Optimization Models for Petroleum Field Exploitation

by

Tore Wiig Jonsbråten

Dissertation submitted for the degree of dr. oecon.

Norwegian School of Economics and Business Administration

May 1998

Preface

I was introduced to the area of petroleum field optimization by my main advisor, Professor Kurt Jörnsten, and I would like to thank Kurt for his support, creative comments and inspiring discussions. Associate Professor Dag Haugland at Stavanger College has been my local advisor, and I am grateful for all advices and encouragement these last years. I also acknowledge support from the third member of my advisory committee, Professor Gunnar Stensland.

In 1996 I had the pleasure of being invited as visiting scholar at University of California at Davis. I am very grateful to Professor Roger Wets for the invitation, and for taking interest in my research. In Davis I also met Associate Professor David Woodruff, and my year in California would not have been the same without Dave. The main result of my stay in Davis is Paper B in this dissertation, a paper co-authored with Roger and Dave. I would also like to thank Director Joel Keizer at the Institute of Theoretical Dynamics, UC Davis, for providing office space and excellent working conditions.

The research has been performed at Department of Business Administration, Stavanger College, where I have been given excellent working conditions in a friendly environment. Associate Professor Hans Jacob Fevang has been project responsible at Stavanger College, and I am grateful for his support. Financial support from Stavanger College and the Norwegian Council of Research is gratefully acknowledged.

I would like to thank my family for their encouragement and interest in my work. Finally, I want to thank Hilde for her patience and support during these years.

Stavanger, May 1998

Tore Wiig Jonsbråten

Summary

The purpose of this dissertation is to present and discuss different models for optimal development of a petroleum field. The objective of the optimization models considered is to maximize the project's expected net present value. In this framework we analyze decisions concerning platform capacity, where and when to drill wells and production strategy for each of the wells. When making these decisions, there is considerable uncertainty about the reservoir properties, future oil-price and technical restrictions, and in this work we focus on how to model and solve problems of petroleum field optimization under uncertainty.

The problems discussed in this dissertation lie on the borderline between petroleum economics and reservoir engineering. When considering the problems within an interdisciplinary framework, it is impossible to use the level of detail that can be employed in each discipline let alone. The analysis presented here is therefore on a more aggregated level, which is in accordance with operations research as an interdisciplinary approach of solving complex planning problems. We have chosen a level of detail where a crude reservoir simulator is represented in the optimization model. This allows us to analyze location problems and to model how different decision variables interact with each other. However, it is important to note that such aggregated models do not exclude the need for detailed analysis in the fields of both engineering and economics.

The dissertation is divided into two parts. The first part give an overview of petroleum field optimization from an operations researcher's point of view. Part II consists of four self-contained articles. Although different in scope, these articles all discuss important problems related to optimal exploitation of a petroleum reservoir.

Part I has the title "Overview of petroleum field optimization", and in addition to giving this overview, it ties together the work presented in the

second part of this dissertation. The two first chapters of Part I delimit the topic under consideration and discuss existing literature in the field. In the third chapter we focus on the reservoir, and reservoir equations for a simple reservoir system are derived. We then discuss how these equations may be discretized in order to solve the problems by finite difference methods. In Chapter 4 these discretized reservoir equations are included in optimization models. We first discuss linear programming models for optimizing production decisions, and we also show how these models can be extended to mixed integer programming models where decisions concerning platform, wells and production strategy are optimized. Different versions of the proposed mixed integer programming model are later used in the Papers A, C and D. The Chapters 5 and 6 treat optimization under uncertainty and petroleum field unitization. The first three papers of Part II deal with different issues of oil field optimization under uncertainty while the last paper treats oil field unitization. The Chapters 5 and 6 give a background for the following papers, as well as they compare and discuss the findings in the four papers. Chapter 7 has concluding remarks and directions for future research are outlined.

The first paper in Part II has the title "Oil Field Optimization under Price Uncertainty", and it is the problem of making optimal development decisions under uncertain oil price that is considered. The uncertain oil price is estimated by a finite set of price scenarios with associated probabilities. Our goal is to find the decision policy that optimizes the project's expected net present value, and this is a policy where future decisions are allowed to depend on revealed price information. This is a stochastic mixed integer programming problem and finding the optimal solution to such a problem is not straightforward. Our approach is to use the scenario and policy aggregation technique developed by Rockafellar and Wets for solving the problem. This technique is developed for the case of continuous variables, and we discuss methods for adapting this approach to the case of mixed integer problems. This is done by utilizing the interaction between the continuous (production) and integer (design) variables. Numerical experiments are reported in the paper, and it is concluded that scenario aggregation may be a suitable technique for solving also mixed integer programming problems.

In the next paper, Paper B, we also consider stochastic optimization problems, but we here focus on problems with decision dependent information discovery. The paper has the title "A Class of Stochastic Programs with Decision Dependent Random Elements", and it is co-authored with Roger J.-B. Wets and David L. Woodruff. The motivation for this paper is our interest in modeling and solving optimization problem under reservoir uncertainty, a problem which is addressed in Paper C. When considering the price uncertainty in Paper A, the information about the price development is revealed independently of the development decisions. In a problem with reservoir uncertainty the situation is different, because future information acquisition depends upon which decisions are made. This leads to more complex models and optimization problems that are significantly more difficult to solve, and the literature dealing with such issues is very sparse.

Paper B identifies a class of such problems that are "manageable", and an implicit enumeration algorithm for finding optimal decision policy is proposed. The numerical experiments are done for a general two-stage production planning problem. A manufacturer faces uncertain production costs for a number of components, and the only way to resolve this uncertainty is by producing the component itself or a rather similar one. The uncertainty is represented by discrete probability distributions, and also the variables controlling the random elements are discrete.

The research in the field of decision dependent information discovery is continued in Paper C, which has the title "Optimal Selection and Sequencing of Oil Wells under Reservoir Uncertainty". We here consider problems where reservoir properties are uncertain, but an initial probability distribution over possible reservoir realizations is given. When production wells are drilled, new information about the reservoir is acquired, and we propose a Bayesian model for updating the probability distribution as test results become available. This decision problem can be modeled in terms of a decision tree, and an implicit enumeration algorithm for solving this sequencing problem is proposed. Numerical experiments are performed by use of the mixed integer field optimization model. The results show that including future information discovery in the models may have influence on optimal drilling decisions. Compared to Paper B, the main difference is that Paper C considers a problem where the information discovery is not complete. Each drilled well provides new test results, but these test results do not completely reveal the reservoir realization. In our view, the work reported in Paper B and C points to interesting topics for further research in modeling and solving problems with decision dependent information discovery. We believe that it is the lack of tools rather than the lack of problems that is the reason for the sparse reported research in this field.

In the last paper, "Nash-Equilibrium and Bargaining in an Oil Reservoir Management Game", we discuss the common pool problem arising when two lease owners have access to the same underlying oil reservoir. Because of the

migratory nature of oil, both lease owners have incentives to drain oil from the competitor's part of the reservoir. They also have incentives to produce the present reserves before so is done by the competitor. Our approach is to demonstrate how a mixed integer programming model can be used in order to clarify unitization negotiations. Unitization is the practice of unifying the ownership and control of the reservoir, such that the field is developed and operated by a single operator. The discussion in Paper D is based on a numerical example. When considered as a non-cooperative game we show that there exists a unique Nash-equilibrium, illustrating the over-investment induced by competitive extraction. Considered as a cooperative game the Nash-equilibrium serves as a threat strategy, and we discuss possible bargaining solutions for such problems.

In our view this dissertation points to several interesting topics for further research. There is a continuous need for improved decision support tools in the petroleum industry, and effort should be made in order to develop models and solution methods for more complex reservoir descriptions. It can also be worthwhile to study if and how available reservoir simulators can be combined with the optimization models proposed here. The dissertation also open for a structured treatment of uncertainty, both traditional decision independent uncertainty but also problems where the random elements are decision dependent. Further research should be undertaken both with such optimization models in general and for application within the petroleum industry. Improved reservoir simulators and optimization models will also benefit further research in the area of petroleum field optimization.

Contents

Part I

Overview of Petroleum Field Optimization

Chapter 1

Introduction

The purpose of the first part of this dissertation is to review previous research in optimal management of petroleum reservoirs and discuss the relationship with the contributions in the second part of this dissertation. The term petroleum field optimization is a wide one, and it is not our aim to review all previous literature in this area, but rather to present different approaches to petroleum field optimization. By reviewing central literature we show which traditions our work is based on and also which aspects we have decided not to include in our analysis.

The problem of optimal oil field development is a complex decision problem, and the starting point for such an analysis is the reservoir itself. It is the value of the reservoir one seeks to maximize, and the decisions are constrained by the reservoir properties and several other technical and economical restrictions. Solving such problems in practice necessitates an interdisciplinary approach which among others involves geologists, engineers of different specialties and petroleum economists. Our approach to this problem follows the tradition of operations research in using quantitative techniques for modeling and solving interdisciplinary decision problems. When solving the problem, we do not necessarily aim at finding *the* optimal solution, but rather to provide insight into the problem and increased knowledge about how the decision variables interact.

When working with decision problems on the borderline between reservoir engineering and petroleum economics, it may be difficult to find an appropriate level of detail in each of the disciplines. We can not possibly perform as detailed analysis as can be done in each discipline alone, and the models need to be at an aggregated level. It must also be emphasized that the models presented here do not exclude the need for more detailed engineering and economic analysis. But our approach is to combine these disciplines, and such models must necessarily represent a trade-off in level of detail. By this is meant how the reservoir properties is represented, and which decision variables and technical restrictions that are included in the model. Moreover, when modeling complex decision problems the challenge is to focus on the variables and constraints that play an important role in the problem under consideration and leave the other variables and constraints for the more detailed analysis.

So far we have been rather general about what is meant by petroleum field optimization, but this term will be more thoroughly discussed in Chapter 2. However, it will become clear that in our research we focus on models which consider platform capacity, where and when to drill production wells and production strategies for each of the wells. In these model equations, description of the fluid flow in the reservoir is included, and this model formulation is implemented for a single phase oil reservoir. This reservoir system may be approximated by linear finite difference equations, and the resulting mixed integer models optimize both design and operating decisions. The implemented models are valid for single phase oil, but our analysis uses petroleum reservoirs in general as a starting point. In this dissertation the terms petroleum and oil are to some extent used interchangeably.

In Chapter 2 we look closer at different aspects of petroleum field optimization, and present a review of selected literature. In Chapter 3 we discuss derivation of reservoir equations and how these equations may be discretized. The focus is here on single phase oil reservoirs, and in Chapter 4 it is discussed how a single phase oil reservoir description may be included in optimization models. First we look at optimal production problems, before the model is extended to include field development decisions. The specific optimization model derived in Chapter 4 is used in three of the papers in Part II of this dissertation.

Chapters 5 and 6 treat topics discussed more closely in the papers in Part II. The papers A, B and C discuss problems related to oil-field optimization under uncertainty, and Chapter 5 gives a general background for these papers, as well as the results in the papers are compared and discussed. Chapter 6 discusses the problem of oil field development when there are several lease owners with access to the same underlying reservoir. This problem is also discussed in Paper D. Chapter 7 gives a summary of Part I, a summary of the findings in Part II, and outlines directions for further research.

Chapter 2

Petroleum Field Optimization

The purpose of this chapter is to discuss the term petroleum field optimization and give an overview of previous research in this area. We first look at different applications of operations research in the oil industry, before we look closer at decision problems in field optimization. In a field development analysis it is the reservoir itself that is the basis for the discussion, and it is the value of this resource one seeks to optimize. A petroleum reservoir is a complex system, and in order to describe its performance it is necessary with knowledge about the rock and fluid properties. Using this as input data one can derive equations describing the relationship between fluid flow and pressure in the reservoir. Derivation of reservoir equations is discussed in the next chapter. But even when development decisions are made, the information about the underlying reservoir is still limited, and the available reservoir data are uncertain. Therefore, how to represent the reservoir in the optimization model is not straightforward, and this question is a central issue in this chapter. We will present models which have an explicit description of reservoir performance as well as more simplified models. By explicit reservoir description we mean models where fluid flow equations are included in some way, thus giving rise to a spatial variation in reservoir pressure. We will also look at problems considering a portfolio of potential petroleum fields, and models for optimal sequencing of field development are presented. In this connection we also look at development of transport networks and how such problems may be modeled.

2.1 Operations Research in the Petroleum Industry

The invention and development of the simplex algorithm by Dantzig [22] in 1947 created the area of linear programming, and together with the invention of digital computers, new possibilities were opened in the field of optimization. Since the end of the 1940s, the petroleum industry has been a large scale user of optimization techniques, and there exists a large amount of literature dealing with application of mathematical programming techniques in the oil industry. A large part of this literature describes optimization of processing and blending in the refinery industry, but many other applications are also present. It was linear programming that sparked the use of mathematical programming in the petroleum industry, but also non-linear programming, dynamic programming and integer programming have been frequently used. In this selected review we focus on applications within the petroleum industry and not on solution techniques. An overview of methods and techniques within mathematical programming can be found in Luenberger [65] and Minoux [72].

By following the framework of Garvin et.al. [36] and Foster [32], the petroleum industry may be divided in four areas:

- Exploration
- Production
- Refining
- Distribution and marketing

In the above mentioned articles by Garvin et.al. [36] and Foster [32], published in 1957 and 1964 respectively, reported applications of operations research in the petroleum industry are surveyed. In those early years this literature was relatively sparse, and thorough surveys comprising all four problem areas mentioned above could be written. To our knowledge, more recently published surveys have a narrower scope, and do not cover the oil industry as a whole. Application of optimization techniques for solving planning problems in the refinery industry represents an area where a lot of research has been done, and it is not uncommon that the petroleum industry and refining industry is considered as being synonyms. An example of this is the paper "A history of mathematical programming in the petroleum industry" by Bodington and Baker [17] published in 1990. One could expect that such a paper would report research in all areas mentioned above, but this paper focuses entirely on applications in the refinery industry.

In **exploration** problems decision making under uncertainty is the key issue. Exploration may be defined as an information acquisition and processing activity where the aim is to assess information about possible hydrocarbon systems. This information will be used for determining the prospectivity of a system and further provide information for deciding if and how a potential petroleum field may be developed. Typical decisions are where to explore for petroleum, what kind of petroleum to emphasize, and one needs to decide what type approach of should be used: Seismic surveys or exploration drilling. Further, all these decisions is dependent upon the available exploration budget and what kind of risk the explorer is willing to accept.

Petroleum exploration is a gamble for high stakes. The risk of using considerable amounts of money without making any profitable discoveries is large, but if a profitable reservoir is found the revenue may be huge. Because of the large extent of uncertainty present when making exploration decisions, simple exploration problems are often found in textbooks as a way to introduce stochastic decision trees and decision making under uncertainty [51, 52]. Optimization under uncertainty is discussed in general in Chapter 5. An overview of problems and methods in petroleum exploration can be found in Harbaugh, Doveton, and Davis [42] and Newendorp [75]. Different approaches for analyzing the problem of allocating limited resources among several possible petroleum prospects are found in Bjørstad, Hallefjord and Hefting [15], Flåm [30], Tjøstheim and Hefting [101] and Walls, Morahan and Dyer [106].

The topic "Petroleum field optimization", which is our main concern, lies within the area of **production**. The point of departure here is that through the exploration activities there are found petroleum reserves that are judged to be profitable. This may either be a single reservoir or a portfolio of reservoirs. Problems included in the area of production optimization are field design, drilling optimization, and strategies for production, injection and enhanced oil recovery. Modeling and solving problems related to optimizing the field design and production strategies for a single reservoir are discussed in Sections 2.2 to 2.4, while decision models for development of several fields are found in Section 2.5. Optimization of the drilling process and strategies for enhanced oil recovery will not be a topic in this dissertation. Examples of literature dealing with optimal drilling are Lummus [66] and Reed [84], while problems in enhanced oil recovery are addressed in Thomas [99] and

Thompson and Blumer [100].

As mentioned do problems in the area of **refining** represent the main part of published literature within applications of operations research in the oil industry. Because of the extensive use of planning tools based on mathematical programming techniques, solving refinery problems is seen as one of the major successes of applied operations research. A survey of models for optimizing refinery operations may be found in Haugland [47]. In this survey the literature is divided in three parts: Models for single refineries, models for clusters of refineries and models for single process units. Early work in the area of optimizing operation of a single refinery is found in Charnes, Cooper and Mellon [20] and Symonds [97], while the problem of optimizing the operation of several refineries are addressed by Manne [68], Langston [59] and Al-Zayer [2]. These models for optimizing the operation of refineries have a rather crude representation of the refinery processes, which again creates a demand for detailed models optimizing the processes. While linear programming models often are sufficient for describing refinery operation, the process models are generally non-linear. Examples of literature dealing with refinery processes are Schrage [91] and Friedman and Pinder [34]. By studying the literature dealing with refinery operations, Haugland reports that while optimizing operation of a single refinery was the main concern in the early years, over the years the interest shifted towards that of optimizing operation of clusters of refineries and optimizing refinery processes.

Many problems in the area of **distribution and marketing** in the petroleum industry, are not significantly different from distribution and marketing problems in general. But in Section 2.5 we will look closer at models for transport of oil and natural gas from offshore petroleum fields to onshore refining and distribution. As will be seen, transport decisions do play an important role in field sequencing problems.

2.2 Decisions and Constraints in Field Optimization

In the optimization models discussed in this section, the starting point is that a petroleum reservoir is found to be profitable and it is decided to develop the field. Once that is done, however, a lot of new decisions have to be made. These are decisions with a large extent of interdependence. Our challenge is to model this interdependence, and use this knowledge when developing the analyzed petroleum fields. Field development involves huge amounts of money, and even if potential improvements may represent a small fraction of the total investments, it is still a considerable amount of money.

In Haugland, Hallefjord and Asheim [48] there is presented a list of decisions that may be included in a field optimization model, and that list is given below.

Decisions concerning design:

- Number of platforms
- Number of wells (for production and injection)
- Number of subsea units
- Platform size (capacity)
- Location of platforms
- Location of wells (for production and injection)
- Assignment of wells to platforms

Design and operation:

- Scheduling of platform operations
- Scheduling of well operations
- Abandonment time

Operation:

- Production rates over time
- Injection rates over time
- Enhanced oil recovery

This is not meant to be a complete list of decisions that can be included in a field optimization model. Which variables to include and how they should be represented in the model, depends on the purpose of the analysis and availability of data. This leads us to the challenge in modeling, namely that of including only variables and constraints that have significant influence on the decision policy and leave the rest for more detailed analysis. The problem of optimizing the development of a petroleum reservoir involves a lot of different decisions in a complex environment, but a larger model is not necessarily a better model.

In an early phase when the development decisions are not yet made, there are still a lot of uncertainty associated with the available data. As discussed, exploration decisions are made under a large extent of uncertainty. During the "life cycle" of a reservoir, from discovery until the production is completed, the amount of information about the reservoir is increasing, but even when a field has been produced and is abandoned, the knowledge about the reservoir is not complete. It is usually of no value performing detailed studies when the input data is uncertain, but that does not mean that no analysis should be performed. Decision making under uncertainty plays a central part in this dissertation, and a more thorough discussion of this is found in Chapter 5. Papers A and C deal with oil field optimization under uncertainty regarding future oil price and reservoir properties.

When optimizing field development we seek to optimize the decisions with respect to some economic criterion, and this criterion is usually the net present value. It must be emphasized that maximizing the net present value leads to quite different decisions than if the objective is to maximize total production. In the oil market there are usually not negotiated long term price contracts, and it is the expected world market oil price that is used when calculating the revenues from a potential field. We will not in this dissertation discuss techniques for estimating the future oil price, but we will in Paper A look at how the development may be optimized given a set of future price scenarios. For gas fields the situation is often different. There may be long term price agreements, and it is also usual that the contracts specify a delivery schedule. Instead of optimizing the net present value, gas field models may therefore have minimization of deviation from a target delivery schedule as its objective. The nature of the gas market will be briefly discussed in Section 2.5.

In an analysis of optimal development decisions, it is the reservoir that is the starting point. Our overall aim is to maximize the value of the reservoir, and thus it is of large importance how the reservoir is represented in the model. As discussed by Asheim and Hallefjord [4], the level of detail should depend on the purpose of the analysis and available data, and not on available computing capability and software. In the next sections we will review selected literature in the area of optimal field development, and as done by Hallefjord, Asheim and Haugland [38] we have grouped the literature according to how the reservoir is represented in the models: Models with an explicit description of reservoir performance and models with a simplified description of reservoir performance. As will become clear later on, it is models in this first group we will focus on in this dissertation. Models with simplified descriptions will be surveyed in the next section. In addition to the equations describing reservoir performance, the decisions in a field optimization model may be restricted by technical, economical or logical constraints. Typical technical constraints are platform capacity and well deliverability, while a limited budget results in economical constraints.

Literature dealing with optimal field development and production planning are surveyed by Durrer and Slater [28]. This twenty years old survey also discusses optimization models for drilling, reservoir modeling and enhanced oil recovery. More resent surveys of literature on field optimization are found in Hallefjord [37], Hallefjord, Asheim and Haugland [38] and Haugland [46].

2.3 Models with a Simplified Description of Reservoir Performance

The first published application of linear programming to oil production seems to be by Lee and Aronofsky [61] and Garvin, Crandall, John and Spellman [36]. Both these papers were published in 1957. Also Charnes and Cooper [19] contributed to this field, followed by Aronofsky and Williams [3]. These early papers established principles for oil field optimization. From assumptions on reservoir performance, restrictions on surface facilities and maximization of profit as objective, a time discretized optimization model was established. Aronofsky and Williams look at single well reservoirs where one seeks to optimize the production decisions, and they also propose models for optimizing drilling decisions in a single reservoir. The model introduced by Lee and Aronsofsky was further refined by Attra, Wise and Black [8], and they introduced additional economical and technical constraints. Rowan and Warren [89] formulated the optimal field development problem in an optimal control framework, and they used the idea of least squares fitting when analyzing the available data. Optimal operation of gas fields was studied by O'Dell, Steubing and Gray [79] and Huppler [54].

Devine and Lesso [24] use mixed integer programming when they investigate the "platform location problem". That is the problem of allocating wells to platforms so as to minimize the sum of platform and drilling costs. The number of wells to be drilled is predetermined, and the general structure of this problem is identical to the "warehouse location problem". Frair and Devine [33] extend the platform location problem by including scheduling of activities and production decisions. Other approaches to solving the platform location problem are found in Dogru [26, 27], and in Hansen, de Luna Pedrosa Filho and Ribeiro [41]. Hansen et.al. propose a tabu search heuristic for solving the problem. Devine and Lesso's approach is also discussed by Lilien [63]. His criticism is that as the wells are drilled sequentially, each well drilled add geological information which can possibly alter the position of future targets. Further he writes:

> "To be operationally meaningful to management, the problem should be considered one of the sequential decision sort, with the information added by each well drilled potentially modifying (updating) the model. The development of such a sequential decision procedure could be a topic for future research."

Paper C in this dissertation follows the direction pointed out by Lilien, and it is modeling of sequential information discovery that is the main topic of that paper.

Models for drilling on multilayer petroleum fields are proposed by Devine [23] and Babayev [10]. Beale [11] proposes a model for offshore gas field development, and present numerical experiments performed for a North Sea gas field. The model takes into account reservoir production, compressor and pipeline capacity. Christiansen and Nygreen [21] propose a model for production planning on offshore petroleum fields. Production on several fields are considered, and here also transport capacities are considered. As mentioned, models including transport decisions are discussed further in Section 2.5.

It can be discussed if the many above mentioned models have a simplified reservoir description. Several of the platform location problems are formulated as integer programming models, and the production variables are not part of the model. However, also such models are to some extent based on an analysis of the underlying reservoir when the well locations are determined. We will now look more closely at a couple of models, where the reservoir and its production capacity are modeled more carefully. If one assumes the reservoir to be homogeneous and the pressure to be the same throughout the reservoir, a tank model may be used for describing the relationship between reservoir pressure and the production rates. Such a model is presented by Wallace, Helgesen and Nystad [105], and they show how this model may be used to generate production profiles for an oil reservoir. The same model is also used by Hallefjord, Haugland and Helgesen [39], and both the model itself and the obtained production profiles are compared to an explicit reservoir description and associated production profiles.

We start out by analyzing a single phase oil reservoir where we have the following equation expressing the relationship between the accumulated production from the field and the reservoir pressure:

$$Q_f(t) = R_f(\frac{p^0 - p(t)}{p^0 - \underline{p}_w}) \tag{2.1}$$

where the following notation is used:

$Q_f(t)$	accumulated production from the field
R_f	the field's technically recoverable resources
$p(t)$	volumetrically weighted average reservoir pressure
p^0	initial reservoir pressure
p_w	minimum well pressure

Equation (2.1) may be rewritten more explicitly expressing how reservoir pressure decreases linearly with the accumulated production:

$$p(t) = p_0 - \frac{Q_f(t)}{R_f}(p_0 - p_w) \tag{2.2}$$

Further we assume the following relationship between production from the field and reservoir pressure:

$$q_f(t) = q_f^i(t)(\frac{p_0 - p(t)}{p_0 - p_w}) \tag{2.3}$$

where

$$q_f^i(t) = N_w(t)q_w^i$$

The additional notation introduced here is:

$q_f^i(t)$	initial field production potential
$q_f(t)$	field production potential
q_w^i	initial well production potential
$N_w(t)$	number of wells on the field at time t

Equation (2.3) describes how the production potential on the field decreases linearly with pressure, and the initial field potential can be expressed as the product of the number of wells at time t, $N_w(t)$, and the initial well potential q_w^i. It is important to note that the initial well potential is not equal to the production capacity of the well. The production from each well may be constrained by a capacity limit, γq_w^i, where γ is a constant, and $0 < \gamma < 1$. In addition the production may be constrained by the processing capacity on the platform, q_f^{max}. By combining equations (2.1) and (2.3) the pressure variable $p(t)$ may be eliminated, and we get:

$$q_f(t) = q_f^i(t)(1 - \frac{Q_f(t)}{R_f}) \tag{2.4}$$

which give us a simple description of the potential production from the field. An example of a typical production profile is illustrated in Figure 2.1, and

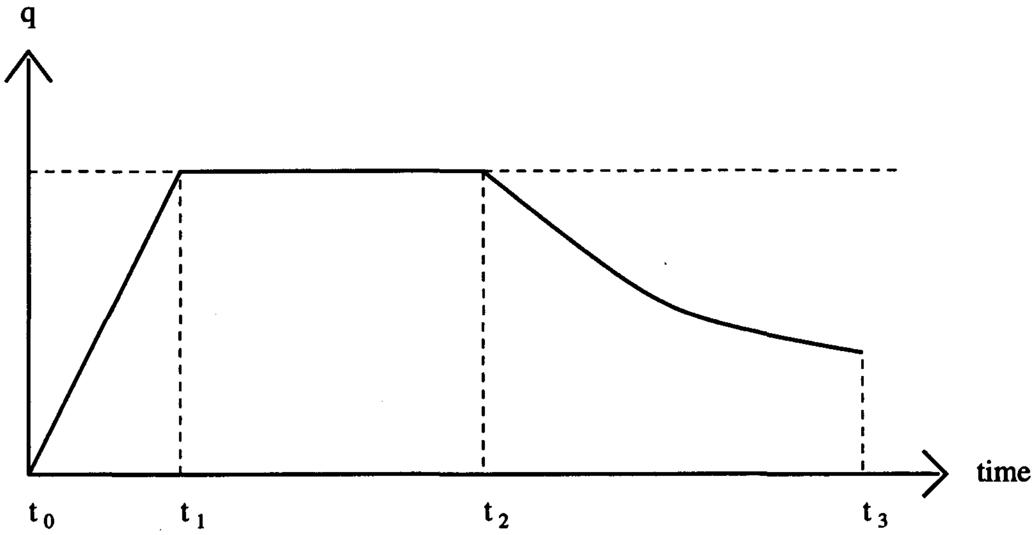

Figure 2.1: Production profile

the rate of production from the field, $q(t)$, may be expressed as:

$$q(t) = \min\{N_w(t)\gamma q_w^i,\ q_f^{max},\ q_f(t)\} \tag{2.5}$$

In the first time interval, $[t_0, t_1)$, new wells are drilled, and each drilled well produces at a rate of γq_w^i. This leads to a stepwise increase in the rate of production as new wells start to produce. We have in Figure 2.1 chosen to approximate this process by a straight line. In the interval $[t_1, t_2)$ the production is constrained by the platform capacity, q_f^{max}, while the production in the interval $[t_2, t_3]$ is described by equation (2.4). At t_3 further production is found to be unprofitable, and the field is abandoned.

Wallace, Helgesen and Nystad [105] also propose an approach for generating production profiles for oil reservoirs with water injection. Also here a tank model is used, and initially one assumes the tank to be filled with oil. As the oil is extracted from the reservoir, water is injected in order to maintain the reservoir pressure. The height of the oil zone is denoted h, and the initial oil zone is of height h_0. The relationship between Q_f and R_f may now be expressed as:

$$Q_f(t) = R_f\left(\frac{h_0 - h(t)}{h_0}\right) \tag{2.6}$$

Furthermore, the production potential for the field is assumed to be:

$$q_f(t) = q_f^i(t)\left(\frac{h(t)}{h_0}\right) \tag{2.7}$$

In words, the production potential is reduced proportionally to the height of the oil zone, h. When equations (2.6) and (2.7) are combined, eliminating h, we see that equation (2.4) also holds for the case of water injection. At a first glimpse, it seems like nothing is achieved by injecting water in the reservoir. However, the technically recoverable resources, R_f, will be larger in a situation with water injection.

Another approach for dealing with reservoir descriptions are found in Odell and Rosing [80]. They develop a mathematical programming model for optimizing offshore oil field development, and the decisions they consider are number and location of platforms, number and location of wells and assignment of wells to platforms. Their main concern is to use this model to analyze the conflicting view between oil companies and the government in respect of optimal field development. However, the model may be used for analyzing optimal field development in general.

While the tank model (0-dimensional) considers the pressure to be homogeneous throughout the reservoir, an opposite view is taken by Odell and Rosing. The reservoir is divided in a number of hexagons, where each hexagon represents a "tank". The size of each hexagon is in accordance with the average spacing of the wells, and the center of each hexagon is considered as a potential well site and platform site. Each platform has a maximum number of wells that can be assigned to it, and the distance between platform and well may not exceed an upper limit. The volume of recoverable oil in each hexagon is found by use of the following formula:

$$a_i = V_i \cdot \phi_i \cdot (1 - B_{oi}S_{wi}) \cdot RF_i \tag{2.8}$$

where

a_i	is the volume of technically recoverable oil in hexagon i
V_i	is the volume of oil saturated sands in hexagon i
ϕ_i	is the porosity
RF_i	is the recovery factor
$B_{oi}S_{wi}$	is a factor that represents the oil/water separation and the expansion factor together

Compared to the tank model, we can say that each hexagon is viewed as a separate reservoir, where the amount of recoverable oil in each hexagon and the associated well productivity are independent of decisions for the neighboring hexagons. Equation (2.8) specifies the amount of recoverable oil in each hexagon, but the production rate over time is not specified by this model. Odell and Rosing have used what they denote "common" production profiles for allocating the total production over the estimated lifetime of the reservoir.

2.4 Models with an Explicit Description of Reservoir Performance

By models with an explicit description of reservoir performance, we mean models where equations describing the fluid flow in the reservoir are included in some way. This does not necessarily mean that fluid flow equations are employed directly in the model, but a reservoir simulator of some kind is a part of the optimization model. In Chapters 3 and 4 we will show derivation of reservoir equations and discuss how equations describing a single phase oil reservoir may be formulated as part of an optimization model. By choosing to include a reservoir simulator in the optimization model, we get rather complex models which are computationally harder to solve, compared to the models reviewed in the previous section. But simplified reservoir models have a quite limited ability to describe the interaction among the decision variables. If location decisions play an important role in the problem under study, it is difficult to use simplified models for describing the reservoir.

The interaction among decision variables are analyzed by Wattenbarger [107], and the problem under investigation is to maximize the total production from a single phase gas reservoir when the demand is subject to seasonal variation. This maximization is constrained by the production potential in each well and the demand for gas. The objective of Wattenbarger's model is to minimize the difference between demand and actual production. However, it is the way the reservoir is represented in the model we will focus on here. The gas reservoir analyzed by Wattenbarger is generally non-linear, but by a priori estimating the average gas density in the reservoir the real gas flow equation may be linearized. (But the average gas density in the reservoir is of course dependent upon the withdrawal rates). For this linearized system the concept of "well interference" is introduced. The idea is that the pressure drop at well i in time period t is dependent upon the withdrawal rates in the wells prior to t, and the pressure drop may be expressed through the following constraint:

$$-\sum_{k=1}^{t}\sum_{j=1}^{N} a_{ij}^{t+1-k} q_i^k \geq \Phi_{\min_i}^t - \bar{\Phi}^t, \quad \text{for all} \quad i = 1, \ldots, N, \; t = 1, \ldots, T \quad (2.9)$$

where T is the number of time steps, N is the number of wells, $\bar{\Phi}^t$ is the real gas pressure at the average reservoir density in time step t, $\Phi^t_{\text{min}_i}$ is the minimum pressure at well i in time step t, a^{t+1-k}_{ij} is a coefficient which represents the pressure drop at well i of a unit production k time steps ago in well j, and q^k_i is the production rate in well i. Wattenbarger's approach to solve this problem is by superposition, which will be discussed in detail in Chapter 4. Instead of introducing a complete reservoir description in the model, production in each of the wells is simulated, and from the results a well interference matrix is constructed.

This idea of using a transient influence matrix as done by Wattenbarger, was proposed by Aronofsky and Williams [3]. But in their paper from 1962 they have not performed numerical experiments for such a reservoir. Also Rosenwald and Green[87] use the principle of superposition, and in addition to production decisions they consider the problem of optimal well location. Their objective is to minimize the difference between demand and production. A set of potential well sites is proposed and the number of wells to drill is specified, and which wells to drill is optimized by the model. The inclusion of drilling decisions necessitates a binary variable, which makes this a mixed integer programming problem. This model is applied to two different reservoirs. The first is a reservoir system with a slightly compressible fluid, and such a system may be described by a linear mathematical model. The second application is a gas reservoir, which is the same kind of reservoir as studied by Wattenbarger, and for this non-linear system the superposition technique is only approximate. Murray and Edgar [74] use the same technique when optimizing operation and design of a gas field. In order to circumvent an integer programming formulation of the problem, it is reformulated as a continuous nonlinear problem. This modeling approach is discussed by Hallefjord, Asheim and Haugland [38]. They state that the minimization problem is still strictly concave, and the computational complexity of this nonlinear problem is comparable to that of the integer formulation.

Haugland, Hallefjord and Asheim [48] use the same approach as Wattenbarger and Rosenwald and Green. Production in single phase oil reservoir is simulated, and from this simulation they get a response matrix that is included in the optimization model. The authors discuss several models for optimizing production and development decisions for an offshore oil field. The objective is to maximize the oil field's net present value, and decisions as platform capacity, which wells to drill and production strategy for each of the wells are considered. Such a model is discussed in detail in Section

4.2. This same model is used by Haugland, Jörnsten and Shayan [49] when studying an oil field exploited by a movable platform, and the model is used for deciding when and where to move the platform on the field.

Lasdon et.al. [60] propose a model for optimizing operating decisions for a dry gas reservoir. Non-linearities in the problem results both from the pressure dependency of the viscosity and density, and from well deliverability constraints on the form:

$$q_i^t \leq c_j \left[(p_j^t)^2 - (pb_j)^2\right]^{n_j}, \quad j = 1, \ldots, N, \quad t = 1, \ldots, T. \tag{2.10}$$

where p_j^t is the pressure at well j, pb_j is the associated back pressure, and c_j and n_j are well specific constants. Given the well flow and the initial reservoir condition, the reservoir pressure may be calculated and thus the optimization problem may be viewed as a problem in production variables only. Lasdon et.al. use the reduced gradient method for solving the non-linear problem. A more detailed discussion of the reduced gradient method may be found in Mantell and Lasdon [69]. Asheim also uses the reduced gradient method when optimizing the production strategies for reservoirs containing slightly compressible fluids [6] and for two-phase oil/water reservoirs [7].

See and Horne [92] use another approach for optimizing production decisions. Their method is performed in two steps. First, a commercially available reservoir simulator is used to perform a series of experiments, and a multiple variable regression analysis is used to fit the experimental data. Expressed by the notation used by See and Horne, a set of equations of the type

$$Y_j = a_{j0} + a_{j1}X_1 + \ldots + a_{jn}X_n, \quad j = 1, \ldots, p \tag{2.11}$$

are defined. Here Y_j is the j-th performance variable while X_i is the i-th decision variable. There are p performance variables and n decision variables. The reservoir simulations result in a sequence of possible decisions and associated performance variables. By use of least squares fit the coefficients $a_{j0}, a_{j1}, \ldots, a_{jn}$ are estimated, and the equations (2.11) is used as constraints in an optimization model. In the problem investigated by See and Horne the development decisions are made, and the production optimization is solved as a linear programming problem. For each timestep an LP problem is solved, and this solution procedure is repeated for each consecutive timestep. As discussed by See and Horne [92] (and Hallefjord, Asheim and Haugland [38]), there are particularly two difficulties that are connected to this method. First, by assuming a linear relationship between decision and performance variables a linearization error is introduced, and care must be taken to analyze this error. Second, by optimizing one timestep at the time, a possibility for arriving at suboptimal solutions is introduced. Several timesteps may be optimized simultaneously, but this will inevitably lead to larger problems.

Another approach for using a reservoir simulator that is external to the optimization model is found in Asheim [5]. Asheim uses an algorithmic value function which is an unconstrained, non-linear representation, with the decision variables as arguments and the corresponding present worth as function value. The use of this algorithmic function avoids the computational problems associated with mathematical programming as we have reviewed. However, Asheim's model may not be very successful in finding the optimal solution to complex optimization problems. Also Nystad [78] uses a reservoir simulator to generate "data points" which are used as input to his optimization model.

2.5 Models Including Sequencing and Transport Decisions

We have so far mainly looked at models for optimizing the development of a single petroleum field, but we will in this section focus on problems involving several reservoirs. The starting point for such an analysis is a portfolio of potential petroleum fields, and questions treated by such a model may be:

- Which of the potential fields should be put into production ?
- When should the selected fields start to produce ?
- Which means of transportation should be chosen for the new fields ?
- Should the means of transport for producing fields be changed ?
- In what order should the selected transport modules be developed ?
- How should the constructed transport system be utilized ?

Even if optimal development solutions and associated production profiles are given for the individual fields, the number of decision variables in such a model can get very large. We will in this section look closer at two different attempts of solving problems with field sequencing and transport decisions. In the first model, given by Jörnsten [57], the emphasis is on the field sequencing decisions, and the transportation network is given in aggregate form. In the other model, proposed by Aboudi et.al. [1], the development of a transport network plays a central role, and the model seeks to answer all of the six above mentioned questions.

As mentioned, the model proposed by Jörnsten seeks to decide which fields to put into production and when they should start to produce. The problem is formulated as a 0-1 integer programming problem, and the decision variables for field development are defined as follows:

$$y_i(k) = \begin{cases} 1 & \text{if field } i \text{ is chosen to start production in time period } k, \\ 0 & \text{otherwise.} \end{cases} \tag{2.12}$$

The objective of the model is to maximize the net present value of the set of potential fields, and the objective function is written as

$$\max \sum_i \sum_{k \in K_i} P_i(k) y_i(k) \tag{2.13}$$

where

i	field index
K_i	is the set of possible starting years for field i
$P_i(k)$	is the net present value if field i starts its production in year k

We further assume that the total production in each period needs to be within a certain range. This can be necessary of various reasons, for example long term agreements about delivery of petroleum products. Jörnsten allows for several petroleum products in his model (oil, natural gas). By defining

$L^q(t)$	lower bound on production of product q in period t
$U^q(t)$	upper bound on production of product q in period t
$a_i^q(k,t)$	the amount of product q produced by field i in period t given that the field starts production in period k.

the constraints on production may now be written:

$$L^q(t) \leq \sum_i \sum_{k \in K_i} a_i^q(k,t) y_i(k) \leq D^q(t), \qquad \text{for all } t \text{ and } q \tag{2.14}$$

A typical configuration of offshore petroleum fields is that a large petroleum field has several smaller fields in its neighborhood. This large field, which is denoted the mother field, is usually developed first. The smaller fields, the satellites, may be developed later using the processing capacity at the mother field, and the production start at the satellites is constrained by the limited processing capacity on the mother field:

$$\sum_{i\in G(m)} \sum_{k\in K_i} a_i^q(k,t)y_i(k) + \sum_{k\in K_i} a_m^q(k,t)y_m(k) \leq H_m^q(t) \text{ for all } t, q \text{ and } m \tag{2.15}$$

In this constraint the following notation has been employed.

- m denotes a particular mother field
- $G(m)$ is the set of satellite fields belonging to m
- $H_m^q(t)$ is the total processing capacity for product q installed at motherfield m in time period t.

In this model the transportation network is only given on aggregated form, and the associated capacity constraints are given as:

$$\sum_{i\in J(s)} \sum_{i\in K_i} a_i^q(k,t)y_i(k) \leq T_s(t), \qquad \text{for all } t \text{ and } s \tag{2.16}$$

where s is the transport system, $J(s)$ is the set of fields that use this transport system, and $T_s(t)$ is the capacity associated with transport system s at time t. Jörnsten also show how constraints due to a limited budget or political regulations may be imposed on the decision policy. In addition, such a model needs some logical constraints in order to be complete. Examples of such constraints are: A field can only be developed once and a satellite field cannot be developed before its associated mother field.

For problems of realistic size, Jörnsten reports that solving them to optimality is hard, and several heuristics for approaching the problem are proposed. Further, uncertainty regarding the future demand is introduced, and solution techniques for maximizing the expected net present value are suggested. Also Haugen [44] presents a model for optimal sequencing of offshore petroleum fields where the transportation decisions are at an aggregated level. Two alternative objective functions are used. The first seeks to minimize the deviation from a given demand profile. The other objective function seeks to maximize the net present value, but deviations from the demand profile are penalized. Haugen introduces uncertainty regarding the amount of recoverable petroleum in the reservoirs, and he arrives at a dynamic stochastic programming formulation for solving the problem.

In the model proposed by Aboudi et.al. [1], the development of a transport system plays an important role. The problem is formulated as a network flow problem where the offshore petroleum fields and onshore terminals are modeled as nodes in the network, while the links (edges) in the network represent means for transporting the petroleum between the nodes. A link may be a pipeline connecting two nodes, but it may also be a transport system with loading buoys and ships. The set of nodes, N, in this network may be divided into three disjoint subsets:

N_E	the set of existing nodes
N_P	the set of potential nodes
N_T	the set of terminal nodes

The nodes N_E and N_P may be petroleum fields, sources in the network, or just nodes connecting links in the transport network. The terminal nodes, N_T, represent onshore terminals which serves as end nodes, sinks, in the network. For the transport links, L, we develop the following notation:

L_E	the set of existing links
L_P	the set of potential links

We further denote

L_n^{in}	the set of links which end in node n
L_n^{out}	the set of links which start in node n

In addition to the decision variables for field development, $y_i(k)$, there are decision variables for development of transport network and for network flow:

$$z_l(k) = \begin{cases} 1 & \text{if link } l \text{ is chosen to start in time period } k, \\ 0 & \text{otherwise.} \end{cases} \tag{2.17}$$

$$x_l(k) = \text{ quantity of petroleum transported on link } l \text{ in time period } k \tag{2.18}$$

As in the model proposed by Jörnsten, also this model can handle several petroleum products, each transported in its associated subnetwork. However, for notational simplicity we will here present the model with just one product.

The objective of this model is to maximize the net present value of the portfolio of petroleum fields, and the net present value is expressed as the difference between the revenue from sale of petroleum, and the costs associated with developing and operating the fields and the transport links. The revenues may be expressed as:

$$\sum_{t=1}^{T} \sum_{n \in N_T} p_n(t) \sum_{l \in L_n^{in}} x_l(t) \tag{2.19}$$

where $p_n(t)$ is the discounted price of petroleum delivered at terminal node n in period t. The investment costs are:

$$\sum_{t=1}^{T}\sum_{k=1}^{T}\left(\sum_{n\in N_P} d_n(k,t)y(n,k) + \sum_{l\in L_P} g_l(k,t)z(l,k)\right) \tag{2.20}$$

where:

$d_n(k,t)$ is the cost in period t if node n starts production in period k
$g_l(k,t)$ is the cost in period t if link l starts in period k

The cost of operating the transport links is given as:

$$\sum_{t=1}^{T}\sum_{l\in L} c_l(t)x_l(t) \tag{2.21}$$

where $c_l(t)$ is the cost of transporting one unit along the link l in period t. The constraints may be grouped into three main categories:

- Conservation of flow constraints
- Budget constraints
- Logical constraints

The budget constraints require that the investment and operating costs each period are lower than a certain maximum amount. We will here look closer at the flow conservation constraints, while a more detailed discussion of budget and logical constraints may be found in Aboudi et al.

Each potential field is given with a fixed production profile:

$a_n(t)$ petroleum produced in an existing node n in time period t
$a_n(k,t)$ petroleum produced in the potential node n in time period t if the node starts its production at time k

For nodes with no production, the coefficients $a_n(t)$ and $a_n(k,t)$ are equal to zero. For potential nodes the production in period t now may be written as:

$$\sum_{k=1}^{T} a_n(k,t)y_n(k) \tag{2.22}$$

The equations for flow conservation in the nodes can then be written:

$$\sum_{l\in L_n^{in}} x_l^t = \sum_{l\in L_n^{out}} x_l^t + a_n(t), \qquad n\in N_E \tag{2.23}$$

and

$$\sum_{l \in L} x_l^t = \sum_{l \in L} x_l^t + \sum_{k=1}^{T} a_n(k,t) y_n k \qquad n \in N_P \tag{2.24}$$

We see that the equations simply state that the flow to a node must equal the flow from the node. In addition, there are capacity constraints imposed on the transport links:

$$x_l^t \leq U_l \tag{2.25}$$

where U_l is the transport capacity on link l, and limited capacity in a node:

$$\sum_{l \in L_n^{in}} x_l^t \leq L_n \tag{2.26}$$

where L_n is the node capacity.

Rather similar models to the one proposed by Aboudi et.al. are discussed in Haugen [43] and Haugen, Bjørkvoll and Minsaas [45]. Earlier literature on optimal pipeline design includes Rothfarb et. al [88], Bohannon [18] and Dogru [27].

Chapter 3

Reservoir Description and Discretization

We will in this chapter give a mathematical description of reservoir properties and use this work to study how the reservoir system may be represented in an optimization model. It is single phase oil reservoirs that will be the most central part of this discussion, but also other reservoir systems will be discussed. A more general discussion of topics in this chapter can be found in Aziz and Settari [9], Peaceman [82] and Thomas [99]. The list below gives a description of central notation:

ρ	fluid density	(kg/m^3)
v	fluid flow velocity	(m/s)
w	source/sink terms	$(kg/m^3/s)$
ϕ	porosity	fraction
t	time	(s)
k	permeability	(m^2)
μ	fluid viscosity	$(bar \cdot s)$
p	pressure	(bar)
c	constant temperature compressibility	(bar^{-1})

3.1 Single Phase Oil Reservoir

By a single phase oil reservoir is meant a reservoir system where only the oil is mobile. The fluid is slightly compressible, and production is made possible as the fluid expands when the reservoir pressure is reduced. The reservoir description is found by combining an equation for mass conservation with Darcy's flow equation that give the relation between the pressure gradient and the fluid flow velocity. Combined with the equation for constant temperature compressibility, these equations give a complete description of reservoir behavior.

3.1.1 Mass Conservation

Let us first consider a cylindrical core of a porous medium with fluid flow in the axial direction. The mass flux vector $\rho\vec{v}$ represents mass flow per unit area per unit time. The area of a crossection of this cylinder is denoted A, and in a time interval Δt, the mass that flow across a crossection at the locations x and $x + \Delta x$ can be expressed as

$$(\rho v)_x A\Delta t \qquad \text{and} \qquad (\rho v)_{x+\Delta x} A\Delta t$$

The difference between inflow and outflow of this control volume is either due to mass accumulation when fluid is compressed or due to a mass sink in the control volume. This control volume is denoted $\Delta V = A\Delta x$, and mass accumulation due to compressibility can for the time interval Δt be expressed as:

$$\frac{\partial(\rho\phi)}{\partial t}\Delta V\Delta t$$

A mass sink w in the volume ΔV in the time interval Δt can be expressed the following way:

$$w\Delta V\Delta t$$

The difference between mass inflow and outflow of the control volume V is then either due mass accumulation or the mass sink:

$$((\rho v)_x - (\rho v)_{x+\Delta x})A\Delta t = \frac{\partial(\rho\phi)}{\partial t}\Delta V\Delta t + w\Delta V\Delta t \tag{3.1}$$

We now divide this equation by $\Delta V\Delta t$:

$$\frac{(\rho v)_x - (\rho v)_{x+\Delta x}}{\Delta x} = \frac{\partial}{\partial t}(\rho\phi) + w \tag{3.2}$$

By taking the limit as $\Delta x \to 0$ we get the following equation:

$$-\frac{\partial(\rho v)_x}{\partial x} = \frac{\partial}{\partial t}(\rho\phi) + w \tag{3.3}$$

We have now found the equation for mass conservation under one directional mass flow. An equivalent analysis can be performed for three dimensional mass flow, and this general description can be expressed as:

$$-\nabla \cdot (\rho\vec{v}) = \frac{\partial(\rho\phi)}{\partial t} + w \tag{3.4}$$

3.1.2 Darcy's Law

Darcy's law gives the relationship between the flow velocity and the pressure gradient. This relationship for single phase flow was discovered by Darcy [53] in 1856, and it is an entirely experimental relationship. The differential form of Darcy's law is

$$\vec{v} = -\frac{k}{\mu}(\nabla p + \rho\frac{\vec{g}}{g_c}) \tag{3.5}$$

where k is the permeability of the porous media, μ is the fluid viscosity, $\vec{g}$ is the gravitational acceleration vector and g_c is a constant. The gravitational term of this equation is in many cases negligible, and Darcy's law can then be written as:

$$\vec{v} = -\frac{k}{\mu}\nabla p \tag{3.6}$$

In the optimization models we have implemented, we have considered two dimensional fluid flow as sufficient for describing the reservoir dynamics, and the fluid flow in the vertical direction is neglected. This will be further discussed in Section 3.4.1.

3.1.3 Constant Compressibility

Under isothermal conditions we assume the fluid to have constant compressibility in the pressure interval under consideration, which gives the following relationship:

$$c = \frac{1}{\rho}\frac{d\rho}{dp} \tag{3.7}$$

By integrating equation (3.7) we get:

$$\rho = \rho^0 \exp[c(p - p^0)] \tag{3.8}$$

In this equation ρ^0 is the fluid density at the reference pressure p^0. Unless the fluid contains a lot of dissolved gas, this is a reasonable approximation. A first order approximation of this equation is

$$\rho = \rho^0(1 + c(p - p^0)) \tag{3.9}$$

3.1.4 Equation for Reservoir Description

By combining Darcy's law (3.6) and the equation for mass conservation (3.4) we get:

$$\nabla \cdot (\frac{\rho k}{\mu}\nabla p) = \frac{\partial(\rho\phi)}{\partial t} + w \tag{3.10}$$

The equation for constant temperature compressibility (3.7) can be expressed as

$$\rho dp = \frac{d\rho}{c}$$

which can be rewritten on a more general form:

$$\rho \nabla p = \frac{1}{c} \nabla \rho \tag{3.11}$$

By combining the equations (3.11) and (3.10) we now get:

$$\nabla \cdot \left(\frac{k}{\mu c} \nabla \rho\right) = \frac{\partial(\rho\phi)}{\partial t} + w \tag{3.12}$$

We now assume that the porosity ϕ is pressure independent, and we use the first order approximation of the equation for constant temperature compressibility. When combining this with equation (3.12) we get the following relationship:

$$\nabla \cdot \left(\frac{k}{\mu} \nabla p\right) = \phi c \frac{\partial p}{\partial t} + \frac{w}{\rho^0} \tag{3.13}$$

This is a complete description of reservoir pressure. Since we have assumed constant temperature compressibility, the coefficients in equation (3.13) are pressure independent. Discretization of this equation will be discussed in Section 3.4.

3.2 Other Reservoir Systems

The proposed models and computational experiments in this dissertation are valid for a single phase oil reservoir, but it is a fact that for most of the petroleum producing reservoirs this is a too simple description. As will be shown, single phase gas reservoirs have non-linear reservoir equations. In oil production it is usual to inject water in the reservoir in order to increase the pressure. For proper modeling of such reservoirs, it is necessary to introduce reservoir models with multiple phases. We will here briefly describe a single phase gas reservoir and reservoir systems with several phases. This is done in order to point out direction for future research, when developing optimization models with more complex reservoir models.

3.2.1 Single Phase Gas Reservoirs

A single phase gas reservoir is a system where only the gas is mobile, and this is a system which gives a good description of most gas reservoirs. One would maybe assume that such a reservoir could be described in the same way as a single phase oil reservoir. However, neither the constant temperature compressibility nor the viscosity of the gas are constant, and thus we get a system of non-linear reservoir equations.

One way of describing this reservoir system, is by introducing the gas equation of state:

$$\rho = \frac{pM}{zRT} \tag{3.14}$$

where

M	averaged gas mole weight	$(kg/kmol)$
T	absolute temperature	(K)
R	universal gas constant	$= 8314 Nm/kmol/K$
z	gas deviation factor	fraction

Except for the z-factor, this is the equation describing the behavior of an ideal gas. In other words, when $z = 1$, the equation describes ideal gas behavior, while at real gas reservoir conditions the z-factor may be in the order of 0.7-0.8. This parameter depends on pressure, temperature and gas molecular composition. When equation (3.14) is combined with equation (3.10), we arrive at the following gas reservoir equation:

$$\frac{M}{RT}\nabla \cdot (\frac{k}{\mu}\frac{p}{z}\nabla p) = \frac{M}{RT}\phi\frac{\partial}{\partial t}\frac{p}{z} + w \tag{3.15}$$

Unlike the coefficients in the single phase oil reservoir equation, the coefficients in (3.15) are pressure dependent. We will later study how the reservoir descriptions may be incorporated in the optimization models, but it is obvious that introduction of non-linearities leads to more difficult models to solve. This topic will be briefly discussed in Section 4.1.5.

3.2.2 Reservoirs with Two and Three Phase Flow

In most of the oil production taking place today the oil is displaced by either water or gas. There is no clear boundary between the phases, but rather a simultaneous flow of several phases through the porous media. We will not give a thorough discussion of such reservoirs, but rather point out the main difference between single and multiple phase reservoirs. This is done in order to identify the challenges when incorporating more complex reservoir systems in optimization models.

We consider two phases, oil and water, represented with subscripts o and w. The saturation of a phase is defined as the fraction of the pore volume that is filled by that phase. For a two phase oil-water system we then have the following phase saturation relationship:

$$S_o + S_w = 1$$

Because of surface tension the pressure in the oil phase will be higher than the pressure in the water phase. The difference, p_c, is denoted the capillary pressure:

$$p_c = p_o - p_w$$

The capillary pressure is a function of saturation, $p_c(S_w)$. In the same way as for the single phase reservoir, we can derive equations describing the mass conservation of the oil phase

$$-\nabla \cdot (\rho_o \vec{v}_o) = \frac{\partial(\rho_o S_o \phi)}{\partial t} + w_o \tag{3.16}$$

and for the water phase

$$-\nabla \cdot (\rho_w \vec{v}_w) = \frac{\partial(\rho_w S_w \phi)}{\partial t} + w_w. \tag{3.17}$$

Before describing Darcy's law for multiphase flow, we have to look closer at the permeabilities when there are simultaneous flow of several fluids through a porous medium. Because of this simultaneous flow, each fluid interferes with the flow of the other, which means that the permeability for each of the fluids is lower than the single fluid permeability, k. The permeability for each of the fluid's is denoted effective permeability, written k_o for the oil phase and k_w for the water phase. It is now possible to define relative permeability as follows:

$$k_{ro} = \frac{k_o}{k} \quad \text{and} \quad k_{rw} = \frac{k_w}{k}$$

The relative permeability is the ratio between the effective permeability and the single phase permeability, and this ratio will always be less than or equal to one. Darcy's law may now be expressed for each of the phases, and when neglecting the gravitational term we get:

$$\vec{v}_o = -\frac{k k_{ro}}{\mu_o} \nabla p_o \tag{3.18}$$

and

$$\vec{v}_w = -\frac{k k_{rw}}{\mu_w} \nabla p_w. \tag{3.19}$$

By combining Darcy's law and the equations for mass conservation, (3.16) and (3.17), we arrive at the following equations:

$$\nabla \cdot \left(\frac{\rho_o k k_{ro}}{\mu_o} \nabla p_o\right) = \frac{\partial(\rho_o S_o \phi)}{\partial t} + w_o \tag{3.20}$$

$$\nabla \cdot \left(\frac{\rho_w k k_{rw}}{\mu_w} \nabla p_w\right) = \frac{\partial(\rho_w S_w \phi)}{\partial t} + w_w \tag{3.21}$$

These equations describe the performance of an oil-water system, but we see that employing them in an optimization model is not straightforward.

In the case of a three phase system, an oil-water-gas reservoir, a set of rather similar equations may be derived (cf. Peaceman [82]). But in such a reservoir system, one must also take the mass transfer between the phases into account. For example, there will usually be gas dissolved in the oil, but when the pressure is reduced this gas will vaporize and become a part of the gas phase.

3.3 Discretization of Partial Differential Equations

We will in this section discuss discretization of partial differential equations in a general setting, and our motivation is the need to discretize the reservoir equations in order to solve them. Only for very simple reservoir systems can an analytic solution to equation (3.13) be found, and this fact makes it necessary to use numerical techniques in order to solve the partial differential equation.

The idea is to replace the original problem with a problem that is easier to solve and has a solution that is close to the solution of the original problem. This is achieved by using the finite difference method, where one seeks to find approximate values of the solution at a finite number of locations in the reservoir at a finite number of points in time. But in order to use such a method, we need to know how close to the real solution the approximated solution is, and what it implies for the discretization in space and time. We will first look at spatial discretization before we discuss discretization in time. We will then briefly discuss the convergence and stability properties of the different methods.

Numerical techniques for solving partial differential equations is the topic in a lot of specialized textbooks and scientific papers, A further treatment of this topic can be found in Aziz and Settari [9], Mitchell and Griffiths [73], Peaceman [82] and Smith [93].

3.3.1 Spatial Discretization

Let us start by considering an ordinary differential equation on the form:

$$AU \equiv \frac{d^2U}{dx^2} - w(x) = 0 \qquad 0 < x < L \tag{3.22}$$

where A represents the differential operator and U represents the function we want to solve. In order to do this we need boundary conditions; e.g. $U(0) = U(L) = 0$. Instead of trying to find a continuous function, $U(x)$, that satisfies equation (3.22), we rather want to find an approximate solution, u, in a finite number of points $x_0, x_1, x_2, ..., x_N$ inside the interval $(0, L)$. The differential equation is replaced by a set of algebraic equations that give the approximated solution $u_0, u_1, u_2, ..., u_N$ at $x_0, x_1, x_2, ..., x_N$. The original problem is given as

$$AU = 0$$

but instead we solve the simpler problem

$$Bu = 0$$

where B is a finite difference operator that approximates the differential operator A. If we define U_i to be the true solution in the point x_i, we have the following relationship:

$$AU_i = BU_i + R_i \tag{3.23}$$

where R_i represents the local discretization error.

We want to find solutions in the points x_i inside the interval $(0, L)$, and the spacing between the points is $\Delta x = x_{i+1} - x_i$. The Taylor series expansion for the points x_{i-1} and x_{i+1} can be written:

$$U_{i+1} = U_i + U_i'\Delta x + U_i''\frac{\Delta x^2}{2} + U_i'''\frac{\Delta x^3}{6} + U_i''''\frac{\Delta x^4}{24} + U_i'''''\frac{\Delta x^5}{120} + \cdots \tag{3.24}$$

$$U_{i-1} = U_i - U_i'\Delta x + U_i''\frac{\Delta x^2}{2} - U_i'''\frac{\Delta x^3}{6} + U_i''''\frac{\Delta x^4}{24} - U_i'''''\frac{\Delta x^5}{120} + \cdots \tag{3.25}$$

Using the above equations we can derive difference approximations for U_i' and U_i''. By solving equation (3.24) for U_i' we get the "forward difference approximation":

$$U_i' = \frac{U_{i+1} - U_i}{\Delta x} + R_i^f \tag{3.26}$$

where R_i^f is the local discretization error. This discretization error consists of all higher order derivatives in equation (3.24):

$$R_i^f = -U_i''\frac{\Delta x}{2} - U_i'''\frac{\Delta x^2}{6} - U_i''''\frac{\Delta x^3}{24} - \cdots \tag{3.27}$$

In the same manner we can find the "backward difference approximation" from equation (3.25):

$$U_i' = \frac{U_i - U_{i-1}}{\Delta x} + R_i^b \tag{3.28}$$

where R_i^b is the local discretization error of this approximation:

$$R_i^b = U_i''\frac{\Delta x}{2} - U_i'''\frac{\Delta x^2}{6} + U_i''''\frac{\Delta x^3}{24} - \cdots \tag{3.29}$$

For finding an approximated solution to equation (3.22) we need an expression for the second derivative, U_i'', and by adding the equations (3.24) and (3.25) we get:

$$U_i'' = \frac{U_{i-1} - 2U_i + U_{i+1}}{\Delta x^2} + R_i^r \tag{3.30}$$

where R_i^r is the error term:

$$R_i^r = -U_i''''\frac{\Delta x^2}{12} - U_i''''''\frac{\Delta x^4}{360} - \cdots \tag{3.31}$$

By replacing U_i'' in equation (3.23) by the expression in equation (3.30) we get:

$$AU_i = \frac{U_{i-1} - 2U_i + U_{i+1}}{\Delta x^2} - w_i + R_i^r \tag{3.32}$$

where $w(x_i)$ is written as w_i. The discretization error is a result of the Taylor series expansion. We have now found a finite difference operator that approximates the differential operator A. As mentioned, we do not know the exact solution U_i, and instead we therefore solve the problem:

$$Bu_i = \frac{u_{i-1} - 2u_i + u_{i+1}}{\Delta x^2} - w_i \tag{3.33}$$

where u_i is the finite difference approximation to U_i. In this discussion of spatial discretization we have only considered a one-dimensional example, but the same principle applies also in higher dimensions.

3.3.2 Discretization in Time

We now include time dependent term in equation (3.22) and get a parabolic equation on the form:

$$\frac{\partial^2 U}{\partial x^2} = \frac{\partial U}{\partial t} + w \tag{3.34}$$

where U and w are both space and time dependent. The reservoir description derived in Section 3.1 is a parabolic equation, where U corresponds to the reservoir pressure and w is the source/sink term. We have so far discussed discretization of the left side of equation (3.34), and we will now look closer at the right hand side. The time can be discretized in time steps of length Δt, and we will seek to find numerical solutions of the equation at discrete points in time: $t_0 = 0, t_1 = \Delta t, ..., t_n = n\Delta t, ...$ The simplest way of approximating the time derivative term is by using a "forward difference approximation":

$$\frac{du_i^n}{dt} \simeq \frac{u_i^{n+1} - u_i^n}{\Delta t} \tag{3.35}$$

The discretized equation can then be written:

$$\frac{u_{i-1}^n - 2u_i^n + u_{i+1}^n}{\Delta x^2} = \frac{u_i^{n+1} - u_i^n}{\Delta t} + w_i \tag{3.36}$$

$$i = 1, 2, ..., N \qquad\qquad n = 0, 1, 2, ...$$

This is the explicit formulation where the equation at time step t_n do only contain one term at next point in time, t_{n+1}. It leads to a solution method where the equation system can be solved explicitly point by point; In each equation there will only be one unknown variable.

Another way of approximating the time derivative is by use of "the backwards difference method". The time derivative can then be expressed as:

$$\frac{du_i^{n+1}}{dt} \simeq \frac{u_i^{n+1} - u_i^n}{\Delta t} \tag{3.37}$$

and the discretized version of equation (3.34) on implicit form is then:

$$\frac{u_{i-1}^{n+1} - 2u_i^{n+1} + u_{i+1}^{n+1}}{\Delta x^2} = \frac{u_i^{n+1} - u_i^n}{\Delta t} + w_i \tag{3.38}$$

$$i = 1, 2, ..., N \qquad\qquad n = 0, 1, 2, ...$$

At each time step there will here be N equations that must be solved simultaneously.

3.3.3 Consistency, Convergence and Stability

We have so far looked at the explicit and implicit method for discretizing partial differential equations, but we have not yet discussed how good approximations to the solutions of the original problem we get by using the presented methods. Another important problem is how the discretization errors depend upon chosen Δx and Δt. For such a discussion we need to look closer at three important notions: Consistency, convergence and stability. We will here briefly discuss each of this notions and how they interact.

Consistency

Consistency is a property of the difference operator. A difference operator is said to be a consistent approximation of the differential operator if the discretization error tends to zero when the distance between the nodes, Δx, tends to zero.

By using the "forward difference operator" in equation (3.26) we found that the local discretization error could be expressed:

$$R_i^f = -U_i''\frac{\Delta x}{2} - U_i'''\frac{\Delta x^2}{6} - U_i''''\frac{\Delta x^3}{24} - \cdots$$

We now want to study the asymptotic behavior of this error term; what happens as Δx tends to zero. For the error term of the "forward difference operator" we get:

$$\lim_{\Delta x \to 0} R_i^f = \lim_{\Delta x \to 0}(-U_i''\frac{\Delta x}{2}) \tag{3.39}$$

We see that the error term proportionally tends to zero at least as fast as Δx tends to zero. This may be expressed as

$$R^f = O(\Delta x) \tag{3.40}$$

The "forward difference approximation" in equation (3.26) can now be written as follows:

$$U_i' = \frac{U_{i+1} - U_i}{\Delta x} + O(\Delta x) \tag{3.41}$$

By looking at the error term in the difference equation for the second derivative we find $R^r = O(\Delta x^2)$, and equation (3.30) can then be expressed as

$$U_i'' = \frac{U_{i-1} - 2U_i + U_{i+1}}{\Delta x^2} + O(\Delta x^2) \tag{3.42}$$

The error term can be made arbitrarily small by choosing a sufficiently small value of Δx. This makes the difference approximation consistent. With reference to equation (3.23), we can say the difference operator B is a consistent approximation to the differential operator A in the point x_i if the error term R_i satisfies

$$R_i = AU_i - BU_i \rightarrow 0 \qquad \text{when} \qquad \Delta x \rightarrow 0 \tag{3.43}$$

As long as approximations of integer orders are used, B is a consistent approximation if $R_i = O(\Delta x^p)$ where $p \geq 1$.

Convergence

We let e_i be the error term in the approximated solution at x_i:

$$e_i = U_i - u_i \tag{3.44}$$

Convergence can now be defined as follows:

A difference operator B is convergent to the differential operator A if $e \rightarrow 0$ when $\Delta x \rightarrow 0$.

The definitions of consistency and convergence are closely related, but a consistent operator is not necessarily a convergent operator. It can be shown that the explicit method in equation (3.35) is consistent but only conditionally convergent [9]. An analysis of convergence properties may be rather complicated, but convergence can be proved via consistency and stability.

Stability

A general definition of stability can be found in Aziz and Settari [9]:

A numerical algorithm is considered stable if any errors introduced at some stage of computation do not amplify during subsequent computations.

This is an important problem when solving time dependent partial differential equations. An error introduced at some level will affect the solutions at all later time levels. The connection between stability and convergence is given by Lax's Equivalence Theorem [85]:

For a consistent approximation, stability is a necessary and sufficient condition for convergence.

We will here use the "Fourier series method" (also called the "von Neumann's method") for examining the stability properties of an approximation. An initial sequence of errors is written as finite Fourier series, and one analyses how a function consisting of these terms at $t = 0$ develop during the course of time. The function is defined inside the interval $(0, L)$. We use Fourier series on complex form, and to avoid to use the same notation as for the index set we denote the complex constant $\mathtt{i}$. The initial errors at $t = 0$ are written as a Fourier series on complex form [73]:

$$E_i = \sum_{n=0}^{N} A_n e^{\mathtt{i}\beta_n i \Delta x} \qquad i = 0, 1, ...N \tag{3.45}$$

where

$$\beta_n = \frac{n\pi}{L}$$

We now want to study how an error term introduced at some stage may influence the solutions computed at later stages. The system under consideration is linear with separable solutions, and therefore it is sufficient look at an arbitrary term: $e^{\mathtt{i}\beta i \Delta x}$. The coefficient A_n is a constant which can be neglected in the following analysis. By writing $t = r\Delta t$, the error term can be expressed as:

$$E_{i,r} = e^{\mathtt{i}\beta x}\, e^{\alpha t} = e^{\mathtt{i}\beta i \Delta x}\, e^{\alpha r \Delta t} = e^{\mathtt{i}\beta i \Delta x}\, \xi^r \tag{3.46}$$

We have here introduced the amplification factor $\xi = e^{\alpha \Delta t}$ where α is a complex constant. When $t = r\Delta t = 0$ we see that $E_{i,r}$ is reduced to $e^{\mathtt{i}\beta i \Delta x}$, which is the initial error term. As long as $|\ \xi\ | \leq 1$ we know that the error term is not amplified during the subsequent computations, which implies that as long as $|\ \xi\ | \leq 1$ we have a stable approximation. As long as the approximation is consistently formulated, the approximation will also be convergent. This method of analyzing stability is valid for linear difference equations with constant coefficients. We will now discuss the stability properties of the implicit and the explicit method given by the equations (3.36) and (3.38).

3.3.4 Stability of the Implicit Method

From equation (3.38) we know that the implicit method in the one dimensional case can be written:

$$\frac{u_{i-1}^{n+1} - 2u_i^{n+1} + u_{i+1}^{n+1}}{\Delta x^2} = \frac{u_i^{n+1} - u_i^n}{\Delta t} + w_i$$

The error term is the difference between the solution of the original problem and solution of the difference equations. The system under consideration is linear, and therefore also the error terms will satisfy the difference equation. By substituting for the error terms we get the following relationship:

$$\frac{1}{\Delta x^2}[e^{\mathrm{i}\beta(i-1)\Delta x}\xi^{r+1} - 2e^{\mathrm{i}\beta i\Delta x}\xi^{r+1} + e^{\mathrm{i}\beta(i+1)\Delta x}\xi^{r+1}] = \frac{1}{\Delta t}[e^{\mathrm{i}\beta i\Delta x}\xi^{r+1} - e^{\mathrm{i}\beta i\Delta x}\xi^{r}] \tag{3.47}$$

By writing $\theta = \Delta t/\Delta x^2$ and dividing the equation by $exp(\mathrm{i}\beta i\Delta x)\,\xi^r$, we get:

$$\begin{aligned}\xi - 1 &= \theta\xi(e^{-\mathrm{i}\beta\Delta x} - 2 + e^{\mathrm{i}\beta\Delta x})\\ &= \theta\xi(2\cos\beta\Delta x - 2)\\ &= -4\theta\xi\sin^2(\frac{\beta\Delta x}{2})\end{aligned}$$

Which can be written:

$$\xi = \frac{1}{1 + 4\theta\sin^2(\frac{\beta\Delta x}{2})} \tag{3.48}$$

We now see that when $\theta \geq 0$ we know that $|\ \xi\ | \leq 1$. This means that the implicit method always is a stable approximation, and because it is a consistent approximation we know that it also is a convergent approximation.

3.3.5 Stability of the Explicit Method

By using equation (3.36) as a starting point, we can also in this case write the equation for the error terms:

$$\frac{1}{\Delta x^2}[e^{\mathrm{i}\beta(i-1)\Delta x}\xi^{r} - 2e^{\mathrm{i}\beta i\Delta x}\xi^{r} + e^{\mathrm{i}\beta(i+1)\Delta x}\xi^{r}] = \frac{1}{\Delta t}[e^{\mathrm{i}\beta i\Delta x}\xi^{r+1} - e^{\mathrm{i}\beta i\Delta x}\xi^{r}] \tag{3.49}$$

As for the implicit method we write $\theta = \Delta t/\Delta x^2$ and divide the equation by $e^{\mathrm{i}\beta i\Delta x}\,\xi^r$:

$$\begin{aligned}\xi - 1 &= \theta(e^{-\mathrm{i}\beta\Delta x} - 2 + e^{\mathrm{i}\beta\Delta x})\\ \xi &= 1 - 4\theta\sin^2(\frac{\beta\Delta x}{2})\end{aligned}$$

In order to have a stable approximation, we have to have $|\ \xi\ | \leq 1$, which implies:

$$-1 \leq 1 - 4\theta\sin^2(\frac{\beta\Delta x}{2}) \leq 1 \tag{3.50}$$

The term in the middle will always be less than 1, so the right part of this inequality is always valid. For the left part we get:

$$\begin{aligned}-2 &\leq -4\theta\sin^2(\frac{\beta\Delta x}{2})\\ \theta &\leq \frac{1}{2\sin^2(\frac{\beta\Delta x}{2})}\end{aligned}$$

The term $\sin^2(\frac{\beta\Delta x}{2})$ is always less or equal to 1, and therefore the denominator of this fraction will be of value 2 or less. This leads to the following sufficient stability condition:

$$\frac{\Delta t}{\Delta x^2} \leq \frac{1}{2} \tag{3.51}$$

It is important to note that for the explicit method, discretization in time and space can not be done independently.

3.3.6 Stability of Problems in Higher Dimensions

As discussed earlier, the finite difference approximations can of course be used also for problems of higher dimensions than one. A two dimensional difference approximation for the explicit method can then be written:

$$\frac{u_{i-1,j}^n - 2u_{i,j}^n + u_{i+1,j}^n}{\Delta x^2} + \frac{u_{i,j-1}^n - 2u_{i,j}^n + u_{i,j+1}^n}{\Delta y^2} = \frac{u_i^{n+1} - u_i^n}{\Delta t} + w_i \tag{3.52}$$

$$i = 1, 2, ..., N \qquad j = 1, 2, ..., M \qquad n = 0, 1, 2, ...$$

For the one dimensional case, we described in equation (3.46) initial errors by use of finite Fourier series, and the same can also be done in the two dimensional case. An arbitrary error term can be written as:

$$E_{p,q,r} = e^{\mathrm{i}\beta x}\, e^{\mathrm{i}\gamma y}\, e^{\alpha t} = e^{\mathrm{i}\beta i \Delta x}\, e^{\mathrm{i}\gamma j \Delta y}\, \xi^r \tag{3.53}$$

By using the same procedure as earlier by replacing the error terms in the difference equation, we get the following inequality:

$$-1 \leq 1 - 4\frac{\Delta t}{\Delta x^2}\sin^2(\frac{\beta\Delta x}{2}) - 4\frac{\Delta t}{\Delta y^2}\sin^2(\frac{\gamma\Delta y}{2}) \leq 1 \tag{3.54}$$

The middle term is always less or equal to one, and the right part of this inequality is always satisfied. With the same reasoning as in the one-dimensional case, we get the following stability criterion:

$$\Delta t(\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2}) \leq \frac{1}{2} \tag{3.55}$$

When going from one to two dimensions, and by letting $\Delta x = \Delta y$, we see that the maximum Δt in the two dimensional case is only half of what it is in the one dimensional case. It can be shown that the implicit method is unconditionally stable also in higher dimensions.

3.3.7 Positive Type Approximations

We have so far only discussed approximations for dimensionless differential equations. The differential equations describing reservoir behavior are more complicated. Implicit method approximations for reservoir equations will always be stable, but it may be a bit more difficult to state if an explicit approximation is stable or not. In Forsythe and Wasow [31] the term "positive type approximations" is introduced, and it is shown that all positive type approximations also are stable approximations. We will here use equation (3.36) as a starting point, but for simplicity the sink/source term w_i is omitted:

$$\frac{u_{i-1}^n - 2u_i^n + u_{i+1}^n}{\Delta x^2} = \frac{u_i^{n+1} - u_i^n}{\Delta t}$$

As done earlier we write $\theta = \Delta t/\Delta x^2$. By isolating the unknown term on the left-hand side of the equation we get:

$$u_i^{n+1} = \theta u_{i-1}^n + (1 - 2\theta)u_i^n + \theta u_{i+1}^n \tag{3.56}$$

If all coefficients of the level n variables are non-negative, this is denoted a positive type approximation and the approximation is stable. In this one dimensional example it means that all coefficients on the right-hand side have to be non-negative, which gives: $\theta = \Delta t/\Delta x^2 \leq \frac{1}{2}$. We see that this is the same stability requirement as found earlier. The value calculated at $n+1$, is a "weighted average" of the values at level n, and if there is no sink/source term, u_i^{n+1} gets a value on the interval between the lowest and the highest value on the right-hand side of the equation. In other words, there are limits for the growth of u_i^{n+1}, which give an intuitive explanation of why the system is stable when $\theta \leq \frac{1}{2}$.

By writing $\lambda = \Delta t/\Delta y^2$, the two dimensional explicit formulation can be expressed

$$u_{i,j}^{n+1} = \lambda u_{i,j-1}^n + \theta u_{i-1,j}^n + (1 - 2\theta - 2\lambda)u_{i,j}^n + \theta u_{i+1,j}^n + \theta u_{i,j+1}^n \tag{3.57}$$

This equation is a positive type approximation if: $\theta + \lambda = \Delta t(1/\Delta x^2 + 1/\Delta y^2) \leq \frac{1}{2}$. This is the same requirement as stated in equation (3.55), derived by use of the Fourier series method. We will in Section 3.4.3 see that the notion of positive type approximations will be very useful when analyzing the explicit method approximations for reservoir equations.

3.4 Discretization of the Reservoir Description

We will now return to the description of the single phase oil reservoir derived in Section 3.1 and use the general methods discussed in Section 3.3 for discretization of the reservoir equations.

3.4.1 Discretization in Space

All reservoirs are of course three dimensional, but in many practical settings it is possible to consider the pressure gradient in one of the directions as negligible compared to the other two directions. In reservoirs that are relatively thin compared to their area extent, it is possible to assume the pressure gradient in the vertical direction to be negligible. In the optimization model that we propose, a two dimensional model is assumed to be sufficient. Small differences in reservoir height may be adjusted for by letting the reservoir height be a function of the $x-$ and $y-$coordinates. Reservoir descriptions with flow only in horizontal directions and with variable reservoir height will not be mathematically correct, but as long as the variation in reservoir height is rather small they will be good approximations [9].

As discussed earlier, when a partial differential equation is discretized and approximated by finite difference equations, the idea is to find approximated solutions at a finite number of locations at a finite number of levels in time. In this case we consider a reservoir which is divided into a finite number of blocks. Each block has the area $\Delta x \Delta y$, and for simplicity we only discuss the case where the block size is uniform over the reservoir. As illustrated in Figure 3.1 the point x_i, y_j is located in the center of the block denoted (i, j), and in the literature this is denoted a block-centered grid. The coordinate $x_{i-\frac{1}{2}}$ corresponds to the left boundary of the block, while $x_{i+\frac{1}{2}}$ corresponds to the right boundary. The law of mass conservation for the block (i, j) gives the following equation:

$$\begin{aligned} & \Delta y(\Delta z \rho v_x)_{i-\frac{1}{2},j} - \Delta y(\Delta z \rho v_x)_{i+\frac{1}{2},j} \\ + \; & \Delta x(\Delta z \rho v_y)_{i,j-\frac{1}{2}} - \Delta x(\Delta z \rho v_y)_{i,j+\frac{1}{2}} \\ = \; & \Delta x \Delta y \Delta z_{i,j} \tfrac{\partial(\phi\rho)_{i,j}}{\partial t} + \Delta x \Delta y \Delta z_{i,j} w_{i,j} \end{aligned} \tag{3.58}$$

The left side of this equation describes the mass flow over the block boundaries, while on the right side we recognize the terms describing the fluid's compressibility and the sink/source term. Both $(\phi\rho)_{i,j}$ and $w_{i,j}$ represents

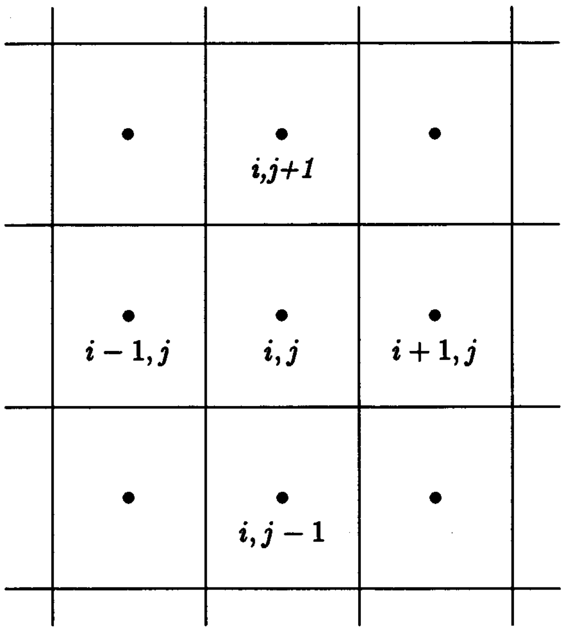

Figure 3.1: Block-centered grid

the block's averaged values. In the same way as done in Section 3.1 we use Darcy's law for describing the relationship between the pressure gradient and the flow velocity, and the flow velocity in x-direction can be written as

$$v_x = -\frac{k}{\mu}\frac{\partial p}{\partial x} \tag{3.59}$$

By doing the same for the flow in the y-direction, and by substituting this in the equation for mass conservation, we get:

$$\begin{aligned} &\Delta y(\Delta z \rho \tfrac{k}{\mu}\tfrac{\partial p}{\partial x})_{i+\frac{1}{2},j} - \Delta y(\Delta z \rho \tfrac{k}{\mu}\tfrac{\partial p}{\partial x})_{i-\frac{1}{2},j} \\ + \; &\Delta x(\Delta z \rho \tfrac{k}{\mu}\tfrac{\partial p}{\partial y})_{i,j+\frac{1}{2}} - \Delta x(\Delta z \rho \tfrac{k}{\mu}\tfrac{\partial p}{\partial y})_{i,j-\frac{1}{2}} \\ &= V_{i,j}\tfrac{\partial(\phi\rho)_{i,j}}{\partial t} + V_{i,j}w_{i,j} \end{aligned} \tag{3.60}$$

We rewrite the equation for constant temperature compressibility and substitute for $\rho\frac{\partial p}{\partial x}$ and $\rho\frac{\partial p}{\partial y}$:

$$\begin{aligned} &\Delta y(\Delta z \tfrac{k}{\mu c}\tfrac{\partial \rho}{\partial x})_{i+\frac{1}{2},j} - \Delta y(\Delta z \tfrac{k}{\mu c}\tfrac{\partial \rho}{\partial x})_{i-\frac{1}{2},j} \\ + \; &\Delta x(\Delta z \tfrac{k}{\mu c}\tfrac{\partial \rho}{\partial y})_{i,j+\frac{1}{2}} - \Delta x(\Delta z \tfrac{k}{\mu c}\tfrac{\partial \rho}{\partial y})_{i,j-\frac{1}{2}} \\ &= V_{i,j}\tfrac{\partial(\phi\rho)_{i,j}}{\partial t} + V_{i,j}w_{i,j} \end{aligned} \tag{3.61}$$

We further use the first order approximation for constant temperature compressibility $\rho = \rho^0(1 + c(p - p^0))$, and as done in Section 3.1, we assume the porosity to be pressure independent. We then get the following equation:

$$\begin{aligned} & \Delta y(\Delta z \tfrac{k}{\mu}\tfrac{\partial p}{\partial x})_{i+\frac{1}{2},j} - \Delta y(\Delta z \tfrac{k}{\mu}\tfrac{\partial p}{\partial x})_{i-\frac{1}{2},j} \\ + \; & \Delta x(\Delta z \tfrac{k}{\mu}\tfrac{\partial p}{\partial y})_{i,j+\frac{1}{2}} - \Delta x(\Delta z \tfrac{k}{\mu}\tfrac{\partial p}{\partial y})_{i,j-\frac{1}{2}} \\ & = \phi_{i,j} c V_{i,j} \tfrac{\partial p_{i,j}}{\partial t} + \tfrac{V_{i,j}}{\rho^0} w_{i,j} \end{aligned} \tag{3.62}$$

The partial derivative of the pressure with respect to x may be written

$$(\frac{\partial p}{\partial x})_{i+\frac{1}{2},j} = \frac{p_{i+1,j} - p_{i,j}}{\Delta x} \tag{3.63}$$

By doing the same operation for the partial derivative with respect to y, and by dividing equation 3.62 by $\Delta x \Delta y$ we get

$$\begin{aligned} & \tfrac{1}{\mu(\Delta x)^2}[(\Delta z k)_{i+\frac{1}{2},j}(p_{i+1,j} - p_{i,j}) + (\Delta z k)_{i-\frac{1}{2},j}(p_{i-1,j} - p_{i,j})] \\ + \; & \tfrac{1}{\mu(\Delta y)^2}[(\Delta z k)_{i,j+\frac{1}{2}}(p_{i,j+1} - p_{i,j}) + (\Delta z k_{i,j-\frac{1}{2}}(p_{i,j-1} - p_{i,j})] \\ & = \phi_{i,j} c \Delta z_{i,j} \tfrac{\partial p_{i,j}}{\partial t} + \tfrac{\Delta z_{i,j}}{\rho^0} w_{i,j} \end{aligned} \tag{3.64}$$

The permeability and the reservoir height will be approximated by using the average values for the two adjoining blocks. For the indexes $i + \frac{1}{2}, j$ we get:

$$(\Delta z k)_{i+\frac{1}{2},j} = \frac{k_{i,j} + k_{i+1,j}}{2} \cdot \frac{\Delta z_{i,j} + \Delta z_{i+1,j}}{2} \tag{3.65}$$

At the outer boundaries of the reservoir we assume the permeability to be zero, which means that there is no mass flow over these boundaries. This discretized reservoir equation is derived in the same way as the continuous equation in Section 3.1 was derived. As earlier, the left-hand side of equation 3.64 represents the mass flow over the block boundary, while the right-hand side describes mass accumulation due to compressibility and due to source/sink terms. By comparing to the general discretized partial differential equation in the previous section, we see that the difference is the coefficients $k, \Delta z$ and μ. What this implicates for the stability properties will be discussed later.

3.4.2 Discretization in Time

The starting point for this analysis is that we assume the reservoir pressure at time t_0 to be known, and we want to find approximated solutions for the reservoir pressure at the levels $t_1 = t_0 + \Delta t, \ldots, t_n = t_0 + n\Delta t, \ldots$ The reason for our interest in the reservoir pressure, is that the reservoir's production capacity depends on the pressure. One way to visualize the relationship between pressure, $\mathbf{p}$, and sink terms (production), $\mathbf{w}$, is as follows:

$$\mathbf{p}^0 \xrightarrow{\mathbf{w}^1} \mathbf{p}^1 \xrightarrow{\mathbf{w}^2} \mathbf{p}^2 \xrightarrow{\mathbf{w}^3} \mathbf{p}^3 \xrightarrow{\mathbf{w}^4} \mathbf{p}^4 \quad \cdots$$

We calculate pressure variables at discrete levels while the production goes on continuously. Above we have indicated the production in the first period, $\mathbf{w}^1$, takes place in the period between t_0 and t_1. From this definition it is not straight forward to define the reservoir pressure in the first period, and that question will be discussed in later sections.

Explicit Method

By using the "forward difference approximation" when discretizing the time derivative in equation (3.64), we get the following equation:

$$\begin{aligned} & \tfrac{1}{\mu(\Delta x)^2}[(\Delta zk)_{i+\frac{1}{2},j}(p^n_{i+1,j} - p^n_{i,j}) + (\Delta zk)_{i-\frac{1}{2},j}(p^n_{i-1,j} - p^n_{i,j})] \\ + & \tfrac{1}{\mu(\Delta y)^2}[(\Delta zk)_{i,j+\frac{1}{2}}(p^n_{i,j+1} - p^n_{i,j}) + (\Delta zk)_{i,j-\frac{1}{2}}(p^n_{i,j-1} - p^n_{i,j})] \\ & = \tfrac{\phi_{i,j}c\Delta z_{i,j}}{\Delta t}(p^{n+1}_{i,j} - p^n_{i,j}) + \tfrac{\Delta z_{i,j}}{\rho^0} w^{n+1}_{i,j} \end{aligned} \tag{3.66}$$

which can be rewritten as

$$\begin{aligned} & \frac{1}{\mu(\Delta y)^2}(\Delta zk)_{i,j-\frac{1}{2}} \quad p^n_{i,j-1} \\ & + \frac{1}{\mu(\Delta x)^2}(\Delta zk)_{i-\frac{1}{2},j} \quad p^n_{i-1,j} \\ & + [\frac{\phi_{i,j}c}{\Delta t}\Delta z_{i,j} - \frac{1}{\mu(\Delta x)^2}((\Delta zk)_{i+\frac{1}{2},j} + (\Delta zk)_{i-\frac{1}{2},j}) \\ & \quad - \frac{1}{\mu(\Delta y)^2}((\Delta zk)_{i,j+\frac{1}{2}} + (\Delta zk)_{i,j-\frac{1}{2}})] \quad p^n_{i,j} \\ & + \frac{1}{\mu(\Delta x)^2}(\Delta zk)_{i+\frac{1}{2},j} \quad p^n_{i+1,j} \\ & + \frac{1}{\mu(\Delta y)^2}(\Delta zk)_{i,j+\frac{1}{2}} \quad p^n_{i,j+1} \\ & - \frac{\phi_{i,j}c}{\Delta t}\Delta z_{i,j} \quad p^{n+1}_{i,j} \\ & - \frac{\Delta z_{i,j}}{\rho^0} \quad w^{n+1}_{i,j} \quad = 0 \end{aligned} \tag{3.67}$$

This equation tells us that from the pressure in block (i, j) and the four neighboring blocks at time level n and the production in the block (i, j) the following period, we can calculate the pressure in the block (i, j) at time level $n+1$.

Implicit Method

When approximating the time derivative in equation (3.64) by the "backwards difference", we get the following expression:

$$\begin{aligned}
&\tfrac{1}{\mu(\Delta x)^2}[(\Delta zk)_{i+\frac{1}{2},j}(p^{n+1}_{i+1,j} - p^{n+1}_{i,j}) + (\Delta zk)_{i-\frac{1}{2},j}(p^{n+1}_{i-1,j} - p^{n+1}_{i,j})] \\
+ \; &\tfrac{1}{\mu(\Delta y)^2}[(\Delta zk)_{i,j+\frac{1}{2}}(p^{n+1}_{i,j+1} - p^{n+1}_{i,j}) + (\Delta zk)_{i,j-\frac{1}{2}}(p^{n+1}_{i,j-1} - p^{n+1}_{i,j})] \\
&= \tfrac{\phi_{i,j}c\Delta z_{i,j}}{\Delta t}(p^{n+1}_{i,j} - p^{n}_{i,j}) + \tfrac{\Delta z_{i,j}}{\rho^0}w^{n+1}_{i,j}
\end{aligned} \tag{3.68}$$

and by sorting the pressure dependent variables this equation can be rewritten as:

$$\begin{aligned}
&\frac{\phi_{i,j}c}{\Delta t}\Delta z_{i,j} \quad p^{n}_{i,j} \\
&+\frac{1}{\mu(\Delta y)^2}(\Delta zk)_{i,j-\frac{1}{2}} \quad p^{n+1}_{i,j-1} \\
&+\frac{1}{\mu(\Delta x)^2}(\Delta zk)_{i-\frac{1}{2},j} \quad p^{n+1}_{i-1,j} \\
&-[\frac{\phi_{i,j}c}{\Delta t}\Delta z_{i,j} + \frac{1}{\mu(\Delta x)^2}((\Delta zk)_{i+\frac{1}{2},j} + (\Delta zk)_{i-\frac{1}{2},j}) \\
&\quad +\frac{1}{\mu(\Delta y)^2}((\Delta zk)_{i,j+\frac{1}{2}} + (\Delta zk)_{i,j-\frac{1}{2}})] \quad p^{n+1}_{i,j} \\
&+\frac{1}{\mu(\Delta x)^2}(\Delta zk)_{i+\frac{1}{2},j} \quad p^{n+1}_{i+1,j} \\
&+\frac{1}{\mu(\Delta y)^2}(\Delta zk)_{i,j+\frac{1}{2}} \quad p^{n+1}_{i,j+1} \\
&-\frac{\Delta z_{i,j}}{\rho^0} \quad w^{n+1}_{i,j} \quad = 0
\end{aligned} \tag{3.69}$$

This equation gives the relation between the pressure in block (i, j) at time level n, the production in the block the following period, and the pressure in block (i, j) and the neighboring blocks at time level $n+1$. If we are considering a reservoir discretized in $M \times N$ blocks, the implicit method will give a system of $M \times N$ equations that need to be solved simultaneously.

3.4.3 Stability

In order to analyze the stability properties of these equations, we will compare the discretized reservoir equations with the general discretized equations for partial differential equations derived in the previous section. We know that the implicit method is an unconditionally stable approximation.

When analyzing the stability properties of the explicit method, we will use equation (3.66) as a starting point. Compared to the general equation (3.52) we see that (3.66) have coefficients describing reservoir height, permeability and porosity, and the fluid viscosity and compressibility. If we simplify by assuming constant height, permeability and porosity we get the following difference approximation:

$$\begin{aligned} &\tfrac{k}{\mu(\Delta x)^2}(p^n_{i-1,j} - 2p^n_{i,j} + p^n_{i+1,j}) + \tfrac{k}{\mu(\Delta y)^2}(p^n_{i,j-1} - 2p^n_{i,j} + p^n_{i,j+1}) \\ &= \tfrac{\phi c}{\Delta t}(p^{n+1}_{i,j} - p^n_{i,j}) + \tfrac{1}{\rho^0} w^{n+1}_{i,j} \end{aligned} \tag{3.70}$$

Since the only difference between the equations are the constants in equation (3.70), we can use the same stability requirement before. By including these constants we get:

$$\frac{k}{\phi c \mu}\Delta t(\frac{1}{\Delta x^2} + \frac{1}{\Delta y^2}) \leq \frac{1}{2} \tag{3.71}$$

A more general result is found by using the definition of positive type approximations as found in Forsythe and Wasow [31]. The requirement derived there is that all coefficients of pressure variables at level n must be positive numbers. This means that the coefficients associated with $p^n_{i,j}$ must be greater than or equal to zero, which leads to the following stability criteria that for all i, j:

$$\begin{aligned} \frac{\phi_{i,j} c}{\Delta t}\Delta z_{i,j} \geq\ & \frac{1}{\mu(\Delta x)^2}((\Delta z k)_{i+\frac{1}{2},j} + (\Delta z k)_{i-\frac{1}{2},j}) \\ & + \frac{1}{\mu(\Delta y)^2}((\Delta z k)_{i,j+\frac{1}{2}} + (\Delta z k)_{i,j-\frac{1}{2}}) \end{aligned} \tag{3.72}$$

3.5 Maximum Well Production

In the derived reservoir descriptions we have introduced the sink term $w^n_{i,j}$ with units $kg/m^3/s$. This term represents an "averaged sink" for the whole block (i, j), while the production takes place in a well located in the block. The relation between the sink term and the rate of production, $q^n_{i,j}$, is as follows:

$$q^n_{i,j} = \frac{1}{\rho_s}\Delta x \Delta y \Delta z_{i,j} w^n_{i,j} \tag{3.73}$$

By multiplying the mass sink by the volume of block i, j, and dividing by the fluid density at surface pressure ρ_s, we get the surface rate of production

$q_{i,j}^n$.

The maximum rate of production in each well depends on the reservoir pressure in the vicinity of the well, and can be approximated by the following relation [48]:

$$q_{i,j}(t) \leq J_{i,j}(p_{i,j}(t) - p_w) \tag{3.74}$$

The rate of production in well (i, j), $q_{i,j}$, is represented by units m^3/s. The pressure in block i, j at time t is given as $p_{i,j}(t)$, while p_w is the minimum well pressure. The well specific productivity index $J_{i,j}$, may either be estimated from the properties of the fluid and the porous media, or it can be found by testing existing wells. We will here use the same productivity index as proposed in [39]:

$$J = \frac{2\pi k \Delta z}{\mu} \frac{1}{\frac{1}{2} \ln \frac{\Delta x \Delta y}{\pi r_w^2}} \tag{3.75}$$

We see that both reservoir height and permeability, fluid viscosity and block area are included in equation (3.75). A more thorough discussion of productivity indexes are given in Peaceman [83].

Equation (3.74) gives the relationship between maximum rate of production and reservoir pressure, but in a time discretized model it may be difficult to define a period's reservoir pressure. The problem is how to define the reservoir pressure in the n-th period, the interval between $\mathbf{p}^{n-1}$ and $\mathbf{p}^n$. We will here look at three different ways of defining this reservoir pressure: Initial-, mid- and end-pressure. These methods are illustrated in the Figures 3.2 - 3.4.

The initial pressure method implies that the reservoir pressure at the **start** of a period defines the reservoir pressure in the following period. With the indexing of variables we have chosen, equation (3.74) becomes:

$$q_{i,j}^n \leq J_{i,j}(p_{i,j}^{n-1} - p_w) \tag{3.76}$$

This is the way the reservoir pressure is defined in Hallefjord, Haugland and Helgesen [39] and Haugland, Hallefjord and Asheim [48].

The opposite way is to define the period pressure by the pressure at the **end** of period n, $\mathbf{p}^n$. Equation (3.74) can then be written as:

$$q_{i,j}^n \leq J_{i,j}(p_{i,j}^n - p_w) \tag{3.77}$$

A third method proposed, is the mid-pressure method where one seeks to estimate a value for the reservoir pressure in the **middle** of period n, so the equation for maximum well production may be expressed as:

$$q_{i,j}^{n} \leq J_{i,j}(p_{i,j}^{n-\frac{1}{2}} - p_w) \tag{3.78}$$

One way of computing this mid-pressure, $\mathbf{p}^{n-\frac{1}{2}}$, is to consider it as an average of $\mathbf{p}^{n-1}$ and $\mathbf{p}^{n}$. Implementations of these methods will be further discussed in the next chapter.

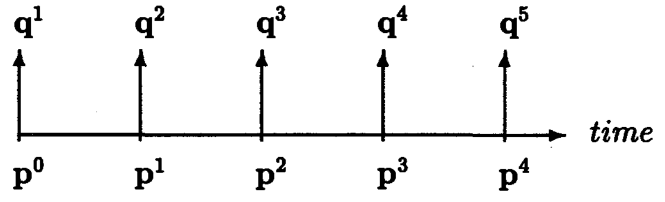

Figure 3.2: Initial pressure

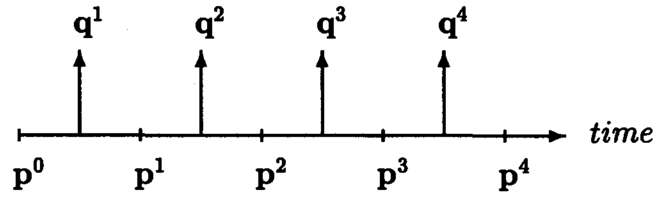

Figure 3.3: Mid-pressure

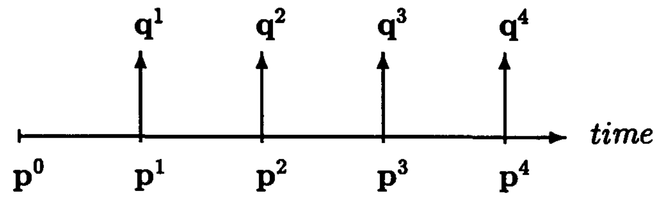

Figure 3.4: End pressure

Chapter 4

Optimization Models and Solution Methods

In this chapter we discuss how the reservoir equations derived in Chapter 3 may be included in models for optimizing oil field exploitation. The problem we want to solve can verbally be formulated as follows: An offshore oil reservoir is found to be profitable, and it is decided to install one platform on the oil-field, but production capacity for this platform is not yet chosen. A set of potential well sites is proposed, but which wells to drill, sequence of drilling operations, and production profile for each well remain to be optimized.

As discussed earlier, it is the reservoir that is the basis for this analysis, and we have chosen a level of detail where fluid flow equations are included in the optimization model. This allow us to study how the decision variables interact, and the idea behind this approach is to simultaneously optimize platform, well and production decisions.

We will first look at pure production optimization models where all the design decisions are already made. These problems have continuous variables, and for a single phase oil reservoir the production optimization can be formulated as a linear programming problem. We will in Section 4.1 discuss two different ways of formulating the problem: Either by employing the reservoir equations directly in the optimization model, or by using superposition for eliminating the pressure variables. In Section 4.2 also design decisions are considered, which results in mixed integer programming models. Some properties of the models are illustrated by use of computational experiments.

4.1 Optimal Production Decisions

The problem considered in this section is as follows: Assume a single phase oil reservoir for which development decisions are already made. That is, the platform capacity is fixed and a set of production wells are drilled. The problem is to find the production decisions that maximizes the net present value of future production from the reservoir:

$$\max \sum_{n=1}^{T} c^n \Delta t^n \sum_{b=1}^{B} q_b^n \tag{4.1}$$

We have here introduced the following notation: T is the number of time periods available for production, c^n is the discounted deterministic oil price in period n, Δt^n is the length of the n-th time period, B is the number of wells, and q_b^n is the production rate in well b in the n-th period. If there are variable production costs associated with the production, we assume that c^n is the net oil-price after the variable costs have been subtracted.

However, as described in Section 3.5, the possible production in each well depends on the reservoir pressure and we will here discuss how the reservoir description may be included in the model. We will first look at models where the discretized reservoir equations are directly included in the constraint matrix, before discussing how the pressure variables may be eliminated by use of superposition.

4.1.1 Models Including a Reservoir Description

When the discretized reservoir equations are directly employed in the optimization model, it means that both the production (q) and pressure (p) are represented as decision variables in the model. However, the reservoir pressure is uniquely determined by the production variables, and the objective function coefficients associated with the pressure variables are zero. It is the equations (3.67) or (3.69) that are used as building blocks in this model; (3.67) if the problem is formulated by use of the explicit method and (3.69) if the implicit method is employed. For each time-step and each block it is necessary with one equation for calculating the reservoir pressure, so for a reservoir discretized in $M \times N$ blocks it is necessary with $M \times N$ equations each time-step. For an arbitrary time-step, n, a simplified version of the system of difference equations on explicit form may be written as

$$A_{i,j}p_{i,j-1}^n + B_{i,j}p_{i-1,j}^n + C_{i,j}p_{i,j}^n + D_{i,j}p_{i+1,j}^n + E_{i,j}p_{i,j+1}^n + F_{i,j}p_{i,j}^{n+1} + G_{i,j}w_{i,j}^{n+1} = 0 \tag{4.2}$$

$$i = 1, 2, ..., M \qquad j = 1, ..., N \qquad n = 0, ..., (T - 1)$$

The constants $A_{i,j}$ to $G_{i,j}$ represents the coefficients given in equation (3.67). For blocks on the reservoir boundary some of the coefficients $A_{i,j}$ to $F_{i,j}$ may be zero, and for blocks in which no well is drilled $G_{i,j}$ is zero. If the time derivative is approximated by the implicit method, the associated equation has only one "time n variable", while the other variables have time index $n+1$. The maximum well production is given by inequalities 3.76, 3.77 or 3.78, depending on how a period's reservoir pressure is defined.

If we are optimizing production decisions over T periods using the initial pressure method, we get a constraint matrix with at least $(T-1) \times M \times N$ equalities and $(T \times B)$ inequalities. If we rather use the mid- or end-pressure method, the number of equalities is increased to $T \times M \times N$. By writing "at least", it is meant that in a production maximization problem the number of constraints will normally be higher in order to keep the production below specified capacity limits. If there is an absolute upper limit on well capacity, this can be formulated as

$$q_b^n \leq S_b, \ \ b = 1, \ldots, B, \ \ n = 1, \ldots, T \tag{4.3}$$

where S_b is the maximum possible production rate in well b. For assuring that the total production each time-step is below the installed platform capacity Q, we introduce the following constraint:

$$\sum_{b=1}^{B} q_b^n \leq Q, \ \ n = 1, \ldots, T. \tag{4.4}$$

For a problem with limited well and platform capacity solved by the mid- or end pressure method we get a problem with $T \times M \times N$ equalities and $T \times (2B+1)$ inequalities.

Computational experiments where the reservoir equations are directly included in the reservoir models are reported in Section 4.1.4.

4.1.2 The Principle of Superposition

We will here look at the system of discretized reservoir equations in a more general setting. A system of discrete linear equations as described in (4.2), may on general matrix form be expressed as [38], [64]:

$$\bar{\mathbf{A}}(t)\mathbf{x}(t) = \bar{\mathbf{C}}(t-1)\mathbf{x}(t-1) + \bar{\mathbf{D}}(t)\mathbf{u}(t) \qquad t = 1, 2, ..., T \tag{4.5}$$

where $\mathbf{x}(0)$ is known. Here $\mathbf{x}(t) \in R^{M \times N}$ is a vector of state variables while $\mathbf{u}(t) \in R^B$ is a vector of control variables. In the problem we are considering, the state variables correspond to pressure variables, while the control variables correspond to the each period's production. If this is a system where the explicit method is used for approximating the time derivative, $\bar{\mathbf{A}}(t)$ is a diagonal matrix. However, by multiplying equation (4.5) by $\bar{\mathbf{A}}^{-1}(t)$, we get the following equation system (it is here assumed that $\bar{\mathbf{A}}^{-1}(t)$ is non-singular):

$$\mathbf{x}(t) = \mathbf{A}(t-1)\mathbf{x}(t-1) + \mathbf{D}(t)\mathbf{u}(t) \qquad t = 1, 2, ..., T \tag{4.6}$$

The matrix $\mathbf{A}$ is referred to as the system matrix, and $\mathbf{D}$ is denoted the distribution matrix. In the problem of reservoir optimization the coefficient matrices are independent of time, and also in this general discussion we will restrict ourselves to time-invariant systems. Equation (4.6) may therefore in the time-invariant case be written as

$$\mathbf{x}(t) = \mathbf{A}\mathbf{x}(t-1) + \mathbf{D}\mathbf{u}(t) \qquad t = 1, 2, ..., T \tag{4.7}$$

An equivalent expression can be written for the time step $t-1$:

$$\mathbf{x}(t-1) = \mathbf{A}\mathbf{x}(t-2) + \mathbf{D}\mathbf{u}(t-1) \qquad t = 2, 3, ..., T \tag{4.8}$$

By substituting the expression for $\mathbf{x}(t-1)$ in equation (4.7) by (4.8), we get:

$$\mathbf{x}(t) = \mathbf{A}^2\mathbf{x}(t-2) + \mathbf{A}\ \mathbf{D}\mathbf{u}(t-1) + \mathbf{D}\mathbf{u}(t) \qquad t = 2, 3, ..., T \tag{4.9}$$

By recursively following this procedure, we get the following expression:

$$\mathbf{x}(t) = \mathbf{A}^t\mathbf{x}(0) + \sum_{l=1}^{t} \mathbf{A}^{t-l}\mathbf{D}\mathbf{u}(l) \qquad t = 1, 2, ..., T \tag{4.10}$$

We see that a unique solution for the state variables $\mathbf{x}(t)$ can be calculated if we know the initial state $\mathbf{x}(0)$, the control variables $\mathbf{u}(l)$ for $l = 1, \ldots, t$, and the matrixes $\mathbf{A}$ and $\mathbf{D}$. The linearity of the system implies that the solution can be computed by the principle of *superposition*. This principle states that the total response due to several inputs is the sum of the individual responses, plus an initial condition term. This is just a verbal interpretation of equation (4.10), where we see that the state variables $\mathbf{x}(1), \ldots, \mathbf{x}(t-1)$ have been eliminated, and where the total solution to the system can be regarded as a sum of free responses initiated at different times.

4.1.3 Reservoir Equations and Superposition

We will now focus on how superposition can be used for solving the reservoir optimization problem. We will first look at a situation with no production, that is $\mathbf{u}(t) = 0$ for all t. In that case the state vector is:

$$\mathbf{x}(t) = \mathbf{A}^t\mathbf{x}(0) \tag{4.11}$$

For the reservoir system, this is a situation without any production/injection, and a reservoir that initially has a homogeneous distribution of pressure will continue to be in equilibrium. So for an arbitrary block in the reservoir, if there is no production or injection, the block continues to have initial pressure, p^0:

$$p_{i,j}^n = p^0$$

For making things simple, we will in the following assume that the reservoir initially has a homogeneous pressure distribution. If the contrary is true, the pressure distribution over time can be calculated by use of the reservoir equations.

The term $\sum_{l=1}^{t} \mathbf{A}^{t-l}\mathbf{D}\mathbf{u}(l)$ represents the responses due to the control variables. When dealing with pressure and production variables, we find it convenient to use summation notation instead the matrix notation in the previous section. Hallefjord, Asheim and Haugland [38] show that the pressure in an arbitrary block may be written

$$p_{i,j}^n = p^0 - \sum_{k=1}^{n}\sum_{b=1}^{B} \alpha_b^{n+1-k}(i,j)q_b^k \tag{4.12}$$

By comparing to equation (4.10) we see that the first term on the right side is due to the initial pressure, which is p^0 throughout the whole reservoir. The second term gives the pressure reduction due to the production from the reservoir. The parameter $\alpha_b^{n+1-k}(i,j)$ has the following interpretation: *If well b produces one production unit in period k, this results in a pressure reduction of $\alpha_b^{n+1-k}(i,j)$ pressure units in block (i,j) in time period n.* If we know all the α-parameters, it is possible to express $p_{i,j}^n$ for an arbitrary production strategy q. Finding all α's can be done by solving the equation system for a set of B linearly independent q-vectors. This operation will be denoted a reservoir simulation.

A convenient choice of q-vectors is to simulate production of one unit in one of the wells and no production in the others. We perform reservoir simulations for all $j = 1, \dots, B$ where

$$q_b^k = \begin{cases} 1, & \text{if } b = j \\ 0, & \text{otherwise} \end{cases} \tag{4.13}$$

After simulating unit production in well b, the pressure in block (i, j) can be expressed:

$$p_{i,j}^n = p^0 - \sum_{k=1}^{n} \alpha_b^k(i,j), \qquad n = 1, 2, ..., T \tag{4.14}$$

This follows from combining equations (4.12) and (4.13). For $n = 1$ we get:

$$\alpha_b^1(i,j) = p^0 - p_{i,j}^1 \tag{4.15}$$

and by substitution we get the following general expression for finding the α-parameters:

$$\alpha_b^k(i,j) = p_{i,j}^{k-1} - p_{i,j}^k, \qquad k = 1, 2, ..., T \tag{4.16}$$

The pressure values used in the equations (4.14) to (4.16) are found by simulating unit production in well b. By performing this procedure for all wells we can calculate a complete response matrix for the reservoir system. We will now discuss what the constraint matrix looks like when using the initial-, mid- or end-pressure method.

Initial Pressure Method

When using the initial pressure method, the pressure at the start of a period is defined to be the period's pressure: $\mathbf{p}^{n-1}$ is the reservoir pressure in period n. The inequality describing maximum well production may then be written as:

$$q_{i,j}^n \leq J_{i,j}(p_{i,j}^{n-1} - p_w) \tag{4.17}$$

and by substituting the expression for $p_{i,j}^{n-1}$ found in equation (4.12) we get:

$$q_{i,j}^n \leq J_{i,j}(p^0 - \sum_{k=1}^{n-1} \sum_{b=1}^{B} \alpha_b^{n-k}(i,j) q_b^k - p_w) \tag{4.18}$$

In Section 4.1.1 we studied models where the reservoir description was directly included in the model, and the resulting constraint matrix had at least $(T-1) \times (M \times N)$ equalities and $(T \times B)$ inequalities. By solving the problem by superposition, the equalities have been eliminated, and the resulting constraint matrix has $(T \times B)$ inequalities. If we consider a problem with 3 wells and 3 time periods, the constraint matrix has a structure as described

q_{b1}^1	q_{b2}^1	q_{b3}^1	q_{b1}^2	q_{b2}^2	q_{b3}^2	q_{b1}^3	q_{b2}^3	q_{b3}^3		
$\frac{1}{J_{b1}}$									$\leq$	$\hat{p}$
	$\frac{1}{J_{b2}}$								$\leq$	$\hat{p}$
		$\frac{1}{J_{b3}}$							$\leq$	$\hat{p}$
$\alpha_{b1}^1(b1)$	$\alpha_{b2}^1(b1)$	$\alpha_{b3}^1(b1)$	$\frac{1}{J_{b1}}$						$\leq$	$\hat{p}$
$\alpha_{b1}^1(b2)$	$\alpha_{b2}^1(b2)$	$\alpha_{b3}^1(b2)$		$\frac{1}{J_{b2}}$					$\leq$	$\hat{p}$
$\alpha_{b1}^1(b3)$	$\alpha_{b2}^1(b3)$	$\alpha_{b3}^1(b3)$			$\frac{1}{J_{b3}}$				$\leq$	$\hat{p}$
$\alpha_{b1}^2(b1)$	$\alpha_{b2}^2(b1)$	$\alpha_{b3}^2(b1)$	$\alpha_{b1}^1(b1)$	$\alpha_{b2}^1(b1)$	$\alpha_{b3}^1(b1)$	$\frac{1}{J_{b1}}$			$\leq$	$\hat{p}$
$\alpha_{b1}^2(b2)$	$\alpha_{b2}^2(b1)$	$\alpha_{b3}^2(b1)$	$\alpha_{b1}^1(b1)$	$\alpha_{b2}^1(b1)$	$\alpha_{b3}^1(b1)$		$\frac{1}{J_{b2}}$		$\leq$	$\hat{p}$
$\alpha_{b1}^1(b1)$	$\alpha_{b2}^1(b1)$	$\alpha_{b3}^1(b1)$	$\alpha_{b1}^2(b1)$	$\alpha_{b2}^2(b1)$	$\alpha_{b3}^2(b1)$			$\frac{1}{J_{b3}}$	$\leq$	$\hat{p}$

Figure 4.1: Constraint matrix, initial pressure method, 3 wells and 3 time periods

$\mathbf{q}^1$	$\mathbf{q}^2$	$\mathbf{q}^3$	$\mathbf{q}^4$	$\mathbf{q}^5$		
$\frac{1}{J}$					$\leq$	$\hat{\mathbf{p}}$
$\boldsymbol{\alpha}^1$	$\frac{1}{J}$				$\leq$	$\hat{\mathbf{p}}$
$\boldsymbol{\alpha}^2$	$\boldsymbol{\alpha}^1$	$\frac{1}{J}$			$\leq$	$\hat{\mathbf{p}}$
$\boldsymbol{\alpha}^3$	$\boldsymbol{\alpha}^2$	$\boldsymbol{\alpha}^1$	$\frac{1}{J}$		$\leq$	$\hat{\mathbf{p}}$
$\boldsymbol{\alpha}^4$	$\boldsymbol{\alpha}^3$	$\boldsymbol{\alpha}^2$	$\boldsymbol{\alpha}^1$	$\frac{1}{J}$	$\leq$	$\hat{\mathbf{p}}$

Figure 4.2: Constraint matrix, initial pressure method, 5 time periods

in Figure 4.1. The right hand side variables are denoted $\hat{p}$, where $\hat{p} = p^0 - p_w$. The constraint matrix is lower triangular and has a density less than 50 % (if $B > 1$). Its exact density depends on B and T. The restriction matrix for $T = 5$ may be written as in Figure 4.2, where $\frac{1}{J}$ and $\boldsymbol{\alpha}$ represent submatrixes of size $B \times B$, and $\mathbf{q}$ and $\hat{\mathbf{p}}$ are vectors of length B. The matrix in Figure 4.2 is of size $T \times T$.

End Pressure Method

When using the end pressure method the pressure in period n is defined to be $\mathbf{p}^n$, and the constraints on production may be written as follows:

$$q_{i,j}^n \leq J_{i,j}(p_{i,j}^n - p_w) \tag{4.19}$$

$$
\begin{array}{ccccccc}
q^1 & q^2 & q^3 & q^4 & q^5 & & \\
(\alpha^1+\frac{1}{J}) & & & & & \leq & \hat{p} \\
\alpha^2 & (\alpha^1+\frac{1}{J}) & & & & \leq & \hat{p} \\
\alpha^3 & \alpha^2 & (\alpha^1+\frac{1}{J}) & & & \leq & \hat{p} \\
\alpha^4 & \alpha^3 & \alpha^2 & (\alpha^1+\frac{1}{J}) & & \leq & \hat{p} \\
\alpha^5 & \alpha^4 & \alpha^3 & \alpha^2 & (\alpha^1+\frac{1}{J}) & \leq & \hat{p}
\end{array}
$$

Figure 4.3: Constraint matrix, end pressure method, 5 time periods

By substituting equation (4.12) in this expression we get:

$$q_{i,j}^n \leq J_{i,j}(p^0 - \sum_{k=1}^{n}\sum_{b=1}^{B}\alpha_b^{n+1-k}(i,j)q_b^k - p_w) \tag{4.20}$$

The constraint matrix for this system is shown in Figure 4.3. Compared to Figure 4.2 we see that all the α parameters are shifted to the right, and this matrix always has a density higher than 50 %.

Mid-pressure Method

When using the mid pressure method we try to find an expression for $p_{i,j}^{n-\frac{1}{2}}$, the pressure in the middle of period n. The production constraint discussed in the previous chapter was written as:

$$q_{i,j}^n \leq J_{i,j}(p_{i,j}^{n-\frac{1}{2}} - p_w) \tag{4.21}$$

This means that the reservoir simulation must be done in a slightly different manner than described in the equations 4.13, 4.14 and 4.16. When simulating unit production in well b, we use the pressure variable after production in half the period, $p_{i,j}^{\frac{1}{2}}$, and we may then write:

$$\alpha_1^{\frac{1}{2}}(i,j) = p^0 - p_{i,j}^{\frac{1}{2}}(1) \tag{4.22}$$

In Section 4.1.4 we discuss how this reservoir simulation is implemented. The complete constraint matrix may now be found by performing the calculations:

$$\alpha_1^{k+\frac{1}{2}}(i,j) = p_{i,j}^{k-\frac{1}{2}}(1) - p_{i,j}^{k+\frac{1}{2}}(1), \quad k = 1,\ldots,T \tag{4.23}$$

$$
\begin{array}{ccccccc}
q^1 & q^2 & q^3 & q^4 & q^5 & & \\
 & & & & & & \\
(\alpha^{\frac{1}{2}} + \frac{1}{J}) & & & & & \leq & \hat{p} \\
\alpha^{\frac{3}{2}} & (\alpha^{\frac{1}{2}} + \frac{1}{J}) & & & & \leq & \hat{p} \\
\alpha^{\frac{5}{2}} & \alpha^{\frac{3}{2}} & (\alpha^{\frac{1}{2}} + \frac{1}{J}) & & & \leq & \hat{p} \\
\alpha^{\frac{7}{2}} & \alpha^{\frac{5}{2}} & \alpha^{\frac{3}{2}} & (\alpha^{\frac{1}{2}} + \frac{1}{J}) & & \leq & \hat{p} \\
\alpha^{\frac{9}{2}} & \alpha^{\frac{7}{2}} & \alpha^{\frac{5}{2}} & \alpha^{\frac{3}{2}} & (\alpha^{\frac{1}{2}} + \frac{1}{J}) & \leq & \hat{p}
\end{array}
$$

Figure 4.4: Constraint matrix, mid-pressure method, 5 time periods

The production constraint in equation (4.21) may be written:

$$q^n_{i,j} \leq J_{i,j}(p^0 - \sum_{k=1}^{n} \sum_{b=1}^{B} \alpha_b^{n+\frac{1}{2}-k}(i,j) q_b^k - p_w) \tag{4.24}$$

The resulting constraint matrix has a structure similar to the end pressure matrix, and it is shown in Figure 4.4.

4.1.4 Numerical Experiments

Numerical experiments for these production optimization models are reported by Jonsbråten [55]. The experiments are done for models where the reservoir description is included, as in Section 4.1.1, and for models where superposition is used for solving the problem. For the models where the reservoir description is included, models with the explicit and the implicit formulation were implemented. The experiments with the explicit formulation showed that the stability requirements made it necessary to chose between a crude spatial discretization or fine discretization in time. For example, when using a spatial discretization with block size $100m \times 100m$ and a set of typical reservoir data, stability requirements made it necessary to use time steps shorter than 6 days. As discussed earlier, the implicit method is always stable, but solving equations which use the implicit method is harder than solving a problem using the explicit method. This was also found when the results for these methods were compared, but because of its stability properties, the implicit method can be used for solving problems where the explicit method showed not to be useful.

However, Jonsbråten found that the model formulations where the reservoir equations are included, are outperformed by the formulations where superposition is used. The main criteria for this statement is the CPU-time used for solving the production optimization problems. When using superpositioning, the response of the reservoir system has to be simulated. This is done by use of the explicit method, and when the production is given this is a rather simple procedure. For a given reservoir, the system has to be simulated just once, and we are therefore not concerned about the CPU-time used by the simulation. The results obtained when the reservoir equations are part of the model are not directly comparable to the results obtained by superpositioning. When the problem is solved by superpositioning the size of the constraint matrix depends on the number of time periods and number of wells, but is independent of the spatial discretization. This influences only the reservoir simulation process. When the reservoir equations are directly employed the size of the constraint matrix depends on spatial discretization, in addition to the number of time periods and wells. But for all situations investigated, it was found that the problem is solved fastest by use of superposition.

In [55] also the initial-, mid- and end-pressure formulations are discussed, and an illustration of the main findings are found in the Figures 4.5 and 4.6. Both figures illustrate calculated production profiles from an oil reservoir, and the total production is not constrained by a certain platform capacity. Further, there are two production wells which both start to produce in the first period. Figure 4.5 shows the production profiles computed with a rather crude discretization in time, 10 time periods of 300 days each. By using the initial pressure method one over-estimates the reservoir pressure in the first period, which again over-estimates the production potential. Because of this over-estimated production, the pressure in the second period is under-estimated, thus giving rise to an oscillating production profile. Also Haugland, Hallefjord and Asheim [48] use the initial pressure method and report problems with oscillating production profiles.

Also Figure 4.6 shows production profiles calculated by use of the initial-pressure method, but these results are calculated by use of 100 time periods of 30 days each. For illustrating the difference these results are shown in the same format as for the 10 time period discretization. We can clearly see that by finer discretization in time, the obtained results are rather close to what we found by use of the mid-pressure method when using the crude time discretization. Another way of seeing it is that even if one uses crude time dicretization, the mid-pressure method gives a more accurate representation of the reservoir's production potential.

Based on these results, the models implemented and used in the Papers

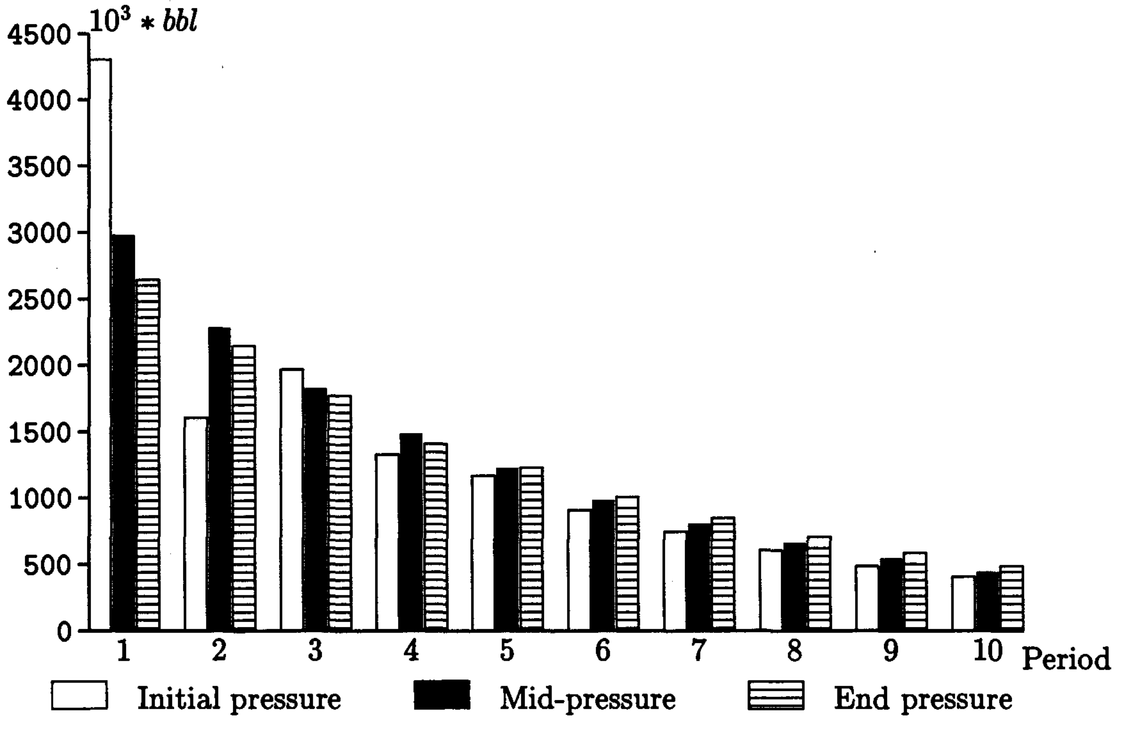

Figure 4.5: Production profile, 10 periods of 300 days each

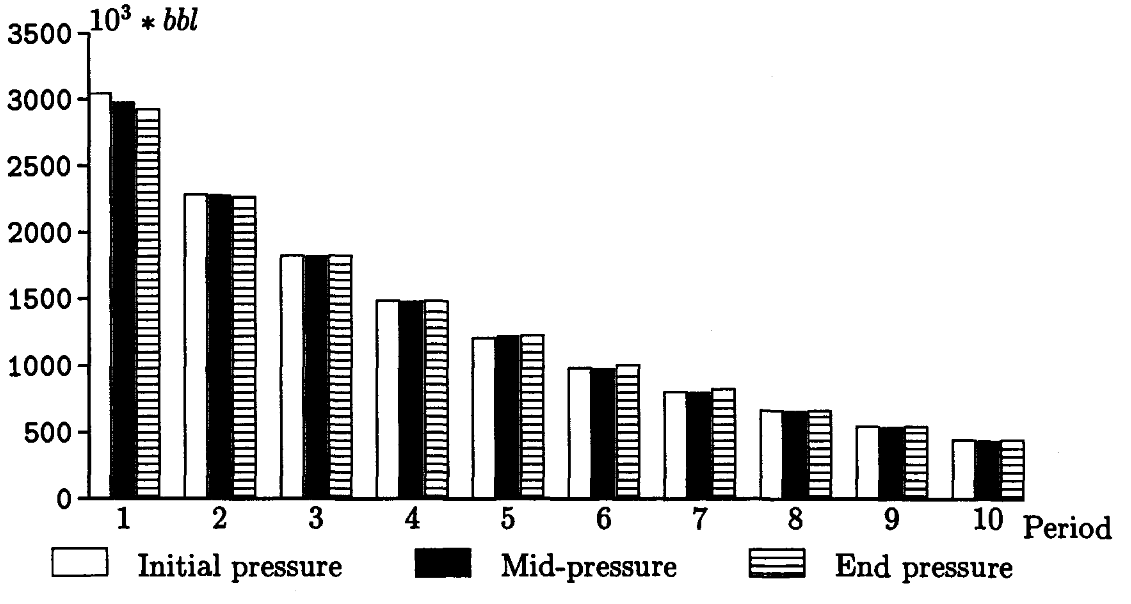

Figure 4.6: Production profile, 100 periods of 30 days each in a 10 period presentation

A, C and D are solved by use of superposition and the mid-pressure method.

4.1.5 Models with Non-linear Reservoir Equations

As pointed out earlier, the models derived in this dissertation is based on a single phase oil reservoir which may be approximated by a set of linear difference equations. In Chapter 3 we found that for describing single phase gas reservoirs and reservoirs with multi-phase flow, it is necessary to use non-linear equations. Several of the models reviewed in Chapter 2 studied gas reservoirs, but in many situations there was made an attempt of linearizing the reservoir description [87, 107]. Another approach of solving such problems are found in Lasdon et al [60], using the reduced gradient technique.

Hallefjord, Asheim and Haugland [38] discuss a possible solution method based on the simplicial decomposition technique. A general discussion of this technique is found in von Hohenbalken [104] and Hearn, Lawphongpanich and Ventura [50]. The idea may be seen as a way of using "superpositioning" also in the non-linear case. In the linear case the response matrix was calculated by simulating unit production in each of the wells. However, this constraint matrix could be calculated using other production vectors, as long as the $B+1$ production vectors are linearly independent. In the technique outlined, $B+1$ such production vectors are used for calculating the reservoir response, and the reservoir description is linearized accordingly. That is, every possible production strategy is seen as a convex combination of these exactly evaluated production strategies. This procedure will be repeated iteratively, by finding updated sets of production vectors defining the new linearized constraints. How successful such a solution technique will be and its convergence properties are left as topics for future research.

4.2 Models for Development and Production Decisions

We will in this section discuss how development decisions may be included in the optimization model. Which development decisions to include and how to do this will necessarily depend on the actual problem under consideration. Once again we will emphasize the fact that the decision variables interact, and because of this interaction our approach is to optimize the decisions simultaneously.

4.2.1 Design decisions

We will here show how decisions concerning drilling sequence, platform capacity and platform operation may be included in the model. Adding integer variables to the model make the resulting problems considerably harder to solve, but due to the nature of these decisions, they are in the proposed model represented by integer variables.

Well Drilling

A set of B potential well sites is proposed, and we consider wells to be drilled for the T_B first periods. We let $C_b^{T_B}$ be the discounted cost of having well b drilled so that production can start in period T_B. We further define $C_b^1, \ldots, C_b^{T_B-1}$ such that the increased cost if we rather choose to drill well b in some period k prior to T_B is: $\sum_{n=k}^{T_B-1} C_b^n$. The well decision variable x_b^n may then be defined as:

$$x_b^n = \begin{cases} 1 & \text{if well } b \text{ is drilled in one of the periods } 1, \ldots, n \\ 0 & \text{otherwise} \end{cases} \tag{4.25}$$

As seen from the definition, x_b^n is non-decreasing for increasing n, and in our model T_B is the latest possible drilling date. The requirement of x_b^n be non-increasing may be written

$$x_b^{n-1} - x_b^n \leq 0, \quad b = 1, \ldots, B, \; n = 2, \ldots, T_B \tag{4.26}$$

The inequality (4.3) assuring that production is below a certain capacity limit S_B, is here reformulated requiring that production can only take place if the well is drilled:

$$q_b^n \leq S_b x_b^n, \quad , b = 1, \ldots, B, \quad n = 1, \ldots, T_B - 1 \tag{4.27}$$

$$q_b^n \leq S_b x_b^{T_B}, \quad , b = 1, \ldots, B, \quad n = T_B, \ldots, T \tag{4.28}$$

In Hallefjord, Haugland and Asheim [38] also other formulations of well decisions are presented, but the above representation of x_b^n is the one they found to be most effective.

Platform Capacity

While the well decisions are discrete, it is not obvious that capacity decisions need to be represented by integer variables. This will of course depend on the situation and the technical restrictions that are given. The platform capacity may be given by discrete alternatives, or it may be specified by a continuous variable with an associated cost. When it comes to solving the problem, it is of course most tractable with a continuous variable with constant cost coefficients. Another possibility is to introduce a piecewise linear cost function. Non-linear cost coefficients or integer variables are less tractable with respect to solving the problem.

However, we here assume a situation where one has to chose between M different platform alternatives with associated capacities and costs. We let Q_1 denote the production capacity of the "smallest" platform, and the total cost of installing this platform is G_1. Further, $Q_2, \ldots, Q_M$ is defined such that if platform g is chosen, the total capacity will be $\sum_{m=1}^{g} Q_m$. The total cost of this platform will be $\sum_{m=1}^{g} G_m$. We see that G_m is defined as the increased cost of expanding the platform capacity with Q_m. The platform decision variable y_m is defined to be non-increasing for increasing m:

$$y_m = \begin{cases} 1 & \text{if a platform with capacity no smaller than } Q_m \text{ is chosen} \\ 0 & \text{otherwise} \end{cases} \tag{4.29}$$

The requirement of y_m being non-increasing is written as

$$-y_{m-1} + y_m \leq 0, \quad m = 2, \ldots, M \tag{4.30}$$

and the inequality requiring production to be below the capacity limit may be formulated as

$$\sum_{b=1}^{B} q_b^n \leq \sum_{m=1}^{M} Q_m y_m, \quad n = 1, \ldots, T \tag{4.31}$$

Operating Decisions

The fixed operating costs of a platform are considerable, and we have therefore decided to include operating decisions in this model. The discounted operating costs in period n are denoted by H_n, and we define the integer operating variable z^n as:

$$z^n = \begin{cases} 1 & \text{if the platform is operating in period } n \\ 0 & \text{otherwise} \end{cases} \tag{4.32}$$

We will not allow the platform to close down for a later restart, and therefore z^n must be non-increasing for increasing n:

$$-z^{n-1} + z^n \leq 0, \quad n = 2, \ldots, T \tag{4.33}$$

To assure that production only can take place when the platform is operated, we introduce the constraint:

$$\sum_{b=1}^{B} q_b^n \leq D^n z^n, \quad n = 1, \ldots, T \tag{4.34}$$

where D^n is a number greater or equal to the platform capacity.

4.2.2 A Deterministic Optimization Model

By combining all the variables, costs and constraints introduced in this section, we get the following deterministic mixed integer model:

$$\max \sum_{n=1}^{T} c^n \cdot \Delta t \sum_{b=1}^{B} q_b^n - \sum_{b=1}^{B} \sum_{n=1}^{T_B} C_b^n x_b^n - \sum_{m=1}^{M} G_m y_m - \sum_{n=1}^{T} H_n z_n \tag{4.35}$$

such that

$$q_b^n \leq J_b\Big(p^0 - \sum_{k=1}^{n} \sum_{l=1}^{B} \alpha_l^{n+\frac{1}{2}-k}(b)\ q_l^k - p_w\Big), \quad b = 1, \ldots, B,\ n = 1, \ldots, T \tag{4.36}$$

$$q_b^n \leq S_b x_b^n \quad b = 1, \ldots, B,\ n = 1, \ldots, T_B - 1 \tag{4.37}$$

$$q_b^n \leq S_b x_b^P \quad b = 1, \ldots, B,\ n = T_B, \ldots, T \tag{4.38}$$

$$x_b^{n-1} - x_b^n \leq 0 \quad b = 1, \ldots, B,\ n = 1, \ldots, T \tag{4.39}$$

$$\sum_{b=1}^{B} q_b^n \leq \sum_{m=1}^{M} Q_m y_m \quad n = 1, \ldots, T \tag{4.40}$$

$$-y_{m-1} + y_m \leq 0 \quad m = 2, \ldots, M \tag{4.41}$$

$$\sum_{b=1}^{B} q_b^n \leq D^n z^n \quad n = 1, \ldots, T \tag{4.42}$$

$$-z^{n-1} + z^n \leq 0 \quad n = 2, \ldots, T \tag{4.43}$$

$$q_b^n \geq 0, \quad b = 1, \ldots, B,\ n = 1, \ldots, T \tag{4.44}$$

$$x_b^n \in \{0, 1\} \quad b = 1, \ldots, B,\ n = 1, \ldots, T_B \tag{4.45}$$

$$y_m \in \{0, 1\} \quad m = 1, \ldots, M \tag{4.46}$$

$$z^n \in \{0, 1\} \quad n = 1, \ldots, T \tag{4.47}$$

The resulting problem has $(B \times T_B + M + T)$ integer variables and $(T \times B)$ continuous variables. We see that the problem size depends on the number of potential wells, drilling periods, number of platform alternatives and the discretization in time. Finer spatial discretization increases the computational effort of the reservoir-simulator, but the size of the optimization problem does not depend upon spatial discretization.

The model presented above, is closely related to the model proposed by Hallefjord, Haugland and Asheim [38], and this model is also used in the Papers A, C and D in Part II of this dissertation.

Chapter 5

Optimization under Uncertainty

We will in this chapter introduce uncertainty in our discussion of optimal field development. First selected literature dealing with petroleum field optimization under uncertainty is reviewed. When uncertainty is introduced in the model proposed in Section 4.2, we end up with a stochastic integer programming problem, and different approaches for modeling and solving such problems are discussed in Section 5.2. The chapter ends with a discussion of the findings in the papers A, B and C.

5.1 Field Optimization under Uncertainty

A petroleum field development project represents a huge investment, and the uncertainty regarding such a project is considerable. In the following section we will look closer at uncertainty regarding:

- the future oil price
- the project costs (both development and operating costs)
- the reservoir

However, in a field development project there are also other factors that play important roles. The timing of project activities, both development and production, has significant impact on project value. Future development of production technology may make it possible to produce petroleum that earlier was considered non-recoverable and to increase the rate of recovery. It is also possible that future technical development may lead to increasing well rates and changed costs. Further do economic factors as currency exchange rates, rate of interest, inflation and governmental regulations have considerable impact on the value of a petroleum field. It is clear that it is very hard, if not impossible, to manage a model with uncertainty in all these factors. Thus the modeling challenge is as always: Focus on the most important uncertainty and leave the rest for more detailed analysis.

Field optimization under **oil price** uncertainty is discussed in Paper A. The uncertainty is there given as a set of future oil price scenarios, and the objective is to maximize the project's expected net present value. Optimization under price uncertainty is further discussed in this chapter and in Paper A, but the topic of oil price forecasting and modeling is beyond the scope of this dissertation. A review of that area is found in Lund [67], and oil price modeling is also discussed by Stensland [94].

The **project costs** may be highly uncertain, especially if the project involves rather new technical solutions. The costs will necessarily depend on the physical conditions where the platform is located and also on the reservoir properties. In other words, if there is large uncertainty regarding the reservoir properties it is also reasonable to assume that there is considerable uncertainty regarding the development and operating costs. Techniques for estimating costs of field development projects are discussed by Nilssen [76]. In the models used in this dissertation, deterministic costs are used, but the models may be revised in order to treat stochastic project costs.

In Chapter 2 we briefly discussed uncertainty regarding the **reservoir**. We discussed which level of detail that should be chosen when representing the reservoir in a deterministic optimization model. We also looked at models for decision support at a stage when available knowledge about the reservoir is rather limited. As pointed out earlier, the available information about the reservoir is increasing throughout its lifetime, from exploration to the field is produced and abandoned. Even when the field is abandoned, the available information about the reservoir system is not complete. By information we here mean knowledge about the physical parameters in a mathematical description of the fluid flow in the reservoir.

How the uncertainty regarding reservoir properties is addressed depends on how the reservoir is represented in the model. In our work we focus on uncertainty regarding the amount of recoverable reserves and the production properties of the reservoir. In the single phase oil reservoir, the amount of oil in place varies proportional to the porosity, and introducing uncertainty regarding the porosity is equivalent to introduce uncertainty regarding the amount of oil in place. The fluid flow in a single phase oil reservoir depends on the reservoir's permeability. As seen in Chapter 3, also the well production rate in our model depends on the permeabilty. Thus, introducing uncertain permeability is a way to introduce uncertainty regarding the reservoir's production capacity and the well rates.

Field development projects take considerable time to be completed, and all decisions do not have to be made at the time the project is initiated. This means that some decisions may be made later on when more information is available. This leads to a sequential decision problem, and modeling the information structure of such a problem becomes an important issue. The idea is that a complete development plan is not made "here and now", but that future decisions may be adapted according to new information. The fundamental issue is to model which information that is available when the decisions have to be made. If we analyze the mechanisms by which the uncertainty is resolved, the uncertainty may be divided in two groups:

- *project exogenous uncertainty:* uncertainty that will be revealed independent of project decisions
- *project endogenous uncertainty:* the date of revelation of random variables depends on project decisions.

Ekern and Stensland [29] denote this project external and project internal uncertainty. Dixit and Pindyck [25] discuss this topic when analyzing project costs and denote it input cost uncertainty and technical uncertainty. The important issue here is that the project endogenous uncertainty can only be resolved by undertaking the project. At a first glimpse this might not seem as a severe difference, but from a modelers point of view it is. When the information structure of the problem depends on decisions, describing the information structure gets more complicated, and this leads to an optimization problem that is significantly harder to solve.

In petroleum field optimization problems both project exogenous and project endogenous uncertainty are present. A typical example of project endogenous uncertainty is the oil price uncertainty. It is reasonable to assume that information about future oil price will be received independent of the project decisions. When it comes to reservoir uncertainty, the situation is completely different. For a potential project where there are uncertainty about reservoir properties, we have the following situation. If no further action is taken, we will not get any more information about the reservoir. Thus, we see that which information that becomes available depends on which decisions that are made, and reservoir uncertainty is an example of project endogenous uncertainty.

We can also think of situations where reservoir information may be acquired independent of decisions. For example, exploration or production activities on neighboring fields or on similar geological structures may provide new information about the reservoir, but in general reservoir uncertainty is project endogenous. Project costs may be both project endogenous and project exogenous. Uncertainty about how much time, effort and materials that are necessary for undertaking a project will not be resolved unless the project is undertaken. On the other hand, prices of labor and materials fluctuate independently of project decisions and are exogenous to the project. A company developing an oil field may hedge against this by letting a contractor deal with the cost uncertainty. However, allocation of this risk depends on the contractual agreement and is not further discussed here.

One approach for dealing with decision making under uncertainty and for calculating the value of such flexibility, is the real options approach. The papers by Black and Scholes [16] and Merton [71] provided a tool for valuation of financial options. These papers also sparked further research in this area, and the development of option pricing theory (contingent claims analysis) for financial investments has also found applications within real investments. For interpretation of the option theory in a real investment framework, the real options theory has been developed. A thorough review of real options literature is found in Trigeorgis [102]. The real options theory provides a tool for valuation of flexibility in a project. This may be options that occur naturally in a project (e.g.defer, contract, shut down or abandon) or it can be options that can can be planned or built in at some extra cost (e.g. expand capacity).

Bjerksund and Ekern [14] consider several problems under oil price uncertainty, and the decision maker may choose between investing "here and now", defer the investment ("wait and see"), or to abandon the project. Olsen and Stensland [81] discuss a problem with both oil price and resource uncertainty. Two reservoirs may be extracted by one fixed production platform or by a movable platform that extracts one reservoir at the time. The main goal of the paper is to put a value on the flexibility inherent in the latter alternative. A real options approach for dealing with reservoir uncertainty is found in Ekern and Stensland [29].

Stensland and Tjøstheim [96] propose a dynamic programming model where there is a gradual reduction of uncertainty over time. A set of production profiles with associated initial probabilities are assumed, and for each period the probabilities are revised according to the information recieved. The uncertainty resolution is interpreted as variance reduction in the probability distribution over production profiles. This model has some similarities to the model proposed in Paper C, where we consider an oil field development project where uncertainty is resolved over time. Stensland and Tjøstheim [95] analyze production data from several fields in order to investigate if stochastic processes may be used when modeling the uncertain production profiles. Tennfjord [98] proposes a decision tree approach for valuing test production on an oil field. The model introduces reservoir uncertainty, and the information provided by the test production may increase the value of the field.

The value of flexibility in an oil field development project is recently investigated in the PhD-dissertation by Lund [67]. He proposes a stochastic dynamic programming model with uncertainty regarding future oil price, amount of recoverable oil and well rates. The reservoir is represented by a tank model, as discussed in Section 2.3, and data from a Norwegian offshore oil field are used when calculating the value of flexibility. Based on these numerical experiments, Lund finds that flexibility in initiation and capacity decisions may add significant value to a field development project.

5.2 Stochastic Programming

For many situations a two stage model is seen as sufficient when modeling decision making under uncertainty. Two stage models have shown to be useful as planning models where the first stage decisions have to be made here and now, while the second stage decisions may be made after the uncertainty is resolved. The first stage decisions may represent the plan (e.g. capacity or location decisions) while operational decisions which are allowed to depend on observed random elements are made at the second stage. Such problems are typically modeled as recourse problems. If we think of a planning model the costs of the plan is incurred in the first period, while the second stage costs, denoted recourse costs, are incurred due to the amount of violation in the constraints.

A general formulation of two stage recourse problems are shown in Paper B. In the area of stochastic programming, there has been great effort on solving two stage stochastic recourse problems. The most cited method for dealing with two stage recourse problems is the L-shaped decomposition method proposed by Van Slyke and Wets [103]. This cutting plane method is a

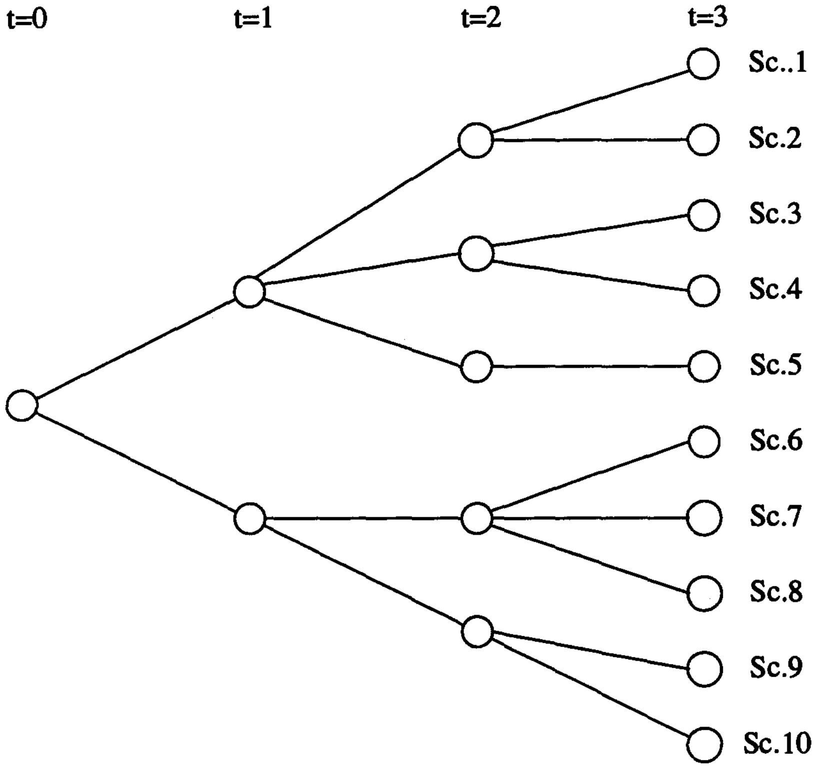

Figure 5.1: Examples of event trees

specialized version of Benders decomposition method [13] and works for programs with continuous variables. An overview of decomposition methods in stochastic programming may be found in Ruszczyński [90].

We will focus on stochastic programming problems with multiple stages. When the random variables have finite discrete probability distributions, the realization of these variables may be illustrated by an event tree. This is a tree that branches of for each value ξ^t taken by the random variable $\tilde{\xi}^t$ at stage t. If decisions are made at times indexed by $0, \ldots, T$, random variables are realized at times $1, \ldots, T+1$. It is our modeling convention that the random variable ξ^t is realized in the t-th time period and is known when the decision at stage t is made. A vector which describes the realizations of the random variables in each time period is denoted a scenario:

$$s = \left(\xi^1, \xi^2, \ldots, \xi^{T+1}\right)$$

We now assume a set of scenarios S which represent the realizations of the random variables. An event tree, as illustrated in Figure 5.1, is also often denoted a scenario tree. The number of leaf nodes on the tree equals the number of scenarios. In traditional scenario analysis it has been usual to solve a deterministic optimization problem for each scenario and analyze the individual solutions in order to see if general trends may be discovered. An "average" solution may also be calculated by assigning probabilities (weights) p^s to each scenario. We let $\boldsymbol{x}^s$ be the solution vector for scenario s, and an averaged solution vector can be found as:

$$\hat{\boldsymbol{x}} = \sum_{s \in S} p^s \boldsymbol{x}^s \tag{5.1}$$

Although the scenario analysis approach gives the decision maker insight into the problem, this averaged solution is not necessarily the optimal one. The stochastic problem we actually want to solve can be written as

$$\begin{aligned} \max \quad & \sum_{s \in S} p^s f^s(\boldsymbol{x}^s) \\ \text{such that} \quad & \boldsymbol{x}^s \in C^s \quad \forall s \\ & \boldsymbol{x} \in \mathcal{N} \end{aligned} \tag{5.2}$$

The progressive hedging algorithm developed by Rockafellar and Wets [86] is a solution procedure where the individual scenario solutions are aggregated in an overall solution, and this algorithm converges to the solution of the stochastic programming problem (5.2). A formal description of this algorithm is found in Paper A. The algorithm is under certain requirements proven to converge for continuous variables, and as discussed in Paper A, the progressive hedging algorithm may not converge to the optimal solution for integer problems.

Another technique for dealing with decision making under uncertainty is the **stochastic decision tree** approach. As for scenario aggregation, this is a method which needs finite discrete probability distributions. While the event tree in Figure 4.2.2 branches off for each random variable, the stochastic decision tree branches off for both decisions and random variables. This means that the stochastic decision tree needs a finite number of possible decisions. Stochastic decision trees are also used for solving problems where it is possible to do additional testing before a final decision is made.

We will here use an oil exploration problem found in Hillier and Lieberman [51] to illustrate the use of stochastic decision trees. A company owns a tract of land that may contain oil, and the basic question is whether the company should drill for oil or sell the land. If oil is found the expected profit is estimated to be \$ 700 000, while a loss of \$ 100 000 occurs if the reservoir is dry.

Another company has offered to by the land for \$ 90 000. The a priori probability of finding oil is estimated to be 0.25, and the expected value of drilling for oil is then found to be $0.25 \times \$700\,000 + 0.75 \times (-\$100\,000) = \$100\,000$. Without the possibility for additional testing, the expected value is maximized by drilling. However, if there is a possibility of performing a seismic survey before the decision is made, the situation is different. The cost of this survey is \$ 30 000, and the test has two possible outcomes: favorable or unfavorable. Based on experience, it is known that if there is oil, the probability of a favorable survey result is 0.6, while there is a 0.4 probability for an unfavorable result. If the reservoir is dry, the probability for a favorable result is 0.2 and the probability for an unfavorable result is 0.8. By use of Baye's theorem, which is discussed in Paper C, the probabilities for "oil" or "dry" given the outcome of the seismic survey may be calculated. These numbers are shown in Figure 5.2. This figure shows the decision tree for this oil exploration problem. The squares represent decision nodes and the circles represent chance nodes. We recognize the decision sequence: First decide if a seismical survey should be undertaken, and then choose between "drill" or "sell".

This figure also illustrates the backwards induction procedure used for solving these problems. At each leaf node an associated payoff is given, and by using the calculated probabilities expected payoff may be calculated for each parenting chance node. By folding back optimal decisions towards the root node, an optimal decision strategy may be found. The example also illustrates how the optimal strategy may depend on the acquired information. The optimal decision strategy is to perform a seismic survey. If the survey result is favorable one should drill for oil and if the result is unfavorable, it is optimal to sell the land. The expected value of this decision policy is \$ 123.000. In other words, the expected value of the information provided by the survey is \$ 53.000, which is \$ 23.000 more than the cost of the test.

The main problem by using the stochastic decision tree is the "curse of dimensionality". If we consider a sequence of decisions, we get a large number of possible sequences and associated test results. In Paper C we consider a stochastic decision tree problem where the decisions themselves provide information for updating probabilities. The approach for solving that problem is discussed in the next session.

The last approach for optimizing stochastic dynamic systems that we will discuss is **stochastic dynamic programming**. As with the other approaches we consider a system with finitely many stages, and the dynamic program-

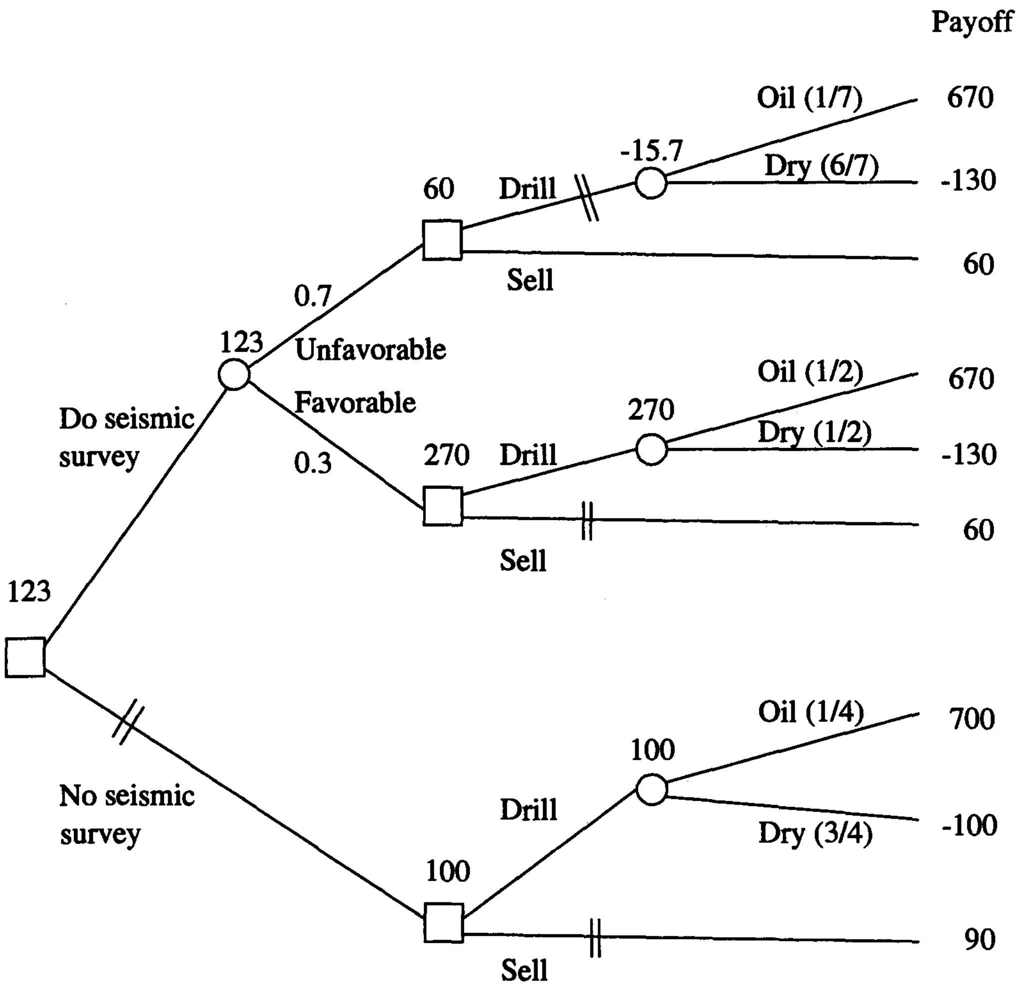

Figure 5.2: Decision tree - The oil exploration problem

ming technique prefers the system at each stage to be in one out of finitely many stages. Theory is also developed for a continuous state space, but for that topic we refer to more specialized literature in the field. While the other methods need a finite decision space and discrete probability distributions, continuous decisions and distributions are acceptable in stochastic programming. The fundamental idea for solving a dynamic programming problem is stated by Bellmann's principle of optimality [12]:

> *An optimal policy has the property that whatever the initial state and initial decisions are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision".*

The solution procedure is to start at the last stage and calculate the optimal decision for each state. Then one moves one step towards the initial stage and again calculate optimal policies for each possible state. Thus solving a stochastic dynamic programming problem can be done through the following recursive relationship:

$$f^n(z^n) = \min_{x^n} \left[r^n(z^n, x^n) + \sum_{z^{n+1}} p_{z^n, z^{n+1}}(x^n) \cdot f^{n+1}(z^{n+1}) \right], \quad \forall z^n, n \in 1, \ldots, T$$

and

$$f^T = \min_{x^n} \left[r^n(z^n, x^n) \right], \quad \forall z^n.$$

where:

z^n	is the state at stage n,
$f^n(z^n)$	is the minimum expected value in state z at stage n,
x^n	is the decision variable at stage n,
$r^n(z^n, x^n)$	is the immediate return by choosing x^n in state z^n,
$p_{z^n, z^{n+1}}$	is the probability of going from state z^n to state z^{n+1} given decision x^n,
T	is the final stage.

5.3 Findings in the Papers A, B and C

We will in this section look closer at the approaches and results in the papers A, B and C. This will be done within the framework of stochastic integer programming techniques outlined in the previous section, and we will to some extent compare the papers.

In Paper A the problem is that of optimizing the development and operation of an offshore oil field under price uncertainty. The starting point for this analysis is the mixed integer programming model described in Chapter 4. The integer (binary) variables in the model represent platform capacity, well drilling decisions and platform operation. By introducing uncertainty regarding the future oil price, the problem becomes a multistage mixed integer stochastic programming problem, and as we have seen, solving such a problem is not a straightforward task. The uncertain oil price is represented by a set of price scenarios, where each scenario is assigned a certain probability. The scenarios may be represented by an event tree, and the solution procedure proposed in the paper is based on scenario aggregation. However, as discussed in the previous section, the progressive hedging algorithm is under certain conditions proved to converge to the optimal solution for problems with continuous variables. The solution procedure applied in Paper A uses progressive hedging on the continuous variables and lets the binary variables adjust automatically. This is the same solution procedure as proposed by Jörnsten and Bjørndal [56]. The reason for using this procedure is the close connection between the design (binary) and production (continuous) variables. Production can only take place from a well that is drilled and the total production may not exceed the installed capacity.

Numerical experiments are reported in the paper, and we conclude that the proposed solution procedure may be a suitable approach for this class of problems. However, it must be emphasized that the convergence properties of this solution technique are not clear. In other words, we have not shown that the algorithm will converge to an implementable solution, and even if the algorithm converge we have not proven that the solution found is optimal. On the other hand, the studied problem is computationally hard, and our approach has been to search for a "good" implementable solution to the problem. How good the solution is, can to some extent be analyzed by looking at what can be obtained under perfect information.

The price uncertainty discussed in Paper A, is a typical example of project exogenous uncertainty. The future oil price will be revealed independently of the project decisions, but if we are considering reservoir uncertainty the situation is different. The paper "A Class of Stochastic Programs with Decision Dependent Random Elements" is a first attempt of modeling and solving stochastic programming problems with both project exogenous and endogenous uncertainty. We look at problems where the probability distribution for the random variables are discrete. For problems with project exogenous uncertainty an event tree representing the realization of random variables may be created, but when the uncertainty is project exogenous this is no longer true. As long as the realization of the random variables depends on decisions, an event tree cannot be specified before the information generating decisions are fixed. In the paper we acknowledge that a scenario is not only characterized by a realization of random variables, but also by the timing of this realization. This means that the scenario aggregation technique used in Paper A is not suitable for solving this kind of problems.

In Paper B we investigate a three stage problem, where information generating decisions are first stage only. In our framework, this means that the decision dependent random variables are either revealed at stage 2 or stage 3. If an event tree is used for representing the random variables, a tree can be constructed as soon as the information generating decisions are specified. When these decisions are specified, the problem is reduced to an ordinary stochastic programming problem with only project exogenous uncertainty. In Figure 5.3 we show examples of different event trees, and which event tree that is realized depends on the first stage decisions. Thus the problem of optimizing programs with both project exogenous and endogenous uncertainty may be seen as a problem of finding the optimal event tree (optimal information generating decisions) and then optimizing the decision policy for that event tree.

The problem of finding the optimal event tree may be solved by:

- Complete enumeration, i.e. solving a stochastic programming problem for each possible event tree
- Implicit enumeration, i.e. finding the optimal solution without explicitly solving the problem for each possible event tree.

In Paper B we propose an implicit enumeration algorithm, which by successively calculating lower and upper bounds on the optimal objective value finds the optimal decision policy.

Also in Paper C we consider a problem with project endogenous uncertainty, and we here return to the area of petroleum field optimization. The basis for our analysis is the mixed integer optimization model described in Chapter 4. We assume the platform capacity to be fixed, and we also assume that it is only possible to drill one production well at the time. The problem is that of selecting where and in which sequence the production wells should be drilled. There are uncertainty regarding the reservoir properties, and information about the reservoir will be obtained as production wells are drilled. There is uncertainty regarding porosity and permeability, and in our numerical experiments we assume these properties to vary proportionally to each other. The reservoir is non-homogeneous and the information received is different for each of the potential wells. Initially the uncertainty is represented by an a priori discrete probability distribution over possible reservoir realizations. When a well is drilled a test result is observed (good, medium or bad). Conditional probabilities for test results given the reservoir realization are estimated a priori, and observed test results are used for calculating an updated probability distribution over possible reservoir realizations.

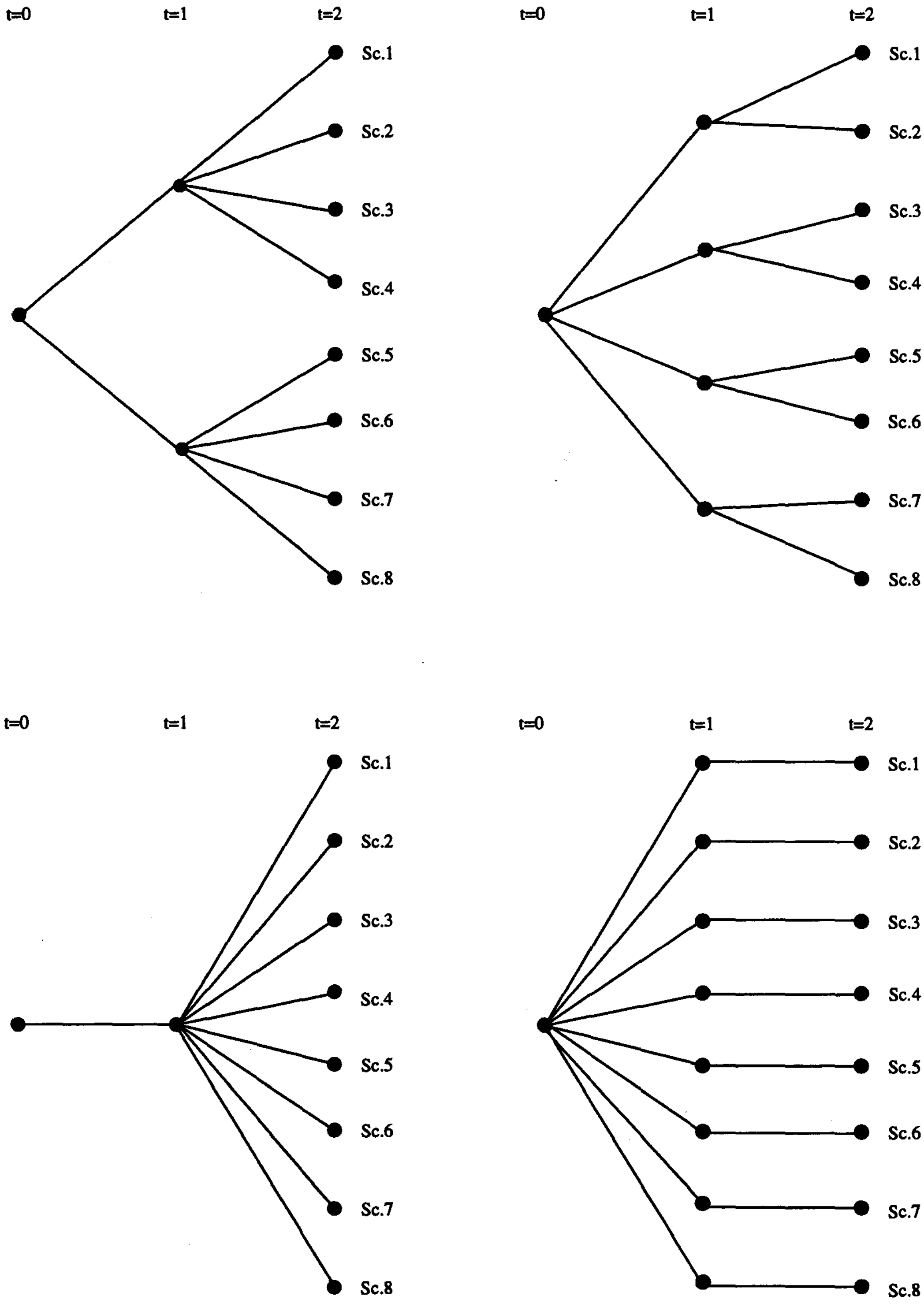

Figure 5.3: Examples of event trees

Compared to Paper B, the main difference is in the way the information is revealed. In the production planning problem analyzed in Paper B, the costs were uncertain, but as soon as a component was produced, the associated costs were known with certainty. For the reservoir problem the situation is different. When a well is drilled and test results observed, we get more information, but all the uncertainty is not resolved. This framework is rather similar to what we find for stochastic decision trees where one has the possibility of additional testing. In the reviewed example from Hillier and Lieberman [51] the main decisions are "drill" or "sell the lease", but it is possible to perform a seismic survey before this decision must be made. The way we see it, the main difference between these approaches is that in the well sequencing problem drilling a well is *both* an investment in production capacity and an investment in information acquisition. In Hillier and Lieberman's problem, it is first decided if information acquisition is profitable, before the main decision is made.

In paper C we also propose a solution procedure that is different from the backwards induction procedure which is the standard solution procedure for stochastic decision trees. The solution algorithm proposed in Paper C works the other way; instead of starting at the leaf nodes and by backwards induction work its way to the root node, the proposed algorithm starts at the root node and works its way toward the leaf nodes. The algorithm is somewhat related to the algorithm proposed in Paper B, because also this algorithm successively calculates upper and lower bounds on the optimal solution. Decision sequences that can be proven to not belong to the set of optimal solutions are bounded out, and the algorithm stops when the optimal decision policy is found. This is a decision policy which is allowed to depend on future information discovery. The complete description of the algorithm is found in Paper C.

Chapter 6

Petroleum Field Unitization

When a reservoir is located on a block boundary, i.e. two lease owners have access to the same reservoir, a common pool problem arises. Because of the oil's migratory nature, both lease owners have an incentive to extract the oil in the reservoir before so is done by the other owner. We will in this chapter discuss how an optimization model may be used for allocating costs and revenues in a coordinated exploitation of the reservoir. This topic is also analyzed in Paper D.

6.1 Introduction to Unitization

So far we have always considered the reservoir as a unit, and the main issue has been to find the optimal development decisions for this reservoir. However, the situation is not always that simple. The Norwegian continental shelf is divided into blocks, and when the government issues exploration rights, these rights are given for each block. Usually several companies get the exploration rights together, and each company's share is decided by the government. Even if the companies share the rights, one of them is appointed to be the one operating the field. Also here we can think of situations where the oil companies disagree about what the optimal development decisions are, but their overall goal is to maximize the value of the petroleum reserves located on the block. But often we find that a petroleum reservoir is located on a block boundary, and thus there are at least two groups of owners that have access to the same reservoir. According to the Norwegian laws the groups of owners are obliged to reach an agreement in which the reservoir is considered as a unit, and an optimal development plan which is independent of the block boundaries has to be chosen [35].

Unitization may be defined as the practice of unifying the ownership and control of a geological unit [70]. A unit operator is given the task of optimizing the value of the field, and shares in this unit have to be negotiated. Norwegian law in this area is based on earlier experiences made when ownership has not been unified. In the United States the property rights to petroleum reserves were regulated by "the rule of capture" - property rights are assigned upon extraction. Because of the oil's migratory nature, each lease owner with access to an oil reservoir has an incentive to dense drilling of production wells in order to drain oil from the neighbours and to take advantage of low extraction costs when the initial reservoir pressure is high. These mechanisms lead to excessive numbers of wells on the field, which clearly reduces the total profitability. Further, the rapid extraction increases extraction costs and the total recovery falls as oil becomes trapped in the geological formations. In Libecap [62] it is reported about oil fields where the economic losses have been large because of competitive exploitation.

The way to overcome these problems is in principle simple, namely by reaching a unitization agreement. Because of the losses which occurs by a competitive exploitation of the reservoir, all parties may receive a part of the profit made by cooperating. However, it is not always that straightforward, and negotiating a unitization agreement may be a very difficult task. The problem under study belongs to the class of common pool problems, and a typical common pool problem is the optimal management of fish resources. When several parties have access to a common fish reserve, one often faces situations where the rate of harvest is larger than the optimal one, and the total catch becomes smaller than what it could have been by coordinated harvesting. In addition we also have the problem of over-investment in fishing vessels. In some cases the fish stocks have been close to extinction because of intensive harvesting. A survey of literature dealing with fishery management in a game theoretic perspective is found in Kaitala [58].

In addition to petroleum reservoirs located on block boundaries, there are also other situations in which the parties in the petroleum industry have to cooperate in order to reach optimal development solutions. Examples of such problems are allocation of costs when several fields use a common transportation system or common processing equipment. In Section 2.5 we looked at optimization models that include transport and processing capacity that are shared among neighbouring fields, but we did not discuss problems that may occur when several owners are involved. Hallefjord, Helming and Jörnsten [40] discuss problems of allocating costs for common processing equipment and transport systems, and unitization of petroleum fields. The problem of cost allocation for common transport networks are also thoroughly discussed by Nilssen[77]. In all these problems the key issue is to allocate the joint costs and benefits, and it is natural to discuss this in the framework of cooperative game theory. From the government's point of view, it is important that development decisions that maximize the value of the portfolio of petroleum reservoirs are chosen. The oil companies involved want a fair allocation of costs and benefits associated with the project. Obviously they will not accept less than what they can achieve on their own, and allocation of the benefits from cooperation is subject to bargaining.

We will here return to the unitization problem. In a bargaining situation with rather homogeneous parties, symmetric distributed information and limited uncertainty about costs and revenues, reaching an agreement do not have to be too difficult. But it is not without reason that it has been difficult to achieve voluntarily unitization of petroleum fields. One obvious reason for this is that the size of the involved leases might be rather different. When development decisions for a petroleum field is made, there is still a lot of uncertainty about the reservoir. Even if more information will become known during the development and production phase, complete information about the underlying reservoir will never become known. What complicates the problem even more, is that which information that will become known depends on the development decisions. This information about the reservoir may also be asymmetrically distributed among the parties. These are all factors that may make it difficult to reach an unitization agreement.

6.2 Findings in Paper D

The main purpose of Paper D has been to illustrate how an optimization model may be used for clarifying unitization negotiations. It is done by use of a numerical example. A reservoir is located on a block boundary, and the model proposed in Chapter 4 is used for finding optimal development and production decisions for this reservoir. When development of the reservoir is modeled as a two player non-cooperative game we find a unique Nash-equilibrium for the game on normal form. In other words, we find the payoff for each of the players if a bargaining solution is not reached. In the paper we find Nash's bargaining solution in which the benefits achieved by cooperating are divided in equal shares among the players. Further we propose possible solutions where the project value is allocated proportionally to amount of recoverable oil in place at each part of the reservoir. We also investigate the project value for each part of the reservoir, in the case of an impermeable wall on the block boundary.

Our aim has not been to recommend a certain procedure for dealing with unitization negotiations in general, but rather to point out how an optimization model may be used for clarifying issues when bargaining for an agreement.

Chapter 7

Conclusions and Directions for Further Research

The main purpose of the first part of this dissertation has been to give a background to the papers in Part II, and to discuss and compare the findings in the papers. These concluding remarks will therefore be based on both this part and the papers in the second part of this dissertation.

7.1 Conclusions

In Chapter 4 we proposed a mixed-integer programming model for optimizing development and operation of an offshore oil field, and different versions of this model are used in Papers A, C and D. The model is formulated as a mixed integer programming problem in which a two-dimensional reservoir simulator is included. In Section 4.1 we showed how the production optimization problem can be formulated as a linear programming problem, while by including field development decisions in Section 4.2 we got a mixed integer problem. Through the literature survey in Chapter 2 we have shown the relationship between the proposed model and previous research. Our work is in the tradition of operations research, where quantitative models are used for solving complex interdisciplinary planning problems. Such an approach combining reservoir engineering and petroleum economics must necessarily be less detailed than for more specialized models. As we have demonstrated, the challenge is to develop models that are manageable and can be solved in a reasonable amount of time and still give a relevant description of the underlying problem. Our approach where fluid flow equations are included in the optimization model can be criticized for leading to problems that are hard to optimize. The model proposed in Chapter 4 is valid for a single phase oil reservoir, and other reservoir systems will inevitably lead to more complex models because of non-linear reservoir equations. However, today's development of computational capability opens for harder problems and for new areas of research.

How uncertainty may be included in optimization models was discussed in Chapter 5, and this is also the topic in Papers A, B, and C. In our review of selected literature treating petroleum field optimization under uncertainty, we especially pointed out the difference between project exogenous and endogenous uncertainty. The project exogenous uncertainty is resolved independent of the project decisions, while project endogenous uncertainty means that when the random variables depend on the project decisions. In a multistage problem this difference is an important issue.

In Paper A we investigate the problem of optimal field development when there are uncertainty regarding future oil price. This is a problem with project exogenous uncertainty. The future oil price is represented by a set of price scenarios with associated probabilities, and our approach is to use the scenario aggregation technique for solving the problem. This technique is developed for continuous variables, while the problem under consideration has both continuous and binary variables. Because of constraints closely relating the binary with the continuous variables, we apply scenario aggregation for the continuous variables and let the binary variables adjust automatically. We conclude that for this class of mixed-integer problems scenario aggregation may be a suitable solution technique.

In Paper B we investigate a class of planning problems with both project exogenous and project endogenous uncertainty. We look at three stage problems where the first stage decisions determine which random variables that will be realized at stage two and will be known when the second stage production decisions are made. In this manner, the selection of optimal first stage decisions and the information refinement are considered as one process. Thus, the first stage production decision is both a production decision and an information acquisition decision as well. We look at problems with discrete probability distributions, and an implicit enumeration algorithm is proposed. The computational experiments are performed for a production planning problem with cost uncertainty, and as such the paper is not an application of petroleum field optimization. But the problem in general, that of modeling and solving problems with project endogenous uncertainty, has applications within petroleum field optimization.

Optimal selection and sequencing of oil wells under reservoir uncertainty is the problem analyzed in Paper C, and this is a problem with project endogenous uncertainty. An a priori probability distribution over reservoir realizations is assumed, and each time a production well is drilled, acquired test results are used for updating the probability distribution. A Bayesian model is proposed for this purpose. The problem has many similarities with the problem analyzed in Paper B, but there are also some important differences. In Paper B, when the random variables are resolved the associated cost is learned with certainty. In Paper C the test results are learned with certainty, but the probability distribution over reservoir realizations is just updated. The problem can be modeled by use of a decision tree, and we propose an implicit enumeration algorithm for finding the optimal solution. Even if solving the problem to optimality turns out be to time consuming, just running a few iterations of the algorithm may give valuable information.

Our work in both Papers B and C show that problems with project endogenous information fight "the curse of dimensionality", and the instances presented in the papers represent problems that have been rather hard to solve to optimality. However, in our view this is an interesting class of problems where the reported research is scarce, and we believe this is a result of lacking tools rather than lack of important problems.

In Paper D we show how a petroleum field optimization model may be a tool for clarifying unitization negotiations. The model is used in a game theoretic perspective where we analyze both non-cooperative and cooperative solutions. When the problem is modeled as a non-cooperative game, we find a Nash-equilibrium to the problem and propose bargaining solutions when the problem is considered as a cooperative game.

7.2 Directions for Further Research

Based on the conclusions of this work, we will here point to three areas of further research: Decision support tools for petroleum reservoirs with a non-linear description, modeling and solving problems with project endogenous uncertainty, and further analysis of petroleum field unitization.

Decision Support Tools and Non-linear Reservoir Descriptions

We have pointed out the importance of decision support tools for field development decisions, and the need for tools based on an interdisciplinary approach. In the models we have proposed, both economical and technical factors are represented on a rather aggregated level. However, the reservoir is the basis for the analysis, and it seems reasonable in future research to focus on how other reservoir models may be included in an optimization model. Other reservoir models will lead to non-linear reservoir descriptions, and how useful such models may be depends on the availability of solution techniques. In Chapter 4 we discussed how simplicial decomposition may be used for solving problems with non-linear system equations. This idea was first proposed by Hallefjord, Asheim and Haugland [38].

It may also be fruitful to investigate if a reservoir simulator that is external to the optimization model may be used. Especially in the uncertainty models the use of commercially available reservoir simulators can be valuable.

Project Endogenous Uncertainty

The research presented in Papers B and C points to several topics for further research. As we pointed out in Paper B, the reported research in this area is very sparse. As a beginning it can be natural to identify different classes of problems, and based on this work analysis of both modeling and solution techniques for these problems can be worked on. In both Papers B and C we propose implicit enumeration algorithms for solving the problems. It may be possible propose generalized algorithms. In Paper B it is pointed to the need for heuristics for generating bounds in the algorithms, and directions for further work are given. Further research on project endogenous uncertainty may be within petroleum related applications, but there are also possibilities for generalization and applications in other areas.

Petroleum Field Unitization

Further research in the area of petroleum field unitization will to some extent depend on the availability and quality of decision support models for field development. The better models, the better reason for the players to agree on the proposed solutions. It may be interesting to include uncertainty about the reservoir properties and asymmetric information distribution in the models. A further analysis of the equilibrium concepts in these dynamic games should also be undertaken.

Bibliography

[1] R. Aboudi, Å. Hallefjord, C. Helgesen, R. Helming, K. Jörnsten, A. S. Pettersen, T. Raum, and P. Spence. A mathematical programming model for the development of petroleum fields and transport systems. *European Journal of Operational Research*, 43:13–25, 1989.

[2] J. A. Al-Zayer. *On Some Mathematical Planning Models for the Management of the Oil, Gas, and Petrochemical Industry in Saudi Arabia.* Ph.D. thesis, University of Texas at Austin, 1986.

[3] J. S. Aronofsky and A. C. Williams. The use of linear programming and mathematical models in underground oil production. *Management Science*, 8:394–407, 1962.

[4] H. Asheim and Å. Hallefjord. Valg av detaljeringsnivå i reservoarmodellering ved tidlig feltevaluering. CMI-report 872330-4, Chr. Michelsen Institute, Fantoft, Norway, 1988.

[5] H. A. Asheim. *Offshore Petroleum Exploitation Planning by Numerical Simulation and Optimization.* PhD-dissertation, University of Texas at Austin, 1978.

[6] H. A. Asheim. Numerical optimization of production strategies for slightly compressible petroleum reservoirs. Technical Report STF28 A81001, Sintef, 7034 Trondheim, Norway, 1980.

[7] H. A. Asheim. Evaluation of the generalized gradient for optimization of combined petroleum production water injection strategies. Technical Report STF28 A81024, Sintef, 7034 Trondheim, Norway, 1981.

[8] H. Attra, W. Wise, and W. Black. Application of optimizing techniques for studying field producing operations. *Journal of Petroleum Technology*, 13:82–86, 1961.

[9] K. Aziz and A. Settari. *Petroleum Reservoir Simulation.* Elsevier Applied Science, London and New York, 1979.

[10] D. A. Babayev. Mathematical models for optimal timing of drilling on multilayer oil and gas fields. *Management Science*, 21(12):1361–1369, 1975.

[11] E. Beale. A mathematical programming model for the long-term development of an off-shore gas field. *Discrete Applied Mathematics*, 5:1–9, 1983.

[12] R. Bellmann. *Dynamic Programming.* Princeton University Press, Princeton, New Jersey, 1957.

[13] J. Benders. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, 4:238–252, 1962.

[14] P. Bjerksund and S. Ekern. Managing investment opportunities under price uncertainty: From "last chance" to "wait and see" strategies. *Financial Management*, Autumn:65–88, 1990.

[15] H. Bjørstad, Å. Hallefjord, and T. Hefting. Decision trees with coupling constraints. CMI-report 30151-2, Chr. Michelsen Institute, Fantoft, Norway, 1988.

[16] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654, 1973.

[17] C. E. Bodington and T. E. Baker. A history of mathematical programming in the petroleum industry. *Interfaces*, 20(4):117–127, 1990.

[18] J. Bohannon. A linear programming model for optimum development of multireservoir pipeline systems. *Journal of Petroleum Technology*, November:1429–1436, 1970.

[19] A. Charnes and W. Cooper. *Management Models and Industrial Applications of Linear Programming.* Wiley, New York, 1961.

[20] A. Charnes, W. Cooper, and B. Mellon. Blending aviation gasolines - a study in programming interdependent activities in an integrated oil company. *Econometrica*, 22(2):135–159, 1952.

[21] M. Christiansen and B. Nygreen. Well management in the North Sea. *Annals of Operations Research*, 43:427–441, 1993.

[22] G. B. Dantzig. *Linear Programming and Extensions.* Princeton University Press, Princeton, 1963.

[23] M. D. Devine. A model for minimizing the cost of drilling dual completion oil wells. *Management Science*, 20(4):532–535, 1973.

[24] M. D. Devine and W. G. Lesso. Models for the minimum cost development of offshore oil fields. *Management Science*, 18(8):378–387, April 1972.

[25] A. K. Dixit and R. S. Pindyck. *Investment under Uncertainty*. Princeton University Press, Princeton, New Jersey, 1994.

[26] S. Dogru. Selection of optimal platform locations. *SPE Drilling Engineering*, 2(December):382–386, 1987.

[27] S. A. Dogru. *The Optimization of Profitability of Offshore Oil Field Operations*. PhD-dissertation, University of Texas at Austin, 1975.

[28] E. J. Durrer and G. E. Slater. Optimization of petroleum and natural gas production - a survey. *Management Science*, 24(1):35–43, September 1977.

[29] S. Ekern and G. Stensland. Realopsjoner og volumusikkerhet ved feltutbyggingsprosjekter. SNF-rapport 106/93, Stiftelsen for samfunns og næringslivsforskning, Bergen, Norway, 1993.

[30] S. D. Flåm and J. Pintér. Selecting oil exploration strategies: Some stochastic program formulations and solution methods. CMI-report 852611-1, Chr. Michelsen Institute, Fantoft, Norway, 1985.

[31] G. E. Forsythe and W. R. Wasow. *Finite-Difference Methods for Partial Differential Equations*. John Wiley & Sons, 1960.

[32] B. L. Foster. Maximization in the oil industry. *Management Technology*, 4(1):26–46, 1964.

[33] L. Frair and M. Devine. Economic optimization of offshore petroleum development. *Management Science*, 21(12):1370–1379, August 1975.

[34] P. Friedman and K. Pinder. Optimization of a simulation model of a chemical plant. *Industrial Engineering Chemistry, Process Design and Development*, 11(4):512–520, 1972.

[35] A. Frihagen. *Unitavtaler*. Bergen, 1985.

[36] W. Garvin, H. Crandall, J. John, and R. Spellman. Applications of linear programming in the oil industry. *Management Science*, 3:407–430, 1957.

[37] Å. Hallefjord. On some operations research models for oil and gas field development. CMI-report 862330-1, Chr. Michelsen Institute, Fantoft, Norway, 1986.

[38] Å. Hallefjord, H. Asheim, and D. Haugland. Petroleum field optimization - comments on models and solution techniques. CMI-report 862331-1, Chr. Michelsen Institute, Fantoft, Norway, 1986.

[39] Å. Hallefjord, D. Haugland, and C. Helgesen. Optimeringsmodeller for olje- og gassutvinning. CMI-report 862325-1, Chr. Michelsen Institute, Fantoft, Norway, 1987.

[40] Å. Hallefjord, R. Helming, and K. Jörnsten. Deling av gevinster ved samordning av petroleumsvirksomhet. CMI-report 30151-1, Chr. Michelsen Institute, Fantoft, Norway, 1988.

[41] P. Hansen, E. de Luna Pedrosa Filho, and C. C. Ribeiro. Location and sizing of offshore platforms for oil exploration. *European Journal of Operational Research*, 58:202–214, 1992.

[42] J. W. Harbaugh, J. H. Doveton, and J. C. Davis. *Probability Methods in Oil Exploration.* Wiley, New York, 1977.

[43] K. K. Haugen. *Possible Computational Improvements in a Stochastic Dynamic Programming Model for Scheduling of Off-shore Petroleum Fields.* Dr.Ing.-dissertation, Norwegian Institute of Technology, Trondheim, 1991.

[44] K. K. Haugen. A stochastic dynamic programming model for scheduling of offshore petroleum fields under uncertainty. *European Journal of Operational Research*, 88:88–100, 1996.

[45] K. K. Haugen, T. Bjørkvoll, and A. Minsaas. Gasslogistikk. Technical Report STF83 A95002, Sintef Anvendt Økonomi, Trondheim, 1995.

[46] D. Haugland. Ei innføring i modellar og metodar for optimering av petroleumsproduksjon. CMI-report A30001, Chr. Michelsen Institute, Fantoft, Norway, 1990.

[47] D. Haugland. *Optimization Methods for Blending Models in Oil Refineries.* Dr.Scient.-dissertation, University of Bergen, 1991.

[48] D. Haugland, Å. Hallefjord, and H. Asheim. Models for petroleum field exploitation. *European Journal of Operational Research*, 37:58–72, 1988.

[49] D. Haugland, K. Jörnsten, and E. Shayan. Modelling petroleum fields with movable platforms. *Applied Mathematical Modelling*, 15:33–39, January 1991.

[50] D. W. Hearn, S. Lawphonngpanich, and J. A. Ventura. Restricted simplicial decomposition: Computations and extensions. *Mathematical Programming Study*, 31:99–118, 1987.

[51] F. S. Hillier and G. J. Lieberman. *Introduction to Operations Research, Sixth Edition.* McGraw-Hill, New York, 1995.

[52] C. A. Holloway. *Decision Making under Uncertainty.* Prentice-Hall, Englewood Cliffs, 1979.

[53] M. K. Hubbert. *The Theory of Ground-water Motion and Related Papers.* Hafner, New York, 1969.

[54] J. D. Huppler. Scheduling gas field production for maximum profit. *Society of Petroleum Engineers Journal*, 14:279–294, 1974.

[55] T. W. Jonsbråten. *Optimeringsmodeller for utvinning av et oljereservoar.* HAS-thesis, Norwegian School of Economics and Business Administration, Bergen, 1994.

[56] K. Jörnsten and M. Bjørndal. Dynamic location under uncertainty. *Studies in Regional and Urban Planning*, Issue 3:163–184, September 1994.

[57] K. O. Jörnsten. Sequencing offshore oil and gas fields under uncertainty. *European Journal of Operational Research*, 58:191–201, 1992.

[58] V. Kaitala. Game theory models of fisheries management - a survey. In T. Basar, editor, *Dynamic Games and Applications in Economics*, pages 252–266. Springer-Verlag, Berlin, 1986.

[59] V. Langston. *An Investment Model for the U.S. Gulf Coast Refining/Petrochemical Complex.* Ph.D. thesis, University of Texas at Austin, 1983.

[60] L. Lasdon, P. E. Coffman, Jr., R. MacDonald, J. W. McFarland, and K. Seehrnoori. Optimal hydrocarbon reservoir production policies. *Operations Research*, 34(1):40–54, 1986.

[61] A. Lee and J. Aronofsky. A linear programming model for scheduling crude oil production. *Journal of Petroleum Technology*, July, 1957.

[62] G. D. Libecap. *Contracting for Property Rights*. Cambridge University Press, Cambridge, 1989.

[63] G. L. Lilien. A note on offshore oil field development and suggested solutions. *Management Science*, 20(4):536–539, 1973.

[64] D. G. Luenberger. *Introduction to Dynamic Systems - Theory, Models & Applications*. John Wiley & Sons, 1979.

[65] D. G. Luenberger. *Linear and Nonlinear Programming*. Addison-Wesley, 1984.

[66] T. Lummus. Drilling optimization. *Journal of Petroleum Technology*, 22(11):1379–1389, 1970.

[67] M. W. Lund. *The Value of Flexibility in Offshore Oil Field Development Projects*. Dr.Ing.-dissertation, The Norwegian University of Science and Technology, Trondheim, 1997.

[68] A. Manne. A linear programming model of the U.S. petroleum refining industry. *Econometrica*, 26:67–106, 1958.

[69] J. Mantell and L. Lasdon. A GRG algorithm for econometric control problems. *Annals of Economic and Social Measurement*, 6(5):581–597, 1978.

[70] S. L. McDonald. *Petroleum Conservation in the United States: An Economic Analysis*. The Johns Hopkins Press, Baltimore and London, 1971.

[71] R. Merton. The theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4:141–183, 1973.

[72] M. Minoux. *Mathematical Programming: Theory and Algorithms*. Wiley, Chichester, 1986.

[73] A. R. Mitchell and D. F. Griffiths. *The Finite Difference Method in Partial Differential Equations*. John Wiley & Sons, New York, 1980.

[74] J. Murray III and T. Edgar. Optimal scheduling of production and compression in gas fields. *Journal of Petroleum Technology*, Januar:109–116, 1978.

[75] P. D. Newendorp. *Decision Analysis for Petroleum Exploration*. Petroleum Publishing Company, Tulsa, 1975.

[76] T. Nilssen. Kostnadsestimering i petroleumsindustrien: En gjennomgang av tre ulike metoder. CMI-report 852310-6, Chr. Michelsen Institute, Fantoft, Norway, 1985.

[77] T. Nilssen. *Kostnadsfordeling ved samordnet transport av petroleum.* HAS-thesis, Norwegian School of Economics and Business Administration, Bergen, 1987.

[78] A. N. Nystad. *Petroleum-Reservoir Management: Reservoir Economics.* Dr.Techn.-dissertation, Norwegian Institute of Technology, Trondheim, 1985.

[79] P. M. O'Dell, N. W. Steubing, and J. W. Gray. Optimization of gas field operations. *Journal of Petroleum Technology*, April:419–425, 1973.

[80] P. R. Odell and K. E. Rosing. *Optimal Development of the North Sea's Oil Fields.* Kogan Page, London, 1976.

[81] T. E. Olsen and G. Stensland. Optimal sequencing of resource pools under uncertainty. *Journal of Environmental Economics and Management*, 17:83–92, 1989.

[82] D. W. Peaceman. *Fundamentals of Numerical Reservoir Simulation.* Elsevier Scientific Publishing Company, Amsterdam, London and New York, 1977.

[83] D. W. Peaceman. Interpretation of well-block pressures in numerical reservoir simulation. *Society of Petroleum Engineers Journal*, June:183–194, 1978.

[84] R. L. Reed. A Monte Carlo approach to optimal drilling. *Society of Petroleum Engineers Journal*, October:423–438, 1972.

[85] R. Richtmyer and K. Morton. *Difference Methods for Initial-Value Problems.* Interscience Publishers, New York, 1967.

[86] R. T. Rockafellar and R. J.-B. Wets. Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of Operations Research*, 16(1):119–147, February 1991.

[87] G. W. Rosenwald. A method for determining the optimum location of wells in a reservoir using mixed-integer programming. *Society of Petroleum Engineers Journal*, February:44–54, 1974.

[88] B. Rothfarb, H. Frank, D. M. Rosenbaum, K. Steiglitz, and D. J. Kleitman. Optimal design of offshore natural-gas pipeline systems. *Operations Research*, 18(6):992–1020, 1970.

[89] G. Rowan. Optimizing the development plan. In *International Petroleum Exhibition and Technical Symposium of the Society of Petroleum Engineers, Beijing, China*, pages 191–199, March 1982. SPE-paper 10017.

[90] A. Ruszczynski. Decomposition methods in stochastic programming. *Mathematical Programming*, 79:333–353, 1997.

[91] R. Schrage. Optimizing a catalytic cracking operation by the method of steepest ascents. *Operations Research*, 6(4):498–515, 1957.

[92] B. A. See and R. N. Horne. Optimal reservoir production scheduling by using reservoir simulation. *Society of Petroleum Engineers Journal*, October:717–726, 1983.

[93] J. Smith. The common pool, bargaining and the rule of capture. *Economic Inquiry*, 25(4):631–644, 1987.

[94] G. Stensland. Generating oil price processes from economic arguments. CMI-report 30152-2, Chr. Michelsen Institute, Fantoft, Norway, 1988.

[95] G. Stensland and D. B. Tjøstheim. Produksjonsusikkerhet. CMI-report 90-A30007, Chr. Michelsen Institute, Fantoft, Norway, 1990.

[96] G. Stensland and D. B. Tjøstheim. Optimal decisions with reduction of uncertainty over time - an application to oil production. In D. Lund and Ø. B., editors, *Stochastic Models and option Value*, pages 267–291. Elsevier Science Publishers B.V., 1991.

[97] G. Symonds. *Linear Programming: The Solution of Refinery Problems*. Esso Standard Oil Company, New York, 1955.

[98] B. Tennfjord. Verdien av prøveproduksjon på eit oljefelt. *Beta*, (1):35–44, 1990.

[99] G. W. Thomas. The role of reservoir simulation in optimal reservoir management. In *SPE 1986 International Meeting on Petroleum Engineering, Beijing, China*, pages 13–22, March 1986. SPE-paper 14129.

[100] R. R. Thompson and D. J. Blumer. Economic optimization of fracture stimulations using mixed integer-linear programming. In *SPE Production Operations Symposium, Oklahoma City, Oklahoma*, pages 161–168, April 1991. SPE-paper 21648.

[101] D. Tjøstheim and T. Hefting. The selection of petroleum exploration strategies and a problem in stochastic portfolio optimization. CMI-report 30800-1, Chr. Michelsen Institute, Fantoft, Norway, 1988.

[102] L. Trigeorgis. Real options and interactions with financial flexibility. *Financial Management*, Autumn:202–224, 1993.

[103] R. Van Slyke and R.-B. Wets. *L*-shaped linear programs with applications to optimal control and stochastic linear programs. *SIAM Journal of Applied Mathematics*, 17:638–663, 1969.

[104] B. von Hohenbalken. Simplicial decomposition in nonlinear programming algorithms. *Mathematical Programming*, 13:49–68, 1977.

[105] S. W. Wallace, C. Helgesen, and A. N. Nystad. Production profiles for oil fields. CMI-report 852310-8, Chr. Michelsen Institute, Fantoft, Norway, 1985.

[106] M. R. Walls, G. T. Morahan, and J. S. Dyer. Decision analysis of exploration opportunities in the onshore US at Phillips Petroleum Company. *Interfaces*, 25(6):39–56, 1995.

[107] R. A. Wattenbarger. Maximizing seasonal withdrawals from gas storage reservoirs. *Journal of Petroleum Technology*, August:994–998, 1970.

Part II

Papers

Paper A

Oil Field Optimization under Price Uncertainty [1]

Abstract

This paper presents a mixed integer programming model for optimal development of an oil field under uncertain future oil price. Based on a two-dimensional reservoir description, the model suggests decisions concerning both design and operation, and the objective is to maximize the expected net present value of the oil field. A finite set of oil price scenarios with associated probabilities is given, and the scenario and policy aggregation technique developed by Rockafellar and Wets is used for solving the problem. This technique is developed for the case of continuous variables, and in this paper we discuss different methods for adapting the scenario aggregation approach to the case of mixed integer problems. This is done by utilizing the interaction between the continuous (production) and integer (design) variables. We present numerical experiments, and conclude that scenario aggregation may be a suitable technique also for mixed integer problems.

[1] Accepted for publication in *Journal of the Operational Research Society*

A.1 Introduction

The purpose of this paper is to develop a model for optimal design and operation of an oil field. The model is meant to be a tool for decision making in an early phase, when there is a lot of uncertainty involved. This is uncertainty regarding future oil price, project costs, and reservoir properties. The current work is based on the model presented in Haugland, Hallefjord and Asheim [8]. Given the input data this model determines optimal platform size, drilling programme and production strategy. The intention of the deterministic model discussed by Haugland, Hallefjord and Asheim "is not to find *the optimal solution*, but to use several versions of input data to generate a number of solutions". In this paper, we will deal with uncertainty in a more structured way. Facing uncertain input data, the objective is to find the decisions that maximize the reservoir's expected net present value.

We restrict our discussion to uncertainty regarding future oil price. One way to describe the uncertainty is by generating a finite set of price-scenarios, where each scenario is given a probability reflecting its relative importance. The problem of maximizing the expected net present value of the oil field can then be formulated as a stochastic programming problem. Such problems easily grow out of hand and become very difficult to solve. In this paper, we will use the progressive hedging algorithm developed by Rockafellar and Wets [20] for solving the stochastic programming problem. This technique can be seen as a decomposition method for stochastic problems. By iteratively solving perturbed versions of the optimization problem for each scenario, we want to find implementable policies, that is policies that do not depend on hindsight. This algorithm is under certain conditions proved to converge to the optimal solution of the stochastic problem for the case of continuous variables, and we discuss how the scenario aggregation approach can be used also for solving mixed-integer problems. This is done by utilizing the interaction between the continuous and the integer variables.

The paper is organized as follows. In the next section, a deterministic mixed-integer optimization model based on a two-dimensional reservoir description is proposed. Later uncertainty is introduced, and we discuss how the progressive hedging algorithm can be applied to mixed integer problems. Results from computational experiments with the resulting stochastic oil field optimization model are presented and discussed, before we, in the final section, give some concluding remarks and outline directions for further research.

A.2 A Deterministic Model

The starting point of our analysis is a reservoir that has been judged to be profitable, and from the size of the reservoir it is decided to develop the field with one platform. We want to make good suggestions for the following decisions:

- platform capacity
- number of wells to be drilled
- well location
- when the wells should be drilled
- production profile for each well

Problems like this are often solved by keeping some of the variables fixed while optimizing the resulting subproblem, but with such a sequential solution procedure, a decision may be fixed before the interaction between the variables are known. The proposed model takes this interaction into account, and it *simultaneously* optimizes platform size, drilling programme and production strategy.

All these decisions will necessarily have the reservoir's production capability as a starting point. The reservoir is a non-renewable resource, and the objective of the optimization problem is to maximize the value of this resource. Therefore, it is of great importance how the reservoir is represented. We study a single phase oil reservoir, a system where only the oil is mobile. The oil is slightly compressible, and production is possible because the oil expands as the pressure is reduced. A reservoir description can be derived by combining an equation for mass conservation, a slightly simplified version of Darcy's flow equation, and an equation for the fluid's constant temperature compressibility. We then arrive at the following reservoir equation where the coefficients do not depend on the pressure:

$$\nabla \cdot (\frac{k}{\mu}\nabla p) = \phi c \frac{\partial p}{\partial t} + \frac{w}{\rho^0} \tag{A.1}$$

In this equation the following notation is used:

k	permeability	(m^2)
μ	fluid viscosity	$(bar \cdot s)$
p	pressure	(bar)

ϕ	porosity	(fraction)
c	constant temperature compressibility	(bar^{-1})
t	time	(s)
w	source/sink terms	$(kg/m^3/s)$
ρ^0	initial fluid density	(kg/m^3)

A more thorough derivation of the reservoir description can be found in Aziz and Settari [1] and Peaceman [18]. The equation (A.1) is based on the assumption that the variation of permeability and porosity with pressure is insignificant, and that the fluid's viscosity is independent of pressure. Further we assume the fluid's compressibility to be constant, and the fluid density's dependence on pressure can then be approximated by a linear term. All these assumptions are common for a single phase oil reservoir over saturation pressure.

A partial differential equation (PDE) like (A.1) is very hard to solve analytically, and common practice is therefore to solve it by finite difference approximations. Because the coefficients in (A.1) are pressure independent, the equation can be approximated by a finite set of linear approximations. In reservoirs that are relatively thin compared to their area extent, it is possible to assume that the flow in the vertical direction (z-direction) is negligible compared to the flow in the other two directions. In our optimization model, we consider a two-dimensional description. Variations in the reservoir height can be adjusted for by letting the height be a function of the x- and y-coordinates.

Another way to represent the reservoir could be by use of a zero-dimensional model, also called tank-model. Hallefjord, Haugland and Helgesen [6] have presented the mathematical formulation of such a model and also its computational results. Another example of zero-dimensional models is presented in McFarland, Lasdon and Loose [17]. In a tank model, the reservoir is assumed to be homogeneous and there is no spatial variation in the reservoir pressure. Such a model gives a very limited description of the reservoir dynamics, and it is not possible to use the model in location analysis. Optimal location of wells is an important question in our analysis, and that is why we have decided to not use this simplified approach.

Previous research on petroleum field optimization includes the above-mentioned work by Haugland, Hallefjord and Asheim [8]. The same authors have also considered this problem in a broader perspective and given a more thorough description of different models and solution techniques for petroleum field optimization [5]. Haugland, Jörnsten and Shayan [9] use the mixed integer model to study fields with movable platforms. They use the model to answer when, if at all, the platform should be moved to another location on the field. Lasdon, Coffman, MacDonald, McFarland and Seehrnoori [15] discuss the problem of finding optimal production profiles for a gas reservoir, and they solve the problem by use of non-linear solution techniques. Devine and Lesso [3], Frair and Devine [4], and Hansen, Filho and Ribeiro [7] also discuss models for optimal development of oil fields. In all these papers, a simplified description of the reservoir performance is used, and the focus is on integer variables and optimal assignment of wells to platforms.

Production Variables

The relation between pressure and production can be approximated in the following way [8]:

$$q_b(t) \leq J_b[p_b(t) - p_w] \tag{A.2}$$

where $q_b(t)$ is the production rate in well b, J_b is a well-specific productivity index, $p_b(t)$ is the reservoir pressure near well b (the pressure in the block where well b is located), and p_w is the minimum well pressure. The productivity index may be estimated from the rock and fluid properties, and it is determined by the permeability, porosity, fluid viscosity, block-area, reservoir height, and well radius [19].

When an explicit description of reservoir performance is used, this does not necessarily mean that a discrete version of equation (A.1) is formulated as a part of the optimization model. We use the same approach as in Wattenbarger [22] and Rosenwald and Green [21], namely the principle of superposition. Initially we assume the reservoir pressure to be p^0, and according to our notation p^n is the pressure after n periods. The estimated pressure in the middle of period n will then be denoted $p^{n-\frac{1}{2}}$. It can be shown for a reservoir described by linear equations that the pressure near well b in the middle of period n can be written [5, 10]:

$$p_b^{n-\frac{1}{2}} = p^0 - \sum_{k=1}^{n} \sum_{l=1}^{B} \alpha_l^{n+\frac{1}{2}-k}(b) \; q_l^k \tag{A.3}$$

The parameters $\alpha_l^{n+\frac{1}{2}-k}(b)$ have the following interpretation: If well l produces one production unit in period k, this results in a pressure drop of $\alpha_l^{n+\frac{1}{2}-k}(b)$ pressure units in period n near well b. Since we want to express p_b^n for a general production q, we have to find all the coefficients α. This can be done by simulation of the system. By solving the discretized reservoir equation for a set of B linearly independent q-vectors, all the α's can be found. This simulation process is discussed in Haugland, Hallefjord and Asheim [8].

Well-drilling

We let $C_b^{T_B}$ be the discounted cost of drilling well b in period T_B, and we define $C_b^1, \ldots, C_b^{T_B-1}$ such that the increased cost if we rather choose to drill well b in some period k prior to T_B is: $\sum_{n=k}^{T_B-1} C_b^n$. The well decision variable x_b^n may then be defined as:

$$x_b^n = \begin{cases} 1 & \text{if well } b \text{ is drilled in one of the periods } 1, \ldots, n \\ 0 & \text{otherwise} \end{cases} \tag{A.4}$$

As seen from the definition, x_b^n is non-decreasing for increasing n, and in our model T_B is the latest possible drilling date.

Platform Capacity

We also want to find optimal platform capacity, and it is possible to choose among M different platform sizes with associated costs. We let Q_1 denote the production capacity of the "smallest" platform, and the total cost of installing this platform is G_1. Further, $Q_2, \ldots, Q_M$ is defined such that if platform g is chosen, the total capacity will be $\sum_{m=1}^{g} Q_m$. The total cost of this platform will be $\sum_{m=1}^{g} G_m$. We see that G_m is defined as the increased cost of expanding the platform capacity with Q_m. The platform decision variable y_m is defined to be non-increasing for increasing m:

$$y_m = \begin{cases} 1 & \text{if a platform with capacity no smaller than } Q_m \text{ is chosen} \\ 0 & \text{otherwise} \end{cases} \tag{A.5}$$

Operating Costs

The fixed operating costs of the platform in period n are denoted by H_n, and we define the integer operating variable z^n as:

$$z^n = \begin{cases} 1 & \text{if the platform is operating in period } n \\ 0 & \text{otherwise} \end{cases} \tag{A.6}$$

We will not allow the platform to close down for a later restart, and therefore z^n must be non-increasing for increasing n.

The complete deterministic mixed integer model can then be written as:

$$\max \sum_{n=1}^{T} c^n \cdot \Delta t \sum_{b=1}^{B} q_b^n - \sum_{b=1}^{B} \sum_{n=1}^{T_B} C_b^n x_b^n - \sum_{m=1}^{M} G_m y_m - \sum_{n=1}^{T} H_n z_n \tag{A.7}$$

such that

$$q_b^n \leq J_b\Big(p^0 - \sum_{k=1}^{n} \sum_{l=1}^{B} \alpha_l^{n+\frac{1}{2}-k}(b)\, q_l^k - p_w\Big),$$
$$b = 1, \dots, B,\ n = 1, \dots, T \qquad (A.8)$$

$$q_b^n \leq S_b x_b^n \qquad b = 1, \dots, B,\ n = 1, \dots, T_B - 1 \qquad (A.9)$$

$$q_b^n \leq S_b x_b^{T_B} \qquad b = 1, \dots, B,\ n = T_B, \dots, T \qquad (A.10)$$

$$x_b^{n-1} - x_b^n \leq 0 \qquad b = 1, \dots, B,\ n = 2, \dots, T_B \qquad (A.11)$$

$$\sum_{b=1}^{B} q_b^n \leq \sum_{m=1}^{M} Q_m y_m \qquad n = 1, \dots, T \qquad (A.12)$$

$$-y_{m-1} + y_m \leq 0 \qquad m = 2, \dots, M \qquad (A.13)$$

$$\sum_{b=1}^{B} q_b^n \leq D^n z^n \qquad n = 1, \dots, T \qquad (A.14)$$

$$-z^{n-1} + z^n \leq 0 \qquad n = 2, \dots, T \qquad (A.15)$$

$$q_b^n \geq 0, \qquad b = 1, \dots, B,\ n = 1, \dots, T \qquad (A.16)$$

$$x_b^n \in \{0, 1\} \qquad b = 1, \dots, B,\ n = 1, \dots, T_B \qquad (A.17)$$

$$y_m \in \{0, 1\} \qquad m = 1, \dots, M \qquad (A.18)$$

$$z^n \in \{0, 1\} \qquad n = 1, \dots, T \qquad (A.19)$$

The objective is to maximize the net present value, and c^n is the discounted oil price in period n and Δt is the length of each period. Constraint (A.8) is the well production capacity, derived by replacing $p_b(t)$ in (A.2) by $p_b^{n-\frac{1}{2}}$ as defined by (A.3). Constraints (A.9) - (A.11) are related to the well decisions. Constraints (A.9) and (A.10) assure that production can take place only if the wells are drilled, while constraint (A.11) requires x_b^n to be non-decreasing for increasing n. The total production may not exceed the platform capacity and this is assured by (A.12). Constraint (A.14) says that production can take place only if the platform is operating, and D^n is an upper bound on the platform capacity. The constraints (A.13) and (A.15) require that y_m and z^n are non-increasing in m and n respectively.

We have a problem with $(B \times T_B + M + T)$ integer variables and $(T \times B)$ continuous variables. The problem size depends on the number of potential wells, drilling periods, number of platform alternatives and the discretization in time. Finer spatial discretization increases the computational effort of the reservoir-simulator, but the size of the optimization problem does not depend upon spatial discretization.

A.3 Scenario Aggregation

As mentioned in the introduction, there is a lot of uncertainty involved in an oil field development project. In this paper, we restrict the uncertainty to the future oil price, and the objective of the stochastic optimization problem will be to maximize the expected value. When solving deterministic multiperiod problems, we want to find a plan for the whole period ahead of time. This is not the case for stochastic optimization problems, because the decisions for all but the first period will depend on what happens in the meantime.

When there is only limited information available about the distribution of the random elements, it may be appropriate to model the uncertainty by generating a finite set of scenarios S:

$$S = \{s^1, s^2, \ldots, s^L\}$$

The scenarios can be illustrated by use of an event tree (scenario tree) as illustrated in Figure A.1. The nodes in the event tree can be seen as decision points while the arcs represent realizations of the random variables. A decision policy for the stochastic optimization problem must be *admissible* and *implementable* [23]. Admissibility means that the decision policy for each $s \in S$ must be feasible. If two scenarios are indistinguishable at time t, then they must yield identical decisions up to and including time t. A decision policy satisfying this at each node in the event tree is called an *implementable* solution.

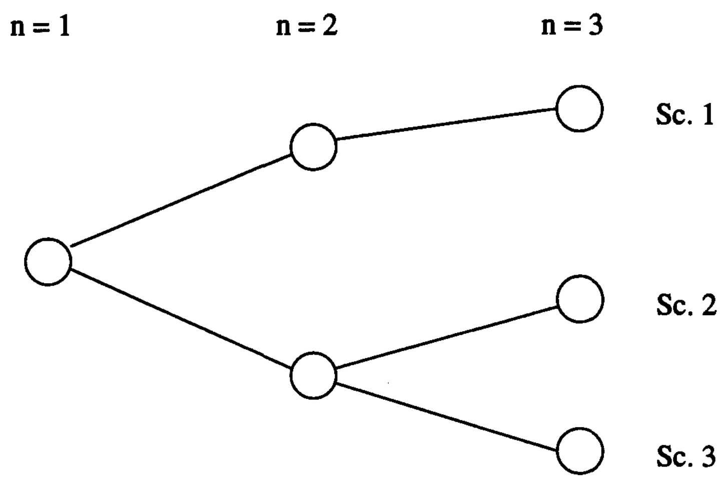

Figure A.1: Event tree

The stochastic problem we actually want to solve can be written as

$$\begin{aligned} \max \quad & \sum_{s \in S} p^s f^s(\boldsymbol{x}^s) \\ \text{such that} \quad & \boldsymbol{x}^s \in C^s \quad \forall s \\ & \boldsymbol{x} \in \mathcal{N} \end{aligned} \tag{A.20}$$

Given the scenarios with associated probability, the objective is to find an admissible and implementable solution, $\boldsymbol{x}$, maximizing the expected value of f. The vector of decision variables, $\boldsymbol{x}$, can be partitioned into one vector for each scenario, so that $\boldsymbol{x} = (\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^L)$. The set C^s contains all solutions that are feasible, given the realizations of s. Hence $\boldsymbol{x}$ is admissible if $\boldsymbol{x} \in C^s$ for all scenarios s. For each period $n = 1, \ldots, T$, the scenario set is partitioned into a finite number of disjoint subsets:

$$S(k^n) \subseteq S, \quad \forall\, k^n \in K^n \tag{A.21}$$

Here k^n is a node in the event tree at stage n, and K^n is the set of all nodes in the tree at that stage. The set $S(k^n)$ consists of scenarios involving the event represented by k^n. The set of implementable solutions thus becomes:

$$\mathcal{N} = \{\boldsymbol{x} \,:\, \boldsymbol{x}^{ns'} = \boldsymbol{x}^{ns''} \;\forall s', s'' \in S(k^n), k^n \in K^n, n = 1, \ldots, T\} \tag{A.22}$$

Solving a problem like (A.20) will often exceed the present computational capabilities. One important feature of (A.20) is that only the implementability constraints, $\boldsymbol{x} \in \mathcal{N}$, connect the scenarios, and it is this fact that is taken advantage of in the scenario aggregation technique.

A.3.1 The Progressive Hedging Algorithm (PHA)

The main idea in the PHA is to decompose the stochastic programming problem into scenario subproblems. The algorithm was developed by Rockafellar and Wets[20], and a simpler version of this paper is found in Wets [23]. In Kall and Wallace [14], the scenario aggregation technique is discussed in relation to other solution techniques for stochastic dynamic programming problems. The problem (A.20) can be rewritten as an augmented Lagrangian where implementability constraints are dualized and removed from the constraint set:

$$\max_{\boldsymbol{x}^s \in C^s} \left[\sum_{s \in S} p^s f^s(\boldsymbol{x}^s) - \sum_{s \in S} p^s w^s(\boldsymbol{x}^s - \hat{\boldsymbol{x}}^{S(k^n)}) - \frac{1}{2}\varrho \sum_{s \in S} p^s [\boldsymbol{x}^s - \hat{\boldsymbol{x}}^{S(k^n)}]^2 \right] \tag{A.23}$$

Here w^s represents multipliers associated with the dualized constraints, while ϱ is a penalty parameter associated with the penalty term of the augmented Lagrangian. The averaged solution for a node in the event tree is written as $\hat{\boldsymbol{x}}^{S(k_n)}$, but because of this term the problem is still not separable in the scenarios. However, in the PHA this term is substituted by the averaged solution from the previous iteration, and thus the problem (A.23) becomes separable. The algorithm can formally be expressed as follows:

Progressive hedging algorithm

Initialize the iteration counter ν, $\nu = 0$.
Solve each individual scenario-problem: $\max_{\boldsymbol{x}^s \in C^s} f^s(\boldsymbol{x}^s)$.
Construct an implementable policy by forming a convex combination of the solutions $\boldsymbol{x}^{ns}(0)$ by use of the weights p^s:

$$\hat{\boldsymbol{x}}^{nS(k^n)}(0) = \frac{\sum_{s \in S(k^n)} p^s \boldsymbol{x}^{ns}(0)}{\sum_{s \in S(k^n)} p^s} \quad \forall\, k^n \in K^n, \quad n = 1, 2, \ldots, T. \tag{A.24}$$

Initialize the price vectors $w^s(0)$ and the penalty parameter $\varrho(0)$.
While the solution is not good enough do:

Increase the iteration counter ν by one.
For each scenario solve:

$$\max_{\boldsymbol{x}^s \in C^s} f^s(\boldsymbol{x}^s) - \boldsymbol{x}^s w^s(\nu - 1) - \frac{1}{2}\varrho(\nu - 1)[\boldsymbol{x}^s - \hat{\boldsymbol{x}}^{S(k^n)}(\nu - 1)]^2 \tag{A.25}$$

Calculate an implementable policy $\hat{\boldsymbol{x}}(\nu)$ using the same technique as in equation (A.24).
Update the multiplier vector:

$$w^{ns}(\nu) = w^{ns}(\nu - 1) + \theta(\nu)[\boldsymbol{x}^{ns}(\nu) - \hat{\boldsymbol{x}}^{ns}(\nu)] \tag{A.26}$$

Adjust the parameters $\theta(\nu)$ and $\varrho(\nu)$.

End

The adjustments of $\theta(\nu)$ and $\varrho(\nu)$ are discussed in a later section. A further discussion of the termination criterion is found in Rockafellar and Wets [20] where it also is shown that this algorithm in the convex case, converges to the global optimal solution of the original stochastic problem for the case of continuous variables. The multipliers, $w^{ns}(\nu)$, can be seen as estimates of the price system associated with the implementability constraints.

A.3.2 The PHA and Mixed Integer Programming

The PHA is constructed for solving optimization problems with continuous variables; but the problem we are concerned with in this article is a mixed integer one. The use of the scenario aggregation approach for solving integer problems is discussed in Jönsson, Jörnsten and Silver [11] and Jörnsten [13]. There a Lagrangian relaxation approach is used for generating upper bounds for the integer maximizing problem. This technique for generating bounds is used together with a heuristic method for obtaining feasible solutions to the integer problem.

Løkketangen and Woodruff [16] use the PHA for solving stochastic multi-stage mixed integer problems. Their approach is to use tabu search for solving the individual scenario problems and the PHA for blending the scenario solutions. Jörnsten and Bjørndal [12] use the PHA for solving a stochastic multi-stage mixed integer problem, and they make use of the close connection between the continuous and binary variables in their problem. They study an uncapacitated dynamic location problem under uncertainty, and the PHA is used on the continuous variables while the binary variables are allowed to adjust automatically. This solution method can be justified because of the close connection between the continuous and binary variables in the problem.

In our case, there is also a close connection between the binary and continuous variables, and our approach will therefore be the same as in Jörnsten and Bjørndal [12]. As seen in the deterministic model (7), there are several constraints that link the production and design decisions. Production can only take place if the well is drilled:

$$q_b^n \leq S_b x_b^n$$

The total production may not exceed the platform capacity:

$$\sum_{b=1}^{B} q_b^n \leq \sum_{m=1}^{M} Q_m y_m$$

The platform must be operated if production is to be possible:

$$\sum_{b=1}^{B} q_b^n \leq D^n z^n$$

All these relations show how the production decisions influence the design decisions. These connections between production and design decisions motivate the following solution method: Apply the PHA on the continuous (production) variables and let the integer (design) variables adjust automatically.

Even if the PHA decomposes the stochastic problem into scenario subproblems, it is not straight-forward to solve these nonlinear subproblems. One way to overcome this problem would be by using a piecewise linear approximation for the quadratic term in equation (A.25). However, this gives another complex problem to solve. The augmented Lagrangian in problem (A.23) can be viewed as a combination of a local duality method and a penalty method. In the augmented Lagrangian the duality and the penalty methods work together to eliminate the slow convergence properties with either method alone. As an approach of making the scenario subproblems fairly easily solvable, we have in the numerical experiments presented here set the penalty parameter, $\rho(\nu)$, to be zero. In other words, the quadratic term of (A.25) is neglected. When the optimal solution is found, both the duality and the penalty term of equation (A.25) equal zero; but the influence of neglecting the penalty term on the algorithm's convergence properties, is not immediately clear. However, we now have subproblems that are solvable, and if the PHA converges to a feasible solution, we get lower bounds on expected maximum profit. These solutions can be compared to what can be achieved under perfect information, and that will give an indication of the quality of the calculated solutions.

A.4 Numerical Experiments

The numerical experiments are performed for a rectangular reservoir with the following rock and fluid properties:

Reservoir length (x-direction): 1000 m

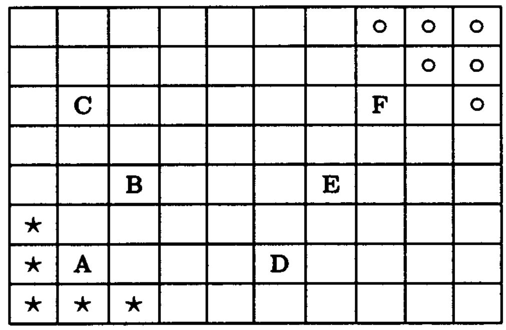

Figure A.2: The reservoir

Reservoir breadth (y-direction):	800 m
Reservoir height (z-direction):	50 m

Viscosity (μ):	$1 \cdot 10^{-7} bar \cdot s$
Constant temp. compressibility (c):	0.001 bar^{-1}
Fluid density at initial pressure (ρ^0):	1150 kg/m^3
Fluid density at surface pressure (ρ^s):	850 kg/m^3
Reservoir porosity (ϕ):	0.2
Reservoir initial pressure (p^0):	300 bar
Lowest pressure for production (p_w):	50 bar
Well radius (r_w):	0.15 m

The lifetime of the field is estimated to be 3000 days, which is discretized into 10 periods of 300 days. We divide the reservoir into 10×8 blocks, each of dimension $100m \times 100m$, as illustrated in Figure A.2. The permeability of the reservoir varies from $1.0 \cdot 10^{-13} m^2$ in the lower-left corner to $0.5 \cdot 10^{-13} m^2$ in the upper-right corner. The most permeable blocks of the reservoir are marked with a ($\star$). The permeability is gradually reduced along the rectangle's diagonal, and reaches its minimum at the blocks marked with a ($\circ$). The 6 potential well sites are in Figure D.1 marked with letters A - F. The drilling cost for each well is set to \$6.25 mill, and it is only for the 3 first periods that the wells may be drilled. The fixed operating costs are \$4.0 mill. each period (300 days). The different platform alternatives with associated costs are given in Table A.1. The discount rate is 3% per period.

We will use 3 scenarios of future oil price (see Table A.2). All the scenarios start with \$17 per barrel in the first period, and scenario 1 has this price in all the 10 periods. In scenario 2, the oil price increases with \$1 every second period, while in scenario 3, the oil price increases with \$1 every period.

Platform	Capacity l/sec	Cost (million \$)
1	10	20.0
2	15	22.5
3	20	25.0
4	25	27.5
5	30	30.0

Table A.1: Platform capacities and costs

Scenario	Oil price development (\$ / bbl)										Prob. I	Prob. II
1	17	17	17	17	17	17	17	17	17	17	0.4	0.2
2	17	17	18	18	19	19	20	20	21	21	0.3	0.2
3	17	18	19	20	21	22	23	24	25	26	0.3	0.6

Table A.2: Oil price scenarios

A.4.1 Scenario Solutions

The scenario problems are solved by use of "CPLEX Mixed Integer Library" [2]. In Step 0 of the PHA, each individual scenario problem is solved to optimality, and the optimal drilling programme for each of the scenarios is shown in the first lines of Table A.3. For all scenarios, it is optimal to drill the same 5 wells; but the timing of the drilling operations differ. The well decisions are,

	Period 1	Period 2	Period 3	Platform	Operation
Scen. 1	C D E F	A		4	8 periods
Scen. 2	C D F	E	A	3	9 periods
Scen. 3	C E	F	A D	2	10 periods
Prob. I	C D F	$E_{1,2}$	A E_3	3	$9_{1,2}$, 10_3 periods
Prob. II	A E	F	C D	2	10 periods

Table A.3: Development decisions

of course, closely connected to the platform capacity, and the three scenarios have all different platform sizes. At first, it may be surprising that scenario 1, the scenario with lowest oil price, has the largest investments in platform capacity, and the "best" scenario leads to the smallest platform. But these decisions have to be judged against the fixed operating costs and the abandonment time. Given scenario 1, it is optimal to produce for 8 periods. The wells' production capabilities depend on the reservoir pressure, and after 8 periods the operating costs exceed the revenue. In scenario 3, the situation is different. Because of increasing oil price, production is most profitable in late periods, and that is why it is optimal to invest in a smaller platform and use all the 3000 days for production.

A.4.2 Multipliers on the Production Variables

With the individual scenario solutions as a starting point, the PHA was applied as discussed earlier, by associating multipliers with the production variables. Implementable production policies $\hat{q}_b^{ns}(\nu)$ are constructed, and the multipliers are calculated as follows:

$$w_b^{ns}(\nu) = w_b^{ns}(\nu - 1) + \theta(\nu)[q_b^{ns}(\nu) - \hat{q}_b^{ns}(\nu)] \tag{A.27}$$

Probabilities I of Table A.2 are used in this experiment. The PHA converged to the decisions presented in line four of Table A.3. The solution obtained is to choose platform 3 and to have 3 of the potential wells drilled for production. For the second period, scenario 1 and 2 are still indistinguishable, and they therefore must have identical decision policies. The subscript in Table A.3 indicates that well E is drilled for the second period if scenario 1 or 2 is observed, and for the third period, if scenario 3 is observed.

An important issue regarding the PHA is the updating of the multiplier vector. The updating is shown in equation (A.26), and the updating parameter, $\theta(\nu)$, has large influence on the convergence properties. Initially the multiplier vector is chosen to be zero. In the experiments reported here, the parameter has been updated in the following way:

$$\theta(\nu + 1) = 0.85 \cdot \theta(\nu) \tag{A.28}$$

During the iterations, the multipliers are adjusted with decreasing steps. The initial updating parameter, $\theta(1)$, depends on how the problem is formulated. In the experiment discussed here, the algorithm converged to implementable design decisions (binary variables) after 30 iterations. After just one more iteration, the algorithm also gave implementable decisions for all the production decisions (continuous variables).

A.4.3 Multipliers on Total Production

By inspection of the optimal multipliers and the iterative process towards the optimal multipliers, we found it possible to speed up the convergence rate.

What the multipliers actually do, is to perturb the price vectors. In the results presented, the oil price in the first period was reduced for scenario 1 and increased for scenario 3. But by using each well's production, this iterative process was rather slow, and we therefore modified the algorithm. While the perturbed scenario solutions do not have the same platform decisions, the multipliers are calculated from each period's total production. When the platform decision is implementable, we use the ordinary approach where the multipliers are calculated from the production in each well.

By using this method for updating the vectors, the algorithm converged to implementable solutions for both design and production decisions after 7 iterations. After only 3 iterations, the platform decision was implementable, and these initial iterations can be seen as a fast way to get the multipliers to their correct level. The last iterations will then only be minor adjustments. The algorithm converged to the same design decisions as in the previous example.

A.4.4 Multipliers on both Production and Platform Variables

In the case of probability distribution II of Table A.3, we were unable to find an implementable solution by using multipliers on the production variables only. The method so far has been to "adjust the oil price"; but in this example, the optimal solution for scenario 1 was to produce nothing in the first period, and the algorithm never gave an implementable platform decision. In the case of divergence, we modified the algorithm by calculating multipliers also for the platform variables. Because of the discreteness of these variables, it may seem meaningless to calculate solutions that are an average of 0-1 solutions. However, the capacity decision is actually one, not m, even if we have m platform alternatives, and as a heuristic approach for obtaining lower bounds, we found it to work here. Platform capacity far from the "average capacity", $\bar{Q}(\nu)$, was more heavily punished than platform capacity close to the "average". The multiplier updating parameter, $\theta(\nu)$, was calculated in the following way:

$$\theta_m(\nu + 1) = |Q_m - \bar{Q}(\nu)| \cdot (0.85)^{\nu} \cdot \theta_m(0) \tag{A.29}$$

Multipliers of the production variables were calculated as done in the first example. Also these results are shown in Table A.3. The algorithm converged to platform 2 after 6 iterations, and after 38 iterations it converged to implementable well decisions. In these numerical experiments, we have used a rather low discount rate, 3% per period. When the discount rate was increased to 4% per period, this last problem could be solved by using multipliers only on the production variables.

A.4.5 The Value of the Model

When analyzing the optimal drilling programmes, we see that the same five wells are drilled in all three scenarios. The differences are in the timing of the drilling operations and in the platform capacity. We have also calculated the expected value of perfect information (EVPI). For both examples discussed here, the EVPI is lower than 1 % of the calculated project value. The EVPI's are $ 1.0 mill. and $ 1.4 mill., respectively. These numbers illustrate the inspiring fact of optimizing oil field development, that even a small improvement of project value measured in percentage represents a considerable amount of money.

A.5 Conclusions and Further Research

The main purpose of this paper has been to develop a model for optimal oil field development under price uncertainty. The uncertain oil price is given as scenarios with associated probabilities. This results in rather complex mixed integer optimization problems, and our approach has been to use the PHA for obtaining lower bounds to these problems. The PHA was used on the continuous variables, and the algorithm converged to an implementable policy both for the integer and the continuous variables. In the numerical experiments, we used both production in each well and the total production in each period for calculating multipliers. The results obtained by using the total production for calculating multipliers until convergence in platform capacity is reached, seem very promising. For some of the problems, it was also necessary to use multipliers on platform variables to obtain convergence to implementable solutions. Further research will be testing of the model on a larger set of input data, and search for parameters that optimizes the algorithm's convergence rate.

When the uncertainty is in the objective function, as it is in the problem studied here, we know that if a solution is feasible for one scenario, it will be feasible for the other scenarios also. Connected to oil field development projects, uncertainty about the reservoir properties may have large impact on the design decisions. Uncertainty with respect to the reservoir leads to uncertainty in the constraints of our programming problem, and then the problem is harder to solve. This information gathering process will also be different from the case with uncertain oil price, because the information revealed about the reservoir will be a result of the actions taken. How such problems can be approached will be addressed in future work.

Acknowledgments

I want to thank professor Kurt Jörnsten, Norwegian School of Economics and Business Administration, and associate professor Dag Haugland, Stavanger College, for valuable ideas and comments. The comments from two anonymous referees are also gratefully acknowledged.

References

[1] K. Aziz and A. Settari. *Petroleum Reservoir Simulation.* Elsevier Applied Science, London and New York, 1979.

[2] CPLEX Optimization, 930 Tahoe Blvd., Bldg 802, Incline Village, NV 89451-9436. *Using the CPLEX Callable Library and CPLEX Mixed Integer Library*, 1993.

[3] M. D. Devine and W. G. Lesso. Models for the minimum cost development of offshore oil fields. *Management Science*, 18(8):378–387, April 1972.

[4] L. Frair and M. Devine. Economic optimization of offshore petroleum development. *Management Science*, 21(12):1370–1379, August 1975.

[5] Å. Hallefjord, H. Asheim, and D. Haugland. Petroleum field optimization - comments on models and solution techniques. CMI-report 862331-1, Chr. Michelsen Institute, Fantoft, Norway, 1986.

[6] Å. Hallefjord, D. Haugland, and C. Helgesen. Optimeringsmodeller for olje- og gassutvinning. CMI-report 862325-1, Chr. Michelsen Institute, Fantoft, Norway, 1987.

[7] P. Hansen, E. de Luna Pedrosa Filho, and C. C. Ribeiro. Location and sizing of offshore platforms for oil exploration. *European Journal of Operational Research*, 58:202–214, 1992.

[8] D. Haugland, Å. Hallefjord, and H. Asheim. Models for petroleum field exploitation. *European Journal of Operational Research*, 37:58–72, 1988.

[9] D. Haugland, K. Jörnsten, and E. Shayan. Modelling petroleum fields with movable platforms. *Applied Mathematical Modelling*, 15:33–39, January 1991.

[10] T. W. Jonsbråten. *Optimeringsmodeller for utvinning av et oljereservoar*. HAS-thesis, Norwegian School of Economics and Business Administration, Bergen, 1994.

[11] H. Jönsson, K. O. Jörnsten, and E. A. Silver. Application of the scenario aggregation approach to a two-stage, stochastic, common component, inventory problem with a budget constraint. *European Journal of Operational Research*, 68:196–211, 1993.

[12] K. Jörnsten and M. Bjørndal. Dynamic location under uncertainty. *Studies in Regional and Urban Planning*, Issue 3:163–184, September 1994.

[13] K. O. Jörnsten. Sequencing offshore oil and gas fields under uncertainty. *European Journal of Operational Research*, 58:191–201, 1992.

[14] P. Kall and S. W. Wallace. *Stochastic Programming*. John Wiley & Sons, Chichester, 1994.

[15] L. Lasdon, P. E. Coffman, Jr., R. MacDonald, J. W. McFarland, and K. Seehrnoori. Optimal hydrocarbon reservoir production policies. *Operations Research*, 34(1):40–54, 1986.

[16] A. Løkketangen and D. L. Woodruff. Progressive hedging and tabu search applied to mixed integer (0,1) multi-stage stochastic programming. *To appear in Journal of Heuristics*, 1996.

[17] J. W. McFarland, L. Lasdon, and V. Loose. Development planning and management of reservoirs using tank models and nonlinear programming. *Operations Research*, 32(2):270–289, 1984.

[18] D. W. Peaceman. *Fundamentals of Numerical Reservoir Simulation*. Elsevier Scientific Publishing Company, Amsterdam, London and New York, 1977.

[19] D. W. Peaceman. Interpretation of well-block pressures in numerical reservoir simulation. *Society of Petroleum Engineers Journal*, June:183–194, 1978.

[20] R. T. Rockafellar and R. J.-B. Wets. Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of Operations Research*, 16(1):119–147, February 1991.

[21] G. W. Rosenwald. A method for determining the optimum location of wells in a reservoir using mixed-integer programming. *Society of Petroleum Engineers Journal*, February:44–54, 1974.

[22] R. A. Wattenbarger. Maximizing seasonal withdrawals from gas storage reservoirs. *Journal of Petroleum Technology*, August:994–998, 1970.

[23] R. J.-B. Wets. The aggregation principle in scenario analysis and stochastic optimization. In S. W. Wallace, editor, *Algorithms and Model Formulations in Mathematical Programming*, pages 91–113. Springer-Verlag, Berlin, 1989.

Paper B

A Class of Stochastic Programs with Decision Dependent Random Elements [1]

Abstract

In the 'standard' formulation of a stochastic program with recourse, the distribution of the random parameters is independent of the decisions. When this is not the case, the problem is significantly more difficult to solve. This paper identifies a class of problems that are 'manageable' and proposes an algorithmic procedure for solving problems of this type. We give bounds and algorithms for the case where the distributions and the variables controlling information discovery are discrete. Computational experience is reported.

[1] Co-authored with Roger J.-B. Wets and David L. Woodruff. Accepted for publication in *Annals of Operations Research*

B.1 Introduction

Models based on stochastic programming lend valuable solutions to many types of problems. In the 'standard' formulation of a stochastic program with recourse, the distribution of the random parameters is independent of the decisions. When this is not the case, the problem is significantly more difficult to solve. This paper deals with a class of such problems that are 'manageable' and proposes an algorithmic procedure for solving problems of this type.

Before getting down to specifics, the issues can best be laid out in terms of a 'simple,' but general, formulation of this class of stochastic optimization problems where the information that will be provided to the decision maker is decision dependent. In order to introduce our new class of problems in the context of the current literature, we first develop the following 'standard' stochastic programming problem:

$$\min_{x \in R^n} E\{f(\boldsymbol{\xi}; x)\} = Ef(x) \tag{B.1}$$

where $f : \Xi \times \mathbb{R}^n \to \overline{\mathbb{R}} = [-\infty, \infty]$ is the 'cost' associated with a decision x when the random variable $\boldsymbol{\xi}$ takes on the value ξ; $\boldsymbol{\xi}$ is a $\mathbb{R}^k$-valued random variable with possible values in $\Xi \subset \mathbb{R}^k$, which is the support of the distribution, μ, of the random variable; $Ef : \mathbb{R}^n \to \overline{\mathbb{R}}$, the function to be minimized, is defined by

$$Ef(x) = \int_{\Xi} f(\xi; x)\, \mu(d\xi).$$

One can recast two stage stochastic programs with recourse so that they are seen as special cases of the problem just formulated. Indeed, for the two stage recourse problem with stage t variable indexes in the set N^t:

$$\begin{array}{ll} \min\limits_{x \in C^1} & f_{10}(x) + E\{\min_{x^2 \in C^2} [\, f_{20}(\xi; x^2) \,\big|\, f_{2i}(\xi; x, x^2) \le 0, i \in N^2 \,] \\ \text{such that} & \qquad f_{1i}(x) \le 0, i \in N^1, \end{array}$$

and the function f is defined as follows:

$$f(\xi; x) = \begin{cases} f_{10}(x) + Q(\xi; x) & \text{if } x \in C^1,\ f_{1i}(x) \le 0 \text{ for } i \in N^1, \\ \infty & \text{otherwise} \end{cases}$$

where

$$Q(\xi; x) = \inf_{x^2 \in C^2} [\, f_{20}(\xi; x^2) \,\big|\, f_{2i}(\xi; x, x^2) \le 0, i \in N^2 \,].$$

Even multistage stochastic programs can be recast so as to fit the general formulation. For example, in the case of a three stage problem, redefine the function Q as follows:

$$Q(\xi;x) = \inf_{x^2 \in C^2} [\, f_{20}(\xi;x^2) + E\{Q_2(\xi,\xi^3;x,x^2)|\xi\} \,\big|\, f_{2i}(\xi;x,x^2) \le 0, i \in N^2 \,],$$

where ξ^3 stands for the observations made after choosing x^2 (but before selecting x^3), $E\{Q_2|\xi\}$ is the conditional expectation of Q_2 given ξ and

$$Q_2(\xi,\xi^3;x,x^2) = \inf_{x^3 \in C^3} [\, f_{30}(\xi,\xi^3;x^3) \,\big|\, f_{3i}(\xi,\xi^3;x,x^2,x^3) \le 0, i \in N^3 \,].$$

The same pattern is followed if there are more than three stages, i.e., in the objective defining Q_2, f_{30} is replaced by $f_{30} + Q_3$, and so on. This 'standard' formulation of the problem implicitly assumes that the measure μ is not affected by the decision x. This assumption, satisfied for all practical purposes by a very rich class of applications, makes possible the design of rather efficient solution procedures for 'real' size problems.

However, there are important decision making problems that do not fit in this mold, namely cases when the distribution of the random quantities will be affected by the decision selected. This can happen in many ways, but it seems the following formulation would cover all such cases:

$$\min E^{\mu} f(x) = \int_{\Xi} f(\xi;x)\, \mu(d\xi) \quad \text{such that} \quad (\mu,x) \in \mathcal{K} \subset M \times \mathbb{R}^n$$

where M is a subset of the probability measures on Ξ and $\mathcal{K}$ are the constraints linking the decision x to the choice of μ. In this formulation, the 'decision' x, affecting the level at which certain activities will be conducted, is viewed as disjoint from the choice of μ which will affect the information received and the time (stage) at which certain realizations of the random quantities will be observed. The constraints linking the choice, or inevitability, of μ to certain decisions x are modeled here through the linking constraints $(\mu,x) \in \mathcal{K}$.

On particular in the literature devoted to discrete event dynamic systems [10, 14, 16], the dependence of the probability measure on the decision(s) has often received the following formulation:

$$\min_{x \in \mathbb{R}^n} \int_{\Xi} f(\xi,x)\, \mu_x(d\xi).$$

In such a situation, the set $\mathcal{K}$ is the graph of the mapping $x \mapsto \mu_x$, i.e., $\mathcal{K} = \{(\mu,x) \,|\, x \in \mathbb{R}^n, \mu = \mu_x\}$; observe that our framework also allows for set-valued dependence between decisions and associated measures. In some other problems the measures might depend on certain parameters that are costly to evaluate and one can (should?) allocate some of the available resources to the estimation of these parameters, see e.g. [1]. If we then include among the decision variables those that will affect the estimation process, the problem is of the same type as here above and thus fits the general framework.

One advantage of our formulation is that it allows for a better classification of problems of this types based on the properties of the set $\mathcal{K}$ of the linking constraints and this should be helpful both in the design of computationally schemes and for purposes of analysis (as in Section B.6, for example). If there are no linking constraints, the problem is relatively simple. Since the 'cost' function is linear in μ, it is known that the solution will occur at a measure μ^{ext} that is an extreme point of the convex hull of the set M. And there are results about generalized moment problems [3, 15], and about optimization techniques for finding optimal submeasures [8, 9] that are useful for dealing with such cases.

But beyond this, i.e., when there are nontrivial constraints linking μ to x, the problem becomes *significantly* more difficult to analyze and solve, and, not surprisingly, the literature dealing with this class of problems is very sparse. In fact, we only know of one such paper dealing with a 'Markovian' case [14]; consult also [5]. The relationship between x and μ can be quite complex, and although one might be able to write down quite general optimality conditions, at this stage we are going to limit our attention to a class of problems that are 'manageable.' In our case the set M will be finite, and its elements can be indexed by a boolean vector $d = (d_1, \ldots, d_q)$, i.e., each $d_j \in \{0, 1\}$, which will indicate if certain options have or have not been selected. The problem can then be reformulated as follows:

$$\min E^d f(x) = \int_\Xi f(\xi; x)\, \mu^d(d\xi) \quad \text{such that} \quad (d, x) \in \mathcal{K} \subset \{0,1\}^q \times \mathbb{R}^n. \tag{B.2}$$

The class of problems covered by this formulation includes those when the choice of d, and consequently of x, will affect the time at which the information about the realizations of certain random elements will become available. In the case of multistage problems, such as those described in the next sections, the choice of d affects the 'timing' of the observation, i.e., at which stage certain random variables will be observed. In contrast with the x-part of the problem which might involve 'staged decisions' that depend on past realizations, the choice of d will be a *first stage* decision only. This is a definite restriction, but it simplifies notation greatly while preserving the important new concepts. The resulting algorithms make inroads into this important class of problems, but full generalization is left as future research.

The following simple example might be helpful in understanding the connection. Suppose a production line must meet the demand for a certain number of products: $P_1, \dots, P_q$, but if demand for a given product can't be satisfied, it's always possible to substitute a better one. Assuming that the products are linearly ordered, with P_q of the highest quality and P_1 of the lowest, it is thus possible to substitute P_k for P_j if $k > j$. However, there is quite a bit of uncertainty about the actual production costs of these different products, and consequently about the potential profits. Moreover, the production capacity available limits the choice to only a few of these items in the first stage. In the first stage, the information available about the production costs is 'probabilistic', i.e., all what is known about the random variable ξ_j is their joint distribution μ. The information available in the second stage is then the actual production costs of the items produced in the first stage, and for the remaining ones it remains 'probabilistic.' In this example, the decision vector x fixes the production levels for $P_1, \dots, P_q$. Each d_j is a boolean variable that indicates if P_j will be produced or not in the first stage, which in turn determines the information that will become available in the second stage.

The overall approach will be one of 'reduced minimization': the overall goal is to choose a 'best' stochastic optimization problem in a certain collection of such problems:

$$\left\{ \min_x E^d f(x) \,\middle|\, (d,x) \in \mathcal{K} \right\} \qquad d \in \{0,1\}^q$$

the one that yields the lowest possible value. Equivalently one could cast the problem as follows,

$$\min_{d \in \{0,1\}^q} \left\{ \min_x E^d f(x) \,\middle|\, (d,x) \in \mathcal{K} \right\}. \tag{B.3}$$

In the remainder of the paper, we are going to be interested in a class of problems where d identifies the time at which certain information is going to become available. More specifically, the production costs of certain items will be revealed either in stage two or stage three depending on d. The choice of a particular $d \in \{0,1\}^q$ determines

$$\mu^d \text{ and } \mathcal{K}^d = \{x \in I\!R^n \mid (\mu^d, x) \in \mathcal{K}\}$$

where this last set describes the corresponding set of feasible x-decisions.

In Section B.2 specific notation is introduced for (stochastic) linear models that is convenient when dealing with algorithmic issues. Bounds given in Section B.3 are employed in an implicit enumeration algorithm that is described in Section B.4. Computational experience is reported in Section B.5. Certain direction of descent are identified in Section B.6 by means of variational analysis and the paper concludes with a summary and directions for further research.

B.2 Linear Models

The abstract definitions given in the previous section are useful for indicating the scope of our new models and for analysis. However, for algorithmic development we need more specific notation. We begin by developing notation for stochastic linear programs, then we extend it to the case where there is decision-dependent information discovery. The bounds in Section B.3 and the implicit enumeration algorithm that we develop in Section B.4 do not rely explicitly on any assumptions of linearity, but linear problems provide a concrete basis for discussion and solvers are available for subproblems.

B.2.1 Stochastic Linear Programs

In order to give concrete definition to the class of problems for which computational results are obtained, we outline a 'standard' scenario-based, multistage, linear (perhaps integer or mixed integer) stochastic programming version of the problem given in expression (B.1). This formulation begins with the assumption that each random variable will be realized at a predetermined time and that scenarios are specified giving a full set of random variable realizations. For each scenario s, we are given the probability of occurrence of (or, more accurately, a realization "near") scenario s as $p(s)$. Each scenario s might actually correspond to the aggregation of a certain number of the possible realizations of the random parameters.

Decisions are made at times indexed by $1, \ldots, T$ and random variables in a scenario are realized at times $1, \ldots, T+1$. Our modelling convention is that information obtained up to time t is available for decisions at time t. Some information is available at time $t = 1$ before the first decision is selected, some information becomes available during the decision process, and some information may not become available until time $T + 1$ after all decisions have been made.

We assume that scenarios are given in the form of a *scenario tree* denoted usually by $\mathcal{S}$. Together with the probabilities p_s attached to its scenarios, a scenario tree completely specifies the stochastic process associated with the random quantities of the problem. The notion of a scenario tree is important for many methods of solving stochastic programs and the notion that new information becomes available only at certain times is central to the construction of 'implementable' solutions. Each (terminal) leaf identifies a particular scenario. The leaves are grouped for connection to nodes at time T. Each leaf is connected to exactly one time T node and each of these nodes represents a unique realization up to time T. The time T nodes are connected to time $T-1$ nodes so that for each scenario connected to the same node at time $T-1$ has the same realization up to time $T-1$. This continues back until information available "now" (at time index 1) constitutes the root node of the scenario tree. Variables that are observed at time t generate the *branches at time* $t-1$.

For each scenario s and each stage t we are given a vector $c^t(s)$ of length n^t, a $m^t \times n^t$ matrix $A^t(s)$ and a column vector $b^t(s)$ of length m^t. For notational convenience let $A(s)$ be $(A^1(s), \dots, A^T(s))$, let $b(s)$ be $(b^1(s), \dots, b^T(s))$, and let $c(s)$ be $(c^1(s), \dots, c^T(s))$. The decision variables are a set of n^t-vectors x^t; one vector for each scenario; notice that we reserve superscripts for the stage (or time) index. Notice also that the solution is allowed to depend on the scenario. Let $X(s) = (x^1(s), \dots, x^T(s))$. If we could know beforehand which scenario would actually occur, call it s, the problem would be solved by minimizing

$$f(s; X(s)) = \sum_{t=1}^{T} \langle c^t(s), x^t(s) \rangle$$

subject to

$$\begin{aligned} A(s)X(s) &\geq b(s), & \text{(B.4)} \\ x_i^t(s) &\in \{0,1\}, \quad i \in I^t,\ t = 1, \dots, T & \text{(B.5)} \\ x_i^t(s) &\geq 0, \quad i \in \{1, \dots, n^t\} \setminus I^t,\ t = 1, \dots, T & \text{(B.6)} \end{aligned}$$

where $AX \geq b$ includes the usual sorts of single period and period linking constraints that one typically finds in multistage linear programming formulations. We use I^t to identify the integer variables in each time stage. For most of the literature, this set is empty. However, approaches for some classes of stochastic programming problems with integer variables in the first stage have been discussed (see, e.g., [2]) or for classes of two stage problems with integers [12, 17, 18, 20]. Recently, there has been some work on various forms of multistage integer stochastic problems [4, 6, 13, 19].

Since humans are not prescient, we must require solutions that do not require foreknowledge and that will be feasible no matter which scenario is realized. We refer to solution systems that satisfy constraints with probability one as *admissible.* We refer to a system of solution vectors as *implementable* if for scenario pairs s and s' that are indistinguishable up to time t, it is true that $x^r(s) = x^r(s')$ for all $1 \leq r \leq t$. For a given scenario tree $\mathcal{S}$, the set of implementable solutions will be denoted by $\mathcal{N}_\mathcal{S}$; one writes $E^\mathcal{S}$ to indicate that expectation is with respect to the measure associated with the scenario tree $\mathcal{S}$. This brings us to the following formulation for stochastic linear programs:

$$\min \sum_{s \in \mathcal{S}} [p(s) f(s; X(s))] = E^\mathcal{S}\{f(\cdot, X(\cdot))\}$$

subject to

$$\begin{aligned} A(s)X(s) &\geq b(s), \quad s \in \mathcal{S} && \text{(B.7)} \\ x_i^t(s) &\in \{0,1\}, \quad i \in I^t,\ t = 1, \ldots, T,\ s \in \mathcal{S} && \text{(B.8)} \\ x_i^t(s) &\geq 0 \ \ i \in \{1, \ldots, n^t\} \setminus I^t,\ t = 1, \ldots, T,\ s \in \mathcal{S} && \text{(B.9)} \\ X(\cdot) &\in \mathcal{N}_\mathcal{S}. && \text{(B.10)} \end{aligned}$$

The expectation here can be expressed as a finite sum since we are dealing with only a finite number of scenarios (in $\mathcal{S}$).

B.2.2 Discoveries as a Result of Decisions

In Section B.1 the 'standard' stochastic programming model was extended to allow for situations when the probability measure is decision dependent. The problems that we are going to consider are stochastic linear programs that fit into the class of problems identified by the formulation in (B.3). In such models the decisions have been split into a (boolean) vector d specifying a probability measure μ^d, and the usual vector X that determines the activity levels; the relationship between these decision being expressed through the constraint $(d, X) \in \mathcal{K}$.

In our formulation of (multistage) stochastic linear programs, the probability measure of the random elements was identified with a scenario tree $\mathcal{S}$ with probabilities p_s attached to each individual scenario $s \in \mathcal{S}$. Since each d corresponds to a different measure, this now translates into (different) scenario trees $\mathcal{S}(d)$, indexed by d with $p(d; s)$ as the probabilities attached to the scenarios s of $\mathcal{S}(d)$. Similarly, the constraints of the stochastic program will depend on the choice of d. For $d \in \{0,1\}^q$, let

$$\mathcal{K}(d) = \{X \mid (d,X) \in \mathcal{K}\}.$$

To stay in the linear framework, we have to assume that for all d, the set $\mathcal{K}(d)$ is itself determined by a finite number of linear equations and linear inequalities, i.e., is a convex polyhedral set. Thus to each choice of d corresponds a (standard) stochastic linear program of the form:

$$\min \sum_{s \in \mathcal{S}(d)} [p(d;s) f(s; X(s))] = E^d\{f(\cdot, X(\cdot))\}$$

subject to

$$\begin{aligned}
A(s)X(s) &\geq b(s), \quad s \in \mathcal{S}(d) && \text{(B.11)}\\
x_i^t(s) &\in \{0,1\}, \quad i \in I^t,\ t = 1,\ldots,T, \quad s \in \mathcal{S}(d) && \text{(B.12)}\\
x_i^t(s) &\geq 0 \;\; i \in \{1,\ldots,n^t\} \setminus I^t,\ t = 1,\ldots,T, \quad s \in \mathcal{S}(d) && \text{(B.13)}\\
X &\in \mathcal{K}(d) && \text{(B.14)}\\
X(\cdot) &\in \mathcal{N}(d) && \text{(B.15)}
\end{aligned}$$

using here the shorthand notation $\mathcal{N}(d)$ to denote $\mathcal{N}_{\mathcal{S}(d)}$. Since implementability depends on the timing at which information will become available, the constraint imposing implementability must also be formulated in terms of the scenario tree $\mathcal{S}(d)$. As already indicated in Section B.1, in the class of problems we have chosen to illustrate our approach, the vectors d —recall that they are first stage decisions— specify at which time information about the values taken by certain random variables becomes available.

B.2.3 Example

In the introduction we illustrated our notation with a small example that is an extension of the model given by Jorjani et al. [11] for the optimal selection of subsets of sizes under demand uncertainty. We continue that example with more detail here. A product is available in a finite number of sizes, and demand for a smaller size can be met by substituting a larger size. However, this substitution comes at a certain cost. The objective is to minimize the expected cost of satisfying the demand for all different sizes. Binary variables model the selection of sizes, while production and substitution quantities are represented by continuous variables. Details for the formulation are summarized in Appendix A.

In addition to the demand uncertainty, we introduce uncertainty regarding the production costs. When we choose to produce a size, we will learn about its production costs. So q matches the number of sizes to be produced and vectors in $\{0, 1\}^q$ have an element for each size and correspond directly to the decision variable that indicates production of the size. In this way a production decision can also be viewed as an investment in information as well as production. In Appendix A, we present the data for an instance with two periods and three different sizes. Because the demand has to be met, we know that the largest size always has to be produced, and in the problem considered here the cost of producing the largest size is deterministic. If we neglect the cost uncertainty and use the expected production costs for the two smallest sizes, the optimal first-stage decision is to produce either size 1 or size 2 (size 3 is always produced). Both these decisions have a total cost of 25140.

When uncertain production costs are introduced, we get different results. The four feasible first-stage production decisions and associated total costs are reported in Table B.1. The vector d gives the production decision for the sizes 1 and 2 where "0" represents not produce while "1" is produce. We

d	Minimum $E^d\{f(\cdot, X(\cdot))\}$
0, 0	\$25180
1, 0	25130
0, 1	25115
1, 1	25220

Table B.1: Production costs, Fixed first stage decisions

see that by producing size 2 (and 3), the expected cost has been reduced to 25115. Even though the reduction in total cost is not large, it is interesting to see that the optimal production policy has been dramatically altered by introducing the cost uncertainty in the problem.

B.3 Bounds

In this section we develop bounds to be used in an implicit enumeration algorithm. There is no universal scenario tree attached to a stochastic (linear) program when the distribution of (some of) the random parameters is decision dependent. However, once a timing of discovery for the values realized by the random variables has been specified (or assumed) a scenario tree can be drawn. Armed with a scenario tree, we can search for solutions for the associated optimization problem and use the result as a bound of some sort.

The bounds that we derive are based on the use of *branching* (or *partial*) vectors, denoted $d^{\#}$, in $\{0,1,\#\}^q$. The components $d_j^{\#}$ of such vectors take on the values 0, 1 or #. The interpretation we attach to these values is as follows:

$d_j^{\#} = 1$ means that information about the j-th (family of) random variable(s) will come as early as possible (here in stage two);

$d_j^{\#} = 0$ means that information about the j-th (family of) random variable(s) will come as late as possible (in our examples, this will be stage three);

$d_j^{\#} = \#$ means that there has been no decision yet about this particular variable.

We refer to vector elements with a 0 or 1 as *determined* and those with # as *undetermined*. To each branching vector $d^{\#}$ is associated a collection, say $\mathcal{D}(d^{\#})$, of decision vectors d obtained by replacing all # entries in $d^{\#}$ by either 1 or 0; these correspond one-to-one with a set of 'standard' stochastic optimization problems that we will call $\mathcal{P}(d^{\#})$. Within the collection $\mathcal{D}(d^{\#})$, let's denote by $(d^{\#})^e$ the vector in $\{0,1\}^q$ obtained by replacing every (undetermined) entry # in $d^{\#}$ by 1, and by $(d^{\#})^l$ the vector obtained by replacing every (undetermined) entry # in $d^{\#}$ by 0. The two vectors correspond respectively to the cases when information associated with the variables d_j for which $d_j^{\#} = \#$ is received as early as possible (the index e stands for early) or late (the index l stands for late). Because discovery at a later time imposes more restrictions on implementable solutions, one has

$$\mathcal{N}((d^{\#})^e) \subseteq \mathcal{N}(d) \subseteq \mathcal{N}((d^{\#})^l) \quad \forall\, d \in \mathcal{D}(d^{\#});$$

recall that $\mathcal{N}(d)$ is shorthand for $\mathcal{N}_{S(d)}$. In the calculation of bounds for the collection $\mathcal{P}(d^{\#})$ of stochastic programs associated with $d \in \mathcal{D}(d^{\#})$, we want to impose the constraints on X from among those generated by the restriction $(d, X) \in \mathcal{K}$ only if they come from the already determined components of $d^{\#}$. So, we let

$$\mathcal{K}(d^{\#}) = \bigcup\nolimits_{d \in \mathcal{D}(d^{\#})} \mathcal{K}(d).$$

We shall only deal with problems where this set is a nonempty convex polyhedral set. So, the linear program,

$$\min \sum_{s \in S^e} [p^e(s) f(s; X(s))]$$

subject to

$$
\begin{aligned}
A(s)X(s) &\geq b(s), \quad s \in \mathcal{S}^e && \text{(B.16)}\\
x_i^t(s) &\in \{0,1\}, \quad i \in I^t,\ t = 1,\dots,T,\ \ s \in \mathcal{S}^e && \text{(B.17)}\\
x_i^t(s) &\geq 0 \ \ i \in \{1,\dots,n^t\} \setminus I^t,\ t = 1,\dots,T,\ \ s \in \mathcal{S}^e && \text{(B.18)}\\
X &\in \mathcal{K}(d^\#) && \text{(B.19)}\\
X(\cdot) &\in \mathcal{N}^e && \text{(B.20)}
\end{aligned}
$$

provides a lower bound for the values of all stochastic programs in $\mathcal{P}(d^\#)$. Here $\mathcal{S}^e$ is used as shorthand for $\mathcal{S}((d^\#)^e)$. Let us denote this *lower bound* by $L(d^\#)$. From this it follows that a lower bound to $L(d^\#)$ also will be a valid lower bound to all stochastic programs in $\mathcal{P}(d^\#)$. This is particularly relevant for integer and mixed integer problems where lower bounds are often readily available via relaxation even when exact minimization might be quite difficult.

If the superscripts e in the linear program above is replaced with superscripts l (late branching), we get a program that provides an *upper bound*, $U(d^\#)$, for the values of all stochastic programs in $\mathcal{P}(d^\#)$ provided that certain restrictions hold on the structure of $\mathcal{K}(d)$. This fact is not exploited in our algorithms because they use full vectors when calculating upper bounds. An upper bound computed by a full decision vector, $U(d)$, is globally valid, at least as tight and requires no greater computational effort.

B.3.1 Example

To illustrate the notation, we continue with the example of Section B.2.3. Upper and lower bounds for branching vectors are displayed in Table B.2. The two vector elements shown correspond to fixing the decisions to produce or not produce the two parts that have an uncertain, but discoverable, cost. We can see that using the best upper bound given in the table we are able to fathom the partial vector (#,0). The upper bounds $U(d^\#)$ are valid because $\mathcal{K}(\cdot)$ is completely separable. When comparing to the results in Table B.1, we see that the upper bound for the vector (#,1) is the same as the minimum cost reported for the full vector (0,1), the optimal d-vector for this problem.

B.4 An Algorithm

The bounds introduced in the preceding discussion can be directly used to create a branch and bound algorithm. In this section we introduce a more

$d^\#$	$U(d^\#)$	$L(d^\#)$
0, #	25140	25080
1, #	25130	25056
#, 0	25140	25130
#, 1	25115	25095

Table B.2: Bounds for the example given in Section B.2.3

memory intensive implicit enumeration algorithm that relies on an ability to store information about each full vector in $\{0,1\}^q$. The algorithm is developed only for situations where there are two decision stages (three realization times). This algorithm is clearly not appropriate for large q, but given present (and projected near term) computer capabilities, exact solutions for such problems will require further developments. The advantage of retaining information in this fashion is that we are able to make good branching decisions and we are able to quickly reduce the search space for good upper bounds.

In order to compress the space required to present the algorithms, we make use of the following notation:

> $\leftarrow$ means assignment. For example, $j \leftarrow j+1$ means that the value of j is incremented.
>
> $d^\#_j \leftarrow i$ means that branching vector $d^\#$ is modified so that element j takes the value i.
>
> $(d^\#_j = i)$ refers to a branching vector for which element j is equal to i and all other elements are undetermined.
>
> $d^\# \subseteq d$ is true if the determined values in $d^\#$ match the corresponding vector elements in d. For example, $(\#,1,\#) \subseteq (0,1,1)$, but $(\#,1,0) \not\subseteq (0,1,1)$.

A conceptual version of the algorithm called I1 is shown in Figure B.2. This is not written to convey the most efficient computer implementation, but to display the concepts. Necessary initialization of variables are done in step 1. The list of lower bounds, $\mathcal{L}[d]$ has all elements set to zero, the branching vector $d^\#$ is assigned undetermined vector elements, the set D^* is assigned the set of all possible d-vectors, J and J' are index sets with q elements and B is an index set for branching decisions. In step 2 the algorithm calculates lower bounds for all possible branching vectors where only one decision is fixed and the others are undetermined. These calculated bounds are incorporated in the list $\mathcal{L}[d]$, that represents the highest lower bound for each full vector, $d \in \{0,1\}^q$, which serves as the index set for the list. The list V is of length q and is used in step 5 for deciding branching order. In step 3 the full vector with lowest lower bound, argmin $\mathcal{L}[d]$, is used for computing a global upper bound. Thus $\mathcal{U}$ represents the best feasible solution found so far to the overall minimization problem. This upper bound is in step 4 used to prune full vectors for which the best lower bound is higher than the upper bound. The set D^* thus represents all full vectors in $\{0,1\}^q$ that have not yet been bounded out.

Step 5 essentially repeats these steps in a slightly more general way to consider branching with a smaller number of undetermined components. While the number of possible full vectors is greater than one or we have not branched on all possible decisions (in the case with non-unique optimal solution), the algorithm continues the branching. The remaining elements in D^* are investigated, and decisions that are identical for all vectors in D^* are fixed. The determined elements of the branching vector $d^\#$, are the decisions proven to be a part of the optimal full vector. The set B represents the variables for which the algorithm branches, and indexes for fixed decisions are removed from B. The index set J' represents all decisions not yet fixed. The order in which the branching is performed (i.e., assignment of j' to B) is determined heuristically, using V_j to estimate the relative efficiency of variable changes. The set B gives the list of indexes for which full factorial branching is to be done. The mechanics of this branching is described by the function Ψ as shown in Figure B.1. The function Ψ returns a set of $2^{|B|}$ branching vectors, and this set represents all possible branching vectors that are undetermined for the elements $(J' \setminus B)$, and determined for the fixed decisions and the branching variables. A lower bound is calculated for the branching vectors $\hat{d}^\#$ and the list $\mathcal{L}$ is updated. An upper bound is calculated for the full vector with lowest lower bound, and all decision vectors, d, for which the lower bound is higher than the global upper bound, is removed from the set D^*. Algorithm I1 terminates when the set D^* has just one element, the optimal full vector, or we have a non-unique optimal solution (i.e., the V's are all negative). Upon termination, D^* must be the optimal set.

$\nu \leftarrow \emptyset$

FOREACH $i \in \{0, \ldots, (2^{|B|} - 1)\}$ (i have a binary representation)

$\tilde{d}^{\#} \leftarrow d^{\#}$

$j \leftarrow 1$

FOREACH $b \in B$

$\tilde{d}_b^{\#} \leftarrow i_j$ (i_j is the j-th digit in i)

$j \longleftarrow j + 1$

$\nu \leftarrow \nu \cup \{\tilde{d}^{\#}\}$

RETURN ν

Figure B.1: Definition of the function $\Psi(d^{\#}, B)$

B.5 Computational Experience

B.5.1 The Subcontracting Problem

The problem of selecting an optimal subset of sizes that we have used as an example has limited usefulness for computational experiments because we can solve only small instances to optimality. This is due to the fact that the problem has both integer and continuous variables and has complicated constraints. In this section we introduce a problem for which we develop a continuous and a pure integer version.

In the *subcontracting problem* we consider a situation where a manufacturer must use a number of different components in the production of some new items. The components may be produced "in house" or purchased from a foreign subcontractor. The subcontractor offers these components at a given price, but in the future it is possible that an import tax will be added to the price. Hence, the subcontract cost is uncertain and the timing of realization is not influenced by the values of the decision variables in this model. Since the manufacturer has no cost history for these new components the in-house production costs are uncertain; however, prior probabilities can be given. The uncertainty for a given component can be resolved by producing that component, or a similar one. Each of the components belongs to some *family* of components, with family membership defined by the property that the discovery of cost information is shared within a family, but not between

1. $\mathcal{L} \leftarrow 0$; $d^{\#} \leftarrow \#$; $D^{*} \leftarrow \{0,1\}^{q}$; $J \leftarrow J' \leftarrow \{1,\ldots,q\}$; $B \leftarrow \emptyset$

2. ForEach $j \in J$

    ForEach $i \in \{0,1\}$

        ForEach $d \in \{0,1\}^{q}$

            If [$((d^{\#}_{j} = i) \subset d)$ and $(\mathcal{L}[d] < L(d^{\#}_{j} = i))$]

                $\mathcal{L}[d] \leftarrow L(d^{\#}_{j} = i)$

    $V_{j} \leftarrow abs(L(d^{\#}_{j} = 0) - L(d^{\#}_{j} = 1))$

3. $\mathcal{U} \leftarrow U(\text{argmin}\,\mathcal{L}[d])$

4. ForEach $d \in \{0,1\}^{q}$

    If $\mathcal{U} < \mathcal{L}[d]$

        $D^{*} \leftarrow D^{*} \setminus \{d\}$

5. While [$(|D^{*}| > 1)$ and $((\max V_{j}) \geq 0)$]

    ForEach $j \in J'$

        ForEach $i \in \{0,1\}$

            If $[(d^{\#}_{j} = i) \subset d \;\forall\; d \in D^{*}]$

                $d^{\#}_{j} \leftarrow i$

                If $(j \subset B)$ $B \leftarrow B \setminus \{j\}$

                $J' \leftarrow J' \setminus \{j\}$

    $\{j'\} = \text{argmax}\, V_{j}$

    $B \leftarrow \{j'\}$

    $V_{j'} = -1$

    ForEach $\hat{d}^{\#} \in \Psi(d^{\#}, B)$

        ForEach $d \in D^{*}$

            If [$(\hat{d}^{\#} \subseteq d)$ and $(\mathcal{L}[d] < L(\hat{d}^{\#}))$]

                $\mathcal{L}[d] \leftarrow L(\hat{d}^{\#})$

    $\mathcal{U} \leftarrow \min(\mathcal{U}, U(\text{argmin}\,\mathcal{L}[d]))$

    ForEach $d \in D^{*}$

        If $\mathcal{U} < \mathcal{L}[d]$

            $D^{*} \leftarrow D^{*} \setminus \{d\}$

Figure B.2: Definition of Algorithm I1.

families. We present two models for this problem: one continuous and the other discrete.

A Continuous Problem

To model a learning curve for costs in a very simple way, we assume that if there is a decision to produce at least a certain threshold fraction, α_f, the in-house production costs for family f will be known. There are n different components, and we define e to be a vector of length n whose elements are all 1. Further e_f is a vector of the same length whose elements are 1 for those indexes that are members of family f and 0 for all other elements. The in-house production is constrained by the availabilty of production capacity, and there are m constrained resources (B.21). There is also a limit on maximum increase in in-house production from one period to the next (B.22). We also include constraints to enforce $X \in \mathcal{K}(d)$ (B.24). The problem we want to optimize can formally be written as:

$$\min_{d\in\{0,1\}^q} \min_X E^d \left\{ \sum_{t=1}^{T} \left[\langle c(s), x^t(s)\rangle + \langle h(s), (e - x^t(s))\rangle \right] \right\},$$

subject to:

$$\begin{aligned}
A\left(e - x^t(s)\right) &\leq b, \quad t = 1,\ldots,T,\ s \in \mathcal{S}(d) && \text{(B.21)}\\
\langle e, x^t(s) - x^{t+1}(s)\rangle &\leq \beta \quad t = 1,\ldots,T-1, s \in \mathcal{S}(d) && \text{(B.22)}\\
0 \leq x_j^t(s) &\leq 1, \quad j = 1,\ldots,n,\ t = 1,\ldots,T,\ s \in \mathcal{S}(d) && \text{(B.23)}\\
\ell_f(d) \leq \langle e_f, (e - x^1(s))\rangle &< u_f(d), \quad f = 1,\ldots,q,\ s \in \mathcal{S}(d) && \text{(B.24)}\\
X &\in \mathcal{N}(d) && \text{(B.25)}
\end{aligned}$$

where we have introduced the following notation:

$1,\ldots,n$	the set of component indexes
$1,\ldots,q$	the set of family indexes
$1,\ldots,m$	the set of resource indexes
$x_j^t(s)$	quantity (a decision) of component j to subcontract in time t under scenario s
$(1 - x_j^t(s))$	quantity (a decision) of component j to produce in-house
$c_j(s)$	the uncertain cost of subcontracting component j
$h_j(s)$	the uncertain, discoverable cost of producing component j
a_{ij}	capacity requirement for component j on resource i
A	a_{ij} matrix
$b_1,\ldots,b_m$	available production capacities
β	maximum increase in in-house production
$[\ell_f(d), u_f(d))$	equals $[0,\alpha_f)$ or $[\alpha_f,\infty)$ if $d_f = 0$ or $= 1$, respectively

f	j	c'_j	c''_j	h'_j	h''_j	a_{1j}	a_{2j}	a_{3j}
1	1	5	6.25	4.8	5.5	1.8	2.0	2.0
	2	6	7.5	5.8	6.5	1.8	1.7	2.0
	3	7	8.75	6.8	7.5	2.0	1.5	1.5
2	4	8	10	7	9.5	1.0	1.3	1.7
	5	9	6.25	4.8	5.5	1.2	1.1	1.1
3	6	8	10	6	10.5	1.0	2.0	1.5
	7	9	11.25	7.5	11	1.2	1.5	1.5
4	8	10	12.5	8.5	12	1.2	1.4	1.9
	9	12	15	11	13.5	1.2	1.7	1.0
5	10	10	12.5	8.5	12	1.2	1.2	2.0
	11	12	15	10.5	14	1.5	1.6	2.0
	12	14	17.5	12	16.5	2.0	1.8	1.4
6	13	6	7.5	5	8	1.2	1.4	1.6
	14	7	8.75	6	8	1.8	1.7	1.3
7	15	8	10	7	10	1.2	1.2	2.0
	16	9	11.25	8	11	1.5	1.6	1.7
	17	10	12.5	9	12	2.0	2.0	1.4

Table B.3: Input data, subcontracting, continuous problem

We consider a problem with two decision stages ($T = 2$). Each element of a vector d is a one if there will be enough in-house production (during the first stage) in the family to learn the cost and zero otherwise. To create the lower bounding problem associated with a branching vector $d^{\#}$, the constraints (B.24) for $X \in \mathcal{K}(d)$ are replaced by:

$$0 \leq \langle e_f, (e - x^1(s)) \rangle < \alpha_f, \quad \text{if } d^{\#}_f = 0, \quad s \in \mathcal{S}(d) \qquad \text{(B.26)}$$

$$\alpha_f \leq \langle e_f, (e - x^1(s)) \rangle, \quad \text{if } d^{\#}_f = 1, \quad s \in \mathcal{S}(d). \qquad \text{(B.27)}$$

Note that there are no constraints associated with the elements of $d^{\#}$ that are 'undetermined', i.e., for those families f for which $d^{\#}_f = \#$ the interval $[\ell_f(d), u_f(d))$ equals $[0, \infty)$.

A Problem with 5 Families and 12 Components

We consider a problem with 12 components that belong to 5 different families, and $\alpha_f = 0.5$ for all $f \in 1, \ldots, q$. The uncertain costs can be found in the first part of Table B.3. The fact that the scenario set depends on the decision values, and furthermore the fact that it is the cross product of the

c and h space, makes it difficult to index the scenarios. We have used prime and double prime to distinguish the (equally likely, in this example) possibilities in each space. The in-house capacity is 9 units of each resource and each component's resource requirements are listed in the $a_{j,k}$-columns in Table B.3.

When ignoring the decision dependent information discovery and using the expected costs for the in-house production, we find the minimum expected total cost to be 226.53. In the first period components 1 and 3, and 2% of component 2 are produced in-house. In other words, all components in the families 2, 3, 4 and 5 are subcontracted.

When the uncertain production costs are included in the model, we find it optimal to produce one member from all of the families in the first period. The expected cost of this decision is 222.41. By taking the uncertain production costs into consideration, the optimal first period decisions have changed dramatically. When solving this problem by use of the I1 algorithm, it was necessary to solve 12 subproblems before the provable optimal solution was found. Complete enumeration would have required solution of 32 subproblems.

A Problem with 7 Families and 17 Components

In this example we consider 17 components belonging to 7 different families. The cost coefficients and capacity requirements are those listed in Table B.3, and there are 9 units of capacity available of each resource. When solving the deterministic problem with expected cost coefficients, we find it optimal to produce all of components 3 and 14, and 2 % of component 1 in-house. This decision policy has an expected cost of 311.38. When the uncertain production costs are included in the model, the optimal policy is to produce a member of all but the first family in the first period. This gives an expected cost of 305.81. By using the I1 algorithm, the optimal solution is found after solving 26 subproblems. Complete enumeration would have required solution of 128 subproblems.

B.5.2 An Integer Problem

In this example we let the production decisions be represented by integer variables, and as soon as one of the family members is produced the production costs for all of the family members are known. The formal model is the same as for the continuous case except that the learning threshold value

$\alpha_f = 1$ for all f, and the constraint (B.23) is written as

$$x_j(s,t) \in \{0,1\}, \quad j \in J, t \in 1,\ldots,T, s \in S(X). \qquad \text{(B.28)}$$

A Problem with 5 Families and 12 Components

Also here we first look at a problem with 12 components that belong to 5 different families. The uncertain costs can be found in the first part of Table B.3. There is one in-house resource and it has a capacity of 8 units; utilizations (b_j) are 2 for parts in families 1 and 5 and 1 for families 2,3, and 4. When using the expected costs for the in-house production, we find an optimal solution with an expected cost of 226.32, and in the first period, all members of the first family are produced in-house, but all others are subcontracted.

When the uncertain production costs are included in the model, the optimal policy is to produce one member from each of the families 2, 3, 4 and 5 in the first period in-house and subcontract all members of family 1. The expected cost of this decision is 222.41. The optimal first period decisions have been altered considerably. None of the components now produced would have been produced if production uncertainty was not included in the model. When solving this problem by use of the I1 algorithm, it was necessary to solve 16 subproblems before the provable optimal solution was found (a 50% savings over enumeration.)

A Problem with 7 Families and 17 Components

When considering 17 components belonging to 7 different families we use the same cost coefficients as used in Section B.5.1. All components require 1 unit of capacity, and the total capacity is 8 units. Also in this example we find it optimal to produce only the first family, if only the expected in-house costs are used. This decision policy has an expected cost of 309.58. When the uncertain production costs are included in the model, the optimal policy is to produce a member from all but the first family. This gives an expected cost of 303.58. By using the I1 algorithm, the optimal solution is found after solving 30 subproblems, which is a large saving over enumeration.

Harder Capacity Constraints

The problem we are considering here is rather similar to the integer problem above with 5 families and 12 components, but here there is a second capacity constraint for a resource with 9 units of capacity. For this resource the utilizations are 1 for parts in families 1, 2 and 2 for parts in families 3,4, and 5. This instance is just beyond the envelope of solution using our implementation that solves subproblems using CPLEX version 3 [7] (the subproblems also could not be solved using version 4). Most of the subproblems generated could not be solved to optimality within a reasonable amount of time so the lower bounds were generated using lower bounds given by the relaxations solved by the branch and bound algorithm used to solve the integer subproblems. At termination algorithm I1 was able to bound out all decisions except in-house production of all five families or in-house production of all families except family 1.

B.6 Variational Analysis

Thus far we have introduced a new modelling concept and outlined methods for finding exact solutions to instances of small to moderate size. In order to enhance our understanding of the model and to facilitate future work on heuristics for it, we describe methods for estimating the effects of perturbations of an optimal d vector, d^*.

It is easier at this point to return to the 'abstract' formulation of the problem featured in the Introduction. More precisely,

$$\min_{x,d} E^d f(x) = \int_\Xi f(\xi; x)\, \mu^d(d\xi) \text{ such that } (\mu^d, x) \in \mathcal{K} \subset M \times \mathbb{R}^n,$$

where $d = (d_1, \ldots, d_q) \in \{0,1\}^q$ is a boolean vector identifying certain options fixing in the process the probability measure μ^d, and $\mathcal{K}$ determines the constraints linking the decision x to the choice of μ. For a given d, or equivalently μ^d, let

$$x^d \in \operatorname{argmin}\{ E^d f(x) \mid x \in \mathcal{K}^d \}$$

where $\mathcal{K}^d = \{x \mid (\mu^d, x) \in \mathcal{K}\}$. Without going through a detailed analysis of the structure of $\mathcal{K}$ it is not possible to obtain computationally useful optimality conditions for x^d that would identify x^d as the optimal solution of the overall problem. Our goal here will be much more limited, viz. to state necessary optimality conditions that can also be used to pass from x^d to a potentially better option/solution combination. Let's begin with the following observation.

Lemma 1

Let M^0 be the space of (nonnegative) measures defined on Ξ, and suppose that for all $\mu \in M^0$, the function

$$\mu \mapsto E^\mu f = \int_\Xi f(\xi;\cdot)\,\mu(d\xi)$$

is well-defined; set $E^\mu f(x) = \infty$ whenever the function $\xi \mapsto f(\xi;x)$ is not bounded above by a μ-summable function. Then $\mu \mapsto E^\mu f$ is linear.

This is an immediate consequence of the properties of the integral. We are going to exploit this as follows: Let M as before denote the space of probability measure on Ξ. This set M is convex, i.e., given any pair of probability measures μ^0, μ^1 and $\lambda \in [0,1]$, one has

$$\mu^\lambda := (1-\lambda)\mu^0 + \lambda\mu^1 \in M$$

and

$$E^{\mu^\lambda} f \text{ is well-defined.}$$

This will allow us to compute a directional derivative of $E^\mu f$ at a point μ^d in a direction $\mu^c - \mu^d$ where μ^c is another probability measure. It is convenient to introduce the following notation:

$$F(\mu, x) := E^\mu f(x)$$

and, with $\mu^\lambda = (1-\lambda)\mu^0 + \lambda\mu^1$,

$$dF(\mu^0, x)(\mu^1 - \mu^0) = \liminf_{\lambda \searrow 0} \lambda^{-1} F(\mu^\lambda, x)$$

as the '*directional derivative*' of F at (μ^0, x) in direction $\mu^1 - \mu^0$. This directional derivative reflects the incremental change in the optimal value of the problem if rather than working with μ^0 we are going to let μ^1 dictate the choice of an optimal x. This (sub) derivative is computed for a fixed x.

Usually it is not too difficult to compute this directional derivative. Indeed, under the integrability conditions specified in Lemma 1,

$$dF(\mu^0, x)(\mu^1 - \mu^0) = \int_\Xi \lim_{\lambda \searrow 0} f(\xi;x)\lambda^{-1}[\,(1-\lambda)\mu^0(d\xi) + \lambda\mu^1(d\xi)\,].$$

If μ^0, μ^1 are absolutely continuous, i.e., one can associate with μ^0, μ^1 density functions h^0, h^1 defined on Ξ, then

$$dF(\mu^0, x)(\mu^1 - \mu^0) = \int_\Xi f(\xi;x)(h^1(\xi) - h^0(\xi))\,d\xi.$$

On the other hand, if μ^0, μ^1 are discretely distributed, say $\Xi = \{\xi^\ell, \ell = 1, \dots\}$ and $\mu^0(\xi^\ell) = p_\ell^0$, $\mu^1(\xi^\ell) = p_\ell^1$, then

$$dF(\mu^0, x)(\mu^1 - \mu^0) = \sum_\ell f(\xi^\ell; x)(p_\ell^1 - p_\ell^0).$$

Thus, if $x^d \in \operatorname{argmin} E^d f$ on $\mathcal{K}^d$ and

$$\text{for all } c \in \{0,1\}^q : \quad dF(\mu^d, x^d)(\mu^c - \mu^d) \geq 0$$

it follows from the linearity of $\mu \mapsto E^\mu f$ that x^d is locally an optimal solution. On the other hand, if

$$\text{for some } c \in \{0,1\}^q : \quad dF(\mu^d, x^d)(\mu^c - \mu^d) < 0$$

there is a potential decrease that could result from going from option 'd' to option 'c'. Note however that this cost reduction might not be realizable because the constraints $(\mu, x) \in \mathcal{K}$ might actually exclude (μ^c, x^d) from the feasible set. Nonetheless, the calculation of the directional derivative suggests directions of descent, and thus can be exploited algorithmically.

B.7 Conclusions and Directions for Further Research

In this paper we have extended the range of models considered by researchers in stochastic programming by explicitly recognizing that a scenario specifies not only the realized values of random variables, but also that the timing of the realization and the timing of information discovery can be influenced by decisions. We have proposed bounds and algorithms for the case where the distributions and the variables controlling information discovery are all zero-one and only affect first stage decisions. We then illustrated that these algorithms can be used to solve instances of integer, mixed integer, and continuous problems of moderate size. The instances also demonstrated using three different examples the intuitively obvious idea that inclusion of the information discovery effects of decisions can have a dramatic qualitative effect on the optimal decision.

Development of good heuristics will improve the performance of exact algorithms such as I1 by providing better upper bounds. Also, heuristics must be used in order to attack larger instances. This is particularly true for integer and mixed integer problems. For the problem introduced in Section

B.2.3 we are not able to solve realistic sized instances to optimality, and for the integer version of the subcontracting problem introduced in Section B.5, only small to moderate sized instances can be solved. The development of heuristics is left as future research, but we have provided assistance in the form of variational analysis that can be used to guide the search.

In terms of applications, one can quickly imagine many possibilities in addition to the production planning examples that we have given here. The abstract model, solution methods, and variational analysis open up these possibilities, which we leave as our primary contribution to modelling stochastic programs.

Acknowledgments

This work was supported in part by grants from the National Science Foundation (Division Mathematical Sciences) and the Norwegian Research Council.

Appendix A — The Sizes Problem

Complete details of the problem and larger instances are provided by Jorjani et al. [11]; here we give a summary and the data that we used. The notation from that paper is retained to the extent possible.

Single Period Deterministic Model

Suppose a product is available in a finite number N of sizes where 1 is the smallest size and N is the largest size. Further, suppose size i is substitutable for size j if $i > j$, i.e., larger sizes may fulfill demand for a smaller size. Let D_i, $i = 1, \ldots, N$ denote the demand for size i.

We introduce a cost structure for the production of the product. Let p_i, $i = 1, \ldots, N$ be the unit production cost for size i. Generally $p_i > p_j$ for $i > j$. Let σ be the set up cost for producing units of any size and ρ be the unit penalty cost of meeting demand for size j with a larger size i. We need the following decision variables:

$$
\begin{aligned}
z_i &= \begin{cases} 1 & \text{if we produce size } i \\ 0 & \text{otherwise} \end{cases} \\
y_i &= \text{number of units of size } i \text{ produced} \\
x_{ij} &= \text{number of units of size } i \text{ cut to meet demand for size } j, \ j < i.
\end{aligned}
$$

Hence, in order to find the optimal subset of the N sizes to produce so as to satisfy demand, we solve the following integer linear program.

$$\min \sum_{i=1}^{N} (\sigma z_i + p_i y_i) + \rho \sum_{j<i} x_{ij}$$

subject to

$$\begin{aligned}
y_i &= D_i - \sum_{k>i} x_{ki} + \sum_{l<i} x_{il} \quad i = 1, \ldots, N, & \text{(B.29)}\\
y_i - M z_i &\leq 0 \quad i = 1, \ldots, N, & \text{(B.30)}\\
z_i &\in \{0, 1\} \quad i = 1, \ldots, N, & \text{(B.31)}\\
x_i, y_i &\in \{\text{non-negative integers}\} \quad i = 1, \ldots, N, & \text{(B.32)}
\end{aligned}$$

Multiperiod, Stochastic Formulation

To produce a multiperiod formulation variables and data are subscripted with a period index t and we add an index for the scenario. To model the idea that items produced in one period can be used as-is or reduced in subsequent periods, we use x_{ijt} to indicate that product i is to be used without reduction in period t if $i = j$ and with reduction otherwise. The y vector gives production quantities for each product in each period without regard to the period in which they will be used (and perhaps reduced for use). The formulation is essentially an extension of the single period formulation except that a capacity constraint must be added in the multiple period formulation.

$$\min \sum_{s \in \mathcal{S}} \Pr(s) \sum_{t=1}^{\tau} \left[\sum_{i=1}^{N} (\sigma z_{its} + p_i y_{its}) + \rho \sum_{j<i} x_{ijts} \right]$$

subject to

$$\begin{aligned}
\sum_{j \geq i} x_{ijts} &\geq D_{its} \quad s \in \mathcal{S},\ i = 1, \ldots, N,\ t = 1, \ldots, \tau & \text{(B.33)}\\
\sum_{t' \leq t} \left[\sum_{j \leq i} x_{ijt's} - y_{it's} \right] &\leq 0 \quad s \in \mathcal{S},\ i = 1, \ldots, N,\ t = 1, \ldots, \tau & \text{(B.34)}\\
y_{its} - M z_{its} &\leq 0 \quad s \in \mathcal{S},\ i = 1, \ldots, N,\ t = 1, \ldots, \tau & \text{(B.35)}\\
\sum_{i=1}^{N} y_{its} &\leq c_{ts} \quad s \in \mathcal{S},\ t = 1, \ldots, \tau & \text{(B.36)}\\
z_{its} &\in \{0, 1\} \quad s \in \mathcal{S},\ i = 1, \ldots, N,\ t = 1, \ldots, \tau & \text{(B.37)}
\end{aligned}$$

Data

For illustration purposed we made use of a small version of this problem, where we consider two decision stages (and a third stage for realization of undiscovered costs) and production of three different sizes. The setup cost for producing each size is $ 453. The deterministic unit production costs for size 3 , the largest size, is $ 0.54. For the smallest size the discrete probability distribution has two values: $ 0.48 and $ 0.52, each with a 0.5 probability. For size number two the uncertain costs are $ 0.50 and $ 0.54, also here each has a 0.5 probability. The unit cutting cost, for substituting a smaller size with a larger, is $ 0.008. The demand in the first period is 7500 for each of the three sizes. The uncertain demand is modeled by use of two scenarios, each with a probability of 0.5. The scenarios are 5000 of each size in the low demand case, and 10000 of each size in the high demand case. The total production capacity in each period is 30000 units.

References

[1] Z. Artstein, and R.J-B Wets, "Sensors and information in optimization under stochastic uncertainty," *Mathematics of Operations Research*, 28, 1993, 523-547.

[2] I.L. Averbakh, "An Additive Method for Optimization of Two-Stage Stochastic Systems with Discrete Variables," *Soviet Journal of Computer and System Sciences*, 28, 1990, 161-165.

[3] J.R. Birge, and R.J-B Wets, "Computing bounds for stochastic programming problems by means of a generalized moment problem," *Mathematics of Operations Research*, 12, 1987, 149-162.

[4] S. Bjørnestad, Å. Hallefjord, K. Jörnsten "Discrete optimization under uncertainty: the scenario and policy aggregation technique," Working Paper 89/06, Chr. Michelsen Institute, Bergen, Norway 1989.

[5] H. Bjørstad, Å. Hallefjord, T. Hefting, "Decision trees with coupling constraints," Report Ref.CMI-No.30151-2, Chr. Michelsen Institute, Bergen, Norway 1988.

[6] C.C. Carøe, and R. Schultz, "Dual Decomposition in Stochastic Integer Programming," Technical Report, Institute of Mathematics, Universitetsparken 5, DK-2100, Copenhagen, Denmark, 1996.

[7] CPLEX Optimization, Inc. *Using the CPLEX Callable Library*, Suite 279, 930 Tahoe Blvd. Building 802, Incline Village, NV, 89451-9436, 1994.

[8] A. Gaivoronski, "Linearization methods for optimization of functionals which depend on probability measures," *Mathematical Programming Study* 28, 1986, 157-181.

[9] A. Gaivoronski, "Stochastic optimization techniques for finding optimal submeasures," in *Stochastic Optimization*, V.I. Arkin, A. Shiraev and R. Wets (Eds.), Springer-Verlag Lecture Notes in Control and Information Sciences 81, 351-363, Berlin, 1986.

[10] Y.-C. Ho, X.-R. Cao, *Perturbation analysis of discrete event dynamic systems*, Kluwer Academic Publisher, Boston, 1991.

[11] S. Jorjani, C.H. Scott, D.L. Woodruff, "Optimal Selection of a Subset of Sizes," working paper, GSM, UC Davis, Davis CA 95616, 1996.

[12] G. Laporte, and F.V. Louveaux, "The Integer L-Shaped Method for Stochastic Integer Programs with Complete Recourse," *OR Letters*, 13, 1993, 133-142.

[13] A. Løkketangen, and D.L. Woodruff, "Progressive Hedging and Tabu Search Applied to Mixed Integer (0,1) Multistage Stochastic Programming," *Journal of Heuristics*, 2, (1996), 111-128.

[14] G.Ch. Pflug, "On-line optimization of simulated Markovian processes," *Mathematics of Operations Research* 15,1990, 381-395.

[15] A. Prékopa, *Stochastic Programming*, Kluwer Academic Publisher, Dordrecht, 1995.

[16] R.Y. Rubinstein, and A. Shapiro, *Discrete event systems: sensitivity analysis and stochastic optimization by the score function*, J. Wiley & Sons, New York, 1993.

[17] R. Schultz, L. Stougie, and M.H. van der Vlerk, "Solving stochastic programs with complete integer recourse: a framework using Gröbner bases, Discussion Paper 9562, CORE, Louvain-la-Neuve, Belgium, 1995.

[18] L. Stougie, and M.H. van der Vlerk, "Stochastic integer programming", in *Annotated Bibliographies in Combinatorial Optimization*, M. Dell'Amico, F. Maffioli and S. Martello (Eds.), Chapter 9, 127-141, Wiley, 1997.

[19] S. Takriti, J.R. Birge, and E. Long, "Lagrangian Solution Techniques and Bounds for Loosely-Coupled Mixed-Integer Stochastic Programs," Technical Report, Department of IEOR, University of Michigan, Ann Arbor, MI, USA, 1994.

[20] S.R. Tayur, R.R. Thomas, and N.R. Natraj, "An algebraic geometry algorithm for scheduling in the presence of setups and correlated demands," *Mathematical Programming* 69, 1995, 369-401.

Paper C

Optimal Selection and Sequencing of Oil Wells under Reservoir Uncertainty [1]

Abstract

In this paper we concider the problem of finding the optimal sequence for drilling of production wells in an oil reservoir. When these sequencing decisions are made, the information about the reservoir is limited, but one gets more information as a result of the drilling and the following production. A Bayesian model for updating the a priori probability distribution over reservoir characteristics is proposed. It is shown how this decision problem can be modelled in terms of a decision tree, and an implicit enumeration algorithm for solving this sequencing problem is proposed. Numerical results are reported. The results are computed by use of a mixed integer opimization model where a reservoir simulator is included. The results show that including future information discovery in the model may have influence on optimal drilling decisions.

[1]Submitted for publication in *European Journal of Operational Research*

C.1 Introduction

In this paper we discuss models for optimal selection and sequencing of drilling operations for production wells in an oil field. Before a decision to develop a field is made, seismic data from the field are collected and closely analyzed. In addition, test wells are drilled to help estimate the amount of recoverable oil and the production properties of the reservoir. However, when one starts to drill the production wells the available reservoir information is still limited, and new information will be acquired both through the development and the production phase. The aim of our work is to take this future information discovery into account, and investigate how it may influence the optimal decision policy.

The problem we want to solve is as follows: It is decided to develop a field and the platform capacity is chosen. A set of potential well sites is proposed, but which wells to drill and in which sequence are not yet decided. The objective of the optimization problem is to find the decisions that maximize the project's expected net present value. If we neglect the value of the information obtained from the drilling process, the drilling decisions can be viewed as an investment in production capacity only. However, when taking the information discovery into account, the drilling can be viewed also as an investment in further information.

The problem can be viewed as a stochastic programming problem where the random elements are decision dependent. This is in contrast with for example oil field optimization under price uncertainty, where we know that future oil price will be revealed independent of the decision to develop the oil field [7]. (We assume the field is small and does not influence the future oil price). For uncertainty regarding the reservoir's properties the situation is different. The reservoir is not subject to random influence, and as such we can say that the amount of recoverable oil is deterministic. But our knowledge about it is limited, and which information we get depends on which wells are drilled and when they are drilled. The models presented in this paper are related to the framework presented in Jonsbråten, Wets and Woodruff [8]. There modelling and solving of stochastic programs with decision dependent random elements is discussed, and an algorithm for solving a certain class of such problems is proposed. However, the information discovery there is different from the problem considered in this paper. There all the associated stochasticity is resolved when a certain decision is made, while this is not the case in the problem we look at in this paper. When a well is drilled we get some information, but there is still uncertainty with respect to the reservoir's characteristics. To model this information discovery process we will propose a Bayesian framework where information from drilling activities is used for revising the probability distribution over possible reservoir realizations. This Bayesian framework can be modeled in terms of a decision tree, and we propose an algorithm for finding the optimal decision sequence in this tree.

In the next section we present the deterministic mixed integer programming model for optimizing reservoir development. In Section 3 uncertainty is introduced and a Bayesian information process where information discovery depend on the drilling decisions is proposed. This process can be modeled in terms of a decision tree, and this approach is presented in Section 4. In Section 5 we propose an implicit enumeration algorithm for finding the optimal decision policy. Numerical experiments are reported in Section 6, while concluding remarks and directions for further research are given in Section 7.

C.2 The Oil Field Optimization Model

We will in this section present the deterministic model that is the starting point for our discussion. This is a mixed integer optimization model where a reservoir description is included, and the decisions to be optimized are which wells to drill and production strategies for each well. The model includes reservoir equations for a single phase oil reservoir, and such a description may be approximated by linear equations. We have here considered a two-dimensional reservoir description as being sufficient. The model was first given in Haugland, Hallefjord and Asheim [2]. The model has been further refined in Jonsbråten [6], and it has also been used when investigating oil field optimization under price uncertainty [7]. In Hallefjord, Haugland and Asheim [1] it is given a broader discussion of models for petroleum field optimization. All these references have a more thorough discussion of the optimization model than given in this paper.

The complete deterministic mixed integer model can be written as:

$$\max \sum_{n=1}^{T} c^n \cdot \Delta t^n \sum_{b=1}^{B} q_b^n - \sum_{b=1}^{B}\sum_{n=1}^{T_B} C_b^n x_b^n - \sum_{n=1}^{T} H_n z_n \quad (C.1)$$

$$\text{s.t.} \quad q_b^n \le J_b\left(p^0 - \sum_{k=1}^{n}\sum_{l=1}^{B} \alpha_l^{n+\frac{1}{2}-k}(b) q_l^k - p_w\right), \quad b = 1,\dots,B,\ n = 1,\dots,T \quad (C.2)$$

$$q_b^n \le S_b \textstyle\sum_{k=1}^{n} x_b^k \quad b = 1,\dots,B,\ n = 1,\dots,T_B - 1 \quad (C.3)$$

$$q_b^n \le S_b \textstyle\sum_{k=1}^{T_B} x_b^k \quad b = 1,\dots,B,\ n = T_B,\dots,T \quad (C.4)$$

$$\textstyle\sum_{n=1}^{T_B} x_b^n \le 1 \quad b = 1,\dots,B, \quad (C.5)$$

$$\textstyle\sum_{b=1}^{B} x_b^n \le 1 \quad n = 1,\dots,T_B, \quad (C.6)$$

$$\sum_{b=1}^{B} q_b^n \le D z^n \quad n = 1,\dots,T \quad (C.7)$$

$$-z^{n-1} + z^n \le 0 \quad n = 2,\dots,T \quad (C.8)$$

$$q_b^n \ge 0, \quad b = 1,\dots,B,\ n = 1,\dots,T \quad (C.9)$$

$$x_b^n \in \{0,1\} \quad b = 1,\dots,B,\ n = 1,\dots,T_B \quad (C.10)$$

$$z^n \in \{0,1\} \quad n = 1,\dots,T \quad (C.11)$$

In the objective function (C.1), it is the project's net present value that is maximized. We consider production over T periods of lengths $\Delta t^1,\dots,\Delta t^n$, and the deterministic oil price in period n is c^n. The production flow from well b in period n is given as q_b^n, and there are B potential well sites. The costs in the objective function are costs of drilling the wells and fixed costs of operating the platform. The index T_B is the number of periods for which we consider the drilling operations. We let C_b^n be the discounted cost of drilling well b in period n, and T_B is the latest possible drilling date. The well decision variable x_b^n can be defined as:

$$x_b^n = \begin{cases} 1 & \text{if well } b \text{ is drilled for production in period } n \\ 0 & \text{otherwise} \end{cases} \quad (C.12)$$

The discounted fixed operating costs of the platform in period n are denoted H_n, and we define the integer operating variable, z^n:

$$z^n = \begin{cases} 1 & \text{if the platform is operating in period } n \\ 0 & \text{otherwise} \end{cases} \tag{C.13}$$

The constraints (C.2) essentially describes the reservoir's performance. In this expression J_b is a well specific productivity index determined by the reservoir characteristics at that site, the fluid's properties and the diameter of the well. The initial pressure in the reservoir is given by p^0 while p_w is the minimum well pressure that make production possible. The parameters $\alpha_l^{n+\frac{1}{2}-k}(b)$ have the following interpretation: If well l produces one production unit in period k, this results in a pressure drop of $\alpha_l^{n+\frac{1}{2}-k}(b)$ pressure units in period n near well b. Since we want to express p_b^n for a general production q, we have to find all the coefficients α. This can be done by simulation of the system. By solving the discretized reservoir equation for a set of B linearly independent q-vectors, all the α's can be found. When introducing uncertainty regarding the reservoir's properties, it is uncertainty regarding the α's that is introduced.

The inequalities (C.3) and (C.4) assure that production can only take place when the well is drilled. The maximum production from each well is given by the parameter S_b. The constraints (C.5) make sure that a well can be drilled only once. It is possible to drill one well every period, and this is assured by constraint (C.6). The constraints (C.7) link the production and platform operation decisions, and production can only take place if the platform is operated. The total platform capacity is given by D. We will not allow the platform to close down for a later restart, and that is taken care of in constraints (C.8) that force z^n to be non-increasing for increasing n. The production decisions are modeled as continuous variables while the drilling and operating decisions are binary, and this is taken care of by the constraints (C.9) - (C.11).

The formulation of the well decisions in the model (C.1 - C.11) is not the same formulation as Haugland, Hallefjord and Asheim found to be most efficient. They ended up with a model where the binary well decision variable is non-decreasing in time and equals 0 before the well is drilled and 1 after drilling. However, we have chosen to define x_b^n as done in (C.12) because it makes the notation simpler when introducing the Bayesian framework.

C.3 The Information Process

We will in this section first present the information process, before we illustrate the process by a numerical example. The proposed model is rather general, but the discussion here will be related to the reservoir optimization problem.

C.3.1 The Bayesian Model

We define a *state space*, Θ, that represents the unknown reservoir. This space consists of a finite set of reservoir realizations labeled by θ. With R different states, the state space can be written:

$$\Theta = \{\theta_1, \ldots, \theta_R\}$$

When considering well drilling decisions, we introduce a slightly different notation than in the previous section. The potential decisions are given as the set X, where each of the B elements in X represents a drilling decision:

$$X = \{x_1, \ldots, x_B\}$$

When a well is drilled one will get more information about the reservoir. This information is given as a test result, v, and when a well is drilled one of K possible test results can be observed. One way to specify these test results can be as good, medium or bad. This sample space can be written as:

$$V = \{v_1, \ldots, v_K\}$$

From the information available prior to the drilling decisions, each state θ is assigned a probability, $P(\theta)$. Based on knowledge and experience there is also estimated a conditional probability measure on the sample space V for given x and θ:

$$P(v|x, \theta)$$

The idea behind this is as follows: When a well is drilled, the observed test result gives some information about the well's production capacity and the properties of the reservoir. This test result, v, will depend on which test is performed (which well that is drilled) and the true reservoir state, θ. For a given reservoir state the test result is not unique but is given as a conditional probability distribution. The probability distribution for the test result v for a given x can now be found as:

$$P(v|x) = \sum_{\theta \in \Theta} P(v|x, \theta) \cdot P(\theta) \tag{C.14}$$

With these conditional probabilities, we have the necessary framework for assigning new probabilities to the possible reservoir states, posterior to knowing the outcome v of the decision x. By use of Baye's theorem the conditional probability for state θ for a given x and v can be written:

$$P(\theta|x,v) = \frac{P(v|x,\theta) \cdot P(\theta)}{P(v|x)} \tag{C.15}$$

Our aim is to develop a multiperiod formulation, where new information is obtained every time a well is drilled. For simplifying the notation we will let the vector $\boldsymbol{x}^t$ represent the sequence of drilling decisions prior to time t. The vector is of length t and its elements are drilling decisions x_b^n. We assume a situation as described in the previous section, where only one well can be drilled in a period. In other words, the vector $\boldsymbol{x}^t$ consists of all non-zero x-variables in the model (1.0 - 1.10) prior to time t. The non-bold x_b^t represents the t-th element of the drilling sequence $\boldsymbol{x}^t$. For periods when there is not drilled any well, the associated vector element is 0. The vector $\boldsymbol{x}^T$ is a full sequence of drilling decisions. Similarly the vector $\boldsymbol{v}^t$ represents the test results associated with the decision vector $\boldsymbol{x}^t$. For a period when there is not drilled a well, there will not be any test result, and the associated element in the $\boldsymbol{v}^t$ vector will be 0. The non-bold v_k^t is the t-th element in the sequence of test results $\boldsymbol{v}^t$.

The conditional probability measure on test results for given x and θ, $P(v|x,\theta)$, can in the multiperiod formulation be written: $P(v_k^t|\boldsymbol{x}^t, \boldsymbol{v}^{t-1}, \theta)$. We have included the sequence of test results $\boldsymbol{v}^{t-1}$ in this formula, because the probability distribution over reservoir realizations prior to the decision x_b^t is dependent upon the sequence of test results $\boldsymbol{v}^{t-1}$. Further, the probability distribution for observing the test result v_k^t with given $\boldsymbol{x}^t$, observations $\boldsymbol{v}^{t-1}$, but unspecified θ can be found as:

$$P(v_k^t|\boldsymbol{x}^t, \boldsymbol{v}^{t-1}) = \sum_{\theta \in \Theta} P(v_k^t|\boldsymbol{x}^t, \boldsymbol{v}^{t-1}, \theta) \cdot P(\theta|\boldsymbol{x}^{t-1}, \boldsymbol{v}^{t-1}) \tag{C.16}$$

The posterior probability for the state θ, after the drilling sequence $\boldsymbol{x}^t$ and getting the results $\boldsymbol{v}^t$, can then be expressed as:

$$P(\theta|\boldsymbol{x}^t, \boldsymbol{v}^t) = \frac{P(v_k^t|\boldsymbol{x}^t, \boldsymbol{v}^{t-1}, \theta) \cdot P(\theta|\boldsymbol{x}^{t-1}, \boldsymbol{v}_k^{t-1})}{P(v_k^t|\boldsymbol{x}^t, \boldsymbol{v}^{t-1})} \tag{C.17}$$

	"B"	"M"	"G"
θ_1	0.10	0.35	0.55
θ_2	0.20	0.45	0.35
θ_3	0.35	0.45	0.20
θ_4	0.55	0.35	0.10

Table C.1: Conditional Probabilities, Well 1

	"B"	"M"	"G"
θ_1	0.25	0.45	0.30
θ_2	0.05	0.35	0.60
θ_3	0.50	0.40	0.10
θ_4	0.40	0.40	0.20

Table C.2: Conditional Probabilities, Well 2

C.3.2 A Numerical Illustration

We will here use a small example to show how the framework presented in the previous section can be used. Let the initial probability distribution be:

$$P(\theta_1) = P(\theta_2) = P(\theta_3) = P(\theta_4) = 0.25$$

In this example only two wells are considered. For each well there is an estimated probability for each test result v, conditional upon well and reservoir. We assume three possible test results: bad - "B", medium - "M" and good - "G". The conditional probabilities are shown in Tables C.1 and C.2. Further we assume that well I is drilled in the first period and well II is drilled in the second period. We also assume that well I gives a test result that is good, whereas well II shows a medium test result. After the first period we then get the following posterior reservoir probabilities:

$$P(\theta_1|x_1^1, \text{"G"}) = \frac{0.55 \cdot 0.25}{0.3} = 0.458$$

$$P(\theta_2|x_1^1, \text{"G"}) = \frac{0.35 \cdot 0.25}{0.3} = 0.292$$

$$P(\theta_3|x_1^1, \text{"G"}) = \frac{0.20 \cdot 0.25}{0.3} = 0.167$$

$$P(\theta_4|x_1^1, \text{"G"}) = \frac{0.10 \cdot 0.25}{0.3} = 0.083$$

The updated probabilities above represents knowledge about the state of the reservoir after the first period. When drilling well II, we know that the probability of test result "M" is:

$$\begin{aligned} P(\text{"M"}|x_1x_2, \text{"G"}) &= (0.458 \cdot 0.45) + (0.292 \cdot 0.35) \\ &+ (0.167 \cdot 0.40) + (0.083 \cdot 0.40) = 0.408 \end{aligned}$$

After the second period, finding that well II is medium, we get the following posterior probabilities:

$$P(\theta_1|x_1^1x_2^2, \text{"G" "M"}) = \frac{0.45 \cdot 0.458}{0.408} = 0.505$$

$$P(\theta_2|x_1^1x_2^2, \text{"G" "M"}) = \frac{0.35 \cdot 0.0.292}{0.408} = 0.250$$

$$P(\theta_3|x_1^1x_2^2, \text{"G" "M"}) = \frac{0.40 \cdot 0.167}{0.408} = 0.163$$

$$P(\theta_4|x_1^1x_2^2, \text{"G" "M"}) = \frac{0.40 \cdot 0.083}{0.408} = 0.082$$

We see that the good test from well I changes the probability distribution over reservoir realizations considerably, while the medium test result in well II only leads to smaller adjustments in the probability distribution.

C.4 The Decision Tree

Our next goal is to use the described information process to generate a decision tree, and in order to do that we will define necessary notation:

$\mathcal{D}$ The complete decision tree consisting of all possible sequences of drilling decisions and associated test results.

$(\boldsymbol{x^t}, \boldsymbol{v^t})$ Decision nodes; decisions prior to t are fixed and the associated test results are observed.

$(\boldsymbol{x^{t+1}}, \boldsymbol{v^t})$ Chance nodes; decisions prior to $t+1$ are fixed and the test results are observed for decisions prior to t.

$(\boldsymbol{x^0}, \boldsymbol{v^0})$ The root node; no decisions are fixed.

$(\boldsymbol{x^T}, \boldsymbol{v^T})$ Terminal nodes; nodes where all uncertainty is resolved.

$P(\boldsymbol{x^t}, \boldsymbol{v^t})$ Node probability; the probability for node $(\boldsymbol{x^t}, \boldsymbol{v^t})$ if decision policy $\boldsymbol{x^t}$ is chosen. $P(\boldsymbol{x^t}, \boldsymbol{v^t}) = \prod_{i=1}^t P(v_k^i|\boldsymbol{x^i}, \boldsymbol{v^{i-1}})$

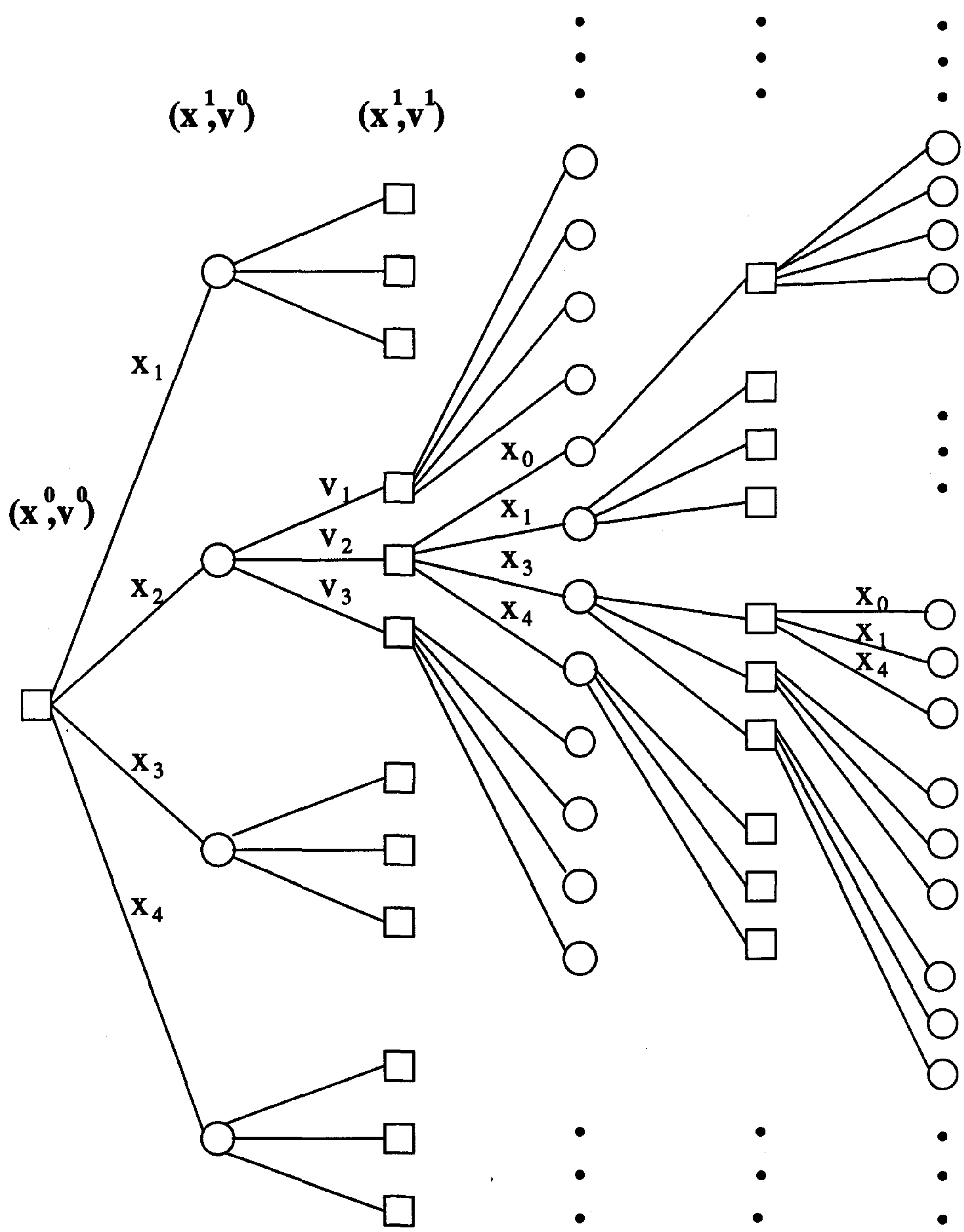

Figure C.1: Decision tree

Figure C.1 illustrates what the tree will look like in a situation with 4 proposed well sites and 3 possible test results. Decisions are made at times $0, \ldots, T-1$, and the associated test results, v, are realized in the following time periods. For making the tree complete, test results v^{T-1} will resolve all uncertainty, so that the reservoir state is known with certainty in the terminal nodes. The decision nodes are represented by squares, while the chance nodes are shown as circles.

One thing that Figure C.1 illustrates clearly is the size of the tree. Even with a moderate number of possible drilling decisions and a few test results, the size of the decision tree grows rapidly. In the problem we are considering here, it is decided to develop the field, and we do not consider the possibility of not drilling any well in the first period. But in the following periods that is a possibility, and as shown in the illustration, if we do not drill we do not get any further information. In the example in Figure C.1 with 4 potential well sites and 3 test results, the tree has 12 decision nodes at $t = 1$. After the second period the number of decision nodes is $12 * ((3 * 3) + 1) = 120$, and after the third there are 876 possible decision nodes. We will denote the number of decision nodes at time t $(1 < t < T)$ as Γ^t which can be formally expressed as:

$$\Gamma^t = B \cdot K \cdot \sum_{i=1}^{t} \left[\binom{t-1}{i} K^i \prod_{j=1}^{i} (B-j) \right] \tag{C.18}$$

As introduced earlier, B is the number of potential wells and K is the number of test results. The first part of this expression, $B \cdot K$, represents the number of decision nodes at $t = 1$. The summation index i essentially represents the number of possible drilled wells. The binomial coefficient give the number of possible sequences when i ordered wells are going to be drilled in $t-1$ periods. The number of test results is given by K^i, while $\prod_{j=1}^{i}(B-j)$ is the number of of possible sequences when drilling i of B possible wells. The number of terminal nodes is not given by this expression. This is because all uncertainty will here be resolved independent of the decision in the last period. The total number of terminal nodes in the the tree, Γ^T can then be written as:

$$\Gamma^T = B \cdot K \cdot \sum_{i=1}^{T-1} \left[\binom{T-2}{i} K^{i-1} \cdot R \cdot (B+1-i) \prod_{j=1}^{i} (B-j) \right] \tag{C.19}$$

What makes this decision tree model different from traditional decision tree models, is that the decisions here are both production decisions and investments in information. One way to introduce decision tree analysis found in textbooks, e.g. Hillier and Lieberman [3], is the oil field problem where the main decision to be made is "drill" or "do not drill". There is uncertainty about the amount of recoverable oil, and a seismic survey can be conducted for helping estimating the amount of oil. The seismic survey has a certain cost, and whether to conduct this survey or not, depends on to which extent it will improve the decisions. The sequence of decisions to be made are "test"/"do not test" and then "drill"/"do not drill". A rather similar, but a bit more sophisticated example can be found in Holloway [5]. However, the important thing to notice is that the decision to invest in further information and the decision to drill are separate decisions, and not the same decision as it is in the problem we are investigating.

C.5 The Implicit Enumeration Algorithm

We will in this section propose an implicit enumeration algorithm for finding the optimal sequencing decisions for the tree described in the previous section. For making things simple, we will introduce the algorithm for an integer programming problem, before we later discuss how the algorithm can be used for solving the mixed integer problem of optimal reservoir development.

C.5.1 Implicit Enumeration for Integer Problems

The integer optimization problem we want to solve may on an aggregated form be written as:

$$\max \sum_{\theta \in \Theta} [P(\theta) f(\boldsymbol{X})]$$

subject to

$$\boldsymbol{A}_{\theta} \boldsymbol{X} \leq \boldsymbol{b} \tag{C.20}$$

$$\boldsymbol{X} \in \mathcal{N} \tag{C.21}$$

$$\boldsymbol{X} \in \{0,1\} \tag{C.22}$$

As before we assume this to be a multiperiod sequencing problem where only one non-zero variable is allowed for each period, and it is the expected value that is maximized. Inequality (C.20) represents the constraint system, and $\boldsymbol{A}_{\theta}$ is the constraint matrix for reservoir realization θ. We let $\boldsymbol{X}$ represent the entire solution system of $\boldsymbol{x}$-vectors. The set of implementable solution systems are represented by $\mathcal{N}$ where

$$\mathcal{N} = \{\boldsymbol{X} : (\boldsymbol{x}^{t+1}, \boldsymbol{v}^{t})' = (\boldsymbol{x}^{t+1}, \boldsymbol{v}^{t})'' \forall (\boldsymbol{x}^{t}, \boldsymbol{v}^{t})' = (\boldsymbol{x}^{t}, \boldsymbol{v}^{t})'', t \in \{0, \ldots, T-1\})\}$$

In words, each decision node must have a unique decision. The optimal solution to this maximization problem will be a solution system where the decisions are allowed to depend upon observed test results. The implicit enumeration algorithm operates on the decision tree described in the previous section. In traditional decision tree analysis the optimal solution is found by use of a backward induction procedure, where one starts at the terminal (leaf) nodes by calculating optimal policies, and then by backwards induction work towards the root node. Such a procedure will provide an optimal solution also for this problem, but as pointed out in the previous section, the decision tree will be very large.

In the algorithm proposed here the problem is approached in a different manner. The algorithm starts its calculations in the root node and then in a breadth first manner works its way towards the terminal nodes. The idea is to successively calculate lower and upper bounds for the decision sequences, and by successively computing tighter bounds we may be able at early stages to bound out decision sequences that are proven to not be optimal. How successful this algorithm will be, depends on the size of the problem, the problem data and our ability to compute tight bounds. However, even if it shows that the problem is of a size that makes it too computationally demanding to find the optimal solution, one will get valuable information about the problem by running just a few iterations of the proposed algorithm.

A lower bound must always be a feasible and implementable solution to the optimization problem, and the problem of computing lower bounds is that of finding a good feasible solution to the optimization problem without too much computational effort. The upper bounds are calculated by allowing for foresight, and the upper bound solutions are therefore not required to be implementable. Obviously there are several ways of calculating bounds, and we do not claim that the way we have done it is the most efficient one. However, we will present the procedures used here, and it will be left as a topic for future research how these procedures may be improved upon.

The proposed algorithm is given in Figure C.7, and the Figures C.2-C.6 show different procedures called by this algorithm. In addition to the notation introduced so far, the algorithm use the following symbols:

$\underline{\mathcal{D}}$ represents the reduced decision tree. Initially all possible drilling sequences are considered as possible optimal solutions, but as soon as a sequence is found to not belong to the optimal solution, it is bounded out, i.e. removed from $\underline{\mathcal{D}}$. At any time the reduced decision tree represents all drilling sequences that are possible optimal solutions. When the algorithm terminates, the reduced decision tree consists of the optimal solution system.

$\mathcal{L}(\boldsymbol{x}^t, \boldsymbol{v}^t)$ Lower bound for the decision node $(\boldsymbol{x}^t, \boldsymbol{v}^t)$

$\mathcal{U}(\boldsymbol{x}^{t+1}, \boldsymbol{v}^t)$ Upper bound for the chance node $(\boldsymbol{x}^{t+1}, \boldsymbol{v}^t)$

In the Figures C.2 - C.7 we use a notation where a left-arrow ($\leftarrow$) means assignment; i.e. the variable to the left is assigned the value on the right side of the arrow. The reduced decision tree $\underline{\mathcal{D}}$ and the lists $\mathcal{L}(\cdot,\cdot)$ and $\mathcal{U}(\cdot,\cdot)$ are all defined as global variables. The procedure for computing lower bounds is formally described in Figure C.2. When procedure `Lower` is called a decision node $(\boldsymbol{x}^\tau, \boldsymbol{v}^\tau)$ is passed to it, and the procedure calculates a lower bound for this decision node. This is done by using the "averaged" reservoir as a starting point; averaged according to the probability distribution over reservoir realizations available at the specific node. However, we also want to take future information discovery into account, and this leads to the recursive structure of procedure `Lower`. As long as it is not a terminal node that is passed to `Lower`, the development decisions for the averaged reservoir realization are optimized, but with decisions prior to τ fixed. The drilling decision for $\tau + 1$ is then added to the sequence of drilling decisions, and procedure `Lower` is now called for each of the associated test results, $v_k^{\tau+1}$, by passing the decision nodes $(\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^{\tau+1})$. In this manner `Lower` works its way towards the terminal nodes. In the terminal nodes all the uncertainty is resolved, and when passed a terminal node `Lower` calculates the objective function value for this deterministic problem. In this way the computed solution system is allowed to depend on future information discovery. The lower bound represents the present best feasible solution found so far:

$$LBD(\boldsymbol{x}^t, \boldsymbol{v}^t) = \max \sum_{\theta \in \Theta} [f(\boldsymbol{X}) | \boldsymbol{x}^t, \boldsymbol{A}_\theta \boldsymbol{x} \le \boldsymbol{b}, \boldsymbol{X} \in \mathcal{N})]$$

The strategy for computing upper bounds is quite different. Upper bounds are calculated by `Upper`, and the formal definition of this procedure is given in Figure C.4. When `Upper` is called a chance node is $(\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^\tau)$ passed to it. The upper bound is generated by optimizing the development decisions for each reservoir realization, but with the decisions prior to $\tau + 1$ fixed. In other words, the upper bounds are calculated under the assumption of perfect information when making the decisions after $\tau + 1$, and the decisions are as such not required to be implementable:

$$UBD(\boldsymbol{x}^{t+1}, \boldsymbol{v}^t) = \sum_{\theta \in \Theta} [\max f(\boldsymbol{x} | \boldsymbol{x}^{\tau+1}, \boldsymbol{A}_\theta \boldsymbol{x} \le \boldsymbol{b})]$$

In the first iteration of the `Implicit Enumeration Algorithm`, $t = 0$, the lower bound is calculated for the root node $\mathcal{L}(\boldsymbol{x}^0, \boldsymbol{x}^0)$. `Upper` calculates an upper bound for each chance node $(\boldsymbol{x}^1, \boldsymbol{v}^0)$, before the procedure `Reduce` removes from the reduced decision tree, $\underline{\mathcal{D}}$, all drilling sequences where $\mathcal{L}(\boldsymbol{x}^0, \boldsymbol{v}^0) > \mathcal{U}(\boldsymbol{x}^1, \boldsymbol{v}^0)$. The procedure `Reduce` is defined in Figure C.6, and its purpose is to remove from $\underline{\mathcal{D}}$ all branches (decision sequences) where the upper bound is lower than the computed lower bound.

These steps are all repeated in the following iterations, but after the initial iteration also the procedures `Update_L` and `Update_U` are included. The procedure `Lower` calculates $\mathcal{L}(\boldsymbol{x}^t, \boldsymbol{v}^t)$ for all decision nodes at level t in $\underline{\mathcal{D}}$. The procedure `Update_L`, defined in Figure C.3, uses these most recently calculated lower bounds to update the lower bounds at the decision nodes at levels $\tau < t$. The procedure `Reduce` uses these updated lower bounds to check if any more decision sequences can be bounded out. The procedure `Upper` calculates upper bounds for the chance nodes $(\boldsymbol{x}^{t+1}, \boldsymbol{v}^t)$, and the procedure `Update_U`, defined in Figure C.5, uses these bounds to update the upper bounds at lower levels in the tree. In this function $X(\boldsymbol{x}^{\tau+1})$ represents the set of possible level $\tau+2$ drilling decisions given the drilling sequence $\boldsymbol{x}^{\tau+1}$

The algorithm terminates after the lower bounds are calculated for the terminal nodes. When the procedure `Upper` is called by the chance node $(\boldsymbol{x}^T, \boldsymbol{v}^{T-1})$, the problem $\max f(\boldsymbol{x}|\boldsymbol{x}^T, \boldsymbol{A}_\theta \boldsymbol{x} \leq \boldsymbol{b})$ is solved for each reservoir realization θ. In the terminal nodes the reservoir realization θ is given with certainty, and when procedure `Lower` is called with the terminal node $(\boldsymbol{x}^T, \boldsymbol{v}^T)$, it is the problem $\max f(\boldsymbol{x}|\boldsymbol{x}^T, \boldsymbol{A}_\theta \boldsymbol{x} \leq \boldsymbol{b})$ that is solved also here. At that point the gap between upper and lower bounds is eliminated, and a unique optimal decision policy is found.

In future research the proposed algorithm should be further evaluated in the context of integer programming techniques. The branch-and-cut algorithm [4, 9] may be a starting point for such analysis.

C.5.2 Implicit Enumeration for the Well Sequencing Problem

As pointed out in Section 2, optimal well sequencing is a mixed integer problem, while we so far in this section, and in the Sections 3 and 4, have treated the problem as an integer problem. The terminal nodes in the decision tree have been considered as nodes where all decisions are fixed and the

---

$\mathcal{L}(\boldsymbol{x}^\tau, \boldsymbol{v}^\tau) \leftarrow 0$

IF $(\tau = T)$

$\mathcal{L}(\boldsymbol{x}^\tau, \boldsymbol{v}^\tau) \leftarrow \max f(\boldsymbol{x}|\boldsymbol{x}^T, \boldsymbol{A_\theta x} \leq \boldsymbol{b})$

ELSE

$\boldsymbol{x}^{\tau+1} \leftarrow \arg\max f(\boldsymbol{x}|\boldsymbol{x}^\tau, \overline{\boldsymbol{A}}\boldsymbol{x} \leq \boldsymbol{b}, \overline{\boldsymbol{A}} = \sum_{\theta\in\Theta} \boldsymbol{A_\theta} \cdot P(\theta|\boldsymbol{x}^\tau, \boldsymbol{v}^\tau))$

FOREACH $v_k^{\tau+1} \in V^{\tau+1}$

`Lower` $(\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^{\tau+1})$

$\mathcal{L}(\boldsymbol{x}^\tau, \boldsymbol{v}^\tau) \leftarrow \mathcal{L}(\boldsymbol{x}^\tau, \boldsymbol{v}^\tau) + \mathcal{L}(\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^{\tau+1}) \cdot P(\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^{\tau+1})$

$\mathcal{L}(\boldsymbol{x}^\tau, \boldsymbol{v}^\tau) \leftarrow \mathcal{L}(\boldsymbol{x}^\tau, \boldsymbol{v}^\tau)/P(\boldsymbol{x}^\tau, \boldsymbol{v}^\tau)$

Figure C.2: Definition of procedure `Lower`$(\boldsymbol{x}^\tau, \boldsymbol{v}^\tau)$

FOREACH $x_b^{\tau+1} \in X \setminus \boldsymbol{x}^\tau$

$\boldsymbol{x}^{\tau+1} \leftarrow \boldsymbol{x}^\tau + x_b^{\tau+1}$

$\hat{L} \leftarrow 0$

FOREACH $v^{\tau+1} \in V$

$\boldsymbol{v}^{\tau+1} \leftarrow \boldsymbol{v}^\tau + v^{\tau+1}$

$\hat{L} \leftarrow \hat{L} + [\mathcal{L}(\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^{\tau+1}) \cdot P(\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^{\tau+1})]$

$\hat{L} \leftarrow \hat{L}/P((\boldsymbol{x}^\tau, \boldsymbol{v}^\tau)$

IF $[\hat{L} > \mathcal{L}(\boldsymbol{x}^\tau, \boldsymbol{v}^\tau]$

$\mathcal{L}(\boldsymbol{x}^\tau, \boldsymbol{v}^\tau] \leftarrow \hat{L}$

Figure C.3: Definition of procedure `Update_L`$(\boldsymbol{x}^\tau, \boldsymbol{v}^\tau)$

$\mathcal{U}(\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^{\tau}) \leftarrow 0$

FOREACH $\theta \in \Theta$

$\quad \mathcal{U}(\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^{\tau}) \leftarrow \max f(\boldsymbol{x}|\boldsymbol{x}^{\tau+1}, \boldsymbol{A}_{\theta}\boldsymbol{x} \leq \boldsymbol{b}) \cdot P(\theta|\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^{\tau})$

Figure C.4: Definition of procedure Upper($\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^{\tau}$)

$\mathcal{U}(\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^{\tau}) \leftarrow 0$

FOREACH $v_k^{\tau+1} \in V$

$\quad \boldsymbol{v}^{\tau+1} \leftarrow \boldsymbol{v}^{\tau} + v_k^{\tau+1}$

$\quad \hat{U} \leftarrow 0$

$\quad$ FOREACH $x_b^{\tau+2} \in X(\boldsymbol{x}^{\tau+1})$

$\quad\quad \boldsymbol{x}^{\tau+2} \leftarrow \boldsymbol{x}^{\tau+1} + x^{\tau+2}$

$\quad\quad$ IF $[\mathcal{U}(\boldsymbol{x}^{\tau+2}, \boldsymbol{v}^{\tau+1}) > \hat{U}]$

$\quad\quad\quad \hat{U} \leftarrow \mathcal{U}(\boldsymbol{x}^{\tau+2}, \boldsymbol{v}^{\tau+1})$

$\quad \mathcal{U}(\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^{\tau}) \leftarrow \mathcal{U}(\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^{\tau}) + [\hat{U} \cdot P((\boldsymbol{x}^{\tau+2}, \boldsymbol{v}^{\tau+1})]$

$\mathcal{U}(\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^{\tau}) \leftarrow \mathcal{U}(\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^{\tau})/P((\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^{\tau})$

Figure C.5: Definition of procedure Update_U($\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^{\tau}$)

$\boldsymbol{x}^{\tau} \leftarrow \boldsymbol{x}^{\tau+1} - x_b^{\tau+1}$

IF $[\mathcal{L}(\boldsymbol{x}^{\tau}, \boldsymbol{v}^{\tau}) > \mathcal{U}(\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^{\tau})]$

$\quad \underline{\mathcal{D}} \leftarrow \underline{\mathcal{D}} \setminus (\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^{\tau})$

Figure C.6: Definition of procedure Reduce($\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^{\tau}$)

$t \leftarrow 0 \quad \underline{\mathcal{D}} \leftarrow \mathcal{D}$

WHILE $(t \leq T)$

    FOREACH $(\boldsymbol{x}^t, \boldsymbol{v}^t) \in \underline{\mathcal{D}}$

        `Lower`$(\boldsymbol{x}^t, \boldsymbol{v}^t)$

    FOREACH $\tau \in \{t-1, \ldots, 0\}$

        FOREACH $(\boldsymbol{x}^\tau, \boldsymbol{v}^\tau) \in \underline{\mathcal{D}}$

            `Update_L` $(\boldsymbol{x}^\tau, \boldsymbol{v}^\tau)$

        FOREACH $(\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^\tau) \in \underline{\mathcal{D}}$

            `Reduce` $(\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^\tau)$

    IF $(t < T)$

        FOREACH $(\boldsymbol{x}^{t+1}, \boldsymbol{v}^t) \in \underline{\mathcal{D}}$

            `Upper`$(\boldsymbol{x}^{t+1}, \boldsymbol{v}^t)$

        FOREACH $\tau \in \{t, \ldots, 0\}$

            FOREACH $(\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^\tau) \in \underline{\mathcal{D}}$

                IF $\tau < t$

                    `Update_U` $(\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^\tau)$

                `Reduce` $(\boldsymbol{x}^{\tau+1}, \boldsymbol{v}^\tau)$

    $t \leftarrow t + 1$

Figure C.7: Definition of the `Implicit Enumeration Algorithm` uncertainty is resolved. Also in the well sequencing problem we assume the uncertainty to be resolved when reaching the terminal nodes, but there are still more operating and production decisions to be made. As such, we can say that each terminal node represents a deterministic mixed integer problem that has to be solved. In the terminal nodes all decisions concerning the T_B first periods are fixed, while decisions for the periods $T_B + 1$ to T are still to be optimized.

In the general framework developed, the implementability constraints play an important role, and it is necessary to look closer at what implementable solutions mean in the oilfield case. Both the drilling and production decisions need to be implementable. We are assured that the drilling decisions are implementable, but so far we have not discussed the continuous production decisions. In the numerical experiments performed here the total platform capacity is larger than the sum of the maximum production capacity for each of the potential wells; $\sum_{b=1}^{B} S_b \leq D$. Because of this, we do not have the problem of allocating production capacity among the producing wells. The implementable production strategy will simply be to produce as much as possible from each well.

If the situation had been that $\sum_{b=1}^{B} S_b \geq D$, it would have been necessary to use a predetermined rule for allocating production capacity among the wells in order to avoid solutions depending on hindsight. In that case the lower and upper bounds for the terminal nodes is not necessarily equal, and we are not guaranteed to be able to bound out all non-optimal decision strategies.

C.6 Numerical Experiments

We will in this section present numerical experiments where the reservoir optimization problem is solved by use of the implicit enumeration algorithm. The example reservoir we consider is illustrated in Figure C.8. Its size is $900m \times 800m$, and 5 different well sites, numbered 1-5, are proposed. In this model the reservoir is completely described by its porosity and permeability. The porosity expresses the amount of oil present in the reservoir, while the permeability gives information about the reservoir's production properties. However, there is uncertainty with respect to the reservoir's properties.

The reservoir uncertainty is expressed by a discrete probability distribution over reservoir realizations. We consider 4 different realizations, and the initial distribution is uniform, i.e.: $P(\theta_1) = P(\theta_2) = P(\theta_3) = P(\theta_4) = 0.25$. The

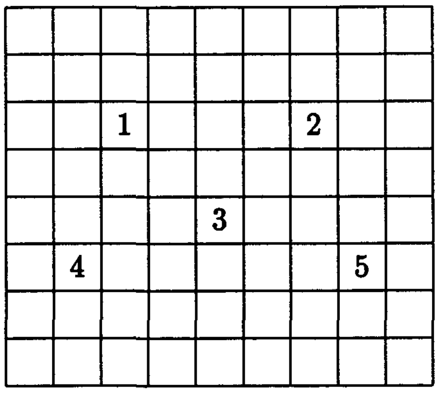

Figure C.8: The example reservoir with potential well locations

reservoir realizations are illustrated in the Figures C.10 - C.13. In the example we are considering, the porosity and permeability vary proportional to each other. For each block in the reservoir there is a number that give the associated properties. To get the porosity the numbers in the Figures C.10-C.13 have to be multiplied by 0.02, while the permeability is found by multiplying by $10^{-14}m^2$. In this example the porosity varies in the order $0.08 - 0.18$ while the permeability varies from $4 \cdot 10^{-14}m^2$ to $9 \cdot 10^{-14}m^2$. As in the introductory example there are estimated conditional probabilities for each of the potential wells; the probability of observing a specific test result conditional upon reservoir realization and which well that is drilled. These estimated probabilities is listed in the Tables C.1 - C.5. We allow wells to be drilled the $T_B = 5$ first periods, and these periods have a length of 100 days each. The next $(T - T_B) = 8$ periods is of length 300 days, so the total lifetime of the oilfield is a period of 2900 days. All the other relevant input data is given in Appendix A.

As mentioned earlier, our main goal is to find the complete optimal solution system, but much is also achieved by finding the optimal first stage solution. When the first stage decision is made and the associated test results are observed, a new analysis with updated probabilities may be performed.

In the first iteration of the implicit enumeration algorithm, $t = 0$, the procedure **Lower** found well 4 to be the optimal first stage solution for the averaged reservoir. Suggested decision vectors are (x_4^1, x_5^2, x_1^3), (x_4^1, x_5^2, x_2^3), (x_4^1, x_2^2, x_2^5), and $(x_4^1, x_2^2, x_5^3, x_1^4)$, depending on observed test results. It is interesting to note that well 3 is the only well that is not included in any of these drilling sequences. The computed lower bound for the root node in this first iteration is \$ 26.59 mill.

When investigating the upper bounds for each of the chance nodes, we also get the optimal decision policies for each of the reservoir realizations. For the reservoir realizations θ_1 and θ_3 it is found to be optimal to drill well 1 first, while for the realizations θ_2 and θ_4, well 2 is the optimal first stage decision. If we at the first stage were in possession of perfect information, the project's expected net present value would be \$ 29.00 mill. The procedure `Upper` gives the following upper bounds on the possible first stage decisions:

$$\mathcal{U}(x_1) = \$\ 28.18 \text{ mill.}$$
$$\mathcal{U}(x_2) = \$\ 27.48 \text{ mill.}$$
$$\mathcal{U}(x_3) = \$\ 28.29 \text{ mill.}$$
$$\mathcal{U}(x_4) = \$\ 28.01 \text{ mill.}$$
$$\mathcal{U}(x_5) = \$\ 28.28 \text{ mill.}$$

The resulting optimal solution found by use of the `Implicit enumeration algorithm` is illustrated in Figure C.9. The solution has an expected net present value of \$ 26.73 mill., and we found drilling of well 3 to be the optimal first stage decision. At the next stage well 4 is drilled if test results "bad" or "medium" are observed, while a "good" test result leads to the decision of drilling well 1. As seen from the illustration the optimal drilling sequences are: (x_3^1, x_4^2, x_2^3), (x_3^1, x_4^2, x_5^3) and (x_3^1, x_1^2, x_5^3). Which sequence that will be realized depends on the observed test results.

C.7 Conclusion and Future Work

The main goal of this work has been to develop a framework for optimizing reservoir development under uncertainty, which is an optimization problem with decision dependent information discovery. To accomplish this we have proposed a Bayesian model that updates the probability distribution over reservoir realizations when new information is acquired from drilling activities. This Bayesian approach can be modeled in terms of a decision tree, and based on this decision tree we have proposed an implicit enumeration algorithm for finding the optimal drilling sequence. More specificly, we find an optimal solution system where the decisions are allowed to depend on the acquired information. By doing this a drilling decision is not only a production decision but also an investment in information acquisition.

Although this information process and the implicit enumeration algorithm

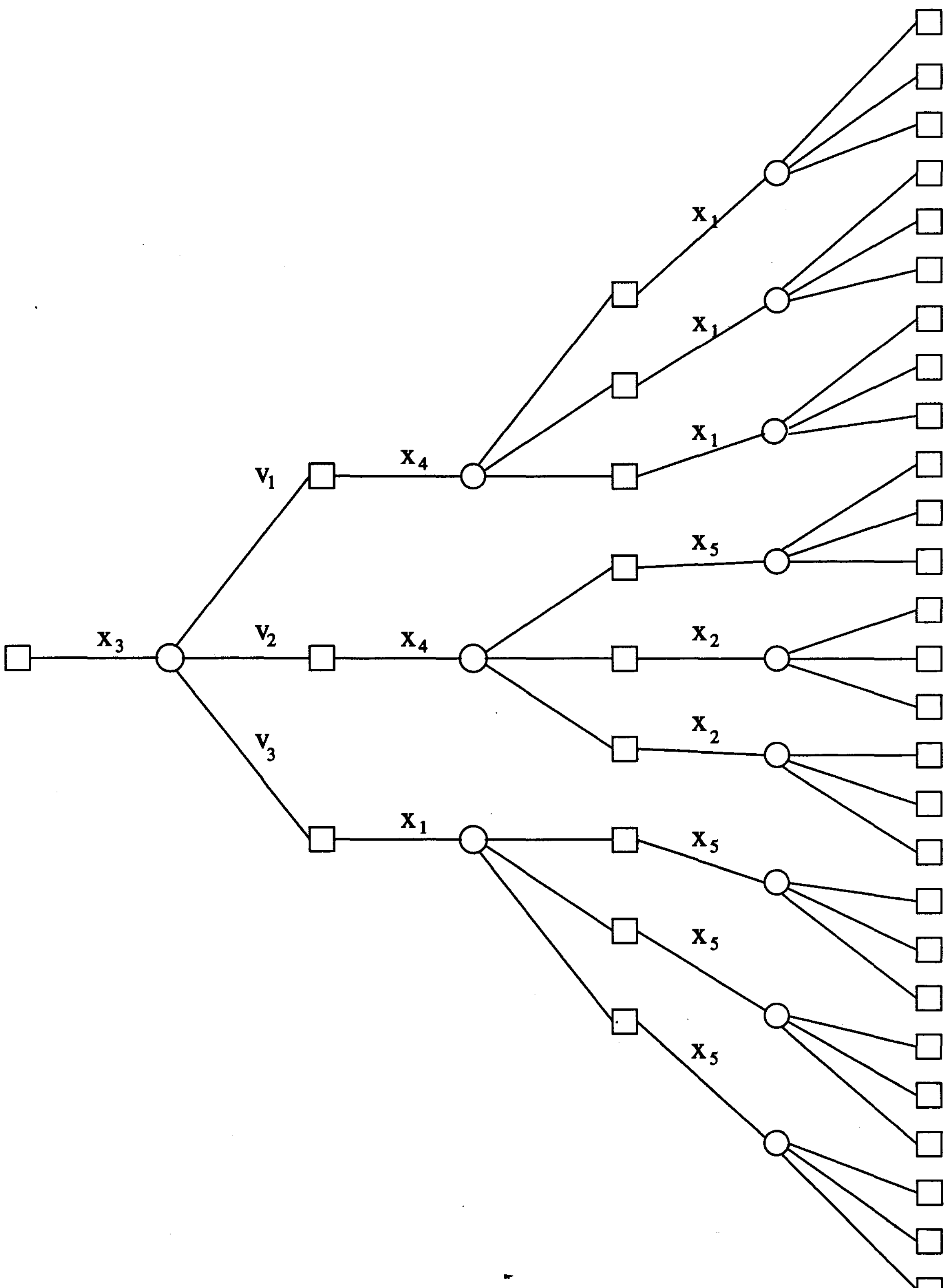

Figure C.9: The optimal solution system

have been developed for optimizing oilfield development, the ideas can hopefully be useful also in other applications. Reservoir uncertainty is an example of an important class of problems where the information discovery is decision dependent. We believe that there exist a lot of planning problems where the information discovery is dependent upon the previous decisions, and we hope that the framework presented here may be an inspiration to further research in this field. The way we see it, this paper points out new topics for further research, both in modelling and solving problems with decision dependent information discovery.

Appendix A - Input data

The numerical experiments presented in this paper are done for a reservoir with the following properties:

Reservoir length (x-direction):	900 m
Reservoir breadth (y-direction):	800 m
Reservoir height (z-direction):	50 m
Viscosity (μ):	$1 \cdot 10^{-7} bar \cdot s$
Constant temp. compressibility (c):	0.001 bar^{-1}
Fluid density at initial pressure (ρ^0):	1150 kg/m^3
Fluid density at surface pressure (ρ^s):	850 kg/m^3
Reservoir porosity (φ):	0.2
Reservoir initial pressure (p^0):	300 bar
Lowest pressure for production (p_w):	50 bar
Wellradius (r_w):	0.15 m
Oil price c^t:	\$ 16 per barrel
Cost of drilling each well:	\$ 10 mill.
Maximum well capacity (p_w):	10 l/sec
Platform cost:	\$ 25 mill.
Maximum platform capacity (D):	50 l/sec
Fixed operating costs (H_t):	\$ 10000 per day

In expected net present values reported in the paper, it is also included the fixed platform cost of \$ 25 mill. The Figures C.10 - C.13 represents reservoir realizations, and the conditional probabilities, probability for a specific test result given which well that is drilled and reservoir realization, are shown in Tables C.1, C.2, C.3, C.4 and C.5.

7	7	8	7	7	6	6	6	6
7	8	8	8	7	7	6	6	6
7	8	8	8	7	7	6	6	6
7	8	8	8	7	7	7	7	6
7	7	8	8	8	8	8	8	7
6	6	7	8	8	8	8	8	7
6	6	6	7	8	8	8	8	7
5	5	6	6	6	7	7	7	6

Figure C.10: Reservoir realization θ_1

7	7	7	7	8	9	8	7	7
7	7	7	7	7	8	9	8	7
7	7	7	7	7	8	9	9	8
8	8	7	7	7	7	8	9	9
8	9	8	7	7	7	7	8	8
8	9	9	8	7	7	7	7	8
7	8	9	8	7	7	7	7	7
7	7	8	8	7	7	7	7	7

Figure C.11: Reservoir realization θ_2

5	5	5	5	4	4	4	4	4
5	5	5	5	5	4	4	4	4
5	5	5	5	5	4	4	4	4
5	5	5	5	5	4	4	4	4
5	5	5	5	5	5	5	5	5
4	5	5	5	5	5	5	5	5
4	4	5	5	5	5	5	5	5
4	4	4	5	5	5	5	5	5

Figure C.12: Reservoir realization θ_3

4	4	4	4	4	5	5	5	5
4	4	4	4	4	5	5	5	5
4	4	4	4	4	4	5	5	5
5	4	4	4	4	4	4	4	4
5	5	4	4	4	4	4	4	4
5	5	5	4	4	4	4	4	4
5	5	5	4	4	4	4	4	4
5	5	5	5	4	4	4	4	4

Figure C.13: Reservoir realization θ_4

	"B"	"M"	"G"
θ_1	0.05	0.25	0.70
θ_2	0.10	0.55	0.35
θ_3	0.35	0.55	0.10
θ_4	0.70	0.25	0.05

Table C.3: Conditional Probabilities, Well 3

	"B"	"M"	"G"
θ_1	0.25	0.45	0.30
θ_2	0.05	0.35	0.60
θ_3	0.45	0.40	0.15
θ_4	0.45	0.40	0.15

Table C.4: Conditional Probabilities, Well 4

	"B"	"M"	"G"
θ_1	0.10	0.35	0.55
θ_2	0.20	0.45	0.35
θ_3	0.35	0.45	0.20
θ_4	0.55	0.35	0.10

Table C.5: Conditional Probabilities, Well 5

References

[1] Å. Hallefjord, H. Asheim, and D. Haugland. Petroleum field optimization - comments on models and solution techniques. CMI-report 862331-1, Chr. Michelsen Institute, Fantoft, Norway, 1986.

[2] D. Haugland, Å. Hallefjord, and H. Asheim. Models for petroleum field exploitation. *European Journal of Operational Research*, 37:58–72, 1988.

[3] F. S. Hillier and G. J. Lieberman. *Introduction to Operations Research, Sixth Edition.* McGraw-Hill, New York, 1995.

[4] K. Hoffman and M. Padberg. Solving airline crew scheduling problems by branch-and-cut. *Management Science*, 39:657–682, 1993.

[5] C. A. Holloway. *Decision Making under Uncertainty.* Prentice-Hall, Englewood Cliffs, 1979.

[6] T. W. Jonsbråten. *Optimeringsmodeller for utvinning av et oljereservoar.* HAS-thesis, Norwegian School of Economics and Business Administration, Bergen, 1994.

[7] T. W. Jonsbråten. Oil field optimization under price uncertainty. To appear in *Journal of Operational Research Society*, 1998.

[8] T. W. Jonsbråten, R. J.-B. Wets, and D. L. Woodruff. A class of stochastic programs with decision dependent random elements. To appear in *Annals of Operations Research*, 1998.

[9] M. Padberg and G. Rinaldi. A branch-and-cut algorithm for the resolution of large-scale symmetric travelig salesman problems. *SIAM Review*, 33:60–100, 1991.

Paper D

Nash Equilibrium and Bargaining in an Oil Reservoir Management Game [1]

Abstract

We discuss the common pool problem faced by two surface lease owners with access to the same oil reservoir. Our approach is to use a mixed integer optimization model for clarifying the unitization negotiations. The model includes a two-dimensional reservoir description, and both development and production decisions are optimized. The discussion is based on a numerical example. We show there exists a unique Nash equlibrium for this non-cooperative game, illustrating the over-investment induced by the competitive extraction. This non-cooperative solution may serve as a starting point in the negotiations for a unitization agreement. We discuss several ways of how the project value in a cooperative game may be shared between the two owners: Nash bargaining solution, recoverable oil and reservoir properties, and the economic value of each of the leases in absence of oil migration.

[1]Paper presented at *Fagkonferansen i Bedriftsøkonomiske Emner, NHH, 1996*

D.1 Introduction

The problem discussed in this paper is optimal development and distribution of costs and revenues for an oil field located on a block boundary. Optimal development of the reservoir requires coordinated actions from the owners, and the non-renewable resource, the reservoir, must be treated as a unit. Unitization is the practice of unifying the ownership and control of an oil reservoir, such that the field is developed and operated by a single operator, representing all the owners. Shares in this unit is to be negotiated, and is determined from surface ownership, recoverable oil and other reservoir characteristics. However, it is not straightforward to reach an agreement, and the idea in this paper is to use an optimization model for clarifying the negotiations.

The reason for this problem is of course the oil's migratory nature. If the extraction is not coordinated, each owner has incentive to extract as much oil as possible, before it is extracted by the competitor. Compared to optimal development of the reservoir, we get over-investment in wells and production capacity. This competitive extraction may also reduce the reservoir's recovery, and compared to coordinated extraction the resulting amount of recoverable oil can be considerable lower under competitive extraction [8]. This oil reservoir management problem has much in common with problems connected to the fisheries [7]. In both cases two or more agents are exploiting a common pool, but in the fisheries we are concerned with a renewable natural resource, whereas the oil field is a non-renewable resource.

The starting point for this work is a mixed-integer optimization model where a two-dimensional reservoir description is included. One of the ideas behind this model is to take the interaction between the different variables into account, and the model simultaneously determines platform capacity, drilling program and production strategy. This model will give the optimal development decisions when the reservoir is developed as a unit. Besides, the model is used for investigation of threat strategies in a non-cooperative game. These threat strategies will then serve as starting points in the negotiations distributing the surplus from cooperative development of the reservoir.

The paper is organized as follows: In the next section the mixed-integer optimization model is presented. In Section 3 we show how this model can be used for finding optimal development decisions for an oil reservoir. Section 4 is devoted to non-cooperative reservoir development, and our aim is to find a stable Nash-equilibrium for this two-agent game. The Nash-solution represents threat strategies that may serve as a starting point in unitization negotiations. Possible solutions to this bargaining problem is discussed in Section 5. Section 6 contain a summary of the findings in this paper and plans for further research.

D.2 The Optimization Model

The starting point for our analysis is an optimization model for development of an oil reservoir, and this model is based on the work presented in Haugland, Hallefjord and Asheim [5]. Input to the model is a set of platform capacities with associated costs, a set of potential well sites with associated costs, operating costs, oil price, discount rate and a description of the reservoir. Decisions supported by the model are:

- Platform capacity
- Number of wells to be drilled
- Location of the wells
- When the wells should be drilled
- Production profile for each well

The aim of the model is to find the decisions that maximize the reservoir value, and in such an analysis the reservoir properties must play an important role. The reservoir studied here is a single phase oil reservoir, which means that only the oil is mobile. The oil is slightly compressible, and production is possible because the oil expands as the reservoir pressure is reduced. A reservoir description can be derived by combining a mass conservation equation, a simplified version of Darcy's flow equation, and an equation for the fluid's constant temperature compressibility. We then arrive at the following reservoir equation where the coefficients do not depend on the pressure:

$$\nabla \cdot (\frac{k}{\mu} \nabla p) = \varphi c \frac{\partial p}{\partial t} + \frac{w}{\rho^0} \tag{D.1}$$

In this equation the following notation is used:

ρ	fluid density	(kg/m^3)
v	flow velocity	(m/s)
w	source/sink terms	$(kg/m^3/s)$
φ	porosity	(fraction)

t	time	(s)
k	permeability	(m^2)
μ	fluid viscosity	$(bar \cdot s)$
p	pressure	(bar)
c	constant temperature compressibility	(bar^{-1})

A more careful derivation of the reservoir description can be found in Aziz [1] and Peaceman [9]. In reservoirs that are relatively thin compared to their area extent, it is possible to assume that the flow in the vertical direction is negligible compared to the flow in the horizontal directions. Equation (D.1) is therefore approximated by a two-dimensional discrete model, where the variations of the reservoir height is a function of the x- and y-variables.

Production Variables

The connection between the pressure and production can be approximated the following way [5]:

$$q_b(t) \leq J_b[p_b(t) - p_w] \tag{D.2}$$

where $q_b(t)$ is the production rate in well b at time t, J_b is a well-specific productivity index, $p_b(t)$ is the reservoir pressure near well b (the pressure in the block where well b is located), and p_w is the minimum well pressure. The productivity index may be estimated from the rock and fluid properties, and it is determined by the permeability, porosity, fluid viscosity, block-area, reservoir height, and well radius [10].

Initially we assume the reservoir pressure to be p^0, and according to our notation p^n is the pressure after n periods. The estimated pressure in the middle of period n will then be denoted $p^{n-\frac{1}{2}}$. It can be shown for a reservoir described by linear equations that the pressure near well b in the middle of period n can be written [4],[6]:

$$p_b^{n-\frac{1}{2}} = p^0 - \sum_{k=1}^{n} \sum_{l=1}^{B} \alpha_l^{n+\frac{1}{2}-k}(b)\; q_l^k \tag{D.3}$$

The parameters $\alpha_l^{n+\frac{1}{2}-k}(b)$ have the following interpretation: If well l produces one production unit in period k, this results in a pressure drop of $\alpha_l^{n+\frac{1}{2}-k}(b)$ pressure units in period n near well b. Since we want to express p_b^n for a general production q, we have to find all the coefficients α. This can be done by simulation of the system. By solving the discretized reservoir equation for a set of B linearly independent q-vectors, all the α's can be found. This simulation process is discussed in Haugland, Hallefjord and Asheim [5].

Well-drilling

We let $C_b^{T_B}$ be the discounted cost of drilling well b in period T_B. The increased cost if we rather choose to drill well b in some period k prior to T_B is: $\sum_{n=k}^{T_B-1} C_b^n$. The well decision variable x_b^n may then be defined as:

$$x_b^n = \begin{cases} 1 & \text{if well } b \text{ is drilled in one of the periods } 1, \ldots, n \\ 0 & \text{otherwise} \end{cases} \tag{D.4}$$

As we see from the definition is x_b^n non-decreasing for increasing n. In our model T_B is the latest possible drilling date, and for each of the T_B first periods we will just allow one well to be drilled.

Platform Capacity

We also want to find optimal platform capacity, and it is possible to choose among M different platform sizes with associated costs. We let Q_1 denote the production capacity of the "smallest" platform, and the total cost of installing this platform is G_1. Further, $Q_2, \ldots, Q_M$ is defined such that if platform g is chosen, the total capacity will be $\sum_{m=1}^{g} Q_m$. The total cost of this platform will be $\sum_{m=1}^{g} G_m$. We see that G_m is defined as the increased cost of expanding the platform capacity with Q_m. The platform decision variable y_m is defined to be non-increasing for increasing m:

$$y_m = \begin{cases} 1 & \text{if a platform with capacity no smaller than } Q_m \text{ is chosen} \\ 0 & \text{otherwise} \end{cases} \tag{D.5}$$

Operating Costs

The fixed operating costs of the platform in period n are denoted H_n. We will let z^n be the integer operating variable and it is defined in the following way:

$$z^n = \begin{cases} 1 & \text{if the platform is operating in period } n \\ 0 & \text{otherwise} \end{cases} \tag{D.6}$$

We will not allow that the platform is closed down for some periods and started up later on, and we therefore require that z^n is non-increasing for increasing n.

The complete deterministic mixed integer model can then be written:

$$\max \sum_{n=1}^{T} c^n \cdot \Delta t^n \sum_{b=1}^{B} q_b^n - \sum_{b=1}^{B}\sum_{n=1}^{T_B} C_b^n x_b^n - \sum_{m=1}^{M} G_m y_m - \sum_{n=1}^{T} H_n z_n \qquad (D.7)$$

$$\text{s.t.} \quad q_b^n \le J_b(p^0 - \sum_{k=1}^{n}\sum_{l=1}^{B} \alpha_l^{n+\frac{1}{2}-k}(b) q_l^k - p_w),$$

$$b = 1,\ldots,B,\ n = 1,\ldots,T \qquad (D.8)$$

$$q_b^n \le S_b x_b^n \qquad b = 1,\ldots,B,\ n = 1,\ldots,T_B - 1 \qquad (D.9)$$

$$q_b^n \le S_b x_b^{T_B} \qquad b = 1,\ldots,B,\ n = T_B,\ldots,T \qquad (D.10)$$

$$\sum_{b=1}^{B} x_b^1 \le 1 \qquad (D.11)$$

$$-\sum_{b=1}^{B} x_b^{n-1} + \sum_{b=1}^{B} x_b^n \le 1 \qquad n = 2,\ldots,T_B \qquad (D.12)$$

$$x_b^{n-1} - x_b^n \le 0 \qquad b = 1,\ldots,B,\ n = 1,\ldots,T \qquad (D.13)$$

$$\sum_{b=1}^{B} q_b^n \le \sum_{m=1}^{M} Q_m y_m \qquad n = 1,\ldots,T \qquad (D.14)$$

$$-y_{m-1} + y_m \le 0 \qquad m = 2,\ldots,M \qquad (D.15)$$

$$\sum_{b=1}^{B} q_b^n \le D^n z^n \qquad n = 1,\ldots,T \qquad (D.16)$$

$$-z^{n-1} + z^n \le 0 \qquad n = 2,\ldots,T \qquad (D.17)$$

$$q_b^n \ge 0, \qquad b = 1,\ldots,B,\ n = 1,\ldots,T$$

$$x_b^n \in \{0,1\} \qquad b = 1,\ldots,B,\ n = 1,\ldots,T_B$$

$$y_m \in \{0,1\} \qquad m = 1,\ldots,M$$

$$z^n \in \{0,1\} \qquad n = 1,\ldots,T$$

The objective is to maximize the net present value, and c^n is the discounted oil price and Δt^n is the length of period n. Constraint (D.8) is the wells' production capacity, achieved by replacing $p_b(t)$ in equation (D.3) by $p_b^{n-\frac{1}{2}}$ as defined by equation (D.2). Constraints (D.9) - (D.13) are related to the well decisions. Constraints (D.9) and (D.10) say that production can only take place in wells that are drilled, while (D.11) and (D.12) ensure that maximum one well is drilled each of the T_b first periods. Constraint (D.13) demands x_b^n to be non-decreasing for increasing n. The total production may not exceed the platform capacity and this is assured by (D.14). Constraint (D.16) says that production can take place only if the platform is operating, and D^n is an upper bound on the platform capacity. The constraints (D.15) and (D.17) assures that y_m and z^n are non-increasing in m and n respectively. We see that the model has $(T \times B)$ continuous (production) variables and $(B \times T_B + M + T)$ binary variables.

D.3 Optimal Reservoir Development

We will in this section use the model for finding optimal development decisions for a specific reservoir. The same reservoir will in the next sections be used in a discussion of non-cooperative and cooperative solutions to this management game. Such numerical experiments are done for a rectangular reservoir, with the following rock and fluid properties:

Reservoir length (x-direction):	800 m
Reservoir width (y-direction):	900 m
Reservoir height (z-direction):	50 m
Viscosity (μ):	$1 \cdot 10^{-7} bar \cdot s$
Constant temperature compressibility (c):	0.001 bar^{-1}
Fluid density at initial pressure (ρ^0):	1150 kg/m^3
Fluid density at surface pressure (ρ^s):	850 kg/m^3
Reservoir porosity (φ):	0.2
Reservoir initial pressure (p^0):	300 bar
Lowest production pressure (p_w):	50 bar
Well radius (r_w):	0.15 m

The lifetime of the reservoir is estimated to be 3000 days. This is discretized into 6 periods of 100 days and 8 periods of 300 days. For each of the 6 first periods, one new well may be drilled. The reservoir is divided into 8×9 blocks, each of dimension $100m \times 100m$, as illustrated in Figure D.1. The reservoir has constant porosity and reservoir height, but the permeability

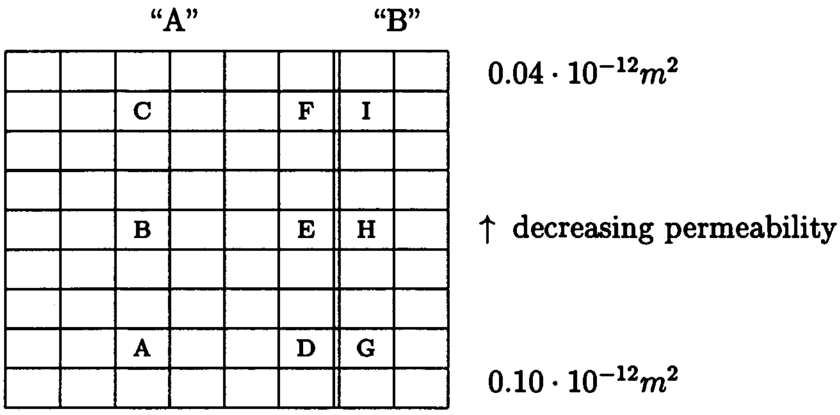

Figure D.1: The example reservoir

Platform	Capacity l/sec	Cost (million $)
1	10	20.0
2	15	22.5
3	20	25.0
4	25	27.5
5	30	30.0

Table D.1: Platform capacities and costs

varies spatially. The permeability, k, is constant in the x-direction, but in the y-direction it varies linearly from $1.0 \cdot 10^{-13} m^2$ in the "lowest" blocks in Figure D.1, to $0.4 \cdot 10^{-13} m^2$ in the "uppermost" blocks.

As indicated in Figure D.1, the reservoir has two owners: Firm "A" owns 75 % of the reservoir and firm "B" 25 %. Because the reservoir in this example is homogeneous in the x-direction, "A" and "B"'s shares of the recoverable oil is also 75 % and 25 %, respectively. The potential well sites in the reservoir is in Figure D.1 marked with letters, and in this example we have 9 potential sites. Owner "A" has 6 potential well sites, A - F, and "B" has 3 potential well sites, G - I, on his part of the reservoir.

The drilling cost is set to $ 8.0 mill per well, and it is only for the 6 first periods the wells will be drilled. The fixed operating costs is $1.0 mill. per 100 days. The different platform alternatives with connected costs are given in Table D.1. The discount rate is 2 % per 100 days. An oil price of $ 16 per barrel is used.

		3			-	-	
		-			-	4	
		2			1	-	

Figure D.2: Optimal drilling program

Given these input data, we can now find the optimal development decisions for the reservoir. The problem is solved by use of "CPLEX Mixed Integer Library" [2]. The field can be developed with one platform, and we will decide which wells to be drilled, independent of ownership. This optimal development strategy is shown in Figure D.2. It is suggested to drill 4 wells, use platform alternative 3 and 3000 days of production. The calculated NPV is \$ 116.7 mill. Obviously this is the decision that maximizes the reservoir's value, but even in this example, with a deterministic model and a rather homogeneous reservoir, it is not straight forward how this value should be allocated between the owners. This problem will be discussed in the following sections.

D.4 Non-cooperative Reservoir Development

We will in this section study non-cooperative reservoir development, with an emphasis on finding stable Nash-solutions to this game. This discussion is based upon the example in the previous section. The motivation for investigating this problem is the following questions: If the parties are unable to reach a bargaining solution, is there a unique disagreement solution that will be played? How will this disagreement solution affect the bargaining process?

In this paper we discuss the *normal* form description of the game:

Definition 1 The *normal form representation* of an n-player game specifies the players' strategy spaces $S_1, \ldots, S_n$ and their payoff functions $u_1, \ldots, u_n$. We denote the game by $G = \{S_1, \ldots, S_n; u_1, \ldots, u_n\}$.

In contradiction to the *extensive* form of the game, the normal form representation suppresses all the informational aspects of the game, and the strategy spaces and payoff functions characterize the game completely. The solution concept we will focus on is the Nash-equilibrium:

Definition 2 In the normal-form game $G = \{S_1, \ldots, S_n; u_1, \ldots, u_n\}$, the strategies $(s_1^\star, \ldots, s_n^\star)$ are a *Nash-equilibrium* if, for each player i, $s_i^\star$ is player i's best response to the strategies specified for the $n-1$ other players, $(s_1^\star, \ldots, s_{i-1}^\star, s_{i+1}^\star, \ldots, s_n^\star)$:

$$u_i(s_1^\star, \ldots, s_{i-1}^\star, s_i^\star, s_{i+1}^\star, \ldots, s_n^\star) \geq u_i(s_1^\star, \ldots, s_{i-1}^\star, s_i, s_{i+1}^\star, \ldots, s_n^\star) \tag{D.18}$$

for every feasible strategy s_i in S_i. That is, $s_i^\star$ solves :

$$\max_{s_i \in S_i} u_i(s_1^\star, \ldots, s_{i-1}^\star, s_i, s_{i+1}^\star, \ldots, s_n^\star) \tag{D.19}$$

Using the normal-form representation of the game give an open-loop equilibrium, where the strategies are functions of time alone. As mentioned earlier, it is the information structure of the game that makes the difference. An alternative would be to search for closed-loop equilibrium, where the players can condition their play at time t on the history of the game until that date. It is clear that the closed-loop strategy space is much larger than the open-loop strategy space, and that is the main reason why we have chosen normal form and open-loop equilibrium.

To be more specific, what we want to find is a stable Nash equilibrium:

Definition 3 A Nash equilibrium solution is **stable** if it can be obtained through an iterative procedure (in strategy space) using the following iteration scheme, regardless of which initial choice starts the iteration:

$$(s_1^\star, \ldots, s_n^\star) = \lim_{\nu \to \infty} s_i^\nu, \ \ \forall s_i^0 \in S_i, \ \ i \in N \tag{D.20}$$

$$s_i^{\nu+1} = \arg\max_{s_i \in S_i} u_i(s_1^{\nu+1}, \ldots, s_{i-1}^{\nu+1}, s_i, s_{i+1}^\nu, \ldots, s_n^\nu) \tag{D.21}$$

This is the discussion framework of the reservoir management game. We will now continue the example in the previous section, and assume the reservoir has two owners, so the management game is a two player game. Although we restrict the analysis to the normal form description of the game, we have a very large number of strategies.

The game is modeled the following way: Both owners may develop their own part of the reservoir, i.e. the reservoir may be developed with two platforms, and the platform alternatives and costs are the same as in the previous section. Player "A" has the potential well sites A - F, and "B" has the well sites G - I. We will start the investigation by observing "A"'s best response to different strategies of "B". Player "B" has three potential well sites, and must choose which wells to drill and the drilling sequence. The scope of the game is to extract as much oil as possible before it is extracted by the opponent, and we therefore assume that "B" will have his wells drilled as early as possible and produce at maximum rate. A number of "B"'s strategies are shown in Table D.2. Another question is how many periods of operation we should specify in these strategies. With the oil price of \$ 16 and fixed operating costs of \$ 1.0 mill per 100 days, the platform production must be 1.15 l/sec to cover the fixed operating costs. For all strategies specified in Table D.2, one more period of production would not have covered "B"'s fixed operating costs. For all the specified strategies, 1500 days of production is chosen, except for 1200 days when only well I is drilled. Because we don't know which platform will be chosen, we let player "B" have free platform capacity.

By investigating "A"'s best responses, we find that "A" always chooses to develop the field with platform 3, and either 4 or 5 wells. Except for the first strategy, where "B" does nothing, well D is always the first well to be drilled. The operating period for player "A" varies from 3000 to 2100 days. Player "A"'s payoff shows how "B"'s strategies affects "A"'s extraction of the reservoir. From these results we expect that "B" will choose to drill all three wells, giving "A" a pay-off from \$ 53.5 mill. to \$ 57.3 mill.

The next step is to investigate "B"'s best response to each of "A"'s calculated strategies. At this point it is also necessary to adjust for the platform capacity. In all of the best responses in Table D.2, the whole platform capacity was utilized in some periods. This is taken into consideration, and this slack is adjusted for in well A. In Table D.3 the best response from "B" is shown for each of "A"'s strategies. As we can see it is always optimal for "B" to have all three wells drilled, and G H I is the optimal drilling sequence. Player "B"'s part of the reservoir is developed with platform 2, and the capacity is fully utilized only in the third period. We also see that "B"'s operating period is 1800 days if "A" drills four wells and 1500 days if five wells are drilled.

Finally we calculate "A"'s best response to "B"'s strategies, now with "B"'s capacity fixed at 15 l/sec. As expected we find "A"'s best answer always to be strategy 11 in Tables D.2 and D.3. Hence we have found a unique Nash

Player "B"					Player "A"							
Str. no.:	Period: 1	2	3	Days of operation	Period: 1	2	3	4	5	6	Days of operation	Pay-off (mill. $)
1				0	A	D	F	B			3000	116.4
2	G			1500	D	A	E	B			3000	83.4
3	H			1500	D	A	E	C			3000	85.8
4	I			1200	D	A	F	B			3000	95.5
5	G	H		1500	D	A	E	B	F		2400	64.5
6	G	I		1500	D	A	E	F	B		2400	65.8
7	H	G		1500	D	A	E	B	F		2400	65.4
8	H	I		1500	D	E	A	F	B		2400	71.5
9	I	G		1500	D	A	E	F	B		2400	69.1
10	I	H		1500	D	E	B	F	A		2400	72.9
11	G	H	I	1500	D	A	E	B	C		2100	53.5
12	G	I	H	1500	D	E	A	C	B		2100	54.3
13	H	G	I	1500	D	A	E	F			2100	56.3
14	H	I	G	1500	D	A	E	B	C		2100	56.1
15	I	G	H	1500	D	A	E	B	C		2100	56.5
16	I	H	G	1500	D	A	E	C	B		2100	57.3

Table D.2: "A"'s best response to "B"'s strategies

Player "A"								Player "B"				
Str.	Period:						Days of	Period:			Days of	Pay-off
no.:	1	2	3	4	5	6	operation	1	2	3	operation	(mill. $)
1	A	D	F	B			3000	G	H	I	1800	22.9
2	D	A	E	B			3000	G	H	I	1800	21.2
3	D	A	E	C			3000	G	H	I	1800	21.7
4	D	A	F	B			3000	G	H	I	1800	22.0
5	D	A	E	B	F		2400	G	H	I	1500	16.7
6	D	A	E	F	B		2400	G	H	I	1500	16.4
7	D	A	E	B	F		2400	G	H	I	1500	16.7
8	D	E	A	F	B		2400	G	H	I	1500	16.1
9	D	A	E	F	B		2400	G	H	I	1500	16.4
10	D	E	B	F	A		2400	G	H	I	1500	16.6
11	D	A	E	B	C		2100	G	H	I	1500	18.4
12	D	E	A	C	B		2100	G	H	I	1500	18.1
13	D	A	E	F			2100	G	H	I	1800	20.1
14	D	A	E	B	C		2100	G	H	I	1500	18.4
15	D	A	E	B	C		2100	G	H	I	1500	18.4
16	D	A	E	C	B		2100	G	H	I	1500	18.5

Table D.3: "B"'s best response to "A"'s responses

		5			-	3	
		4			3	2	
		2			1	1	

Figure D.3: The reservoir

equilibrium for this reservoir management game, a solution from which none of the players want to deviate. The drilling decisions in this non-cooperative solution is illustrated in Figure D.3. In addition we have found that "A" chooses platform alternative 3 and 2100 days of operation, while "B" develops his lease with platform 2 and 1500 days of operation.

The calculated NPV is \$ 53,5 mill. for "A" and \$ 18,4 mill. for "B". Compared to the cooperative solution, a surplus of \$ 44.8 mill. is lost by realizing these threat strategies. The loss is a result of developing the field with two platforms and eight wells. We will in the following sections discuss how this surplus may be shared between the to owners. This non-cooperative solution will be a lower bound on what the parties may expect in the negotiations. If not forced to it by external institutions, none of them will accept a bargaining solution giving less than they get from their threat strategies.

D.5 Bargaining Solutions

The aim of the bargaining is to reach a unitization agreement which specifies each owner's share when the field is developed as a unit. We will in this section discuss different ways of reaching such an agreement. First we will look at Nash's bargaining solution, an axiomatic approach in which the threat strategies play an important role. The usual way to deal with unitization negotiations is to calculate unitization formulas from the reservoir properties. We will also show how the optimization model can be used to calculate each lease's value in absence of oil migration between the two parts of the reservoir.

D.5.1 Nash Bargaining Solution

One way to treat bargaining problems is to conclude that theories of bargaining can do nothing more than just specify a range in which an agreement may be found. With such a view it is natural to focus on the bargaining process and the bargainers interaction. A contrary view is that by sufficient information about the bargainers and the structure of the bargaining problem, it is possible to predict a unique outcome. This view was advocated by John Nash, and he developed what has come to be called an "axiomatic model of bargaining". The bargainers are described by von Neuman-Morgenstern utility functions, and Nash proposed the following four axioms for "reasonable" behavior [11]:

1. Independence of equivalent utility representations - the solution is not affected by linear transformations of the bargainers utility function.

2. Symmetry - if the parties are indistinguishable, then so should be their payoffs.

3. Independence of irrelevant alternatives - enlarging the set of feasible payoffs should either cause one of the new payoffs to be selected or leave the original outcome unaffected.

4. Pareto optimality - no other feasible outcome is preferred by all of the players

Nash's theorem says that there exists a unique solution to the bargaining problem that obeys these four properties. This unique outcome happens to maximize the geometric average of the gains which the players realize by reaching an agreement instead of getting the disagreement outcome. Also in Smith [12] is the Nash bargaining solution used for distributing project value between the owners, but his model do not include a reservoir description. The optimal development plan is given as a number of wells for each owner, and the non-cooperative plan is assumed to be a constant multiple of the optimal plan for both firms.

The disagreement outcome will in our case be the non-cooperative equilibrium calculated in the previous section. This is a threat that will be implemented if the negotiations fail. Each threat must be credible, i.e. it must be individually rational for each of the players to carry out their threats if an agreement is not reached. We will also assume that the bargainers have equivalent utility representations, and the payoff will be measured in money.

The bargaining problem is illustrated in Figure D.4. The disagreement outcome, point D, gives the solution (d_A, d_B). The set of feasible payoffs to this bargaining game is the triangle DFG, and the Pareto optimal solutions is on the line FG. The Nash bargaining solution, Q, with payoffs (q_A, q_B) is the unique outcome maximizing the product $(q_A - d_A)(q_B - d_B)$, and the solution obviously gives $(q_A - d_A) = (q_B - d_B)$. The surplus generated by cooperating is distributed in two equal parts, and in our example this gives player "A" $ (53.5 + 22.4) mill. = $ 75.9 mill. and player "B" $ (18.4 + 22.4) mill. = $ 40.8 mill.

D.5.2 Recoverable Oil and Reservoir Properties

The usual way to deal with unitization negotiations is to reach agreement on a unitization formula, distributing the shares of the project according to amount of recoverable oil and other reservoir properties [3]. This is not to complicated in a very homogeneous reservoir with complete information, like the one in our example. But in real life the early phase information about the reservoir is rather limited. In addition the bargaining may be further complicated because of asymmetric information [13].

In this example the reservoir has uniform porosity, which means that the recoverable oil is uniformly distributed in the reservoir. The permeability is constant in the x-direction, and varies linearly in the y-direction. The well's productivity index varies proportionally to the permeability, and as such the permeability is a measure of the production capacity. Because the permeability is identical distributed in the two leases and the poroesity is constant in the whole reservoir, it seems reasonable to distribute the project value according to amount of recoverable oil (here identical to surface ownership). This give owner "A" 75 % and "B" 25 % of the value, $ 87.5 mill. and $ 29.2 mill. respectively. This solution is shown as point R in Figure D.4. We see that this solution give higher payoff to "A", compared to the Nash bargaining solution Q, but both parties are better off than the disagreement solution D.

D.5.3 The Economic Value of the Leases

Another way to look at the problem is this: What is the value of each of the leases if they had been separated by an impermeable zone. Obviously we are leaving the physical realities, and as such the result is more of theoretical interest. But it might be an interesting starting point for the negotiations, knowing the value of the leases "on their own", when drainage of the neigh-

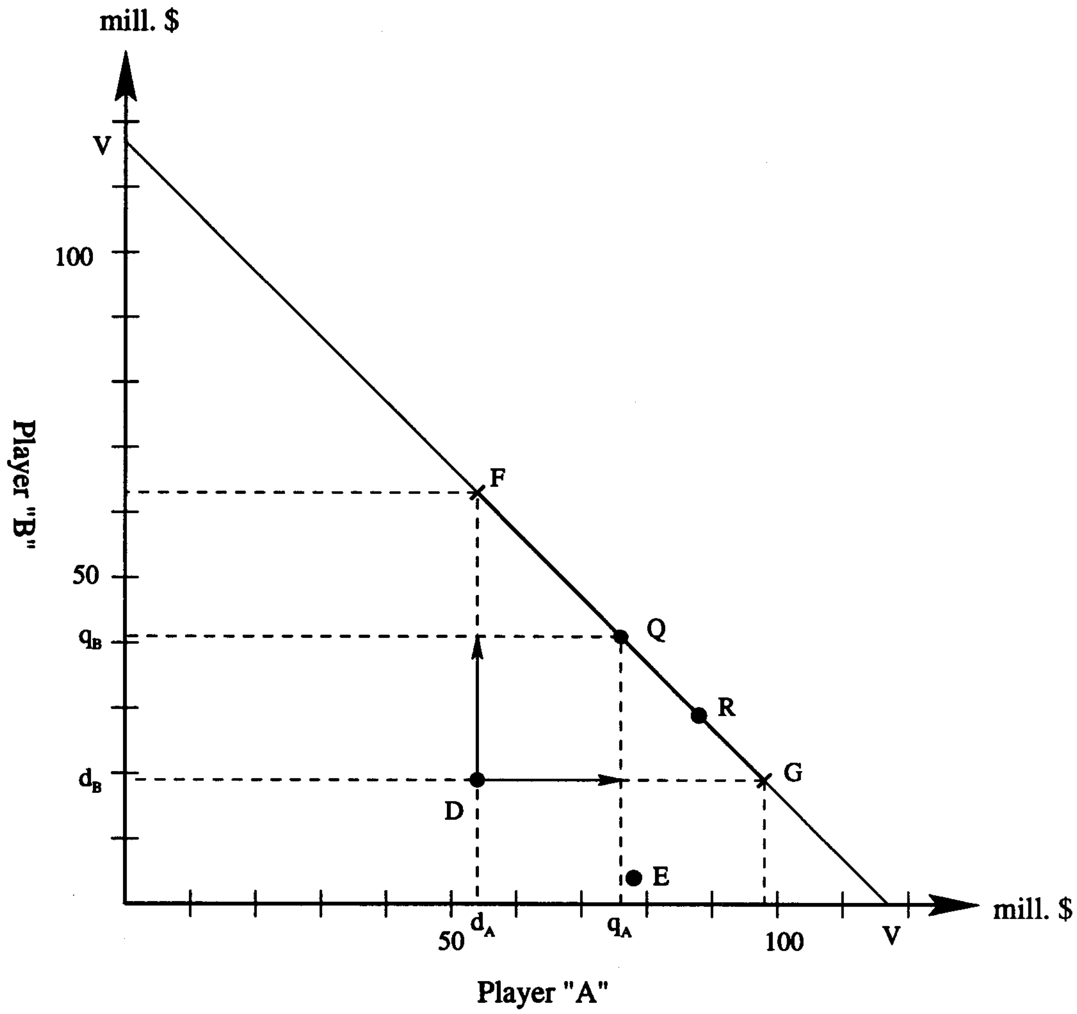

Figure D.4: Distribution of the project value

		-			-	2	
		2			3	-	
		1			-	1	

Figure D.5: Optimal development with an impermeable zone

boring lease is impossible. The drilling program for each of the leases is shown in Figure D.5. For "A"'s part of the reservoir the NPV is \$ 78.1 mill, and it is optimal to drill 3 wells, use platform 2 and 3000 days of production. The NPV of "B"'s part of the reservoir is only \$ 3.8 mill, and it is optimal to drill two wells, choose platform 1 and 1200 days of production. This solution is illustrated by point E in Figure D.4.

As in the non-cooperative case we assume the oil field is developed with two platforms, and the total value will therefore be considerable lower than in the cooperative case. But it is interesting to see that without oil migration between the leases, the total NPV of the reservoir is \$ 81.9 mill, i.e. \$ 10.0 mill higher than the NPV of the two threat strategies. The main reason for this is that only 5 wells are drilled, and both platforms have lower capacity than in the non-cooperative game. When there no longer is competitive extraction of the reservoir, the total investments in production capacity is lower, and the time of operation is longer.

Compared to the non-cooperative solution, owner "A" has increased his payoff with \$ 24.6 mill while "B"'s payoff has decreased with \$ 14.6 mill. This result illustrates that if the project value should be distributed according to economic value, this would be of benefit for the owner with the larger lease. If the economic value is the starting point for the negotiations, it natural to think of two ways to distribute the surplus created by cooperating: distribute the surplus in equal shares or according to recoverable oil. If the surplus (\$ 44.8 mill.) is distributed in equal shares "A" gets \$ 95.5 mill and "B" gets \$ 21.2 mill. Compared to the disagreement solution D, both players are better off, but player "B" only gets \$ 2.8 mill. more by cooperating. If the surplus is shared according to recoverable oil, "B" gets only $ 12.5 mill. Because this is less than he gets from the disagreement solution, this is not a feasible payoff in the cooperative game.

D.5.4 Discussed Bargaining Solutions

We will here list the different bargaining solutions discussed in this section:

- **Nash bargaining solution** - Each owner get their threat plus 50 % of the surplus:
  "A": $ (53.5 + 22.4) mill = $ 75.9 mill
  "B": $ (18.4 + 22.4) mill = $ 40.8 mill

- **Recoverable oil** - Each owner's share of the total NPV is determined from the amount of recoverable oil:
  "A": $ 87.5 mill
  "B": $ 29.2 mill

- **Economic value and equal shares** - Each owner get the isolated lease' value and the surplus is divided in two equal shares:
  "A": $ (78.1 + 17.4) mill = $ 95.5 mill
  "B": $ (3.8 + 17.4) mill = $ 21.2 mill

- **Economic value and recoverable oil** - Each owner get the isolated lease' value and the surplus is shared according to recoverable oil:
  "A": $ (78.1 + 26.1) mill = $ 104.2 mill
  "B": $ (3.8 + 8.7) mill = $ 12.5 mill.

D.6 Summary and Further Research

We have in this paper discussed the common pool problem arising when two surface lease owners have access to the same oil reservoir. The discussion is based on a numerical example. By use of a mixed-integer optimization model we have found the optimal development decisions for the whole reservoir. Obviously this solution gives the highest total payoff, but disagreement about how the total value should be allocated between the two parties, may make it impossible to reach a cooperative solution. We have shown that there exists a unique Nash-equilibrium for the non-cooperative normal form game. This disagreement solution also illustrates the resulting over-investment by the two parties. The non-cooperative solution is a starting point in our analysis of possible bargaining solutions and how the negotiations is affected by the possibility of disagreement.

The analysis in this paper is of course restricted to the specific numerical example, and further research may be to examine a larger group of reservoirs and different locations of the potential wells. These experiments have been performed under the assumption of complete information about the reservoir and future oil price. In a real world situation the information about the reservoir is limited, and it is reasonable to believe that this uncertainty will affect the negotiations. In addition will often the information be asymmetric distributed between the players. How uncertainty will affect both the non-cooperative and cooperative solutions of this reservoir management game seem like interesting topics for further research.

References

[1] K. Aziz and A. Settari. *Petroleum Reservoir Simulation.* Elsevier Applied Science, London and New York, 1979.

[2] CPLEX Optimization, 930 Tahoe Blvd., Bldg 802, Incline Village, NV 89451-9436. *Using the CPLEX Callable Library and CPLEX Mixed Integer Library*, 1993.

[3] T. R. Douglass. Linear programming optimizes unit negotiations. *Journal of Petroleum Technology*, pages 330–336, March 1986.

[4] Å. Hallefjord, H. Asheim, and D. Haugland. Petroleum field optimization - comments on models and solution techniques. CMI-report 862331-1, Chr. Michelsen Institute, Fantoft, Norway, 1986.

[5] D. Haugland, Å. Hallefjord, and H. Asheim. Models for petroleum field exploitation. *European Journal of Operational Research*, 37:58–72, 1988.

[6] T. W. Jonsbråten. *Optimeringsmodeller for utvinning av et oljereservoar.* HAS-thesis, Norwegian School of Economics and Business Administration, Bergen, 1994.

[7] V. Kaitala. Game theory models of fisheries management - a survey. In T. Basar, editor, *Dynamic Games and Applications in Economics*, pages 252–266. Springer-Verlag, Berlin, 1986.

[8] G. D. Libecap and S. N. Wiggins. The influence of private contractual failure on regulation: The case of oil field unitization. *Journal of Political Economy*, 93(4):690–714, 1985.

[9] D. W. Peaceman. *Fundamentals of Numerical Reservoir Simulation.* Elsevier Scientific Publishing Company, Amsterdam, London and New York, 1977.

[10] D. W. Peaceman. Interpretation of well-block pressures in numerical reservoir simulation. *Society of Petroleum Engineers Journal*, June:183–194, 1978.

[11] A. E. Roth. *Axiomatic Models of Bargaining.* Springer-Verlag, Berlin, 1979.

[12] J. Smith. The common pool, bargaining and the rule of capture. *Economic Inquiry*, 25(4):631–644, 1987.

[13] S. N. Wiggins and G. D. Libecap. Oil field unitization: Contractual failure in the presence of imperfect information. *The American Economic Review*, 75(3):368–385, June 1985.