovery factor (i.e. the proportion of gas produced compared to the total volume of initial gas in place). Notably, and intuitively, for a given transmissivity, a higher volume of initial gas in place leads to a higher final cumulated gain. Moreover, for a given volume of initial gas in place, a higher transmissivity also leads to a higher final cumulated gain. However, for a given volume of initial gas in place, a higher transmissivity does not lead to a higher final recovery factor, as we optimize the Net Present Value, not the total production.

We then compare the trajectory of the cumulated gains and the recovery factor to the optimal trajectory of the fully observed case, as presented in Figures 6.8 and 6.9. Note that the final cumulated gains are higher in the fully observed case (between 0.2 % and 4 %). However, the final recovery factors of the fully observed cases tend to be lower (between 0.2 % and 3 %). We would hence be able to more efficiently pilot the reservoir if we had more information on the initial state of the reservoir. The results are compiled in Table 6.1.

We also study the computation time and the number of reachable beliefs depending on the number of states and the cardinality of the support of the initial belief, as seen in Table 6.2. Note that, until a certain point, it is faster to apply Dynamic Programming (i.e.

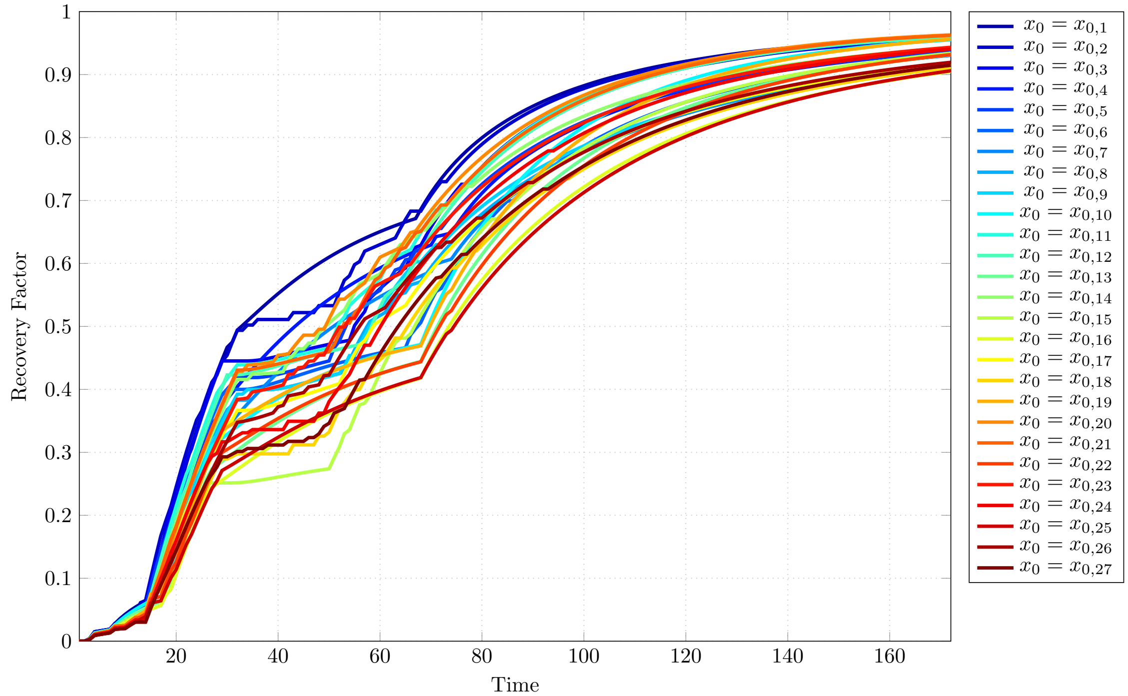


FIGURE 6.7: Evolution over time of the recovery factor when applying the optimal policy for all the possible initial states. Each color represents a given initial state

	CPU time (h)	Value (M€)	$\left\lvert \mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![0,T]\!]}(b_0) \right\rvert$
Partially observed	172	683	11,518,532
Fully Observed	0.5	699	26,830

TABLE 6.1: Comparison with regards to CPU time, values, and number of reachable beliefs between the partially observed and the fully observed two tanks problem

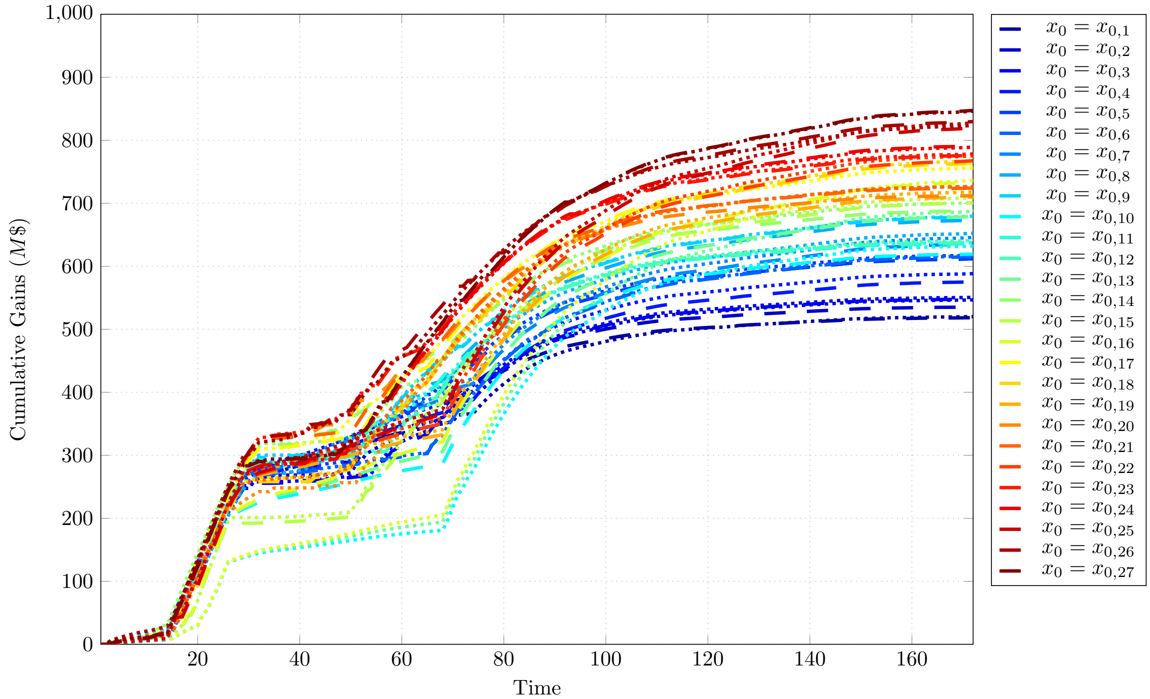


FIGURE 6.8: Trajectory of the cumulative gains when applying the optimal policy for all the possible initial states, compared with applying the optimal policy of the fully observed case. Each color represents a given initial state. The dotted curves represent the fully observed cases, while the dashed curves represent the partially observed case

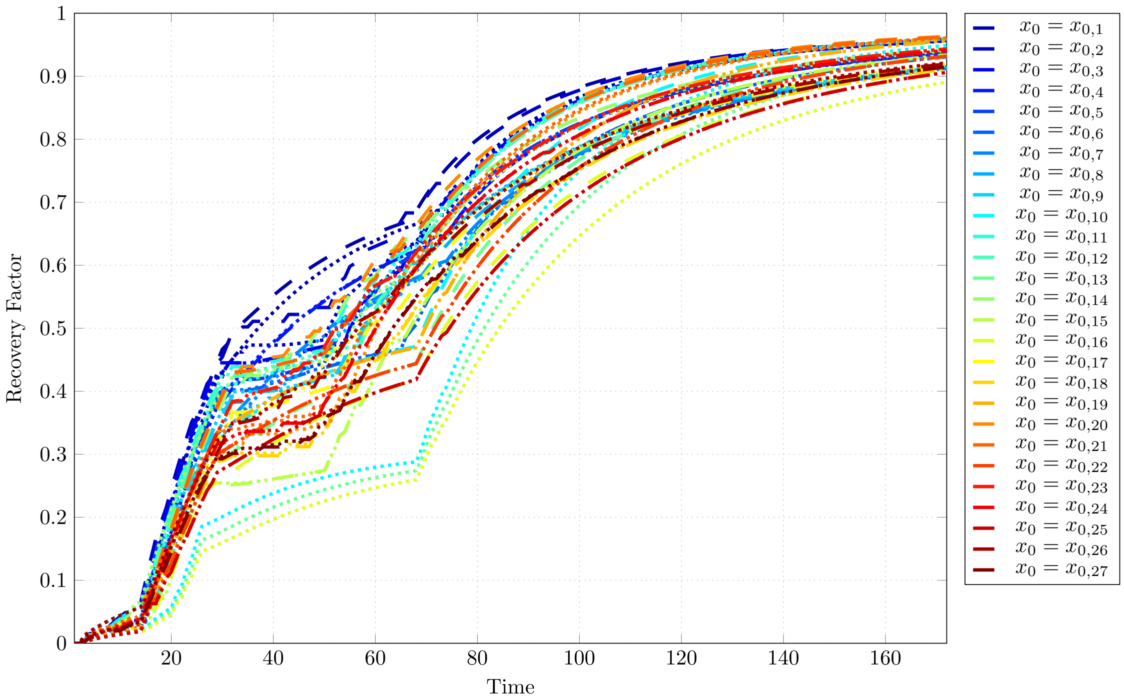


FIGURE 6.9: Trajectory of the recovery factor when applying the optimal policy for all the possible initial states, compared with applying the optimal policy of the fully observed case. Each color represents a given initial state. The dotted curves represent the fully observed cases, while the dashed curves represent the partially observed case

Algorithm 4) than it is to compute the reachable beliefs. Indeed, at a certain point, the problem became too big for the RAM of the system, hence we used the computer memory on disk to hold part of the problem.

$\lvert\mathbb{X}\rvert$	$\lvert\mathrm{supp}(b_0)\rvert$	CPU time (generation) (h)	$\lvert\mathbb{B}^{\mathrm{R},\mathcal{D}}_{[\![0,T]\!]}(b_0)\rvert$	CPU time (DP) (h)
$100 \times 100 \times 1 \times 1 \times 3$	3	1	27,038	0.5
$200 \times 200 \times 1 \times 1 \times 3$	3	1.5	88,284	0.8
$400 \times 400 \times 1 \times 1 \times 3$	3	4	423,204	1.5
$100 \times 100 \times 3 \times 3 \times 1$	9	1	85,243	1
$200 \times 200 \times 3 \times 3 \times 1$	9	5	1,294,860	4
$400 \times 400 \times 3 \times 3 \times 1$	9	32	9,025,697	[1]
$100 \times 100 \times 3 \times 3 \times 2$	18	1.5	682,106	1
$200 \times 200 \times 3 \times 3 \times 2$	18	8	2,254,659	8
$100 \times 100 \times 3 \times 3 \times 3$	27	7.5	853,569	6
$150 \times 150 \times 3 \times 3 \times 3$	27	36	3,157,322	32
$200 \times 200 \times 3 \times 3 \times 3$	27	62	11,518,532	110

TABLE 6.2: Evolution of the CPU time and number of reachable beliefs depending on the number of states and the cardinality of the support of the initial belief. "CPU time (generation)" correspond to the time to compute the sets of reachable beliefs, while " CPU time (DP)" is the computation time of Algorithm 4

6.3.2 Second application: oil reservoir with water injection

We consider the case of an oil reservoir where the pressure is kept constant by re-injecting water in the reservoir. The deterministic version of that problem was treated in §4.4.2. We now add a partial observation of the content of the reservoir.

The state is reduced to the vector $x = (V^{\mathrm{W}}, V^{\mathrm{P}})$, whereas the control is the bottom-hole pressure $u = P$. We inject enough water to keep the pressure constant, hence the amount of water injected is not a control itself, but is deduced from the bottom-hole pressure P.

The observation is the water-cut w^{CT}.

[1] Algorithm 4 was not launched for this instance.

Full formulation

$$\max_{(\boldsymbol{V}_t^{\mathrm{W}},\boldsymbol{P}_t,\boldsymbol{w}_t^{\mathrm{CT}})} \mathbb{E}\left[\sum_{t=0}^{T-1}\left[\rho^t r_t \alpha \frac{P^{\mathrm{R}}-\boldsymbol{P}_t}{B_{\mathrm{O}}(P^{\mathrm{R}})}\left[1-\boldsymbol{w}_t^{\mathrm{CT}}\right]-\rho^t c_t \alpha \frac{P^{\mathrm{R}}-\boldsymbol{P}_t}{B_{\mathrm{W}}(P^{\mathrm{R}})}\right]\right] \tag{6.7a}$$

$$s.t.\ \mathbb{P}_{\boldsymbol{V}_0^{\mathrm{W}},\boldsymbol{V}_0^{\mathrm{P}}} = b_0\ , \tag{6.7b}$$

$$\boldsymbol{w}_t^{\mathrm{CT}} = \mathrm{W}^{\mathrm{CT}}\left(\frac{\boldsymbol{V}_t^{\mathrm{W}} B_{\mathrm{W}}(P^{\mathrm{R}})}{\boldsymbol{V}_0^{\mathrm{P}}}\right)\ ,\ \forall t \in \mathcal{T}\ , \tag{6.7c}$$

$$\boldsymbol{V}_{t+1}^{\mathrm{W}} = \boldsymbol{V}_t^{\mathrm{W}} - \alpha \frac{P^{\mathrm{R}}-\boldsymbol{P}_t}{B_{\mathrm{W}}(P^{\mathrm{R}})}\left[\boldsymbol{w}_t^{\mathrm{CT}}-1\right]\ ,\ \forall t \in \mathcal{T}\ , \tag{6.7d}$$

$$F_{min}^{\mathrm{W}} \leq \alpha \frac{P^{\mathrm{R}}-\boldsymbol{P}_t}{B_{\mathrm{W}}(P^{\mathrm{R}})}\boldsymbol{w}_t^{\mathrm{CT}} \leq F_{max}^{\mathrm{W}}\ ,\ \forall t \in \mathcal{T}\ , \tag{6.7e}$$

$$F_{min}^{\mathrm{O}} \leq \alpha \frac{P^{\mathrm{R}}-\boldsymbol{P}_t}{B_{\mathrm{O}}(P^{\mathrm{R}})}\left[1-\boldsymbol{w}_t^{\mathrm{CT}}\right] \leq F_{max}^{\mathrm{O}}\ ,\ \forall t \in \mathcal{T}\ , \tag{6.7f}$$

$$\boldsymbol{P}_t \geq 0\ ,\ \forall t \in \mathcal{T}\ , \tag{6.7g}$$

$$\sigma(\boldsymbol{P}_t) \in \sigma(\boldsymbol{w}_0^{\mathrm{CT}},\ldots,\boldsymbol{w}_t^{\mathrm{CT}},\boldsymbol{P}_0,\ldots,\boldsymbol{P}_{t-1})\ ,\ \forall t \in \mathcal{T}\ . \tag{6.7h}$$

Optimization of the production When considering $|\mathbb{X}| = 55{,}885$, $|\mathbb{U}| = 10$, $|\mathbb{O}| = 10$, $|\mathrm{supp}(b_0)| = 10$, $T = 100$, we obtain Table 6.3. The bounds we obtain with Theorems 5.9 and 5.15 are, respectively, 2.9×10^{47} and 57.2×10^6. We are therefore far lower than any of the two bounds presented (by a factor of 10^{41} for the general Det-Pomdp bound, and of around 50 for the Separated Det-Pomdp bound).

The size of the problem is such that it is solved in a reasonable time: the computation of the reachable beliefs of the problem was made in 3200 seconds, while the solution time was of 4200 seconds (applying Algorithm 4).

Set considered	Cardinal of the set
$\mathbb{X}$	55,885
$\mathbb{B}$	809,665

Table 6.3: Size of the sets of the oil reservoir with water injection

6.4 Conclusion

In this chapter, we have formulated the management of a petroleum production system under partial observation as Problem (6.1), which is a Det-Pomdp optimization problem. Moreover, if the initial consistent Assumption 6.1 is satisfied, Problem (6.1) is equivalent, by Proposition 6.2, to a Separated Det-Pomdp optimization problem. We therefore apply the results of Chapter 5 to Problem (6.1). Notably, under the finite set Assumption 5.1, we use Algorithm 4 to solve Problem (6.1).

We have presented in §6.3 two numerical applications: a gas reservoir with two tanks; an oil reservoir where pressure is kept constant through water injection. We have managed to solve these problems with an initial belief with a support containing 27 elements (in §6.3.1), although it took a large amount of time (around a week) as predicted by the exponential bounds we have obtained in Chapter 5. This demonstrates that the method can be used in real cases, and has the potential to be used in complex cases. Of course, for substantial real cases, one has to improve the code, notably the memory management.

6.A Technical lemmata on Separated DET-POMDP for the oil and gas case

In this section, we present technical lemmata that are useful to prove that there is a Separated DET-POMDP optimization problem equivalent to Problem (6.1). For this purpose, we define a DET-POMDP optimization problem:

$$
\begin{aligned}
\mathcal{J}^\star(b_0, o_0) = \min_{\boldsymbol{X},\boldsymbol{O},\boldsymbol{U}} \ & \mathbb{E}\Big[\sum_{t=0}^{T-1} \mathcal{L}_t(\boldsymbol{X}_t, \boldsymbol{U}_t) + \mathcal{K}(\boldsymbol{X}_T)\Big] && \text{(6.8a)}\\
s.t.\ \ & \mathbb{P}_{\boldsymbol{X}_0} = b_0 \ , && \text{(6.8b)}\\
& \boldsymbol{X}_{t+1} = f(\boldsymbol{X}_t, \boldsymbol{U}_t) \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , && \text{(6.8c)}\\
& \boldsymbol{O}_0 = o_0 \ , && \text{(6.8d)}\\
& \mathrm{supp}(b_0) \subset h^{-1}(o_0) \ , && \text{(6.8e)}\\
& \boldsymbol{O}_t = h(\boldsymbol{X}_t) \ , \ \ \forall t \in \mathcal{T} \setminus \{0, T\} \ , && \text{(6.8f)}\\
& \boldsymbol{U}_t \in \mathcal{U}_t^{ad}(\boldsymbol{X}_t) \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ , && \text{(6.8g)}\\
& \sigma(\boldsymbol{U}_t) \subset \sigma(\boldsymbol{O}_0, \dots, \boldsymbol{O}_t, \boldsymbol{U}_0, \dots, \boldsymbol{U}_{t-1}) \ , \ \ \forall t \in \mathcal{T} \setminus \{T\} \ . && \text{(6.8h)}
\end{aligned}
$$

Note that this problem is quite similar to a standard DET-POMDP optimization problem (defined as Problem (5.2)), with a few distinctions. First, for all time $t > 0$, the evolution functions and observation functions are stationary (i.e. they are not time-dependent). Second, the observation function only depends on the state of the dynamical system, and not on the controls previously applied. Third, we have an added admissibility constraint, Equation (6.8e), which ensures that the support of the initial belief is consistent with the initial observation.

6.A.1 Notations

Here, we recall some of the notations previously used in Chapter 5. First, we recall how we represent the dynamics of a DET-POMDP as pushforward measures. Indeed, we have defined mappings F and $\mathcal{R}$ (respectively in Equations (5.18) and (5.19)) which give, thanks to Lemma 5.4,

$$\tau_t(b, u, o) = \mathcal{R} \circ (F_t^{u,o})_\star(b) \ .$$

Now, in the case of Problem (6.8), we obtain the following expressions for the mappings F and F_0 (which take into account constraint (6.8e)):

$$F^{u,o} : \overline{\mathbb{X}} \to \overline{\mathbb{X}} \ , \ \ x \mapsto \begin{cases} f^u(x) & \text{if} \quad x \neq \partial \text{ and } f^u(x) \in h^{-1}(o) \ , \\ \partial & \text{otherwise} \ . \end{cases} \tag{6.9}$$

and

$$F_0^{u,o} : \overline{\mathbb{X}} \to \overline{\mathbb{X}} \ , \quad x \mapsto \begin{cases} F^{u,o}(x) & \text{if} \quad x \in h^{-1}(o_0) \ , \\ \partial & \text{otherwise} \ . \end{cases} \tag{6.10}$$

We also recall that we denote, for any time $t \in \mathcal{T} \setminus \{T\}$, by $\mathbb{F}_t^{\mathcal{D}}$ the set of mappings $F_t^{u,o}$ for all pairs $(u, o) \in \mathbb{U} \times \mathbb{O}$, and by $\mathbb{F}^{\mathcal{D}}$ the set of the composition of mappings F, i.e.

$$\mathbb{F}_t^{\mathcal{D}} = \left\{ F_t^{u,o} \,\middle|\, u \in \mathbb{U}, o \in \mathbb{O} \right\} \subset \mathbb{L}(\overline{\mathbb{X}}; \overline{\mathbb{X}}) \ , \tag{6.11}$$

$$\mathbb{F}^{\mathcal{D}} = \bigcup_{t \in \mathcal{T} \setminus \{T\}} \mathbb{F}_{0:t}^{\mathcal{D}} \subset \mathbb{L}(\overline{\mathbb{X}}; \overline{\mathbb{X}}) \ . \tag{6.12}$$

Now, we recall that a Separated Det-Pomdp is a Det-Pomdp such that the set of mappings $\mathbb{F}^{\mathcal{D}}$ is a (∂)-separated mapping set (see Definition 5.13 for the definition of a Separated Det-Pomdp, and Definition 5.12 for the definition of a (∂)-Separated Mapping Set).

6.A.2 Evolution functions and Separated Det-Pomdp

Here, we present conditions such that Problem (6.8) is a Separated Det-Pomdp. We first present a general result true for any Det-Pomdp (i.e. for Problem (5.2)), hence also true for Problem (6.8).

Lemma 6.3. *Let $t > 0$. Let $\{u_i\}_{i \in [\![0, t-1]\!]} \in \mathbb{U}^t$ and $\{o_i\}_{i \in [\![1,t]\!]} \in \mathbb{O}^t$ be two sequences of t controls and observations. For all state $x \in \mathbb{X}$, either:*

- $F_{0:t-1}^{u_{0:t-1}, o_{1:t}}(x) = \partial$, *or*
- *for all time $t' \in [\![1, t]\!]$, $h_{t'}\big(f_{0:t'-1}^{u_{0:t'-1}}(x), u_{t'-1}\big) = o_{t'}$.*

Proof. We prove this lemma by induction on t.

Initialization. Let $t = 1$. Let $u_0 \in \mathbb{U}$ and $o_1 \in \mathbb{O}$. For all state $x \in \mathbb{X}$, we have, by the definition of $F_0^{u_0, o_1}$ (see Equation (5.18)) that,

- either $F_0^{u_0,o_1}(x) = \partial$,
- or $h_1\big(f_0^{u_0}(x), u_0\big) = o_1$.

Induction step. We assume that the result holds up to time $t > 0$. Let us prove it also holds true for time $t + 1$. Let $\{u_i\}_{i \in [\![0,t]\!]} \in \mathbb{U}^{t+1}$ and $\{o_i\}_{i \in [\![1,t+1]\!]} \in \mathbb{O}^{t+1}$ be fixed be fixed. We have

$$F_{0:t}^{u_{0:t}, o_{1:t+1}} = F_t^{u_t, o_{t+1}} \circ F_{0:t-1}^{u_{0:t-1}, o_{1:t}} \ .$$

Thus, for all state $x \in \mathbb{X}$ such that $F_{0:t}^{u_{0:t}, o_{1:t+1}}(x) \neq \partial$, we have

$$h_{t+1}\big(f_{0:t}^{u_{0:t}}(x), u_t\big) = o_{t+1} \ , \quad \text{(using the Definition of } F_t^{u_t, o_{t+1}} \text{ in Equation (5.18))} \tag{6.13}$$

and, as $F_t^{u_t,o_{t+1}}(\partial) = \partial$, that

$$F_{0:t-1}^{u_{0:t-1},o_{1:t}}(x) \neq \partial \; . \tag{6.14}$$

Thanks to Equation (6.14), we can apply the induction assumption at time t; thus, for all time $t' \in [\![1,t]\!]$, $h_{t'}\big(f_{0:t'-1}^{u_{0:t'-1}}(x), u_{t'-1}\big) = o_{t'}$. Combined with Equation (6.13), it hence holds true for all time $t' \in [\![1,t+1]\!]$. This ends the proof. □

Now, we present lemmata regarding the (∂)-separation. Note that – instead of the framework presented in Chapter 5, where we showed that a problem is a Separated DET-POMDP optimization problem if its evolution functions form a separated mapping set– we directly study the sets $\mathbb{F}^{\mathcal{D}}$, and check that they are indeed (∂)-separated mappings sets.

Lemma 6.4. *Let $\mathbb{X}_1$ and $\mathbb{X}_2$ be two given sets, and let $\mathbb{G}_1 \subset \mathbb{L}(\overline{\mathbb{X}}_1;\overline{\mathbb{X}}_1)$ and $\mathbb{G}_2 \subset \mathbb{L}(\overline{\mathbb{X}}_2;\overline{\mathbb{X}}_2)$ be two (∂)-separated mappings sets. Then, the set of mappings*

$$\mathbb{G} = \Bigg\{ g \in \mathbb{L}((\mathbb{X}_1 \times \mathbb{X}_2) \cup \{\partial\}; (\mathbb{X}_1 \times \mathbb{X}_2) \cup \{\partial\}) \Bigg| \; \exists (g_1,g_2) \in \mathbb{G}_1 \times \mathbb{G}_2, g(x_1,x_2) = \begin{cases} (g_1(x_1), g_2(x_2)) \; \textit{if } g_1(x_1) \neq \partial \textit{ and } g_2(x_2) \neq \partial \\ \partial \; \textit{otherwise} \end{cases} \Bigg\} \tag{6.15}$$

is a (∂)-separated mappings sets.

Proof. Let $(g,g') \in \mathbb{G}^2$. By Equation (6.15), there exists $(g_1,g_1') \in \mathbb{G}_1^2$ and $(g_2,g_2') \in \mathbb{G}_2^2$ such that

$$g(x_1,x_2) = \begin{cases} (g_1(x_1), g_2(x_2)) \text{ if } g_1(x_1) \neq \partial \text{ and } g_2(x_2) \neq \partial \\ \text{or } \partial \end{cases} ,$$

and

$$g'(x_1,x_2) = \begin{cases} (g_1'(x_1), g_2'(x_2)) \text{ if } g_1'(x_1) \neq \partial \text{ and } g_2'(x_2) \neq \partial \\ \text{or } \partial \end{cases} .$$

Let us assume that there exists $x = (x_1,x_2) \in \mathbb{X}_1 \times \mathbb{X}_2$ such that $g(x) = g'(x) \neq \partial$. Then, for all $x' = (x_1',x_2') \in \mathbb{X}_1 \times \mathbb{X}_2$, we have either $g(x') = \partial$, or $g'(x') = \partial$, or, as $\mathbb{G}_1$ and $\mathbb{G}_2$ are (∂)-separated mappings sets,

$$g(x') = \big(g_1(x_1'), g_2(x_2')\big) = \big(g_1'(x_1'), g_2'(x_2')\big) = g'(x') \neq \partial \; .$$

Hence, by Definition 5.12, $\mathbb{G}$ is (∂)-separated. □

We now present lemmata that link the mathematical expression of the mappings F to the (∂)-separation of the set $\mathbb{F}^{\mathcal{D}}$.

Lemma 6.5. *Assume that there exists $g : \mathbb{O} \times \mathbb{U} \to \mathbb{X}$ such that, for all pairs of controls and observation $(u,o) \in \mathbb{U} \times \mathbb{O}$, the mappings $F^{u,o}$ are given by*

$$F^{u,o} : \overline{\mathbb{X}} \to \overline{\mathbb{X}} \; , \quad x \mapsto \begin{cases} x + g\big(h(x), u\big) & \text{if} \quad x \neq \partial \text{ and } \big(x + g\big(h(x), u\big)\big) \in h^{-1}(o) \; , \\ \partial & \text{otherwise} \; . \end{cases} \tag{6.16}$$

Then, the set $\mathbb{F}^{\mathcal{D}}$ is a (∂)-separated mapping set.

Proof. In order to prove that the set $\mathbb{F}^{\mathcal{D}}$ is a (∂)-separated mapping set, we need to prove that all pairs of mappings $(F, F') \in (\mathbb{F}^{\mathcal{D}})^2$ are (∂)-separated.

First, we explicit the form of the mappings F. As $\mathbb{F}^{\mathcal{D}}$ is given by $\mathbb{F}^{\mathcal{D}} = \cup_{t \in \mathcal{T} \setminus \{T\}} \big(\mathbb{F}^{\mathcal{D}}_{0:t}\big)$ (see Equation (6.12)), for all elements F of $\mathbb{F}^{\mathcal{D}}$, there exists $t > 0$ and two sequences of controls and observations $u_{0:t-1}$ and $o_{1:t}$ such that $F = F^{u_{0:t-1}, o_{1:t}}_{0:t-1}$. To compute the composition of the mappings $F_{t'}$ given by Equation (6.16) (for $t' > 0$) and Equation (6.10) (for $t' = 0$), we need to determine the observation at time t' (i.e. the value of $h(x)$ in Equation (6.16)). But it is known thanks to the sequence $o_{0:t}$: by Lemma 6.3, either we attain ∂; or, for all time $t' \in [\![1, t]\!]$, the observation at time t' is $o_{t'}$, and at time 0, Equation (6.10) states that the observation is o_0.

Hence, combined with Equation (6.16), we obtain that (without specifying the conditions such that we do not attain the cemetery state ∂)

$$F^{u_{0:t-1}, o_{1:t}}_{0:t-1} : \overline{\mathbb{X}} \to \overline{\mathbb{X}} \; , \quad x \mapsto \begin{cases} x + \sum_{t'=0}^{t-1} g\big(o_{t'}, u_{t'}\big) \; , \\ \text{or } \partial \; . \end{cases} \tag{6.17}$$

Second, let $\big(F^{u_{0:t-1}, o_{1:t}}_{0:t-1}, F^{u'_{0:t'-1}, o'_{1:t'}}_{0:t'-1}\big) \in (\mathbb{F}^{\mathcal{D}})^2$. In order to prove that the pair $\big(F^{u_{0:t-1}, o_{1:t}}_{0:t-1}, F^{u'_{0:t'-1}, o'_{1:t'}}_{0:t'-1}\big)$ is (∂)-separated, we need to consider two cases.

1. For all state $x \in \mathbb{X}$, we have either $F^{u_{0:t-1}, o_{1:t}}_{0:t-1}(x) = \partial$, or $F^{u_{0:t-1}, o_{1:t}}_{0:t-1}(x) \neq F^{u'_{0:t'-1}, o'_{1:t'}}_{0:t'-1}(x)$. Then, the pair $\big(F^{u_{0:t-1}, o_{1:t}}_{0:t-1}, F^{u'_{0:t'-1}, o'_{1:t'}}_{0:t'-1}\big)$ is (∂)-separated.

2. There exists a state $x \in \mathbb{X}$ such that $F^{u_{0:t-1}, o_{1:t}}_{0:t-1}(x) = F^{u'_{0:t'-1}, o'_{1:t'}}_{0:t'-1}(x) \neq \partial$. Thus, we have by Equation (6.17)

$$x + \sum_{i=0}^{t-1} g\big(o_i, u_i\big) = x + \sum_{i=0}^{t'-1} g\big(o'_i, u'_i\big) \; ,$$

i.e.

$$\sum_{i=0}^{t-1} g\big(o_i, u_i\big) = \sum_{i=0}^{t'-1} g\big(o'_i, u'_i\big) \; .$$

Hence, for all states $x' \in \mathbb{X}$, either

$$F_{0:t-1}^{u_{0:t-1},o_{1:t}}(x') = F_{0:t'-1}^{u'_{0:t'-1},o'_{1:t'}}(x') = x' + \sum_{i=0}^{t-1} g\big(o_i, u_i\big) \neq \partial \ , \quad \text{or } F_{0:t-1}^{u_{0:t-1},o_{1:t}}(x') = \partial \ ,$$

or $F_{0:t'-1}^{u'_{0:t'-1},o'_{1:t'}}(x') = \partial$. Hence, the pair is (∂)-separated.

In both cases, the pair is (∂)-separated. As this holds true for all pairs of the set $\mathbb{F}^{\mathcal{D}}$, the set $\mathbb{F}^{\mathcal{D}}$ is a (∂)-separated mapping set. $\square$

Lemma 6.6. *Assuming that* 0 *does not belong to set the* $\mathbb{X} \subset \mathbb{R}$ *and that there exists* $a \in \mathbb{R}$ *such that, for all pairs of controls and observation* $(u, o) \in \mathbb{U} \times \mathbb{O}$*,* $\mathbb{O} \subset \mathbb{R}$*, the mappings* $F^{u,o}$ *are given by*

$$F^{u,o} : \overline{\mathbb{X}} \to \overline{\mathbb{X}} \ , \quad x \mapsto \begin{cases} x \times \Big(1 + a\big(h(x) - o\big)\Big) & \textit{if } x \neq \partial \textit{ and } \bigg(x \times \Big(1 + a\big(h(x) - o\big)\Big)\bigg) \in h^{-1}(o) \\ \partial & \textit{otherwise} \ , \end{cases} \tag{6.18}$$

then the set $\mathbb{F}^{\mathcal{D}}$ *is a* (∂)*-separated mapping set.*

Proof. The proof is similar to the proof of Lemma 6.5: we prove that all pairs of mappings $(F, F') \in (\mathbb{F}^{\mathcal{D}})^2$ are (∂)-separated.

First, for all elements F of $\mathbb{F}^{\mathcal{D}}$, there exists $t > 0$ and two sequences of controls and observations $u_{0:t-1}$ and $o_{1:t}$ such that $F = F_{0:t-1}^{u_{0:t-1},o_{1:t}}$. Moreover, by Lemma 6.3, either we attain ∂, or, for all time $t' \in [\![1, t]\!]$, the observation at time t is $o_{t'}$, and at time 0 Equation (6.10) states that the observation is o_0.

Hence, by combining this with Equation (6.18) the expression of the composition of mappings $F_{0:t-1}^{u_{0:t-1},o_{1:t}}$ is given by (without specifying the conditions such that we do not attain the cemetery state ∂)

$$F_{0:t-1}^{u_{0:t-1},o_{1:t}} : \overline{\mathbb{X}} \to \overline{\mathbb{X}} \ , \quad x \mapsto \begin{cases} x \times \prod_{t'=0}^{t-1} \big(1 + a(o_{t'} - o_{t'+1})\big) \ , \\ \text{or } \partial \ . \end{cases} \tag{6.19}$$

Second, let $\big(F_{0:t-1}^{u_{0:t-1},o_{1:t}}, F_{0:t'-1}^{u'_{0:t'-1},o'_{1:t'}}\big) \in (\mathbb{F}^{\mathcal{D}})^2$. In order to prove that the pair $\big(F_{0:t-1}^{u_{0:t-1},o_{1:t}}, F_{0:t'-1}^{u'_{0:t'-1},o'_{1:t'}}\big)$ is (∂)-separated, we consider two cases.

1. For all state $x \in \mathbb{X}$, we have either $F_{0:t-1}^{u_{0:t-1},o_{1:t}}(x) = \partial$, or $F_{0:t-1}^{u_{0:t-1},o_{1:t}}(x) \neq F_{0:t'-1}^{u'_{0:t'-1},o'_{1:t'}}(x)$. Then, the pair $\big(F_{0:t-1}^{u_{0:t-1},o_{1:t}}, F_{0:t'-1}^{u'_{0:t'-1},o'_{1:t'}}\big)$ is (∂)-separated.

2. There exists a state $x \in \mathbb{X}$ such that $F^{u_{0:t-1},o_{1:t}}_{0:t-1}(x) = F^{u'_{0:t'-1},o'_{1:t'}}_{0:t'-1}(x) \neq \partial$. Thus, we have by Equation (6.19)

$$x \times \prod_{i=0}^{t-1} \big(1 + a(o_i - o_{i+1})\big) + \sum_{i=0}^{t'-1} g\big(o_i, u_i\big) = x \times \prod_{i=0}^{t'-1} \big(1 + a(o'_i - o'_{i+1})\big) \ ,$$

which leads to, as $x \neq 0$,

$$\prod_{i=0}^{t-1} \big(1 + a(o_i - o_{i+1})\big) = \prod_{i=0}^{t'-1} \big(1 + a(o'_i - o'_{i+1})\big) \ .$$

Hence, for all states $x' \in \mathbb{X}$, either

$$F^{u_{0:t-1},o_{1:t}}_{0:t-1}(x') = F^{u'_{0:t'-1},o'_{1:t'}}_{0:t'-1}(x') = x' \times \prod_{t'=0}^{t-1} \big(1 + a(o_{t'} - o_{t'+1})\big) \neq \partial \ , \quad \text{or}$$

$$F^{u_{0:t-1},o_{1:t}}_{0:t-1}(x') = \partial \ , \quad \text{or } F^{u'_{0:t'-1},o'_{1:t'}}_{0:t'-1}(x') = \partial \ .$$

Hence, the pair is (∂)-separated.

In both cases, the pair is (∂)-separated. As this holds true for all pair of set $\mathbb{F}^{\mathcal{D}}$, set $\mathbb{F}^{\mathcal{D}}$ is a (∂)-separated mapping set. □

Lemma 6.7. *Assume that the set $\mathbb{X} = \mathbb{X}_1 \times \mathbb{X}_2$, and that there exists three mappings $g_1 : \mathbb{O} \times \mathbb{U} \to \mathbb{X}_1$, $g_2 : \mathbb{O} \times \mathbb{U} \to \mathbb{X}_2$, and $m : \mathbb{O} \to \mathbb{X}_2$, such that, for all pairs of controls observation $(u, o) \in \mathbb{U} \times \mathbb{O}$, the mappings $F^{u,o}$ are given by*

$$F^{u,o} : \overline{\mathbb{X}} \to \overline{\mathbb{X}} \ , \quad x \mapsto \begin{cases} \begin{pmatrix} x^{(1)} + g_1(h(x), u) \\ \\ x^{(2)} + g_2(h(x), u) + \\ m(o) \times \Big(x^{(1)} + g_1(h(x), u)\Big) \\ \qquad - m\big(h(x)\big) \times x^{(1)} \end{pmatrix} & \text{if } x \neq \partial \text{ and } \begin{pmatrix} x^{(1)} + g_1(h(x), u) \\ \\ x^{(2)} + g_2(h(x), u) + \\ m(o) \times \Big(x^{(1)} + g_1(h(x), u)\Big) \\ \qquad - m\big(h(x)\big) \times x^{(1)} \end{pmatrix} \in h^{-1}(o) \ , \\ \partial & \text{otherwise} \ . \end{cases} \tag{6.20}$$

Then, the set $\mathbb{F}^{\mathcal{D}}$ is a (∂)-separated mapping set.

Proof. First, for all elements F of $\mathbb{F}^{\mathcal{D}}$, there exists $t > 0$ and two sequences of controls and observations $u_{0:t-1}$ and $o_{1:t}$ such that $F = F^{u_{0:t-1},o_{1:t}}_{0:t-1}$. Moreover, by Lemma 6.3, either we attain ∂ or, for all time $t' \in [\![1, t]\!]$, the observation at time t is $o_{t'}$, and at time 0 Equation (6.10) states that the observation is o_0.

Hence, by combining this with Equation (6.20) the expression of the composition of mappings $F^{u_{0:t-1},o_{1:t}}_{0:t-1}$ is given by (without specifying the conditions such that we do not attain the cemetery ∂)

$$F_{0:t-1}^{u_{0:t-1},o_{1:t}} : \overline{\mathbb{X}} \to \overline{\mathbb{X}}\ ,\quad x \mapsto \begin{cases} \begin{pmatrix} x^{(1)} + \sum_{i=0}^{t-1} g_1\big(o_i, u_i\big) \\ x^{(2)} + \sum_{i=0}^{t-1} \Big\{ g_2(o_i, u_i) + m\big(o_{i+1}\big)(x^{(1)} + \sum_{j=0}^{i} \{g_1(o_j, u_j)\}) - \\ \qquad m(o_i)\big(x^{(1)} + \sum_{j=0}^{i-1} g_1(o_j, u_j)\big) \Big\} \end{pmatrix} \\ \text{or } \partial\ , \end{cases} ,$$

which can be simplified to

$$F_{0:t-1}^{u_{0:t-1},o_{1:t}} : x \mapsto \begin{cases} \begin{pmatrix} x^{(1)} + \sum_{i=0}^{t-1} g_1\big(o_i, u_i\big) \\ x^{(2)} + \sum_{i=0}^{t-1} \{g_2(o_i, u_i)\} + m\big(o_t\big)(x^{(1)} + \sum_{j=0}^{t-1} \{g_1(o_j, u_j)\}) - m(o_0)\big(x^{(1)}\big) \end{pmatrix} \\ \text{or } \partial\ . \end{cases} , \tag{6.21}$$

According to Lemma 6.5, the first component of $F_{0:t-1}^{u_{0:t-1},o_{1:t}}$ is (∂)-separated, as it is of the form $x + g(h(x), u)$. By Lemma 6.4, we hence we only need to check that the second coordinate is also (∂)-separated. Let $\big(F_{0:t-1}^{u_{0:t-1},o_{1:t}}, F_{0:t'-1}^{u'_{0:t'-1},o'_{1:t'}}\big) \in (\mathbb{F}^{\mathcal{D}})^2$. In order to prove that the pair $\big(F_{0:t-1}^{u_{0:t-1},o_{1:t}}, F_{0:t'-1}^{u'_{0:t'-1},o'_{1:t'}}\big)$ is (∂)-separated, we consider two cases.

1. For all state $x \in \mathbb{X}$, we have either $F_{0:t-1}^{u_{0:t-1},o_{1:t}}(x) = \partial$, or $F_{0:t-1}^{u_{0:t-1},o_{1:t}}(x) \neq F_{0:t'-1}^{u'_{0:t'-1},o'_{1:t'}}(x)$. Then the pair $\big(F_{0:t-1}^{u_{0:t-1},o_{1:t}}, F_{0:t'-1}^{u'_{0:t'-1},o'_{1:t'}}\big)$ is (∂)-separated.

2. There exists a state $x \in \mathbb{X}$ such that $F_{0:t-1}^{u_{0:t-1},o_{1:t}}(x) = F_{0:t'-1}^{u'_{0:t'-1},o'_{1:t'}}(x) \neq \partial$. Thus, we have, by Equation (6.19), that

$$x^{(2)} + \sum_{i=0}^{t-1} \{g_2(o_i, u_i)\} + m\big(o_t\big)(x^{(1)} + \sum_{j=0}^{t-1} \{g_1(o_j, u_j)\}) - m(o_0)\big(x^{(1)}\big) =$$
$$x^{(2)} + \sum_{i=0}^{t'-1} \{g_2(o'_i, u'_i)\} + m\big(o'_{t'}\big)(x^{(1)} + \sum_{j=0}^{t'-1} \{g_1(o'_j, u'_j)\}) - m(o_0)\big(x^{(1)}\big)\ .$$

Moreover, in order to have $F_{0:t-1}^{u_{0:t-1},o_{1:t}}(x) = F_{0:t'-1}^{u'_{0:t'-1},o'_{1:t'}}(x) \neq \partial$, we need to have $o_t = o'_{t'}$ (as we must have $o_t = h\Big(F_{0:t-1}^{u_{0:t-1},o_{1:t}}(x)\Big) = h\Big(F_{0:t'-1}^{u'_{0:t'-1},o'_{1:t'}}(x)\Big) = o'_{t'}$). Combined with $\sum_{i=0}^{t-1} g\big(o_i, u_i\big) = \sum_{i=0}^{t'-1} g\big(o'_i, u'_i\big)$ (see proof of Lemma 6.5), this leads to

$$\sum_{i=0}^{t-1} \{g_2(o_i, u_i)\} = \sum_{i=0}^{t'-1} \{g_2(o'_i, u'_i)\}\ .$$

Hence, for all state $x' \in \mathbb{X}$, either

$$F_{0:t-1}^{u_{0:t-1},o_{1:t}}(x') = F_{0:t'-1}^{u'_{0:t'-1},o'_{1:t'}}(x') = \\ x^{(2)} + \sum_{i=0}^{t-1}\{g_2(o_i,u_i)\} + m\big(o_t\big)(x^{(1)} + \sum_{j=0}^{t-1}\{g_1(o_j,u_j)\}) - m(o_0)\big(x^{(1)}\big) \neq \partial \ ,$$

or $F_{0:t-1}^{u_{0:t-1},o_{1:t}}(x') = \partial$, or $F_{0:t'-1}^{u'_{0:t'-1},o'_{1:t'}}(x') = \partial$. Hence, the pair is (∂)-separated.

In both cases, the pair is (∂)-separated. As this holds true for all pairs of the set $\mathbb{F}^{\mathcal{D}}$, the set $\mathbb{F}^{\mathcal{D}}$ is a (∂)-separated mapping set. □

We now present a lemma based on the previous four lemmata 6.4 to 6.7.

Lemma 6.8. *Assume that the set $\mathbb{X} = \prod_{i\in[\![1,5]\!]}\mathbb{X}_i \subset \mathbb{R}^5$, and that there exists $a \in \mathbb{R}$ and four mappings $g_1 : \mathbb{O}\times\mathbb{U} \to \mathbb{X}_1$, $g_2 : \mathbb{O}\times\mathbb{U} \to \mathbb{X}_2$, $g_3 : \mathbb{O}\times\mathbb{U} \to \mathbb{X}_3$, and $m : \mathbb{O} \to \mathbb{X}_2$, such that, for all pairs of controls observation $(u,o) \in \mathbb{U}\times\mathbb{O}$, the mappings $F^{u,o}$ are given by*

$$F^{u,o}, \overline{\mathbb{X}} \to \overline{\mathbb{X}}\ , \quad x \mapsto \begin{cases} \begin{pmatrix} x^{(1)} + g_1(h(x),u) \\ x^{(2)} + g_2(h(x),u) + \\ \quad m\big(o\big)(x^{(1)} + g_1(h(x),u)) - m(h(x))x^{(1)} \\ x^{(3)} + g_3(h(x),u) \\ x^{(4)}\big(1 + a(o - h(x))\big) \\ o^{(1)} \end{pmatrix} & \text{if } h(f(x,u)) = o\ , \\ \partial \ \text{otherwise}\ . \end{cases}$$

Then, the set $\mathbb{F}^{\mathcal{D}}$ is a (∂)-separated mapping set.

Proof. We can apply Lemma 6.5 on the first and third components (they are of the form $x + g(h(x),u)$), which means that we indeed obtain a (∂)-separated mapping set when considering a restriction to those components. Similarly, we can also apply Lemma 6.6 on the fourth component. Finally, we can apply Lemma 6.7 on the first and second component. By Lemma 6.4, since all components are (∂)-separated mapping set, then the set $\mathbb{F}^{\mathcal{D}}$ is a (∂)-separated mapping set. □

Chapter 7

Conclusion

Contents

As a conclusion, we summarize our main contributions in §7.1, before outlining some possible extensions of our work in §7.2.

7.1 Main contributions

In this dissertation, we have first presented in Chapter 4 a mathematical formulation for the optimal management over time of an oil production network as a multistage optimization problem. In this formulation, the reservoir is modeled as a controlled (non-linear) dynamical system derived from material balance equations and the black-oil model. In the resulting dynamic system, the state is of dimension five, which is quite large for numerical resolution via Dynamic Programming algorithm. However, we were able to use Dynamic Programming to numerically solve the management optimization problem for specific cases of interest with either oil or gas, both presenting a reduced dimensionality of the state. We have also shown that our mathematical formulation is an improvement over decline curves formulation. First, as predicted by the theory, we replicated results from decline curve formulations when considering the first recovery of a one tank system (as seen in §4.4.1.1). Second, in more complex cases with inter-connected tanks, as described in §4.4.1.2, we have shown that we can surpass the Net Present Value returned by the decline curve formulation. Third, we have gone beyond the first recovery of hydrocarbons (as seen in §4.4.2, where we studied a case with water injection, hence studied secondary oil recovery).

Second, in Chapter 5 we made a theoretical detour and studied a class of problem that is of interest for taking into account the partial observation of the reservoir: POMDP. More specifically, we studied a subclass of POMDPs, that we named Separated DET-POMDPs, which has properties that contribute to push back the curse of dimensionality for Dynamic Programming. Indeed, the conditions on the dynamics for Separated DET-POMDP improve the bound on the cardinality of the set of reachable beliefs: the bound is reduced from $\left(1+|\mathbb{X}|\right)^{|\text{supp}(b_0)|}$ (in the case of Det-Pomdp, see theorem 5.9) to $2^{|\text{supp}(b_0)|}|\mathbb{X}|$ (theorem 5.15). In the case of Separated Det-Pomdp, more specifically, the improvement of the bounds on the cardinality of the set of reachable beliefs are derived using a new representation of the beliefs dynamics using the notion of push-forward. We enumerated the number of mappings used in the push-forwards in order to get bounds on the cardinality of the set of reachable beliefs. Indeed, there cannot be more reachable beliefs than mappings used in the representation of the beliefs dynamics as push-forward. Then, the tighter bound allows Dynamic Programming algorithms to efficiently solve Separated Det-Pomdp problems, especially when the supports of the initial state distributions are assumed to be of small cardinality. Moreover, we have shown that the bound is tight (see proposition 5.18). We thus validated that the Separated Det-Pomdp class of problems is indeed an interesting framework for oil and gas applications, which motivated this theoretical detour.

Third, we applied in Chapter 6 the theoretical framework presented in Chapter 5 to the management of a petroleum production system under partial observation. Indeed, we have formulated the management of a petroleum production system under partial observation as Problem (6.1), which is a Det-Pomdp optimization problem. Moreover, if the initial consistent Assumption 6.1 is satisfied, i.e. if the initial belief considered in Problem (6.1) is consistent with the initial observation, then Problem (6.1) is equivalent, by Proposition 6.2, to a Separated Det-Pomdp optimization problem. We therefore apply the results of Chapter 5 to Problem (6.1). Notably, under the finite set Assumption 5.1, we use Algorithm 4 to solve Problem (6.1).

Finally, we illustrated our works on two numerical applications presented in both Chapter 4 and Chapter 6: a gas reservoir with two tanks; an oil reservoir where pressure is kept constant through water injection. Those cases were studied with perfect and partial state observations, in Chapter 4 and 6. Notably, as shown in Chapter 6, we have managed to solve these problems with an initial belief with a support containing 27 elements (in §6.3.1), although it took a large amount of time (around a week) as predicted by the exponential bounds we have obtained in Chapter 5. This demonstrates that the method can be used in real cases, and has the potential to be used in complex cases.

This thesis led to the publication of one paper, Vessaire et al. [2022], of which Chapter 4 is a transcript. Moreover, it also led to the submission of a patent. An article based on Chapter 5 is currently being finalized.

We now present possible extensions of the present thesis.

7.2 Perspectives

The first possible extension to this thesis would be to add uncertainties on prices to both Problem 4.1 and Problem 6.1. Indeed, we assumed that oil prices are known in advance. This is obviously false, but reflect accurately real industrial studies on petroleum production systems made with fixed (known) prices. Adding stochastic prices could add flexibility to future industrial studies. Indeed, some prices models preserving the use of Dynamic Programming algorithm (leading thus to Stochastic Dynamic Programming) could be taken into account to extend both Problem 4.1 and Problem 6.1.

The second possible extension regards Det-Pomdps. Indeed, as shown in Chapter 5, we can find improved complexity bounds for sub-classes of Det-Pomdps when the resulting set of mappings used in the representation of the belief dynamics as push-forward is smaller than general sets of mappings on the state space. Thus, finding how more characteristics of the dynamics and observations influence the cardinality of that set of mappings, hence on the set of reachable beliefs, could be a new avenue for future applied problems belonging to the class Det-Pomdp. This was only partially done in this thesis, as we focused on the oil and gas applications presented in Chapter 6. This method could certainly be used on other Det-Pomdps applications.

List of Figures

List of Tables

Bibliography

S. Al-Rbeawi. Pseudo-steady state inflow performance relationship of reservoirs undergoing multiphase flow and different wellbore conditions. *Journal of Natural Gas Science and Engineering*, 68:102912, Aug. 2019. doi: 10.1016/j.jngse.2019.102912.

R. Archer, Y. Del Castillo, and T. Blasingame. New Perspectives on Vogel Type IPR Models for Gas Condensate and Solution-Gas Drive Systems. In *All Days*, pages SPE–80907–MS, Oklahoma City, Oklahoma, Mar. 2003. SPE. doi: 10.2118/80907-MS.

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. *Dynamic programming and optimal control. Vol. 1: ...* Number 3 in Athena scientific optimization and computation series. Athena Scientific, Belmont, Mass, fourth ed edition, 2016.

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.

J. M. Bohannon. A Linear Programming Model for Optimum Development of Multi-Reservoir Pipeline Systems. *Journal of Petroleum Technology*, 22(11):1429–1436, Nov. 1970. doi: 10.2118/2626-PA.

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.

B. Bonet and H. Geffner. Planning with Incomplete Information as Heuristic Search in Belief Space. In *AIPS'00: Proceedings of the Fifth International Conference on Artificial Intelligence Planning Systems*, pages 52–61, Breckenridge, CO, 2000. AAAI Press.

J. A. Caballero and I. E. Grossmann. An algorithm for the use of surrogate models in modular flowsheet optimization. *AIChE Journal*, 54(10):2633–2650, Oct. 2008. doi: 10.1002/aic.11579.

E. Camponogara, A. F. Teixeira, E. O. Hulse, T. L. Silva, S. Sunjerga, and L. K. Miyatake. Integrated Methodology for Production Optimization from Multiple Offshore Reservoirs in the Santos Basin. *IEEE Transactions on Automation Science and Engineering*, 14(2): 669–680, Apr. 2017. doi: 10.1109/TASE.2016.2640240.

Y. Chow, A. Tamar, S. Mannor, and M. Pavone. Risk-Sensitive and Robust Decision-Making: a CVaR Optimization Approach. *arXiv:1506.02188 [cs, math]*, June 2015. arXiv: 1506.02188.

L. Dake. *Fundamentals of Reservoir Engineering*. IISSN. Elsevier Science, 1983.

A. Danesh. *PVT and Phase Behaviour Of Petroleum Reservoir Fluids*. IISSN. Elsevier Science, 1998.

A. M. Daoud, M. Abdel Salam, A. A. Hashem, and M. Sayyouh. Inflow Performance Relationship Correlation for Solution Gas-Drive Reservoirs Using Non-Parametric Regression Technique. *The Open Petroleum Engineering Journal*, 10(1):152–176, Aug. 2017. doi: 10.2174/1874834101710010152.

A. Davis. Optimal field development and production design for unconventional reservoirs: A case study from Central Sub-Basin, Permian Basin, New Mexico. *Petroleum Research*, page 11, 2021.

E. Delage and S. Mannor. Percentile Optimization for Markov Decision Processes with Parameter Uncertainty. *Operations Research*, 58(1):203–213, Feb. 2010. doi: 10.1287/opre.1080.0685.

E. I. Epelle and D. I. Gerogiorgis. A computational performance comparison of MILP vs. MINLP formulations for oil production optimisation. *Computers & Chemical Engineering*, 140:106903, Sept. 2020. doi: 10.1016/j.compchemeng.2020.106903.

L. Frair and M. Devine. Economic Optimization of Offshore Petroleum Development. *Management Science*, 21(12):1370–1379, August 1975. doi: 10.1287/mnsc.21.12.1370.

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

D. I. Gerogiorgis, M. Georgiadis, G. Bowen, C. C. Pantelides, and E. N. Pistikopoulos. Dynamic oil and gas production optimization via explicit reservoir simulation. In W. Marquardt and C. Pantelides, editors, *Computer Aided Chemical Engineering*, volume 21 of *16th European Symposium on Computer Aided Process Engineering and 9th International Symposium on Process Systems Engineering*, pages 179–184. Elsevier, Jan. 2006. doi: 10.1016/S1570-7946(06)80043-X.

V. Gupta and I. E. Grossmann. An efficient multiperiod MINLP model for optimal planning of offshore oil and gas field infrastructure. *Industrial & Engineering Chemistry Research*, 51(19):6823–6840, 2012. doi: 10.1021/ie202959w.

V. Gupta and I. E. Grossmann. Multistage stochastic programming approach for offshore oilfield infrastructure planning under production sharing agreements and endogenous uncertainties. *Journal of Petroleum Science and Engineering*, 124:180 – 197, 2014. doi: 10.1016/j.petrol.2014.10.006.

G. Hepguler, S. Barua, and W. Bard. Integration of a Field Surface and Production Network With a Reservoir Simulator. *SPE Computer Applications*, 9(03):88–92, May 1997. doi: 10.2118/38937-PA. Publisher: OnePetro.

B. Hong, X. Li, S. Song, S. Chen, C. Zhao, and J. Gong. Optimal planning and modular infrastructure dynamic allocation for shale gas production. *Applied Energy*, 261:114439, Mar. 2020. doi: 10.1016/j.apenergy.2019.114439.

R. R. Iyer, I. E. Grossmann, S. Vasantharajan, and A. S. Cullick. Optimal planning and scheduling of offshore oil field infrastructure investment and operations. *Industrial & Engineering Chemistry Research*, 37(4):1380–1397, 1998. doi: 10.1021/ie970532x.

C. S. Khor, A. Elkamel, and N. Shah. Optimization methods for petroleum fields development and production systems: a review. *Optimization and Engineering*, 18(4):907–941, Dec. 2017. doi: 10.1007/s11081-017-9365-2.

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.

G. Lei, T. L. Silva, and M. Stanko. Compact formulations for efficient early-phase field development optimization of multi-reservoir fields. *Computers & Chemical Engineering*, 150:107319, July 2021. doi: 10.1016/j.compchemeng.2021.107319.

G. Lei, M. Stanko, and T. L. Silva. Formulations for automatic optimization of decommissioning timing in offshore oil and gas field development planning. *Computers & Chemical Engineering*, 165:107910, Sept. 2022. doi: 10.1016/j.compchemeng.2022.107910.

K. Li and Y. Li. Effect of initial water saturation on crude oil recovery and water cut in water-wet reservoirs: Effect of initial water saturation on oil recovery. *International Journal of Energy Research*, 38(12):1599–1607, Oct. 2014. doi: 10.1002/er.3182.

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

R. Marmier, U. Awasthi, and I. Grossmann. Multiperiod optimization model for oilfield production planning: bicriterion optimization and two-stage stochastic programming model. *Optimization and Engineering*, 07 2019. doi: 10.1007/s11081-019-09455-0.

A. Moolya, A. Rodríguez-Martínez, and I. E. Grossmann. Optimal producer well placement and multiperiod production scheduling using surrogate modeling. *Computers & Chemical Engineering*, 165:107941, Sept. 2022. doi: 10.1016/j.compchemeng.2022.107941.

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. T. Rockafellar and R. J.-B. Wets. The optimal recourse problem in discrete time: L_1-Multipliers for Inequality Constraints. *SIAM Journal on Control and Optimization*, 16(1):16–36, Jan. 1978. doi: 10.1137/0316002.

R. T. Rockafellar and R. J. B. Wets. *Variational Analysis*, volume 317 of *Grundlehren der mathematischen Wissenschaften*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1998. doi: 10.1007/978-3-642-02431-3.

P. Sarma, L. J. Durlofsky, K. Aziz, and W. H. Chen. Efficient real-time reservoir management using adjoint-based optimal control and model updating. *Computational Geosciences*, 10(1):3–36, Mar. 2006. doi: 10.1007/s10596-005-9009-z.

A. Satter and G. M. Iqbal. 13 - decline curve analysis for conventional and unconventional reservoirs. In A. Satter and G. M. Iqbal, editors, *Reservoir Engineering*, pages 211–232. Gulf Professional Publishing, Boston, 2016. doi: 10.1016/B978-0-12-800219-3.00013-9.

L. Silva and C. Guedes Soares. Oilfield development system optimization under reservoir production uncertainty. *Ocean Engineering*, 225:108758, Apr. 2021. doi: 10.1016/j.oceaneng.2021.108758.

D. Silver and J. Veness. Monte-carlo planning in large pomdps. In J. Lafferty, C. Williams, J. Shawe-Taylor, R. Zemel, and A. Culotta, editors, *Advances in Neural Information Processing Systems*, volume 23, page 10. Curran Associates, Inc., 2010.

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.

A. Somani, N. Ye, D. Hsu, and W. S. Lee. Despot: Online pomdp planning with regularization. In C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Weinberger, editors, *Advances in Neural Information Processing Systems*, volume 26. Curran Associates, Inc., 2013.

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.

C. Vessaire, J.-P. Chancelier, M. De Lara, P. Carpentier, A. Rodriguez-Martinez, and A. Robert. Multistage optimization of a petroleum production system with material balance model. *Computers & Chemical Engineering*, 167:108005, Nov. 2022. doi: 10.1016/j.compchemeng.2022.108005.

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.

K. 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.