FEDERAL UNIVERSITY OF SANTA CATARINA
GRADUATE PROGRAM IN AUTOMATION AND SYSTEMS ENGINEERING

Eduardo Otte Hülse

ROBUST PRODUCTION OPTIMIZATION OF GAS-LIFTED OIL FIELDS

Florianópolis

2015

Eduardo Otte Hülse

ROBUST PRODUCTION OPTIMIZATION OF GAS-LIFTED OIL FIELDS

Dissertation presented to the Graduate Program in Automation and Systems Engineering in partial fulfillment of the requirements for the degree of Master in Automation and Systems Engineering.

Advisor: Prof. Eduardo Camponogara

Florianópolis

2015

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Hülse, Eduardo Otte
Robust production optimization of gas-lifted oil fields / Eduardo Otte Hülse ; orientador, Eduardo Camponogara - Florianópolis, SC, 2015.
150 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Engenharia de Automação e Sistemas.

Inclui referências

1. Engenharia de Automação e Sistemas. 2. Robust optimization. 3. Production optimization. 4. Gas-lift systems. I. Camponogara, Eduardo. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia de Automação e Sistemas. III. Título.

Eduardo Otte Hülse

ROBUST PRODUCTION OPTIMIZATION OF GAS-LIFTED OIL FIELDS

This Dissertation is recommended in partial fulfillment of the requirements for the degree of "Master in Automation and Systems Engineering", which has been approved in its present form by the Graduate Program in Automation and Systems Engineering.

Florianópolis, April 29th 2015.

Prof. Rômulo Silva de Oliveira, Dr.
Graduate Program Coordinator
Federal University of Santa Catarina

Dissertation Committee:

Prof. Eduardo Camponogara, Ph.D.
Advisor
Federal University of Santa Catarina

Geraldo Afonso Spinelli Martins Ribeiro, Ph.D.
PETROBRAS

Prof. Bjarne Anton Foss, Ph.D.
Norwegian University of Science and Technology

Prof. Agustinho Plucenio, Dr.
Federal University of Santa Catarina

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to the Federal University of Santa Catarina and to Petrobras for letting me fulfill my dream of achieving a Master degree. Through the course of these two years many people contributed somehow to this final dissertation, I thank them all.

To my colleagues and now friends from GOS, our research group, especially Thiago, Marco, Caio, Leonardo, Lauvir and Angelo, I am extremely grateful for the assistance, long discussions, teachings, the trips and also for making the working environment so enjoyable and fun. This is for sure what I will remember the most from this time.

To all my colleagues and friends from other labs, especially Igor, Rodrigo, Fernando and Patrick thanks for providing a nice scape in the least enlightened moments.

To my friends who encourage me taking this step in life in special to Felipe and the Inpulse family.

Also many thanks to my family, Sergio, Elke, Augusto and Leticia, for all the support they have given to me throughout my returning time at the University.

Most of all, I am fully indebted to Eduardo Camponogara, my advisor and friend, for without his early inspiration, coaching, trust and enthusiasm none of this would have happened.

To be uncertain is to be uncomfortable,
but to be certain is to be ridiculous.
(Chinese Proverb)

RESUMO ESTENDIDO

Com a crescente demanda por energia fóssil as operadoras petrolíferas têm buscado determinar planos operacionais que otimizam a produção dos campos em operação para satisfazer a demanda do mercado e reduzir os custos operacionais. Neste contexto, a pesquisa operacional tem se mostrado uma importante ferramenta para determinação dos planos de produção de curto prazo para campos de petróleo complexos. Alguns trabalhos já desenvolveram estratégias para a otimização integrada da produção que visam auxiliar engenheiros de produção e operadores a atingir condições de operação ótimas. Estes avanços científicos atestam o potencial da área de otimização integrada da produção de campos, justificando a busca por estratégias de otimização global e integradas de ativos. Contudo, a incerteza dos parâmetros que caracterizam o reservatório, os poços, fluidos e os diversos processos de produção não vem sendo considerada pelos modelos e algoritmos de otimização da produção diária. Considerando os modelos de produção de curto prazo, estas incertezas podem ser atribuídas a erros de medição, comportamento oscilatório dos sistemas, modelos imprecisos, entre outros. A influência da incerteza dos parâmetros em problemas de otimização tem, desde tempos, sido foco da comunidade de programação matemática. E já foi verificado que soluções de problemas de otimização podem apresentar significativa sensibilidade à pertubações nos parâmetros do dado problema, podendo levar a soluções não factíveis, subótimas ou ambas. Assim, buscando tornar as abordagens de otimização existentes mais confiáveis e robustas às incertezas intrínsecas dos sistemas de produção, esta dissertação investiga a modelagem e tratamento de incertezas na otimização diária da produção e propõe formulações em programação matemática para otimização robusta da produção de poços operados por gas-lift. As formulações representam curvas amostradas através de dados simulados ou medidos que refletem as incertezas dos sistemas de produção. Estas representações levam a formulações robustas em programação matemática inteira mista obtidas pela aproximação das curvas de produção através de linearização por partes. Além disso, este trabalho apresenta os resultados de uma analise computacional comparativa da aplicação da formulação robusta e da formulação nominal a um campo de petróleo em ambiente de simulação, porém considerando simuladores multifásicos amplamente empregados pela indústria do petróleo e gás, que representam a fenomenologia muito próximo da realidade.

O primeiro capítulo apresenta a problemática em que estão

envolvidos os desenvolvimentos realizados nesta dissertação e um resumo dos capítulos subsequentes.

No segundo capítulo alguns conceitos fundamentais são apresentados para a compreensão do trabalho desenvolvido. Este capítulo é dividido em três partes. A primeira parte inicia apresentando brevemente a indústria de petróleo e gás com uma perspectiva histórica, econômica e dos processos envolvidos. Na sequência são expostos conceitos básicos de engenharia de petróleo necessários para o entendimento do sistema de produção utilizado ao longo a dissertação – i.e. gas-lift. Finalmente, o problema de otimização da produção é situado dentro do problema maior, que é o gerenciamento completo das operações de um campo de petróleo, seguido de uma revisão da literatura no que se refere a abordagens clássicas para otimização da produção de campos operados por gas-lift. A segunda parte é uma descrição compacta sobre modelagem de problemas de otimização utilizando programação matemática e na menção dos métodos de solução deste tipo de problema utilizados na parte experimental desta dissertação. A terceira parte começa com uma revisão sobre incerteza em problemas de otimização e sobre as decisões de modelagem enfrentadas quando na presença de problemas de otimização incertos. Na sequência o paradigma de otimização robusta é introduzido e é apresentada uma compilação de alguns dos principais resultados da área de otimização robusta linear. Além disso, ao fim, alguns pontos específicos da teoria de otimização robusta são apresentados pela suas relevâncias para o desenvolvimento da teoria dos capítulos seguintes.

O terceiro capítulo inicia com uma discussão sobre as origens das incertezas nos modelos de produção para então prover uma revisão bibliográfica dos poucos trabalhos que mencionam ou lidam com incerteza em sistemas de produção. Na sequência, a incerteza é examinada na perspectiva do problema de otimização. Um sistema simples é usado para exemplificar a metodologia de otimização robusta desenvolvida nesta dissertação.

O quarto capítulo apresenta dois problemas padrões de otimização da produção, um contendo poços satélites e outro com poços e completação submarina. Para ambos uma formulação em programação linear inteira mista é descrita considerando valores nominais para todos os parâmetros. Então, para cada problema uma reformulação robusta é implementada considerando incerteza nas curvas de produção do poço. A metodologia utilizada para o primeiro problema é a mesma detalhada no capítulo três, e para o segundo uma extensão da metodologia é proposta para poder lidar

com restrições de igualdade incertas.

No quinto capítulo são apresentados resultados experimentais de um problema de otimização da produção de um campo com poços satélites. Os resultados obtidos com otimização clássica (nominal) e com otimização robusta são então comparados em um campo de produção sintético instanciado em um simulador multifásico comercial. A solução robusta se mostrou indicada para cenários de operação mais críticos onde factibilidade e segurança são prioridade.

No capítulo final uma análise dos resultados obtidos na dissertação é feita sob a perspectiva do possível emprego das técnicas desenvolvidas na indústria de óleo e gás. Apesar de à primeira vista os resultados serem conservadores e de sua utilização parecer limitada, existe potencial para a metodologia ser empregada no caso de situações que priorizam segurança. Além disso a metodologia aqui desenvolvida pode servir como ponto inicial para pesquisas e desenvolvimentos futuros. Uma breve descrição de possíveis trabalhos futuros é feita ao final deste capítulo. O apêndice traz a descrição de algoritmos de amostragem de curvas côncavas desenvolvidos para os experimentos numéricos realizados na dissertação.

Palavras-chave: Otimização robusta. Otimização da produção de petróleo. Sistemas de gas-lift.

ABSTRACT

Managing production of complex oil fields with multiple wells and coupled constraints remains a challenge for oil and gas operators. Some technical works developed strategies for integrated production optimization to assist production engineers in reaching best operating conditions. However, these works have neglected the uncertainties in the well-performance curves and production processes, which may have a significant impact on the operating practices. The uncertainties may be attributed to measurement errors, oscillating behavior, and model inaccuracy, among others. To this end, this dissertation investigates how uncertainty might be considered in daily production optimization and proposes formulations in mathematical programming for robust production optimization of gas-lifted oil fields. The formulations represent system-measured and simulated sample curves that reflect the underlying uncertainties of the production system. The representations lead to robust Mixed-Integer Linear Programming (MILP) formulations obtained from piecewise-linear approximation of the production functions. Further, this work presents results from a computational analysis of the application of the robust and nominal formulations to a representative oil fields available in simulation software.

Keywords: Robust optimization. Production Optimization. Gas-lift systems.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

CO_2 carbon dioxide 32, 75

H_2S hydrogen sulfide 32, 75

CC convex combination 19, 90

DCC disaggregated convex combination 19, 90

DLog logarithmic version of DCC 90

DoC Declaration of Commerciality 31

E&P Exploration and Production 28

FDP field development plan 29, 31

GLR gas-liquid ratio 15, 33, 34, 83, 85, 87, 93–95, 98, 106, 123, 124

GOR gas-oil ratio 33

ID inside diameter 83, 84, 106, 123

Inc incremental 90

IPR inflow performance relationship . 33, 34, 37, 38, 83, 85, 106

LHS left hand side 62, 63, 96, 97

Log logarithmic version of CC 90

LP Linear Programming 15, 48, 49, 61–72, 98

MC multiple choice 90

MILP Mixed-Integer Linear Programming 13, 45, 48, 49, 70, 73, 74, 87, 92, 103, 109, 110, 119, 120

MINLP Mixed-Integer Nonlinear Programming 45, 48, 49, 82, 87, 103, 109

NP-complete Non-deterministic Polynomial-time complete . 48

NP-hard Non-deterministic Polynomial-time hard 48, 73

NPV net present value 29

PI productivity index 85, 87, 123

PWL piecewise-linear 44, 87–91, 103, 110, 143

RC robust counterpart 59, 63, 64, 67–69, 71, 73, 96

RDO robust design optimization 57, 63, 71

RHS right hand side 61, 62, 97

SAA sample average approximation 55

SG specific gravity 33, 123

SOCP second-order cone programming 69

SOS2 special ordered sets of type 2 . 45, 90–92, 100, 105, 107, 108, 110, 112, 114, 119

TPR tubing performance relationship 33, 34

WC water cut 15, 83, 85, 87, 94, 95, 98, 106, 123, 124

WPC well performance curve 37, 38, 43, 44, 46, 79, 82

CONTENTS

1 INTRODUCTION

This chapter presents an overview of the purpose and focus of the study, its significance, and how it was conducted. Each of the following chapters is outlined at the end.

1.1 PROBLEM STATEMENT

In the daily operation of an oil field many decisions have to be taken that affect the volume of fluids produced. A decision made by a production engineer or field operator takes into account the capacities of the surface facility in processing, storing, and exporting fluids, the pressures and fluid handling limits in subsea equipment, the restrictions coming from reservoir management, and all these are linked by production models that predict the production of the wells. Many studies have been carried out to propose mathematical tools that help the decision-makers to select the best production plan. A particular type of oil field operation is required when a gas-lift system is used, and there are several works that deal with this problematic including [3, 4, 5, 6, 7, 8, 9, 10]. Each of these studies suggests an approach to solve the daily production optimization problem considering an specific set of variables and constraints, among the many possible scenarios of optimization that arise when gas-lift is present. Although those approaches can consider variation in equipment operating conditions (e.g. failures and valves alignments) they all considered only nominal operating conditions of the wells which, despite being valid for a short time horizon, may vary significantly to the extent of compromising and even invalidating a nominal solution.

Uncertainty in production optimization problems could be found in the definition of the system capacities as well as in the production models. In the latter, the lack of accuracy to predict the system production arise from measurement errors, unmodeled oscillating behavior, and system trends evolving dynamically in time, which hinders the sampling of informative data. All happening in a time scale that could affect a daily production optimization solution. Few works have investigated manners of dealing with uncertainty in the scope of daily production optimization. The problem is in fact twofold: quantifying uncertain data [11], and handling the uncertainty in the optimization problems in order to provide a solution that is at least to some extent immune to data perturbation [12, 13, 14]. Besides the small number of studies on the latter issue, there is only one that to some extent considered uncertainty explicitly in the optimization problem [13], but only for few parameters.

To this end, this work presents a formulation for production optimization which can account explicitly for uncertainty that are inherent to production wells.

1.2 OBJECTIVES AND CONTRIBUTIONS

The research purpose is to develop production optimization models that can produce practical and robust solutions when the operative scenario faces uncertainty in the parameters that characterize reservoirs, wells, or equipment.

The proposed production optimization models are designed based on the theory developed for robust linear optimization [15, 16, 17, 18], extending and adapting it to the specific requirements of this application. The robust production optimization models have their solutions compared to standard production optimization models, which are based on nominal (i.e. expected value) parameter values, in order to highlight the benefits and drawbacks of the optimal robust solutions and to demonstrate the impact of using standard optimal solutions in an uncertain scenario. Experiments are performed by using synthetic but representative oil fields instantiated in a commercial simulator.

The main contributions of this work can be synthesized as:

- The development of a robust optimization methodology that can be applied to several instances of gas-lift optimization problems;

- An analysis of the performance of standard and robust production optimization to oil fields operating under uncertainty, using their optimal solutions in multiphase simulation softwares.

One central assumption of this work is that each uncertain parameter can be modeled as a range of possible values, not requiring a complicated description. Intuitively this provides an easier approach for modeling uncertainty, however, even finding relevant bounds for the parameter values remains a practical and theoretical challenge.

1.3 ORGANIZATION OF THE DISSERTATION

This dissertation is divided in six chapters and one appendix. Chapter 2 reviews fundamental topics that are important for understanding the developments described in the following chapters. It covers concepts of the oil and gas industry, with emphasis on gas-lift systems and daily production optimization, with a review of the main works in this area. It also presents a short overview on optimization modeling and optimization under uncertainty. Still in this chapter, the main results from the theory of robust linear optimization are described. The focus is on selected topics and on important remarks about the theory that are important for the robust production optimization models proposed in this work.

Chapter 3 has a discussion on uncertainty in production systems, with a review of the few works that consider uncertainty in short-term production optimization. It also details some core results of this dissertation. A simplified gas-lift production system is used explain how to model an uncertain production optimization problem and from this uncertain model how to produce the proposed robust optimization model.

Chapter 4 presents the results of the same methodology when applied to two production optimization problems with routing decisions, one for a group of satellite wells, and another to a group of wells with subsea manifolds. Chapter 5 has experimental results comparing standard and robust solutions of one of the models developed in the previous.

A conclusion of the dissertation is presented in Chapter 6, with comments on the results achieved and a perspective of further developments from this work.

Finally, Appendix A presents a proposed algorithm for sampling the production curves that produces a piecewise-linear approximation curve with a given maximum error.

2 FUNDAMENTALS

This chapter provides a compact but comprehensive view of the core areas encompassed in this dissertation. The reader familiar with the subject addressed in one of the sections may skip to the following.

2.1 OIL PRODUCTION

The challenges present in the integrated operations of the production of oil and gas reservoirs motivated the development of this work. This section presents a short overview of the oil and gas industry. The focus is kept on the production of oil and gas wells with special attention to oil fields operated by gas-lift, which are the ones considered in the subsequent chapters.

2.1.1 Petroleum industry overview

By the beginning of the twenty-first century petroleum is still one of the main traded natural resources. It is a very versatile commodity, being used as raw material for many products such as fuels for internal combustion engines, fertilizers, pharmaceuticals, and plastics, to name a few.

Petroleum is a fossil fuel, thus a non-renewable energy source, that consists of hydrocarbons of various molecular weights. Many theories exist about how this hydrocarbons were formed. One well-known perspective is the *organic theory*, which states that oil and gas came from large quantities of organic matter buried underneath sedimentary rock and subjected to intense heat and pressure. Duration and intensity, as well as the type of organic matter, determine the characteristics of the hydrocarbons formed in this process. After oil and natural gas were formed in the source rock, they tended to migrate to the surface by density-related mechanisms. Migration occurs through the pores interconnections in the sedimentary rock or through open fractures within non-porous rocks. Occasionally, the crude hydrocarbons oozed to the surface in form of a seep, or spring. Sometimes structural or stratigraphic traps, covered by impermeable rock layers, trapped the hydrocarbons thousand of meters below the surface. These formations are what we know today as oil and gas reservoirs [19].

Oil seeping out of shallow accumulations is a common, worldwide phenomenon. In fact, even though historical records of petroleum (in an unrefined state) usage go back to the ancient Babylonians and Egyptians, until the nineteenth century, oil and natural gas seeps were the main source of petroleum [20]. The modern history of petroleum began in the mid-1800s with the formulation of its first refining process followed by the first oil-well drilled. The refined oil was initially used to replace whale oil for lighting. However, it was just in the early twentieth century, with the introduction of the internal combustion engine, that demand skyrocketed and petroleum became one of the most valuable commodities traded on world markets.

The growing demand for hydrocarbons over the last century has pushed the exploration of oil and natural gas reservoirs in almost every world location they appear. Nowadays even remote locations have been exploited. From the rainforests of central Africa to the deserts of Libya, from the Arctic region to the deep waters more than 200 km from the coast. Moreover the demand has driven the exploration of even deeper oil reservoirs, as the ones in the *pre-salt* layers in the Brazilian coast, and the exploitation of unconventional petroleum deposit as the shale gas and the bituminous sands.

In addition to that, the more restrictive governmental policies and regulations and more severe environmental laws of today have imposed tougher operating conditions to the entire oil industry. All these new challenges have been pushing the need of operational cost reduction, the better inventory and assets management and the investment in new technologies for the entire productive chain.

Modern oil and gas industry can be divided in three major sectors: *upstream*, *midstream* and *downstream*. Upstream is involved with exploration and production of oil and gas fields, as well as drilling the wells. Midstream and downstream refer to oil processing, storing, transporting, marketing and any oil operation after the production phase through the point of sale.

The present dissertation relates to the upstream sector, more precisely to the production of oil and gas fields.

2.1.2 Exploration and Production

In the petroleum industry Exploration and Production (E&P) is another name for the upstream sector. Next the two main phases of E&P are briefly described.

Exploration

The first step of exploration is based on extensive geophysical and geological studies (specially seismic surveys) performed in order to find the location and evaluate the preliminary production potential of oil or gas deposits. Just after these studies that an exploration, or wildcat, well is drilled. The purpose of this type of well is to determine whether oil or gas exist in a subsurface rock formation and to further estimate reserves and potential production. This is the most expensive phase of the exploration process [21]. If a wildcat well discovers oil or gas, the company might drill several confirmation wells to verify the well tapped a rock layer with sufficient hydrocarbons for the company to develop it economically [20]. The analysis of the oil or gas field potential is also made through a reservoir characterization in terms of *porosity* (size of reservoir rock pores) and *permeability* (level of reservoir rock pores interconnectivity), sedimentary basin analysis and studies of structural geology. There are several tests available in the industry to collect data to infer the aforementioned and other reservoir characteristics [21].

Once sufficient knowledge of the field is gathered, a hydrocarbon recovery feasibility study is carried out. This study considers economical feasibility (net present value (NPV) and capital versus operating cost profiles), field production strategies (e.g., to maximize oil and gas recovery, or fast oil and gas recovery), production equipment options (accounting for operational flexibility and scalability), logistics for the production transportation, the underlying risks to exploit the field (technical, operational, financial, and geopolitical), and the environmental impact [22]. Based on these studies a field development plan (FDP) is then prepared in order to achieve a desired production and the economical objective. This plan comprises all activities and processes necessary to develop a field. It is important to notice that the FDP is continuously updated during the lifetime of an oil field. The reasons are: newer and more accurate data is constantly been gathered from the field, different market conditions are imposed, new technologies appear, among others. Some of the activities and processes are:

- Drilling plans:
  - include drilling schedules, well placement and trajectories, and casing and cementing programs to have the wells at the desired location;
  - they shall be elaborated for the development wells and for

the injection wells. A development well designed to produce oil (usually along with some gas) is called an **oil well**. While a well that is planned to produce mainly gas is termed a **gas well**. Injection wells are the wells drilled to implement *improved recovery techniques*, e.g., waterflooding, immiscible or miscible gas injection, chemical flooding and thermal recovery [22];

- Design and sizing of other equipment and facilities:
  - in order to achieve target well rates and handle the stream variation of produced and injected fluids. Also providing plans for upgrades and additions when required;
  - surface facilities and its equipment, which may include artificial lift surface equipment (e.g., gas-lift compressors, variable speed drives), power generators, multi-phase separators and tanks;
  - subsea equipment that make up the distribution pipeline network connecting the wellheads to the separation facility. Some examples are flowlines, subsea manifolds, headers, risers and umbilicals;
  - downhole equipment that are installed inside the wellbore. They are usually the casing, liners, production tubing and wellhead, but may also include, artificial lift equipment (e.g., gas-lift valves, electrical submersible pump string), packers, and various valves [21];

- Operational policies:
  - are related to long-term strategic decision making, mid-term strategic or operational decision making and even to short-term operational decision making. In this way it comprises the expertise of different management levels, from executive managers, to reservoir engineers and production engineers, on how to operate the oil field;

- Reservoir depletion plan:
  - is designed by a reservoir management team and defines how the reservoir should be exploited over time;
  - is closely related to the drilling plan to conform the schedule with the available facility capacity and injection requirements. It also provides well rate forecasts to aid well design, and multiple production scenarios to aid equipment and facility design.

The cycle of gathering fresh (field, economic, etc.) data to be analyzed, evaluating that analysis and amending the FDP accordingly, and finally implementing the required changes made in the plan and measuring their impact, is a decision-making process. Observe that each part of the FDP has a different cycle frequency. This is directly related to the time constant of the various processes [23]. The operation of an oil field considering all these long, medium and short-term decision-making processes together is here referred as *integrated operations* and is further developed in Section 2.1.4.

When a field is considered suitable for exploitation, and all paper work and licenses required by the competent authorities are submitted and acquired, all initial actions of the FDP can be executed. For instance, according to Brazilian rules, this moment is defined by the Declaration of Commerciality (DoC) of the field. It marks the end of the **exploration phase** and the beginning of the **production phase**.

Production

Before the first drop of oil can be produced the first wells of the drilling schedule have to be constructed. A well is normally drilled using a drilling rig and applying a variety of techniques depending on the location, design and geological conditions [20]. After drilling is finished a careful formation evaluation is performed to determine whether the specific well is economically productive. If the test is positive the borehole is cased and cemented, otherwise is plugged and abandoned.

At this point, if the surface facilities are build, the surface equipment are installed, and the components that make up the distribution pipeline network are laid down, the field is ready to start the production.

In the petroleum industry, production is the phase of operation that deals with bringing well fluids to the surface and preparing them for transportation to the refinery or processing plant [22]. Production comprises the following activities:

(i) preparing and maintaining the borehole for production/injection;

(ii) managing reservoir injections and drainage;

(iii) bringing fluids to the surface facilities;

(iv) separating into oil, gas, water and contaminants streams that are measured for quantity an quality.

where all these activities must be carried on within very strict standards for health, safety, environmental and social responsibility.

Activity (i) are the operations of *completion*, *well services* and *workover*. Completion consist of procedures that enables the well fluid to be produced. For the development wells this can include preparing the bottom hole, perforating and stimulating the production zones, running in the production tubing along with the required downhole tools and installing the wellhead and Christmas tree [21]. After this point the well can be connected to the surface facilities and it is ready to produce. On the other hand, well services and workovers are well interventions. These operations are usually intended to perform routine maintenance or to correct downhole equipment failure, excessive gas or water production, sand production or any problem affecting productivity [21].

Reservoir management (ii) is involved with the reservoir depletion plan which defines how to use primary drive mechanisms to exploit hydrocarbon resources and how, when, or if these mechanisms should be supplemented for additional recovery. This plan is developed with the input from several engineering disciplines.

The activity (iii) is made through a *lifting method*, which allows bringing well fluids to the surface. It also include the routing and choking operations within the distribution pipeline network connecting the outlet valves of the production trees to the tanks and multiphase flow separators in surface facilities. More on lifting methods in Section 2.1.3

Well fluids, usually a mixture of oil, gas, water, and also CO_2, H_2S and sand, must be separated as they reach the surface. Water must be treated before re-injection or disposal, while oil after treated is transported from the well site. The gas can be either compressed and exported, re-injected, used as lift-gas, serve as fuel, or flared by exception. The content of contaminants is continuously measured and actions are taken accordingly. All of these belong to activity (iv).

Among all areas of production, the present work relates to the operations of well fluid flowing, more precisely on how to operate **lifting methods** in the context of **integrated operations**.

2.1.3 Lifting methods

For the well fluid to flow from the inner parts of the reservoir to the surface facilities it has first to move through the reservoir pores interconnections until it finds the reservoir-wellbore interface. There, the fluid enters the wellbore by perforations. In order to flow from the bottomhole to the wellhead it has to overpower both, the pressure resulting from the fluid weight in the production tubing and the pressure drop caused by frictional forces in the well. From this point until it enters the separation equipment the flow still has to overcome additional pressure drops due to friction in pipelines and chocking effect in the choke valves.

The flow rate/pressure relation taking place from the inside of the reservoir to the wellbore interface is represented by the inflow performance relationship (IPR), or just *inflow*. Complementarily, the term tubing performance relationship (TPR), or more generically *outflow* performance, is used for the flow rate/pressure relation existing in between the same interface and the surface equipment [1].

Both inflow and outflow are extremely affected by the fluid characteristics – gas-oil ratio (GOR), GLR, water cut (water-liquid ratio), oil and mixture viscosities, water, oil and gas specific gravity (SG), bubble point, etc. The IPR is also influenced by rock properties (i.e. porosity and permeability), flow regime, the structure of the perforations, formation damage or stimulation, the reservoir internal energy source (drive mechanism and recovery methods), etc [24]. On the other hand, what further affects the outflow are the well geometry, the production tubing size and material (friction effects), the choking valves opening, the surface or subsea pipeline network design, gas and liquid velocities, vertical and horizontal flow regimes, the surface equipment operating pressures and any energy source added in the outflow path.

For example, the relationship between the liquid flow rate, q, and the bottomhole flowing pressure, p_{wf}, for a well with high water cut (greater than 80%), can be approximated as a straight line. However, for wells with lower water cut, with a considerable GOR, and flowing pressures below the bubble point pressure, gas starts coming out the solution already inside the formation. The lower the p_{wf}, the higher the quantity of free gas flowing through the pores and narrow the space to the liquid to flow. In this case, the IPR is commonly approximated by a *Vogel curve* [25], as depicted in Figure 2.1. In the same Figure, a TPR curve is plotted. It represents

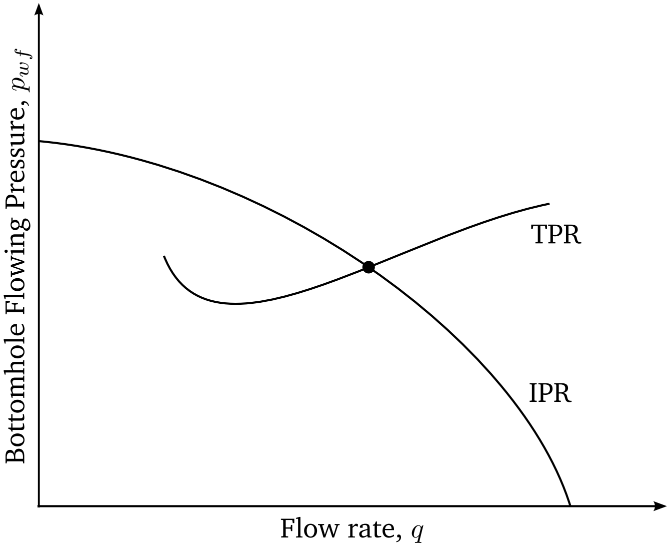


Figure 2.1: Well deliverability.

the curve of a high-GLR fluid (or gas wells). In this scenario, for low liquid flow rates, q, the hydrostatic pressure drop decreases with increasing flow rate. Yet, as q is incremented, frictional pressure drop starts having a more important play and after the minimum p_{wf} it becomes dominant, with a relatively small positive slope. For lower GLR values the curve is likely to be almost flat, with only a very small slope [1]. Another component of the TPR curve is the wellhead pressure, which is usually a design requirement, thus, it has constant value for every flow rate. The IPR and TPR curves are combined in graphics, as it is presented in Figure 2.1, to estimate the well deliverability (expected production rate and bottomhole flowing pressure), which is the intersection of the two curves.

The energy required to initiate and maintain a well flowing can come from within the reservoir (by natural and artificial manners), but it can also be added inside the wellbore.

For recently drilled wells, it may be the case that the energy source is exclusively provided by reservoir drive mechanism, such as depletion drive, water drive, gravity drainage or a combination of them. Reservoirs which the driving mechanisms are inefficient and by consequence retain large quantities of hydrocarbons, may employ a series of process known as improved recovery methods. These are classified as conventional methods – e.g. miscible gas injection and waterflooding; or unconventional methods – e.g. steam injection, or chemical flooding. The former helps to maintain a sufficient formation pressure, while the latter enables reducing the friction between the fluid and the reservoir rocks [22].

A well is called a natural flowing well, if it only uses the internal reservoir energy (commonly referred as *reservoir pressure*) to produce the well fluids. In this case the energy source can come either from driving mechanisms or recovery methods. This is the simplest and basic lifting method to be used in a well. In a natural flowing well the primary elements to control fluid flow are the chocking valves. However, as the reservoir pressure drops and fluid production cannot reach the surface an **artificial lift method** has to be used. Artificial lift methods can also be applied to natural flowing wells in order increase fluid flow rate. Unlike driving mechanism and recovery methods, which only directly affect the inflow performance, the artificial lift methods affect exclusively the outflow performance. The basic function of an artificial lift method is to create the required differential pressure to drive fluids from the reservoir into the wellbore, which is known as pressure drawdown.

There are many artificial lift methods available. Choosing an artificial lift method is a matter of numerous factors like: fluid viscosity, fluid composition and presence of solids; wellbore temperature, depth and geometry (vertical, deviated, or horizontal); well location, whether onshore or offshore production and number of installations; well productivity, installation and operational costs, and accessible infrastructure. Some of the most adopted methods in the industry are as follows.

Subsurface Sucker-Rod Pumps consist of the usual simple combination of a cylinder and piston or plunger with a suitable intake valve and discharge valve for displacing the well fluid into the tubing and to the surface [26]. Thus, it produces the well fluids in batches.

Electrical Submersible Pump is a system composed of multistage centrifugal downhole pumps driven by a downhole electrical motor and auxiliary equipment. Its operating principle is to use the centrifugal pump to generate the required bottomhole pressure drawdown that allows a desired flow rate. The use of a variable-speed drive allows a more flexible production

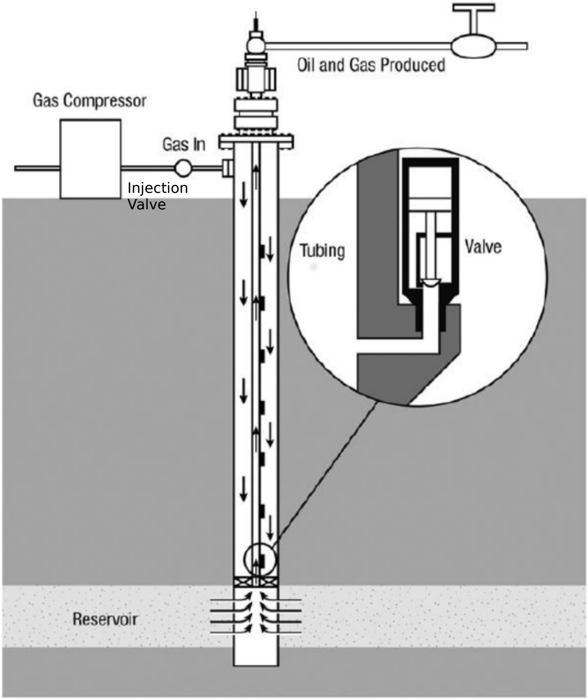


Figure 2.2: Gas-lift concept (Economides et al. [1]).

operation. It is perhaps the most versatile of the major oil-production artificial lift methods [26].

Gas Lift System is a method where an external source of high-pressure gas is injected continuously to the annular space of the well at a given rate and enters the production tubing through a gas-lift valve installed as deep as possible, given the gas compression pressure. The injection gas mixes with the stream of well fluids and decreases the mixture density. Hence, the lower flowing pressure gradient reduces the flowing bottom-hole pressure to establish the drawdown required for attaining a design production rate for the well (Figure 2.2) [26].

The explanation given above is for a continuous gas lift system, which is the lifting method considered in this dissertation. Nevertheless, some wells, usually low production wells, also operate using an intermittent gas lift system, which has a slightly different operating principle [27].

2.1.3.1 Gas Lift System

Gas lift is the artificial lift method in which a high-pressure gas is injected in the production tubing providing energy to lift the well fluids. Since the cost of compression far exceeds the cost of downhole gas-lift equipment, the primary consideration in the selection of a gas-lift system to lift a well, a group of wells, or an entire field is the availability and compression cost of gas [26].

Figure 2.2 illustrates a well operated by gas lift. In this type of system the production rate is controlled – using an injection valve – by the rate of injection gas entering the lowest gas-lift valve. Notice that the choking valves in the production tree, or sometimes at the production unit, can also be used to control the well flow as previously mentioned. The upper gas lift valves are only necessary to help removing the completion fluid in start-up operations, a process which is called unloading a well. Thus, they remain closed during well production.

There are many aspects in designing continuous gas lift systems. Availability of high-pressure injection-gas, well fluid data, perforations depth, completion fluid properties, expected IPR and required wellhead flowing pressure are some of the information used to choose the size, material, amount and setting depth of the gas-lift valves. Although the design of a gas-lifted system plays a key role in the potential well deliverability [1], in the present work, instead, we are interested in the problem of operating an existing system. In the following we focus on operational aspects of already installed gas lift systems.

Bertuzzi et al. (1953 apud Economides et al.[1], 2012) in developing design procedures for gas-lift systems, introduced the *gas-lift performance curve* (a.k.a.[1] well performance curve (WPC)), which is the steady state relationship between the gas-lift rate and the rate of well fluid at the outlet. This makes the WPC the main mean in finding operating points for a gas-lifted well.

Figure 2.3(a) illustrates typical WPCs. Curve **A** represents a natural flowing well. As can be seen, any amount of gas injection alters its production. Curve **C**, on the other hand, represents a well, which requires a minimum gas injection in order to start flowing.

[1] abbreviation for “also known as”

Although curves **B** and **D** refer respectively to the same wells of curves **A** and **C**, the former are subjected to a higher pressure drop in the chocking valve of the production tree.

The characteristic shape of a WPC is depicted in Figure 2.3(b). The shape depends on both, the well IPR and the different outflow performance relationships generated at each gas-lift injection rate. For instance, each point of the curve in Figure 2.3(b) is a mapping of a different well deliverability (see Figure 2.1) at the corresponding constant lifting gas injection rate. The varying outflow is a combination of the hydrostatic pressure head and the friction pressure drop. While gradually increasing the gas injection results in a fluid density reduction with modest friction pressures, continued increase in the lift gas supply makes friction pressure losses in the tubing dominant and the production rate starts declining. This effect divides a WPC in two regions that are highlighted in Figure 2.3(b). In region **(I)** an increase in gas injection produce an increase of production, while in region **(II)** the production decreases as more gas is injected. The transition point between the two regions is the point of optimum injection rate. When not considering production constraints this is the point of maximum production. In reality though, operators do not have an unlimited cost-free supply of lift gas, unpredicted gas slugging may happen in deviated wells, and excessive pressure drawdown may cause formation or perforations damage. All this limits the optimum lift gas injection rate that is then somewhat lower than the peak value, located in the so called economical region. Also note that in region **(I)**, close to point of maximum production, $q_{\text{oil}}^{\text{max}}$, a large increase of gas injection grants only a small gain in oil flow rate.

Deriving inflow and outflow relationships using only physical equations is difficult, as many phenomenological model parameters are unknown and also because of the limited understanding of some of the physics involved. Therefore, pure physical WPC models are uncommon and well performance curves are usually generated numerically by multiphase simulation software. These simulators approximate the real curve using available data (formation, fluid and composition data, wellbore and production string parameters), along with several models (phenomenological and empirical [24, 29, 30]), and applying appropriated calculation methods (e.g. nodal analysis [31]). After a well is producing, well tests are also an important source of information to derive models for the WPC. By the use of operating data curve fitting techniques may be applied.

Many of these simulators have embedded tools for finding the

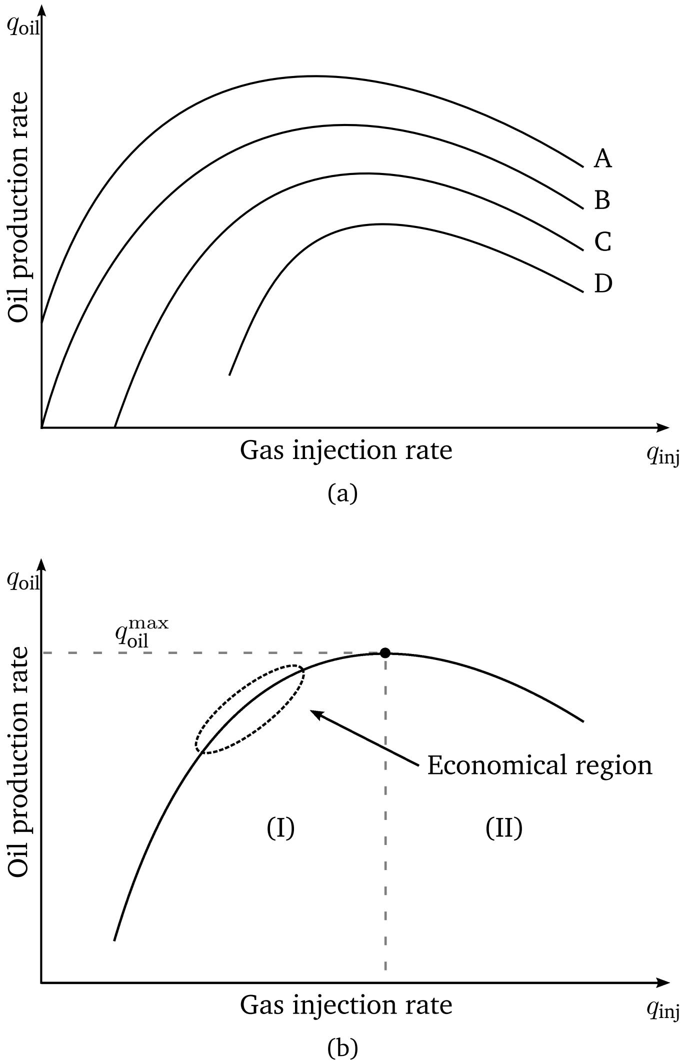


Figure 2.3: Typical well performance curves.

optimal injection point for individual wells in an unconstrained situation. Nevertheless, for optimization under constrained production and considering multiple wells sharing the gas lift source or the distribution pipeline, other tools have to be used. Section 2.1.4.1 gives an overview of such methods in the context of integrated operations.

2.1.4 Integrated operations

The integrated operation of an oil field in the course of the production phase is a complex task. During this phase an operator company has a main goal that usually is to optimize profitability while satisfying a number of constraints (physical, financial, regulatory and human). However, looking at each activity within the production phase, there are many other opposing and consonant goals.

The list of goals includes: (i) minimize capital expenditures and operational costs by efficient investment strategies and optimal assets and supply chain management; (ii) maximize recovery using improved recovery techniques, optimal well location and design, and optimal production/injection targets; (iii) Sustain or increase oil and gas production through well allocation, testing and recompletion, and providing optimal setpoints for surface, subsea and downhole equipment; (iv) reduce the overall risk of decisions; (v) among others [32].

Solving this huge problem is very challenging. Technological innovations in hardware and software in the recent decades, such as tools for digital measurement and data processing, brought the concept of smart fields [33, 6] and real-time optimization [23, 34] to the context of integrated operation of oil and gas producing systems. Saputelli et al. [23] define this concept as a process of measure-calculate-control cycles at a frequency which maintains the system's optimal operating conditions within the time-constant constraints of the system.

The various decision-making processes associated with each objective listed previously can be alternatively organized by their different planning horizon as proposed by Saputelli et al. [35]. This division of the integrated operation problem into subproblems on different time scales limits complexity, allowing the problem to be tractable. Similar hierarchical structure is found in Foss and Jensen [2]. They proposed a time scale division in long-term, medium-term and short-term. The decisions in operating an oil field are then

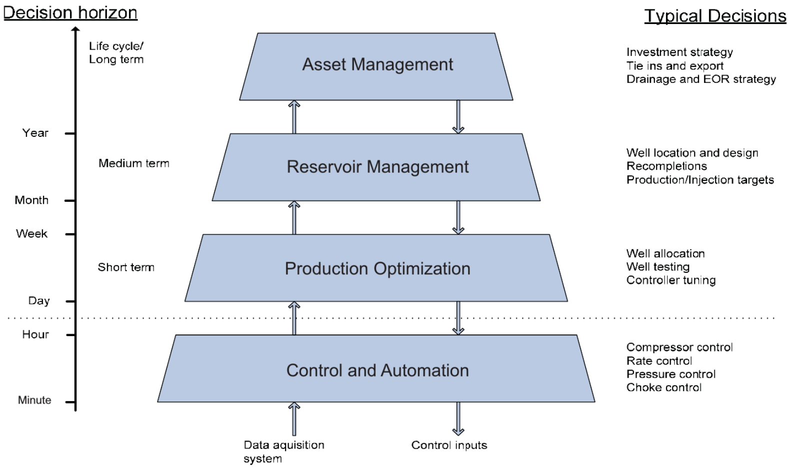


Figure 2.4: Multilevel control hierarchy (Foss and Jensen [2]).

grouped in four levels (asset management, reservoir management, production optimization, and control and automation) within this time frame. Figure 2.4 depicts this idea in what is called a multilevel production control hierarchy. The arrows connecting adjacent layer remind that decisions do influence other levels. Foss and Jensen [2] point that a set of decisions in an upper-level imposes constraints on a lower-level optimization problem to avoid a short-term production strategy harming long-term recovery. Further, lower-level information is used to align upper-level optimization with relevant operational information. It is also mentioned that decisions below the dotted line are taken automatically, while decisions above have human participation, thus being semiautomatic.

In this work, we are interested in the production optimization layer. In this layer usually a production engineer is the one responsible for taking many daily decisions in order to optimize the production of an oil field. The aim of many of the researches found in the area of production optimization is to provide the production engineers a tool that helps them make the best decisions to operate their field. Besides helping optimize production under normal conditions, this kind of tool is particularly important to assist operation under unusual situations, e.g. determining setpoints for the whole field when one equipment fails.

In general, decisions are: determining the setpoints for compressors, choke valves, pressure controllers, and separation equipment, setting manifold alignment, well allocation and establishing downhole equipment operating points, among others. In order to find the set of decisions that optimize production, the field operator must have first a reliable production model. There are also many constraints affecting the scope of this layer. Those arise from the reservoir management layer, e.g. production targets, or from physical capacities of the surface, subsea and downhole equipment.

Notice that each field will have a distinct set of decision variables and a different set of constraints, thus a different production model. In this dissertation we concentrate on typical gas-lifted oil field operative scenarios as the background to develop our work.

2.1.4.1 Production optimization of gas-lifted oil fields

At first sight the physical structure and components in a gas lift system, as presented in Section 2.1.3.1, may look basically the same for each gas-lifted oil field. However, every gas-lift application has its particularities and the variation between them may be rather significant if considering production optimization. Some of the components and their possible variations:

- Separation units:
  - single or multiple units, allowing routing or not;
  - variable or fixed separation pressure;
  - liquid handling capacity;

- Processing, Storage and Export units:
  - water treatment, storage and discharge capacities;
  - platform gas handling capacity;
  - compositional requirements and limits, e.g. contaminants;

- Gas lift compressor:
  - compression capacity;

- Subsea manifolds:
  - could be fixed routing or allow flexible routing of the fluid stream;
  - it may or may not admit flow splitting;
  - pressure and flow rate limits;

- Flow lines:
  - could be individual for each well (i.e. *satellite* or *platform* wells) or could be shared by a group of wells (usually the case when the production of *subsea wells* is directed to a subsea manifold);

- Choke valves:
  - could be installed in any point of the distribution pipeline network. Typically each well has one located at the production tree, or in offshore production a surface choke at the platform;
  - could have fixed or ranged opening;

- Well:
  - many flow assurance issues are related to well characteristics. These may constraint the production to a maximum or a minimum value. Examples are:
    - flow control to avoid gas or water coning;
    - some wells may need a minimum production to ensure a continuous flow;
    - production may be limited to avoid formation and perforations damage;
    - injection valves may require minimum and maximum lift-gas flow;
    - flow limits to avoid slugging.

Additionally, the production of individual wells may be limited due to different reservoir management reasons. Examining the list given above one can see that some of the items are related to equipment settings – we will refer to them as decision variables, while others are related to constraints in the operation of gas-lifted oil fields. The purpose of production optimization is to determine the optimal setpoints for the relevant subset of decision variables, given some criteria (e.g. maximum amount oil produced) and respecting the system constraints.

Production optimization also requires a **production model** which establishes the relationship between the system variables and the production of fluid from the wells. As it was previously mentioned in Section 2.1.3.1 the most important model for gas lift systems is the well performance curve (WPC). Moreover, in the presence of subsea completion, models of pressure drop versus multiphase flow are also necessary because the production stream of different wells can be sharing the same flow line. Models used in the production optimization of gas-lifted wells are normally based on commercial multiphase flow simulators, either by querying the simulator directly or by building tables of simulator predictions, so called proxy-models [11]. Since the surrogate models allow usage of a broader range of mathematical tools to aid solving the production optimization problem, many have been proposed. In the case of the WPC, for example, it is possible to model it as third order polynomial, or with a second order polynomial plus a logarithmic term as proposed by Alarcón et al. [36]. Nakashima and Camponogara [12] fit it with the composition of two exponential curves. Another approach is to use piecewise-linear (PWL) curves as in Camponogara and De Conto; Gunnerud and Foss; Codas and Camponogara; Silva and Camponogara [37, 7, 8, 9], which dismisses the fitting step required to other proxy-models.

Several works have provided different approaches for solving the problem of production optimization of gas-lifted oil fields. The strategies used vary from sensitivity analysis using simulation tools, to heuristic methods and mathematical programming methods. Those works also consider different gas-lift system scenarios. That means, while some may consider gas compression unlimited others do not; some consider that all wells are satellite while others may consider subsea completion, thus contemplating pressure constraints; Some may account for routing decisions and others not and so on. Following are some of the works that have applied mathematical programming methods to obtain optimal production plans.

Redden et al. [3] developed a calculation technique for determining the most profitable distribution of gas to wells in a continuous gas-lift system. The procedure uses production data and vertical flow correlations to predict the production of the individual gas-lifted wells. The algorithm accounts for the constraints on gas compression by iteratively reducing injection based on a ranking of wells. Their results were used to assist production in two operating fields.

Buitrago et al. [4] were among the first in proposing a global optimization algorithm for lift-gas allocation considering compression restrictions. Their algorithm combines a stochastic domain exploration and a heuristic calculation of a descent direction to avoid find suboptimal solutions. They paid special attention to wells that do not respond instantaneously to gas injection (as curve C and D in Figure 2.3(a)).

Kosmidis et al.; Kosmidis et al. [5, 38] addressed a more general production system consisting of wells, manifolds, and separators, and modeling pressure in the flow lines and facilities. The production optimization problem is cast as an Mixed-Integer Nonlinear Programming (MINLP) problem that models pressure as a nonlinear function balancing the momentum of the flow lines. They proposed a method for finding a local optimum of the MINLP program which solves a sequence of MILP problems. While their work is focused on oil field modeling, this paper develops efficient models and algorithms to find a globally optimal operating point for a simpler oil field.

Camponogara and De Conto [37] derived valid inequalities, for which exact and approximate lifting procedures yield stronger inequalities. This was used to solve the problem of distributing a limited rate of lift-gas, while respecting injection bounds and activation precedence constraints. Numerical results show that these cover-based cuts can reduce the number of nodes explored in a branch-and-bound search.

Rashid [39] solved a lift-gas allocation problem with gas constraints. He addresses the effects of interactions between wells by developing an algorithm that iterates until convergence on wellhead pressures. A simulator was used in the loop to validate results, test pressures, and generate curves. Separation constraints are not taken into account.

Campos et al. [6] established the main requirements for integrated production optimization of large-scale oil fields. In the paper they stressed the importance of accurate well models to predict coning effects and integration with real-time optimization algorithms to reach optimal operating conditions. The importance of optimization algorithms capable of dealing with flow routing is also recognized.

Gunnerud and Foss [7] presented an MILP formulation for oil fields in a decentralized structure. This is achieved by modeling clusters of independent wells, manifolds, and pipelines while the separation facilities are centralized in a platform. Piecewise linearization techniques based on special ordered sets of type 2 (SOS2) constraints are used to approximate nonlinear functions. The contribution lies in efficient decomposition of the optimization problem by the application and assessment of the Lagrange relaxation and the Dantzig–Wolfe decomposition.

Codas and Camponogara [8] designed two MILP formulations as alternative to the real MINLP problem of optimal lift-gas allocation with well-separator routing. The first (compact) was obtained using piecewise linearization of the nonlinear curves using binary variables to represent the linearization and the routing decisions. The second formulation (integrated) combines routing and linearization variables together. Both structures are explored to design extended cover cuts. Results showed that cutting planes were more efficient for the integrated formulation, however the compact formulation has less variables.

Silva and Camponogara [9] solved a global optimization problem for gas-lift operations using different piecewise-linear models for WPCs and pressure drop curves with domains spliced in hypercubes and simplexes. The problem consists of wells connected to subsea manifolds which can be routed to different separation units. They focus on a comparison of seven different formulations of the problem.

All the aforementioned works deal with the problem of production optimization of gas-lifted oil fields in a particular way. In this dissertation we take advantage of piecewise-linear models [40, 9] to propose and study formulations of robust production optimization. More specifically, in Chapter 4 the formulation developed in Section 4.1 is based on models of Codas and Camponogara [8] and the formulation from Section 4.2 is based on the formulations of Silva and Camponogara [9].

Production optimization models that to some extent try to handle model uncertainty are not very common. In Chapter 3 we present a literature review in this topic.

2.2 OPTIMIZATION MODELING

Mathematical optimization, or just optimization, is a branch of applied mathematics that deals with the problem of, for a given criteria, selecting the best element within a set of available alternatives. An optimization problem has the following elements:

Decision variables are the variables whose values determine a solution to the problem.

Objective function is an expression that relates the decision variables for an specific goal. This is the function for which a maximum or a minimum value is being sought.

Constraints are expressions relating the decision variables. They define the space of feasible solutions for the problem.

Many applications of science, including optimization, make use of models. The term *model* is usually employed for either concrete or abstract structures used to represent a real object or system. These structures are built for a specific purpose (i.e. help one know, understand, or simulate a subject the model represents). Thus, depending upon the use to which a model is to be put, it can exhibit as many and as precise features and characteristics of the subject as required.

An *optimization model* is designed in order to represent an optimization problem. These models are usually mathematical, in that algebraic symbolism (such as equations, inequalities and logical dependencies) are used to mirror the internal relationships of the problem (i.e. decision variables, objective functions and constraints). Often, practical problems can be modeled in more than one standard or non-standard way. That is because modeling an optimization problem is not a straightforward task, and depends on the model builder's experience, intuition, creativity and power of synthesis [41].

Many optimization models take standard forms. Examples of standard alternatives are simulation models, constraint programming, and mathematical programming. The latter is probably the most commonly used standard type of model [41] and was chosen to represent the optimization problems throughout this dissertation.

2.2.1 Mathematic programming

Mathematic programming is a modeling language largely employed to express optimization problems. It uses mathematical relationships to represent the elements of an optimization problem.

A general optimization problem P can be written in mathematic programming as:

$$\begin{aligned} P: \text{Maximize} \;\; & f(\boldsymbol{x}) \\ \text{Subject to:} \;\; & g_i(\boldsymbol{x}) \leq 0, \;\; i = 1, \ldots, m \\ & \boldsymbol{x} \in \mathcal{X} \end{aligned}$$

where $\boldsymbol{x}$ is the vector of decision variables, $f : \mathbb{R}^n \to \mathbb{R}$ is the objective function, $g_i : \mathbb{R}^n \to \mathbb{R}, i = 1, \ldots, m$ are the constraints of the problem, which limit the space of feasible solutions, and $\mathcal{X}$ is a continuous subset in $\mathbb{R}^n$. Depending on the nature of objective function, constraints, and decision variables an optimization model can be classified within specific classes of problems. Below a list of a few of these classes is presented:

Linear Programming (LP) is a class of optimization problems for which both objective function and constraints are linear in the decision variables. A standard mathematical programming form of this class of problems is given by:

$$\begin{aligned} P_L : \text{Maximize} \;\; & \mathbf{c}^\intercal \boldsymbol{x} \\ \text{Subject to:} \;\; & \mathbf{A}\boldsymbol{x} \leq \mathbf{b} \\ & \boldsymbol{x} \in \mathbb{R}^n_+ \end{aligned}$$

where $\mathbf{b} \in \mathbb{R}^m$ and $\mathbf{c} \in \mathbb{R}^n$ are vectors[2,3] and $\mathbf{A} \in \mathbb{R}^{m \times n}$ is a matrix. This sort of problem is widely used and has efficient algorithms to solve it, as the *Simplex* algorithm and *interior point* methods [42].

Mixed-Integer Linear Programming (MILP) has a similar formulation to Linear Programming, since objective function and constraints must be linear. Nevertheless, in this class of problems some of the decision variables belong to discrete domain.

$$\begin{aligned} P_{MI} : \text{Maximize} \;\; & \mathbf{c}^\intercal \boldsymbol{x} + \mathbf{h}^\intercal \boldsymbol{y} \\ \text{Subject to:} \;\; & \mathbf{A}\boldsymbol{x} + \mathbf{G}\boldsymbol{y} \leq \mathbf{b} \\ & \boldsymbol{x} \in \mathbb{R}^n_+ \\ & \boldsymbol{y} \in \mathbb{Z}^p_+ \end{aligned}$$

where $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{h}$ are vectors and $\mathbf{A}$ and $\mathbf{G}$ are matrices with convenient dimensions. This class of problem is Non-deterministic Polynomial-time hard (NP-hard), because any problem in the Non-deterministic Polynomial-time complete (NP-complete) class can be reduced to a P_{MI} in polynomial time, for example, the satisfiability problem [43]. However, there are advanced algorithms for solving mixed-integer programs including, branch-and-bound, branch-and-cut, branch-and-price, and cutting-plane methods [44].

Mixed-Integer Nonlinear Programming (MINLP) is a more gen-

[2]In this dissertation we consider vectors as column matrices.
[3]$\mathbf{A}^\intercal$ denotes the transpose of matrix $\mathbf{A}$.

eral form of optimization problem.

$$
\begin{aligned}
P_{NL} : \text{Maximize} \quad & f(\boldsymbol{x}, \boldsymbol{y}) \\
\text{Subject to:} \quad & g_i(\boldsymbol{x}, \boldsymbol{y}) \leq 0, \ i = 1, \ldots, m \\
& \boldsymbol{x} \in \mathcal{X} \\
& \boldsymbol{y} \in \mathcal{Y}
\end{aligned}
$$

where f and g_i, $i = 1, \ldots, m$, are nonlinear real-valued functions on $\mathbb{R}^{n+p}$ and not necessarily continuous. $\mathcal{Y}$ is a finite subset in $\mathbb{Z}_+^p$ and $\mathcal{X}$ is a continuous subset in $\mathbb{R}^n$. There is no general method to solve problems of this class and one must cope with each problem individually in order to find an efficient algorithm, when it exists. A difficulty of any algorithm proposed to solve this class of problems – and any subclass containing a nonlinear function – is the guarantee that it can be solved to global optimality and not only to provide an optimal solution within a near neighborhood (local optimum) [45]. A particular subclass is the quadratic programming, which has a relative easy solution when problem is convex.

Keep in mind that the same optimization problem may be modeled into distinct mathematical programming models, which could even belong to different mathematic programming classes. This illustrates the importance of modeling in optimization problems.

2.2.2 Solving optimization problems

The solution to an optimization problem lies within the feasible space delimited by its constrains. Yet, for the majority of the optimization problems, enumerating all feasible solutions is not a realistic strategy to find the optimal solution. Therefore more sophisticated alternatives have been developed. The task of finding a strategy to solve an optimization problem modeled in mathematical programming is directly related to its classification (LP, MILP, MINLP, etc.). A set of operations for solving a particular class of problem or model is known as an *algorithm*. Linear programming and integer programming, for example, have well-known algorithms for solving each problem. For instance, the most efficient general methods apiece are the revised simplex algorithm and the branch-and-bound algorithm. It is beyond of the scope of this dissertation to get into the details of these algorithms. They are well described in Vanderbei [46] and Wolsey and Nemhauser [44] respectively.

Solvers

For frequently used algorithms there are very mature computer programs available off-the-shelf. Implementation of many of these algorithms are commonly grouped into a computational package and are commercial available for solving mathematical programming models. These packages, also referred as *solvers*, offer some advantages. They are generally flexible to use, providing many procedures and options (e.g. automatic problem reduction) to the user [41]. State-of-the-art solvers represent many person-years of programming effort. In addition to algorithms that provide optimal solutions (or at least certificated suboptimal solution) they also include heuristic that help finding quality solutions efficiently. Though this eases the model builder from the task of programming the computer to solve the model, it may complicate a fair comparison between different model approaches for some computational indexes, particularly regarding efficiency as time to solve an optimization problem. Since the aim of this work is to confront optimization models, but mainly to compare its solutions – variables and objective values at optimality, all results are obtained using commercial solvers for the sake of simplicity.

Modeling languages

A number of special purpose high level programming languages, known as algebraic modeling languages, exist to aid users to structure and instantiate their model of an optimization problem into solvers. They function as an interface between the model builder and the solver, greatly reducing the effort and increasing the reliability of formulation and analysis [47]. Modeling languages can provide many advantages for the user as: a more natural input format, higher level data structures, easier debugging, syntax based on conventional mathematical notation, etc. For those reasons all optimization models used in the experiments are implemented through a modeling language. For reference, we chose the AMPL® modeling language [48], but any other similar one could have been used.

2.3 ROBUST OPTIMIZATION

From the second half of the twentieth century, mathematical optimization has found widespread applications over different areas as engineering design, planning and scheduling of production systems, logistic problems, the study of physical and chemical systems, and more. Already at the beginning, many researchers in applying optimization techniques to those problems identified that uncertainty is almost always present and that its effects can be a key aspect in finding an optimal solution for real world problems. This section provides a short overview of optimization under uncertainty leading to an examination of one of its subfields, that is robust optimization. Finally important results specific for robust linear optimization are introduced.

2.3.1 Optimization under uncertainty

The paradigm of making a decision without knowing what their full effects will be, distinguishes the field of optimization under uncertainty from other branches of optimization. In decision theory, *uncertainty* is defined as a state that limits the capability of predicting a future outcome, or describing the current state of a system. What makes a system uncertain is particular to each system. Natural and technological systems are almost always confronted with uncertainty, and its possible sources may include the lack of proper measurement/estimation systems, measurement/estimation errors (in physical systems, technological process, environmental conditions ...), imperfect manufacturing process and material properties, implementation errors (coming, for example, from the impossibility to implement a solution exactly as it is computed), oversimplified models, etc [16, 49]. In real-world applications of optimization it is common to replace uncertain data by its nominal value – e.g. an approximated or expected value. However, studies like the one carried out by Ben-Tal and Nemirovski [16], show that optimal solutions for those problems may become severely infeasible or completely meaningless from a practical viewpoint, if the nominal data is slightly perturbed. Consequently, in optimization, there is a real need for methodologies that can account for the effects of uncertainty.

The pioneering work in the area of optimal decision-making under uncertainty emerged in the nineteen fifties with the works of Dantzig; Bellman; Charnes and Cooper (1955, 1957, 1959 apud Sahinidis[53], 2004), which are the seminal works for, respectively, stochastic programming, dynamic programming and optimization under probabilistic constraints. Since then, new procedures and techniques, for both modeling and algorithmic, have been developed in the field of optimization under uncertainty to support the many areas of application where such problems appear.

At the modeling side, one wants to consider the effects introduced by uncertain elements into the underlying optimization models and evaluate how the unknown values affect the quality of the solution obtained. In this sense, a variety of peculiar modeling philosophies were established to deal with the paradigm needs. Hence, when modeling under uncertainty additional *concepts* have to be considered. Below four of these concepts are presented. They concern with how uncertainty is expressed in the optimization problem, how to model uncertain elements, how to deal with uncertain objective functions, and also with how to treat uncertain constraints.

Initially let's take a basic deterministic optimization problem, P, in the $\mathbb{R}^n$:

$$\begin{aligned} P : \text{Maximize } \; & f(\boldsymbol{x}) \\ \text{Subject to: } \; & g_i(\boldsymbol{x}) \leq 0, \; i = 1, \ldots, m \\ & \boldsymbol{x} \in \mathcal{X}, \end{aligned}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, and $\mathcal{X} \subset \mathbb{R}^n$. Problem P does not regard for the presence of uncertainty in its model. This bring us to the **first** concept, that is, the modeler has to somehow represent the effects of uncertainty in the decision-making model. This in itself is a challenging subject, particular to the model structure of each application. There is no general rule to help introducing uncertainty in an optimization model. Let's consider for now that the designer can come up with a representation $\tilde{P}$, alternative to P. Such formulation would have an objective function $f(\boldsymbol{x}, w)$ and a feasible set given by $\{\boldsymbol{x} : \boldsymbol{x} \in \mathcal{X} \text{ and } g_i(\boldsymbol{x}, w) \leq 0, \; i = 1, \ldots, m\} \subset \mathbb{R}^n$, in which w stands for something uncertain – a parameter ranging over a space Ω representing its possible states.

This generic description leads us to the **second** concept. It concerns the characteristics of the aforementioned space Ω, which is directly related with the model of the uncertain parameter. Rockafellar [54] describes three modeling approaches for w.

Deterministic modeling could refer to: (i) a mathematical formulation in which uncertainty plays no role as in model P; (ii) a mathematical formulation that replace the random element w by some particular estimate $\widehat{w}$ (commonly the expected value of w, $\mathbb{E}[w]$). This produces a deterministic version of $\tilde{P}$ as,

$$\begin{aligned} \tilde{P}_{\text{det}} : \text{Maximize} \quad & f(\boldsymbol{x}, \widehat{w}) \\ \text{Subject to:} \quad & g_i(\boldsymbol{x}, \widehat{w}) \le 0, \ i = 1, \ldots, m \\ & \boldsymbol{x} \in \mathcal{X}. \end{aligned}$$

Those are widely used alternatives to cope with uncertainty (in fact, to avoid it) in practice. The reasons to use such models can be the impossibility to add the effects of uncertainties in the model – option (i), or the lack of adequate data to infer w, or even to software limitation for getting numerical solutions for problem $\tilde{P}$ – option (ii).

Two common approaches to deal with uncertainty are sensitivity analysis and scenario optimization. Both use model $\tilde{P}_{\text{det}}$, but in different ways.

- *Sensitivity analysis* is performed to study the effect of changes in the parameter values *after* a solution is obtained from a problem in the form of $\tilde{P}_{\text{det}}$. Thus, fixing $\boldsymbol{x}$ to be $\boldsymbol{x}^\star$ (solution to $\tilde{P}_{\text{det}}$) and varying w in its full range of possible values to see how the objective function value is perturbed. This analysis is very limited to observe full implications of uncertainty in the underlying problem. Some drawbacks and limitations are discussed by Wallace [55].
- *Scenario optimization* uses a number of scenarios (certain representations of how the future might unfold) that can be generated by simulation or using some kind of probabilistic model. Problem $\tilde{P}_{\text{det}}$ is then solved for every scenario and their results – optimal solution, $\boldsymbol{x}^\star$, and respective objective function value – are grouped in a heuristic manner. Also based on heuristics, these groups of solutions are inspected to decide the most appropriate solution. Nevertheless, the approach can easily lead to faulty conclusions [54].

Stochastic modeling is applied when uncertain elements, i.e. w, in a problem can be modeled as random variables. That means Ω is a probability space with a known probability distribution. Notice that this distribution could be obtained from statistical data or simply chosen by educated guess. Either way, the mathematical treatment is the same.

Range modeling is an alternative modeling approach to the uncertain elements. Instead of assigning them a probability distribution, which is often unavailable, one can view its values as restricted to a given interval (*range*). In this case Ω is a closed uncertainty set with no particular distribution. This is the type of modeling used in robust optimization.

$\tilde{P}$ is a problem comprising uncertainty, but it is not yet an "optimization problem" [54]. To achieve that, one has to provide a prescription of exactly *what* should be maximized over *what*. In order to make this point clearer, let's take each "what" at a time.

The first one relates to the objective function of problem $\tilde{P}$ and brings us to the **third** concept. When the decision has to be made before any information about w is known other than the range it belongs or its probability distribution (depending on the type of uncertainty modeling), one might even know the space over which the maximization is being performed – namely $\mathbb{R}^n$ or a subset – anyhow the objective is still not well defined. For instance, choosing a feasible $\boldsymbol{x}$ does not provide a single objective function value, but an uncertain value $f(\boldsymbol{x}, w)$, since $w \mapsto f(\boldsymbol{x}, w)$ on Ω. There are two prevailing approaches to transform the objective function and thereafter be able to maximize it over $\boldsymbol{x}$. The idea is to change its result from a *function* depending on $\boldsymbol{x}$ to a *number* depending on $\boldsymbol{x}$ [54].

Worst-case approach considers only the worst outcome resulting for choosing a given $\boldsymbol{x}$. It does not weight for the likelihood of the events. This approach can be used along with stochastic or range modeling of the uncertainties, nevertheless it is usually related to the latter. Therefore, the objective function $f(\boldsymbol{x}, w)$ is replaced and the problem becomes of maximizing

$$\hat{f}(\boldsymbol{x}) := \inf_{w \in \Omega} f(\boldsymbol{x}, w).$$

This has the same format of Wald's minimax (maximin) model [56] and it is the approach used in robust optimization.

Stochastic approach can be used when Ω is a probability space. In this case, for a given $\boldsymbol{x}$, $f(\boldsymbol{x}, w)$ can be understood as a random variable inheriting its distribution from w. There are different manners to convert a random variable (dependent on $\boldsymbol{x}$) to a numerical quantity (also dependent on $\boldsymbol{x}$). The most common alternative is to use the "average outcome", that means, replacing the random variable by its expected value. Then, the objective function takes the form,

$$\hat{f}(\boldsymbol{x}) := \mathbb{E}_w[f(\boldsymbol{x}, w)],$$

where the expectation is given by a weighed sum or an integral, if the underlying probability space is discrete, or continuous, respectively [4]. Notice that more complex schemes may be required to have a more realistic and reliable model. For instance, it is possible to have an alternative objective function given by $\varphi(\boldsymbol{x}) := \Phi(\hat{f}(\boldsymbol{x}), \hat{\hat{f}}(\boldsymbol{x}))$, where $\hat{\hat{f}}(\boldsymbol{x})$ denotes the variance of $f(\boldsymbol{x}, w)$, $\hat{f}(\boldsymbol{x})$ the mean, and Φ provides a formula for balancing the two quantities. It may also be appropriate in some situations, for example, to consider that $\boldsymbol{x}$ affects the probability distribution of w.

Well, we are still left to clarify over "what" we are maximizing in problem $\tilde{P}$. So, the **forth** concept is associated with the definition of the feasible region of the problem. Here it is worth taking a more general formulation for the constraints specifying the feasible region. Conventionally, a system of constraints can be represented by inequality constraints, equality constraints, and a subset of $\mathbb{R}^n$, as described below.

$$\begin{cases} g_i(\boldsymbol{x}, w) \leq 0, \ i = 1, \ldots, m \\ h_j(\boldsymbol{x}, w) = 0, \ j = 1, \ldots, q \\ \boldsymbol{x} \in \mathcal{X} \end{cases}$$

In other branches of optimization, constraints can be modeled in various equivalent forms. It is common to transform equalities to equivalent inequalities and the opposite. However, in the presence of uncertainty, conversions not always result in equivalent formulations and should be avoided whenever possible. Ben-Tal et al. [17] provide an explanatory example of this phenomenon. Following, it is presented how uncertainty affects the modeling of equality and inequality constraints.

Equality constraints can be used to model different aspects of a system. For example, they can be employed to express a state

[4] A problem involving continuous random variables can be discretized in various ways to avoid the multi-variate integration when evaluating the objective function. One possibility is to use *simulation*, e.g. sample average approximation (SAA).

variable through a corresponding state equation, or to model physical-based constraints, such as $F = ma$ (in these cases at least one variable is dependent and one is independent). They are also used to represent physical laws, as the principle of mass conservation, or as a designer-imposed constraint, such as dimensional constraints (for these cases all variables are independent). Looking at the equality constraints in the set of equations above, for $j = 1, \dots, q$ it is being demanded that $h_j(\boldsymbol{x}, w) = 0$. However, requiring $\boldsymbol{x}$ to make $h_j(\boldsymbol{x}, w)$ assume a constant value zero is in conflict with the fact that the value $h_j(\boldsymbol{x}, w)$ could vary with respect of the unknown value of w. This gives an image of how tricky are equality constraints in such scenario.

In general these constraints are very problematic when contaminated with uncertainty. Ben-Tal et al.; Rockafellar [17, 54] suggest to refrain from this type of structure when possible. But this is not always the case, so one way to get rid of equality constraints is through substitution. For instance, usually state variables (and its state equations) can be substituted into inequalities constraints, thus eliminating these constraints. Substitution cannot always be applied, so an alternative is to use the constraint relaxation approach. Relaxation can be implemented in different manners, as using slack variables to relax the equality constraints, or replacing the equality constraints by two inequalities. Another possible approach is satisfying the equality constraint at its mean value. Although in a different (but related) context, Rangavajhala et al. [49] provide an interesting perspective on how to classify and deal with equality constraints.

Inequality constraints, as presented previously, are generally the same as requiring that,

$$\hat{f}_i(\boldsymbol{x}) \leq 0, \text{ where } \hat{f}_i(\boldsymbol{x}) := \sup_{w \in \Omega} f_i(\boldsymbol{x}, w).$$

From this perspective the inequality is seen to have a "worst-case" aspect. This is a conservative approach that reveals one end side of the intrinsic dilemma when optimizing under uncertainty, that is, the trade-off between optimality and robustness.

When the restrictions are not so severe, thus a worst-case condition is not truly intended and specially if uncertainty has a stochastic modeling, there are alternative modeling options for the inequality constraints. One of them is to work with finite penalty expressions incorporated to the objective function in substitution to the hard constraints, with their infinite penalty aspect. Another approach is to use chance constraints. The idea, then, is to relax the hard constraints probabilistically [54]. The general structure is in the form,

$$\text{prob}[f_i(\boldsymbol{x}, w) \leq 0] := \beta_i,$$

where β_i is the probability that the inequality holds for a given $\boldsymbol{x}$. For latter modifications the reformulated optimization problem is called to be over soft constraints.

Each alternative approach to restructure the underlying constraints has its own drawbacks and advantages, in both, technical ramifications and modeling aspects.

The four aforementioned concepts give a sense of how models of decision-making under uncertainty may differ from deterministic ones. Notice, however, that the modeling philosophies used when optimizing under uncertainty bring in also other concepts. For example, throughout this dissertation we contemplate only single-stage uncertain problems, thus, logically, we consider that all the uncertainty reveals itself afterwards, when all decisions are already made – otherwise they would not be *uncertain* problems after all. However, for two-stage and multi-stage problems the concept of precedence between fixing the decision and making an observation, which make the uncertainty dissipate, is key to reformulate and solve the problem.

It is also important to notice that the choice between the different alternatives for each concept is a responsibility of the decision-maker when analyzing the basic problem; there is no general recipe. Mixing the various formulations presented earlier to manage the uncertainties produce families of problems with totally different structures. It is to handle this families of problems that different subfields of optimization under uncertainty have been developed. Some have a broader applicability such as stochastic programming (linear, integer, nonlinear), probabilistic (chance-constraint) programming, stochastic dynamic programming, fuzzy programming [53], and robust optimization (linear, convex) [17], while others are application-specific as robust design optimization (RDO) which is used in engineering design problems [49].

In this work we use the concepts behind robust linear optimization to design the models for production optimization of gas-lifted oil fields considering the presence of uncertainty.

2.3.2 Robust optimization

The idea of robustness of a solution is shared by many areas in applied science, as robust statistics, robust control and robust optimization. The connection between them are not always explicit, but they all rely on approaches that provide solutions which, at least in some level, are immune, when the scenario of the problem is uncertain. The conceptual roots behind robust optimization trace back to the first half of the twentieth century, precisely to Wald's contributions to modern decision theory, with his minimax model [56]. However, the field of study established today as robust optimization is consider to have started only in 1973 with the work of Soyster[57], though it was just with the terminology used in researches made in the 1990s – for convex programming [16, 58, 15, 59, 60] and integer programming [61] – that optimization for worst-case value of parameters within a set has become effectively known as "robust optimization" [62]. Now, let's briefly examine the main concepts of the robust optimization paradigm.

From the previous section one can extract that uncertainty in an optimization problem, threatens the relevance of its solution in two facets:

(i) on one hand, when implemented, the calculated optimal solution, which is thought to be feasible, might actually violate the problem constraints;

(ii) and on the other hand, when indeed feasible, the solution might be far from the true optimal value.

As it was pointed out in Section 2.3.1, there are many techniques available to deal with these issues. Also in the previous section, it was discussed that for each of these techniques, different assumptions are made on the characteristics of the uncertainty and on how uncertain elements are structured in the optimization problem. The list below present the key assumptions made for robust optimization [17].

A.1. The uncertain data in a problem is modeled in terms of ranges. The data can assume any value within a given uncertainty set, which has an unknown probability distribution.

A.2. All the decision variables of the problem represent "here and now" decisions – the problem must be solved, and their values must be chosen, before the actual data "reveals itself".

A.3. The constraints of the problem are "hard". Meaning that violation of the constraints cannot be tolerated, when data is within the prespecified uncertainty set. Any solution of the problem is then called robust feasible.

A.4. The objective is also worst-case oriented. So the optimal solution is the least worst for all possible uncertain data realization.

By itself, the robust optimization methodology presented above can be applied to every uncertain optimization problem where one can separate numerical data (that can be partly uncertain) from the structure of a problem (that is known in advance and common for all instances of the uncertain problem).

In robust optimization, an uncertain optimization problem is defined as a collection $\{\max_{\boldsymbol{x}}\{f(\boldsymbol{x},\boldsymbol{w}) : g_i(\boldsymbol{x},\boldsymbol{w}) \leq 0, \forall i\} : \boldsymbol{w} \in \mathcal{W}\}$ of optimization problems of a common structure with data $(\boldsymbol{w})$ varying in a given uncertainty set $\mathcal{W}$, with its characteristics given by A.1. The solution of this collection of problems should be a fixed vector that must remain feasible for all constraints, regardless of which values reveal later as real for $(\boldsymbol{w})$, given only the values are within $\mathcal{W}$. Thus, A.2 and A.3 will hold. In order to assure that the objective is worst-case oriented (A.4), it is sufficient to evaluate, within the robust feasible solutions $\boldsymbol{x}$, its smallest value $\inf\{f(\boldsymbol{x},\boldsymbol{w}) : \boldsymbol{w} \in \mathcal{W}\}$ over all realizations of the data from the uncertainty set. Thus, the best possible robust feasible solution is the one that solves the optimization problem,

$$\max_{\boldsymbol{x}} \left\{ \inf_{\boldsymbol{w}\in\mathcal{W}} f(\boldsymbol{x},\boldsymbol{w}) : g_i(\boldsymbol{x},\boldsymbol{w}) \leq 0, \forall i,\ \forall \boldsymbol{w} \in \mathcal{W} \right\},$$

or, which is equivalent, by epigraph reformulation, to the optimization problem,

$$\max_{\boldsymbol{x},t} \{t : t - f(\boldsymbol{x},\boldsymbol{w}) \leq 0, g_i(\boldsymbol{x},\boldsymbol{w}) \leq 0, \forall i,\ \forall \boldsymbol{w} \in \mathcal{W}\}.$$

The latter problem is called the robust counterpart (RC) of the original uncertain optimization problem. The feasible/optimal solutions to the RC are called robust feasible/robust optimal solutions to the uncertain problem. The robust optimization methodology, in its simplest version, proposes to associate with an uncertain problem its robust counterpart and to use, as “real life” decisions, the associated robust optimal solutions [17].

Notice that the robust counterpart problem may have infinite constraints, i.e., is a semi-infinite problem [63], and as such it looks computationally intractable. In fact, in general, the robust problem is intractable, however, many interesting classes of problems admit efficient solution and much of the literature in the 1990s and 2000s has focused on specifying classes of functions g_i, coupled with different types of uncertainty sets $\mathcal{W}$, that yield tractable problems [62].

In their book, Ben-Tal et al. [17], split the research questions related to robust optimization in three main categories.

(i) Extensions of the robust optimization paradigm. These might be necessary because the assumptions A.1-A.4, which point to a robust counterpart of the original problem, while meaningful in various applications, in some others might not be suitable to reflect the relation between the problem and the uncertain data.

(ii) Both, investigating tractability issues of robust counterparts, and identifying the cases where the robust counterpart of an uncertain problem admits a computationally tractable equivalent reformulation. This is considered the main theoretical challenge in the field of robust optimization.

(iii) Applications of robust optimization to specific optimization problems.

The main objective of this dissertation falls exactly in Category (iii). Insights of the importance of the research made in Category (ii) are given in the next section. Category (i) is disregarded in this work and hereafter we continue using the standard assumptions stated earlier for the robust optimization paradigm, although extensions to the paradigm might provide key mathematical tools for further developments of the present work (see Chapter 6).

To further explain the methodology offered by robust optimization and because the framework used herein for optimizing gas-lift oil fields is of linear programming programs, in the next section, the particular case of robust linear optimization is discussed.

2.3.3 Robust Linear Optimization

Linear programs are perhaps the most applied modeling structure in optimization problems. The canonical LP has the form,

$$\begin{aligned} P_{\text{can}} : \text{Maximize } & \mathbf{c}^\intercal \boldsymbol{x} \\ \text{Subject to: } & \mathbf{A}\boldsymbol{x} \leq \mathbf{b} \\ & \boldsymbol{x} \geq \mathbf{0}, \end{aligned} \tag{2.1}$$

However, in a general form, LPs can contain inequalities, equalities and box constraints as,

$$\begin{aligned} P_{\text{gen}} : \text{Maximize } & \mathbf{c}^\intercal \boldsymbol{x} \\ \text{Subject to: } & \mathbf{A}\boldsymbol{x} \leq \mathbf{b} \\ & \mathbf{E}\boldsymbol{x} = \mathbf{e} \\ & \mathbf{l} \leq \boldsymbol{x} \leq \mathbf{u}. \end{aligned} \tag{2.2}$$

where $\mathbf{c},\ \mathbf{l},\ \mathbf{u} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{e} \in \mathbb{R}^p$, $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{E} \in \mathbb{R}^{p \times n}$ are data vectors and matrices, and $\boldsymbol{x} \in \mathbb{R}^n$ is the vector of decision variables. As commented in the previous sections, for a number of reasons, real-world problems might have a mixture of uncertain and well-known elements in its data. The uncertain entries in the data matrices and vectors of Formulation (2.2) would then render an uncertain LP problem,

$$\begin{aligned} \tilde{P}_{\text{gen}} : \text{Maximize } & \tilde{\mathbf{c}}^\intercal \boldsymbol{x} \\ \text{Subject to: } & \tilde{\mathbf{A}}\boldsymbol{x} \leq \tilde{\mathbf{b}} \\ & \tilde{\mathbf{E}}\boldsymbol{x} = \tilde{\mathbf{e}} \\ & \tilde{\mathbf{l}} \leq \boldsymbol{x} \leq \tilde{\mathbf{u}}, \end{aligned} \tag{2.3}$$

where all $(\tilde{\mathbf{c}}, \tilde{\mathbf{b}}, \tilde{\mathbf{A}}, \tilde{\mathbf{e}}, \tilde{\mathbf{E}}, \tilde{\mathbf{l}}, \tilde{\mathbf{u}})$[5] could have uncertain data within its values. Notice that some of the data might be certain.

Further, without lost of generality, it can be assumed that the objective function and the right hand side (RHS) of the inequalities are not subjected to uncertainty, since Formulation (2.3) can always be manipulated and these uncertain elements inserted in an

[5]As a convention, we indicate that a matrix or vector has at least one uncertain entry by adding a tilde ($\sim$) over it.

augmented matrix $\tilde{\mathbf{A}}$ as,

$$
\begin{aligned}
\tilde{P}_{\text{gen}} : \text{Maximize} \quad & \begin{bmatrix} \mathbf{0}^{\intercal} & 1 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ t \\ v \end{bmatrix} \\
\text{subject to:} \quad & \begin{bmatrix} \tilde{\mathbf{A}} & \mathbf{0} & -\tilde{\mathbf{b}} \\ -\tilde{\mathbf{c}}^{\intercal} & 1 & 0 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ t \\ v \end{bmatrix} \leq \mathbf{0} \\
& \begin{bmatrix} \tilde{\mathbf{E}} & 0 & 0 \\ \mathbf{0} & 0 & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{x} \\ t \\ v \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{e}} \\ 1 \end{bmatrix} \\
& \tilde{\mathbf{l}} \leq \boldsymbol{x} \leq \tilde{\mathbf{u}} \\
& t \in \mathbb{R},
\end{aligned}
\tag{2.4}
$$

where v and t are auxiliary variables used as an artifice, only to remove uncertain coefficients from the RHS of the inequalities and from the objective function and move them all to the left hand side (LHS) of the inequality constraints. Therefore the previous model for $\tilde{P}_{\text{gen}}$ can actually be seen as,

$$
\begin{aligned}
\tilde{P}_{\text{gen}} : \text{Maximize} \ & \mathbf{c}^{\intercal}\boldsymbol{x} \\
\text{Subject to:} \ & \tilde{\mathbf{A}}\boldsymbol{x} \leq \mathbf{b} \\
& \tilde{\mathbf{E}}\boldsymbol{x} = \tilde{\mathbf{e}} \\
& \tilde{\mathbf{l}} \leq \boldsymbol{x} \leq \tilde{\mathbf{u}},
\end{aligned}
\tag{2.5}
$$

where now only $(\tilde{\mathbf{A}}, \tilde{\mathbf{e}}, \tilde{\mathbf{E}}, \tilde{\mathbf{l}}, \tilde{\mathbf{u}})$ would have uncertain elements.

The strategy performed with the uncertain data structures in the objective ($\tilde{\mathbf{c}}$) and in the RHS of inequality constraints ($\tilde{\mathbf{b}}$), by inserting them inside matrix $\tilde{\mathbf{A}}$, cannot be used with uncertain data structures in equality constrains ($\tilde{\mathbf{E}}$, $\tilde{\mathbf{e}}$) or box constraints ($\tilde{\mathbf{l}}$, $\tilde{\mathbf{u}}$). In fact, for $\tilde{\mathbf{e}}$, it is possible to get it into an augmented $\tilde{\mathbf{E}}$, in the same fashion that it is done with $\tilde{\mathbf{b}}$, and thus having a certain RHS $\mathbf{e}$ for equality constraints.

Although for certain LPs it is straightforward to transform a model in the format of Formulation (2.2) into (2.1), thus including information of $\mathbf{l}$, $\mathbf{u}$, $\mathbf{e}$, $\mathbf{E}$ into augmented versions of $\mathbf{c}$, $\mathbf{b}$, $\mathbf{A}$, the same is not true for uncertain LPs.

The complication with both transformations, the one of equality and the one of box constraints, is related with a general issue of uncertain equality constraints (the first is explicitly an uncertain equality constraint, while for the second, uncertain equalities would appear at the end of the transformation).

Let's take the uncertain equality constraints in Formulation (2.5). The trouble comes from $\boldsymbol{x}$ having to be chosen before the real values of $(\tilde{\mathbf{E}}, \tilde{\mathrm{e}})$ are observed, thus forcing the uncertain equations in the LHS to be a constant value. As commented by Rockafellar, the more you think about it, the more you wonder whether this even makes sense. Consequently, for most results in optimization under uncertainty, and in all the major results in robust linear optimization, equality constraints are treated as being certain. Unless, of course, when considering constraints on recourse actions taken after the real values for $\tilde{\mathbf{E}}$ are observed. But this assumption do not follow the classic paradigm of robust optimization presented earlier and so this case is not treated here.

However, for some real-world problems, we cannot neglect that equality constraints will have uncertain data. Later in this section, insights of how to handle the uncertain equalities are discussed. Those are borrowed from a remotely related area called RDO. Thereby, until that point, and when it is not explicitly mentioned, all equality constraints are considered to be constituted of certain data with the same occurring for the box constraints.

For now, by disregarding uncertain entries in matrix $\tilde{\mathbf{E}}$ and in vectors $\tilde{\mathrm{e}}$, $\tilde{\mathbf{l}}$, $\tilde{\mathbf{u}}$, it is possible to transform Formulation (2.5) to include their information inside matrix $\tilde{\mathbf{A}}$ and arrive at a canonical uncertain LP formulation,

$$\begin{aligned} \mathcal{P} : \text{Maximize} \;\; & \mathbf{c}^{\intercal}\boldsymbol{x} \\ \text{Subject to:} \;\; & \tilde{\mathbf{A}}\boldsymbol{x} \leq \mathbf{b} \\ & \boldsymbol{x} \geq \mathbf{0}, \end{aligned} \tag{2.6}$$

where uncertainty is associated only with matrix $\tilde{\mathbf{A}}$ and with $\mathcal{P}$ being the family of all LP programs (P) defined by $\mathbf{c}$, $\mathbf{b}$ and some realization of $\tilde{\mathbf{A}}$. Each program of this type is called an instance of the uncertain LP program, including the broadly used program with nominal values $\bar{\mathbf{A}} \in \tilde{\mathbf{A}}$,

$$\begin{aligned} (\mathrm{P_{nom}}) : \text{Maximize} \;\; & \mathbf{c}^{\intercal}\boldsymbol{x} \\ \text{Subject to:} \;\; & \bar{\mathbf{A}}\boldsymbol{x} \leq \mathbf{b} \\ & \boldsymbol{x} \geq \mathbf{0}. \end{aligned} \tag{2.7}$$

Now let's take the uncertain LP $\mathcal{P}$ in the form given in Formulation (2.6). The first step in order to establish its robust counterpart

(RC) formulation, which will provide the robust solution, is to define an uncertain and bounded set $\mathcal{U}$ that encompasses all possible realizations of uncertainty in matrix $\tilde{\mathbf{A}}$. A robust feasible solution to the robust counterpart of $\mathcal{P}$ should, by definition, satisfy all realizations of the constraints induced by the uncertainty set $\mathcal{U}$, and the robust optimal solution is a robust feasible solution with the best possible value of the objective. Since Formulation (2.6) has a certain objective function, its RC is simply defined to be the optimization problem,

$$\begin{aligned} P_{\mathcal{U}} : \text{Maximize } \ & \mathbf{c}^{\intercal}\boldsymbol{x} \\ \text{Subject to: } & \tilde{\mathbf{A}}\boldsymbol{x} \leq \mathbf{b}, \ \forall \tilde{\mathbf{A}} \in \mathcal{U} \\ & \boldsymbol{x} \geq \mathbf{0}. \end{aligned} \tag{2.8}$$

Before going any further, let's depict some of the concepts develop heretofore. Figure 2.5 illustrates the nominal instance $(\mathrm{P}_{\mathrm{nom}})$ of an uncertain LP with a decision vector $\boldsymbol{x} \in \mathbb{R}^2$, five constraints and a linear objective given by $\mathbf{c}$. The yellow area $F_{(\mathrm{P}_{\mathrm{nom}})}$ represents the feasible region of the nominal instance and $\boldsymbol{x}^{\star}_{(\mathrm{P}_{\mathrm{nom}})}$ is the optimal decision vector.

In Figure 2.6 the limit points of the feasible region (dashed line) and optimal solution $\boldsymbol{x}^{\star}_{(\mathrm{P}_{\mathrm{nom}})}$ of the nominal instance $(\mathrm{P}_{\mathrm{nom}})$ are kept the same as in Figure 2.5. An uncertain LP, associated with $(\mathrm{P}_{\mathrm{nom}})$, $\mathcal{P}$, is added to Figure 2.6. $\mathcal{P}$ has four of the five constraints having uncertain parameters taking their values from the convex and closed uncertainty set $\mathcal{U}$. Note that the uncertain problem is in reality an infinite set of instances (P), with the same form of $(\mathrm{P}_{\mathrm{nom}})$, each taking a different set of parameters values from $\mathcal{U}$. So the hatched region $F_{\mathcal{P}}$ represents the region containing the limit points for all possible instances (P) induced by $\mathcal{U}$. Clearly, as depicted, the limit points of the nominal instance are contained in the region $F_{\mathcal{P}}$. It is also evident that in this case the optimal solution for the nominal instance is not optimal for other instances of the problem. Most importantly, this nominal optimal solution is in fact infeasible for many realizations of the data in the uncertainty set $\mathcal{U}$. Finally, the khaki region $F_{\mathcal{U}}$ has the set of robust feasible solutions, where solutions are always feasible, regardless of the values assumed by the constraint matrix $\tilde{\mathbf{A}}$. $\boldsymbol{x}^{\star}_{P_{\mathcal{U}}}$ is the robust optimal solution for this uncertain problem.

It is not at all clear when Formulation (2.8) is efficiently solvable, since as written $P_{\mathcal{U}}$ may have infinitely many constraints. However, despite the robust problem normally being intractable, depend-

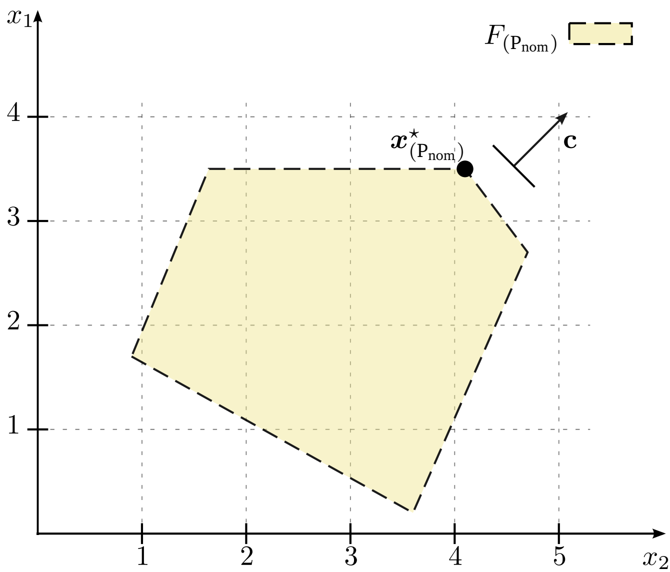


Figure 2.5: Instance (P) of an uncertain LP with nominal data.

ing on the geometries of the uncertainty set $\mathcal{U}$ it is possible to lead $P_{\mathcal{U}}$ to an explicit robust counterpart of nice analytical structure and which can be solved by high-performance optimization algorithms.

Uncertainty models

The uncertainty set models investigated in the literature can be classified in two frameworks which are referred as *column-wise* uncertainty models and *row-wise* (a.k.a. *constraint-wise*) uncertainty models.

Column-wise uncertainty was first considered by Soyster. In this model each column $\tilde{\mathbf{a}}_j$ of the constraint matrix $\tilde{\mathbf{A}} \in \mathbb{R}^{m \times n}$ is only known to belong to a given subset $\mathcal{U}_j \subset \mathbb{R}^m$ (uncertainty set). If $\mathcal{U}_j$ is a singleton, then the parameters associated with the corresponding decision variable x_j has no uncertainty. The robust

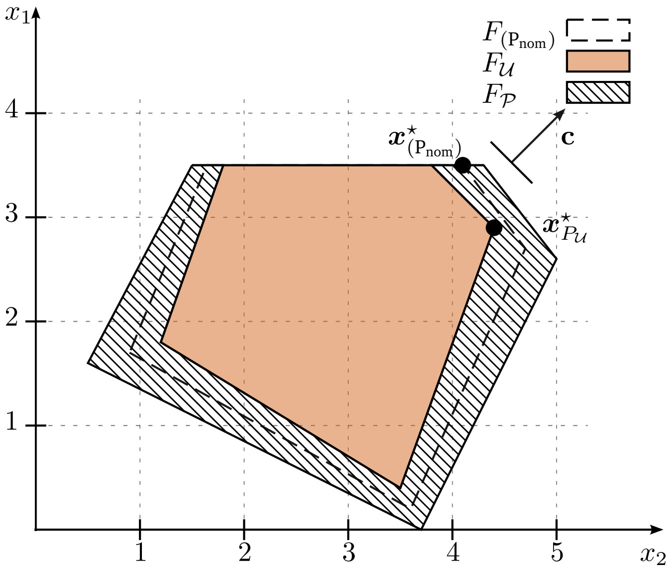


Figure 2.6: Uncertain LP, nominal instance and robust counterpart.

counterpart would then be,

$$
\begin{aligned}
P_{\mathcal{U}} : \text{Maximize } \ & \mathbf{c}^\intercal \boldsymbol{x} \\
\text{Subject to: } & \sum_{j=1}^{n} x_j \tilde{\mathbf{a}}_j \leq \mathbf{b}, \ \forall (\tilde{\mathbf{a}}_j \in \mathcal{U}_j, \ j = 1, \dots, n) \\
& \boldsymbol{x} \geq \mathbf{0}.
\end{aligned}
\tag{2.9}
$$

As it is shown in [57], solving Formulation (2.9) is equivalent to solve an ordinary LP in the form,

$$
\begin{aligned}
P_{\mathcal{U}} : \text{Maximize } \ & \mathbf{c}^\intercal \boldsymbol{x} \\
\text{Subject to: } & \mathbf{A}^* \boldsymbol{x} \leq \mathbf{b} \\
& \boldsymbol{x} \geq \mathbf{0}.
\end{aligned}
\tag{2.10}
$$

where each entry $a^*_{ij} = \sup_{\tilde{\mathbf{a}}_j \in \mathcal{U}_j} (\tilde{\mathbf{a}}_j)_i$, since from Formulation (2.9) we are restricting all variables to be nonnegative. For negative variables, it would be the case to get the infimum of the parameters. Note that this supremum values can be efficiently determined if the uncertainty set $\mathcal{U}_j$ is either closed convex or of finite cardinality. As observed by many authors, column-wise uncertainty models lead to extremely conservative solutions. Each constraint of the robust counterpart (2.10) corresponds to the case when every coefficient in the uncertain constraint matrix assumes its worst-case value at the same time, regardless the geometry of $\mathcal{U}_j$.

On the other hand, row-wise uncertainty has received attention only in more recent decades. In this class of models, a vector $\tilde{\mathbf{a}}_i$ represents the parameters of the i^{th} constraint of matrix $\tilde{\mathbf{A}} \in \mathbb{R}^{m\times n}$ and takes its values in the uncertainty set $\mathcal{U}_i \subset \mathbb{R}^n$. For a row-wise uncertainty model the RC is,

$$\begin{aligned} P_{\mathcal{U}} : \text{Maximize } & \mathbf{c}^\intercal \boldsymbol{x} \\ \text{Subject to: } & \tilde{\mathbf{a}}_i^\intercal \boldsymbol{x} \le b_i, \ \forall\, \tilde{\mathbf{a}}_i \in \mathcal{U}_i, \ i = 1, \dots, m \\ & \boldsymbol{x} \ge \mathbf{0}. \end{aligned} \tag{2.11}$$

Intuitively, $\tilde{\mathbf{a}}_i^\intercal \boldsymbol{x} \le b_i, \ \forall\, \tilde{\mathbf{a}}_i \in \mathcal{U}_i$, if and only if, $\max_{\{\tilde{\mathbf{a}}_i \in \mathcal{U}_i\}} \tilde{\mathbf{a}}_i^\intercal \boldsymbol{x} \le b_i$. Those are called subproblems and their underlying structure determine the complexity for solving the robust counterpart formulation, which now has the general form,

$$\begin{aligned} P_{\mathcal{U}} : \text{Maximize } & \mathbf{c}^\intercal \boldsymbol{x} \\ \text{Subject to: } & \max_{\{\tilde{\mathbf{a}}_i \in \mathcal{U}_i\}} \tilde{\mathbf{a}}_i^\intercal \boldsymbol{x} \le b_i, \ \forall i \\ & \boldsymbol{x} \ge \mathbf{0}. \end{aligned} \tag{2.12}$$

Notice that this formulation differs significantly from the results of Soyster. The equivalence between (2.9) and (2.10) is particular to column-wise uncertainty and do not hold for row-wise uncertainty. In fact, typically, the RC of the problem with row-wise uncertainty is not an LP program and will depend on the geometry of $\mathcal{U}_i$. Row-wise uncertainty is a more general case of uncertainty set than column-wise uncertainty, in the sense that with the former the RC is capable to reflect the fact that regularly the elements of the constraints cannot be simultaneously at their worst values.

Following the robust counterpart, Problem (2.12), is reformulated for three different uncertainty set geometries: box, polyhedral and ellipsoidal uncertainty.

Box Uncertainty: A box uncertainty set $\mathcal{U}_i$ of a particular row i can be defined as the Cartesian product of the intervals of validity for all coefficients $\tilde{a}_{ij}$ in that row. With each entry $\tilde{a}_{ij}$ being modeled as a bounded random variable that takes values in a interval

$[a^{\mathrm{l}}_{ij}, a^{\mathrm{u}}_{ij}]$ defined by its upper and lower bounds (where for the certain parameters $a^{\mathrm{l}}_{ij} = a^{\mathrm{u}}_{ij}$), the RC (2.12) becomes,

$$\begin{aligned} P_{\mathcal{U}} : \text{Maximize} \;\; & \mathbf{c}^{\intercal}\boldsymbol{x} \\ \text{Subject to:} \;\; & \sum_{j=1}^{n} a^{\mathrm{u}}_{ij} x_j \leq b_i, \;\; \forall i \\ & \boldsymbol{x} \geq \mathbf{0}. \end{aligned} \tag{2.13}$$

Which is an LP program, with the same size of the original uncertain problem. Notice that boxed row-wise uncertainty set produce equivalent robust results as convex closed column-wise uncertainty sets.

Polyhedral Uncertainty: An uncertainty set with polyhedral geometry encloses the possible values of uncertain vector of coefficients $\tilde{\mathbf{a}}_i$ in row i to $\mathcal{U}_i = \{\mathbf{a}_i : \mathbf{D}_i \mathbf{a}_i \leq \mathbf{d}_i\}$. The robust counterpart can be written as,

$$\begin{aligned} P_{\mathcal{U}} : \text{Maximize} \;\; & \mathbf{c}^{\intercal}\boldsymbol{x} \\ \text{Subject to:} \;\; & \max_{\{\mathbf{D}_i \mathbf{a}_i \leq \mathbf{d}_i\}} \mathbf{a}_i^{\intercal}\boldsymbol{x} \leq b_i, \;\; \forall i \\ & \boldsymbol{x} \geq \mathbf{0}. \end{aligned} \tag{2.14}$$

Observe that $\boldsymbol{x}$ is not a variable of optimization in the internal optimization problems, so the dual of the subproblem for each constraint i is,

$$\begin{aligned} \max \; & \mathbf{a}_i^{\intercal}\boldsymbol{x} \\ \text{s.t.} \; & \mathbf{D}_i \mathbf{a}_i \leq \mathbf{d}_i \end{aligned} \quad \Longleftrightarrow \quad \begin{aligned} \min \; & \mathbf{d}_i^{\intercal}\boldsymbol{p}_i \\ \text{s.t.} \; & \mathbf{D}_i^{\intercal}\boldsymbol{p}_i = \boldsymbol{x} \\ & \boldsymbol{p}_i \geq 0 \end{aligned}$$

and therefore, as presented by Bertsimas et al. [18], the robust counterpart (2.14) can be rewritten as,

$$\begin{aligned} P_{\mathcal{U}} : \text{Maximize} \;\; & \mathbf{c}^{\intercal}\boldsymbol{x} \\ \text{Subject to:} \;\; & \mathbf{d}_i^{\intercal}\boldsymbol{p}_i \leq b_i, \;\; \forall i \\ & \mathbf{D}_i^{\intercal}\boldsymbol{p}_i = \boldsymbol{x}, \;\; \forall i \\ & \boldsymbol{p} \geq \mathbf{0} \\ & \boldsymbol{x} \geq \mathbf{0}, \end{aligned} \tag{2.15}$$

where $\boldsymbol{x}$ and $\boldsymbol{p}$ are, respectively, the vector of primal and dual variables, and $(\mathbf{D}_i, \mathbf{d}_i)$ define the uncertain polyhedron of the parameters in row i. Although the size of such problems grows polynomially in size of the nominal problem and the dimensions of the uncertainty set, the robust counterpart remains a LP problem.

Ellipsoidal Uncertainty: In the late 90s, Ben-Tal and Nemirovski [15, 16], and El Ghaoui and Lebret [59, 60] proposed the use of ellipsoidal uncertainty sets. Those works were a significant step forward for developing a more complete theory for robust optimization. The ellipsoidal geometry enables flexibility in modeling the relationship between the uncertain parameters, and so it allows to remove the most unlikely outcomes from consideration. Also, controlling the size of the ellipsoidal sets can be interpreted as tuning the trade off between robustness and performance. A drawback of modeling $\mathcal{U}_i$ as an ellipsoidal set is the increased complexity of the resulting problem. For example, for a given $\mathcal{U}_i = \{\mathbf{a}_i : \mathbf{a}_i = \bar{\mathbf{a}}_i + Q_i \boldsymbol{v}_i, \text{ with } ||\boldsymbol{v}_i||_2 \leq \rho_i\}$, where $\bar{\mathbf{a}}_i$ denotes the nominal value for parameters in row i, ρ_i the ellipsoids radius, and Q_i is a positive definite matrix defining the shape of the ellipsoid, the RC turns into a second-order cone programming (SOCP) problem in the form,

$$\begin{aligned} P_{\mathcal{U}} : \text{Maximize } & \mathbf{c}^\intercal \boldsymbol{x} \\ \text{Subject to: } & \bar{\mathbf{a}}_i^\intercal \boldsymbol{x} + \rho_i ||Q_i^\intercal \boldsymbol{x}||_2 \leq b_i, \ \forall i \\ & \boldsymbol{x} \geq \mathbf{0}. \end{aligned} \tag{2.16}$$

Although more demanding computationally than LP models, this nonlinear, but convex, model can be solved efficiently by interior-point methods [15, 18].

The three uncertainty model geometries presented above are perhaps the most used ones in robust linear optimization, but they are not the only ones that lead to tractable versions of the robust counterpart problem. Other examples are, norm uncertainty and cardinality constrained uncertainty sets [64]. For further information, Bertsimas et al. [18] provide a compact overview of the current state-of-the-art for robust optimization (linear and convex) with respect to tractability issues.

Additional Remarks

The theory of robust linear optimization presented throughout this section relates to the problem statement itself, specially to show how uncertainty models play a key role in designing a solvable robust version of the nominal/uncertain problem. This is by far the main concern of the field. However, there are other specific aspects of robust problems that are relevant to our application. Below, three important aspects to linear robust optimization are considered. The first one relates to duality in robust problems. The second is on how uncertain equality constraints could be handle in robust optimization. Finally, the third is on the extensions of the above theory to Mixed-Integer Linear Programming (MILP) problems.

Duality: From the two frameworks of uncertainty set models presented, column-wise and row-wise, the former is at least as conservative as the latter, regardless the geometry proposed for the uncertainty set. As a result of the lack of modeling flexibility and the overly conservative solutions produced by column-wise models, frequently row-wise uncertainty models are preferred. Nevertheless, some problems have its parameters intrinsically related in a column-wise fashion. For those problems, a first reasonable idea would be to use duality theory to generate the dual of the column-wise uncertain linear problem, which would be a row-wise uncertain linear program – assuming, of course, the same uncertainty model for the columns of the given linear program and for the corresponding rows in the dual [65],

$$
\begin{aligned}
P_{\mathcal{U}}: \quad & \text{Maximize } \mathbf{c}^{\intercal}\boldsymbol{x} \\
& \text{Subject to: } \sum_{j=1}^{n} x_j \tilde{\mathbf{a}}_j \leq \mathbf{b},\ \forall \tilde{\mathbf{a}}_j \in \mathcal{U}_j,\ \forall j \\
& \qquad\qquad\quad \boldsymbol{x} \geq \mathbf{0}
\end{aligned}
$$

$$\Updownarrow$$

$$
\begin{aligned}
D_{\mathcal{U}}: \quad & \text{Minimize } \mathbf{b}^{\intercal}\boldsymbol{y} \\
& \text{Subject to: } -\tilde{\mathbf{a}}_j^{\intercal}\boldsymbol{y} \leq -c_j,\ \forall \tilde{\mathbf{a}}_j \in \mathcal{U}_j,\ \forall j \\
& \qquad\qquad\quad \boldsymbol{y} \geq \mathbf{0}.
\end{aligned}
$$

The problem is that, even though $D_{\mathcal{U}}$ being a row-wise uncertain LP as (2.11), strong duality, in general, does not hold for uncertain LPs. That is intuitive, since the primal and dual problems are seeking for a robust solution, both are paying a price for uncertainty that will reflect in their objective value. I.e., in maximizing the primal, its robust optimal solution value will be (in general) less than the primal "nominal" optimal solution value (for the LP instance with nominal data), and in minimizing the dual, the solution will be larger when compared to the nominal dual solution; so a gap will exist.

Thus, there is no advantage in converting column-wise to row-wise uncertain problems in order to solve robust linear optimization problems, as the solutions of the primal and dual are not equivalent [65].

Equality constraints: It has been pointed out that uncertain equality constraints are in general problematic, and therefore they have being considered certain for the results presented so far (in accordance with the main results in robust linear optimization). Still, one cannot ignore the fact that uncertain equality constraints can occur in optimization problems.

In order to increment the robust optimization methodology presented earlier in this section to account for uncertain equalities, we address their issues using some of the insights mentioned in Section 2.3.1. Those insights were taken mainly from [49, 66] and the references therein, which focus on a distinct area of robust optimization called robust design optimization (RDO). Rangavajhala and Messac [66] consider three prevailing equality constraint formulation approaches in the literature of RDO:

(1) equality constraint relaxation approach

(2) satisfying the equality constraint at its mean value

(3) elimination of the equality constraint through substitution.

Based on some of those ideas we enhance the robust linear optimization methodology presented here by adding the following steps:

(i) Whenever possible, the first step in finding a RC should be to eliminate uncertain "state equations" (which are equality constraints) by the substitution of the corresponding "state variables" into inequality constraints or the objective function. An example of this procedure is given below.

$$\begin{aligned} &\max\ c_2x_2 \\ &\quad \text{s.t. } \tilde{a}_{11}x_1 + a_{12}x_2 \leq b_1 \\ &\qquad \tilde{e}_{11}x_1 + x_2 = e_1 \\ &\qquad x_1, x_2 \geq 0 \end{aligned} \quad \Longleftrightarrow \quad \begin{aligned} &\max\ (-\tilde{e}_{11}c_2)x_1 \\ &\quad \text{s.t. } (\tilde{a}_{11} - a_{12}\tilde{e}_{11})x_1 \leq \dots \\ &\qquad\qquad \dots(b_1 - a_{12}e_1) \\ &\qquad x_1 \geq 0 \end{aligned}$$

(ii) For the uncertain equality constraint that cannot be eliminated (usually, e.g., physics-based constraints as mass balance equations), we propose to relax the constraint by replacing it by two inequalities that keep the original constraint satisfied to the maximum extent possible. Therefore, for a give uncertain LP that has passed through step (i) and has the format,

$$\begin{aligned} \tilde{P}_{\text{gen}} : \text{Maximize }\ & \mathbf{c}^\intercal \boldsymbol{x} \\ \text{Subject to: }\ & \tilde{\mathbf{A}}\boldsymbol{x} \leq \mathbf{b} \\ & \tilde{\mathbf{E}}\boldsymbol{x} = \mathbf{e} \\ & \boldsymbol{x} \geq \mathbf{0}, \end{aligned} \tag{2.17}$$

using the equality relaxation approach the uncertain problem will have the form,

$$\begin{aligned} \tilde{P}_{\text{gen}} : \text{Maximize }\ & \mathbf{c}^\intercal \boldsymbol{x} \\ \text{Subject to: }\ & \tilde{\mathbf{A}}\boldsymbol{x} \leq \mathbf{b} \\ & \begin{bmatrix} \tilde{\mathbf{E}} \\ -\tilde{\mathbf{E}} \end{bmatrix} \boldsymbol{x} \geq \begin{bmatrix} \mathbf{e} \\ -\mathbf{e} \end{bmatrix} \\ & \boldsymbol{x} \geq \mathbf{0}, \end{aligned} \tag{2.18}$$

Then the robust counterpart of Formulation (2.18) can be produced using any of the theory presented up to here.

Figure 2.7 illustrates the concept of equality relaxation. $F_{\mathbf{E}_=}$ is the feasible region for a certain equality constrain (e.g. with nominal values for $\mathbf{E}$). The intersection of regions $F_{\tilde{\mathbf{E}}_\geq}$ and $F_{-\tilde{\mathbf{E}}_\geq}$ is the feasible region of a relaxed uncertain equality constraint given by $\tilde{\mathbf{E}} = \mathbf{e}$. Remember that certain equality constraints can instead be converted to two inequalities and be added to matrix $\tilde{\mathbf{A}}$.

Integer variables: The robust optimization paradigm set in Section 2.3.2 and 2.3.3 has no restriction in considering part of the decisions variables to be integer. Thus, it is possible to apply the robust

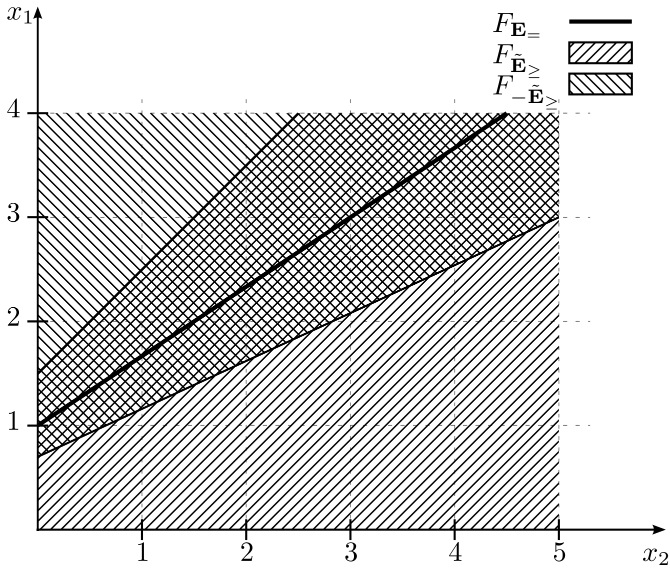


Figure 2.7: Uncertain equality constraint relaxation.

optimization methodology to an uncertain MILP and render a robust counterpart of the problem. However, tractability issues arising from uncertain linear optimization and uncertain mixed-integer linear optimization problems need quite different approaches [17]. For example, the resulting robust formulations involving conic quadratic problems, when data vary in ellipsoidal sets, cannot be directly applied to discrete optimization [67].

Kouvelis and Yu [61] propose a framework specific for discrete optimization problems, which uses a set of possible scenarios for the data instead of continuous uncertainty sets. Under their approach, the RC to a number of polynomially solvable discrete optimization problems becomes NP-hard. Averbakh [68], in considering interval representation for uncertain data (the same idea as in the previously referred box uncertainty), show that polynomial solvability is preserved.

In a framework similar to the one used in Section 2.3.3, Bertsimas and Sim [67] present for cardinality constrained uncertainty sets that the robust counterpart of an uncertain MILP problem is still tractable. Since box uncertainty set are a special case of cardinality constrained uncertainty sets, the results in [67] can promptly be extended to a box uncertainty model. Thus, the robust formulation (2.13) is tractable, even when coming from an uncertain MILP problem.

For further developments on robust discrete optimization see [69] and the references therein.

2.4 SUMMARY

This chapter brought in an overview of fundamental topics related to the core work of this dissertation. First, the oil and gas industry was briefly presented, followed by basic concepts from petroleum engineering that are required for understanding the production system – i.e. gas-lift, which is used as our object of study. Then the production optimization problem was situated within the entire operation of an oil field, and a literature review of classical approaches for production optimization of gas lifted oil fields is made. The second part was a compact description of modeling optimization problems using mathematical programming and on the means chosen for solving the numerical experiments of this dissertation. The final part, starts with a discussion of uncertainty in optimization problems and on modeling decisions that one has to face when dealing with uncertain optimization problems. Then the classic paradigm of robust optimization is introduced. Finally, a compilation of some of the main results in the area of robust linear optimization is presented, along with some important remarks that relates with the work that is developed in the following chapters of this dissertation.

3 ROBUST PRODUCTION OPTIMIZATION METHODOLOGY

This chapter details the steps required to design robust optimization models. It starts describing how uncertainty appears in oil production systems. Then it analyzes the effects of uncertainty in a simplified production optimization problem. Finally it presents a methodology to design a robust production optimization model to this simplified example, which can be extended to more complex production systems.

3.1 UNCERTAINTY IN OIL PRODUCTION SYSTEMS

Production systems in the context of oil and gas fields include all elements from the reservoir interface to the export lines. The production itself is a mass flow mixture of hydrocarbon with different molecular weights, as well as water, sand, H_2S, CO_2 and possibly other components. The production in a production system travels as a multiphase flow from the reservoir to the well inlet, then from the bottomhole through the production tubing and flow lines until it reaches the processing facilities for separation and later to be stored or exported, as illustrated in Figure 3.1. In order to enable or enhance the potential hydrocarbon recovery from the reservoir, improved recovery methods as waterflooding can be used. On the other hand, artificial lift methods, as gas-lift systems, can be employed to allow a greater use of this reservoir potential, by artificially increasing the pressure difference in the reservoir-wellbore interface.

An operative oilfield requires the integrated operation of several areas. Operational and strategical decisions from different subsystems and in different planning horizons impact each other. Due to the high complexity of the entire integrated operation, it has been suggested a hierarchical division of the decisions [2]. Those are grouped in layers according to their "time constant" as depicted in Figure 2.4. In this way the problem of each layer can be solved independently, although a higher level typically specifies constraints on a lower level to link these problems.

Here we are interested in the short-term production optimization layer. At this level the aim is to optimize the daily/weekly production. Decisions at this level are mostly semiautomatic, thus petroleum engineers, aided by recommendation tools, decide the best operation plan according to the scenario of operation for that day. The objective is usually to maximize oil/gas production or some

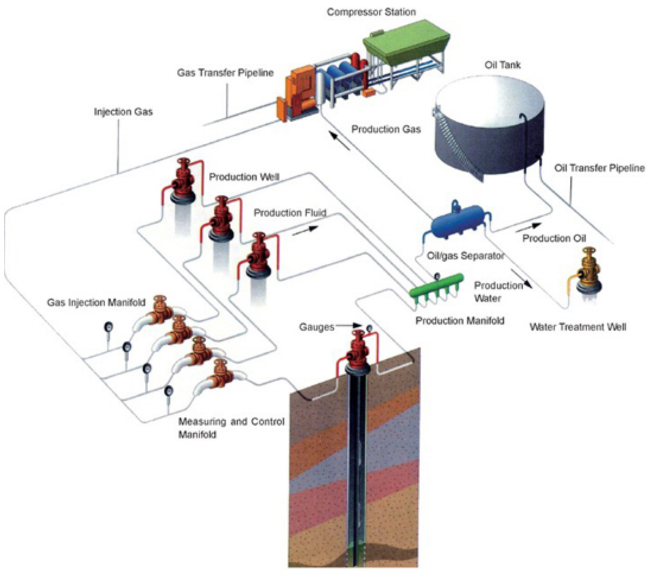


Figure 3.1: Snapshot of an operational oilfield.

other specific criteria. The possible decision variables of this layer vary according to the different production equipment installed and their available corresponding setpoints. However, in general, the decision variables include well allocation, manifold alignment, production choke settings, and when existent, artificial lift equipment operating points. In the scope of production optimization various factors can restrict production. Constraints could come from downhole, for example, limits on pressure drawdown to avoid formation damage (skin effect), or from subsea equipment, as fluid handling capacities and pressure limits in the flow lines and manifolds, or from the surface equipment capacities, as limits on compression capacity for lift gas or fluid handling capacity in separators and tanks. The reservoir management layer (see Figure 2.4) can also establish constraints to the production optimization layer, e.g. production targets.

A production model that predicts the effects of the decision variables on the production is required to perform a production optimization. Usually production optimization uses steady-state models of the production system. These models are generally based on multiphase flow simulators, either by querying the simulator directly or by using simulator data and fitting proxy-models with them. Simulators use physical equations and empirical relationships (derived in laboratories), to establish the relation between decision variables and production. The simulator models normally include many parameters. Some of these parameters are well established physical properties that tend not to vary in time, as pipeline dimensions, that are accurately and easily available from the installation projects. Other parameters are measurements from the production system, which may or may not have an easy/costless availability. For some other parameters it is required to perform laboratory experiments or to use the experience in the field to find values that make the model fit the production data. System identification is also an alternative to models based on simulators, though this also rely on data that might not be promptly and frequently available to update the models.

Elgsæter et al. [11] suggest three reasons why production models may not describe production accurately.

- One reason is model related, in which the underlying structure of the model is insufficient to predict correctly the production. For instance, unmodeled disturbances which may significantly influence production but are not accounted for.

- A second reason may be measurement uncertainty, since any model to some extent relies on measured data. The causes for this type of uncertainty are many. For example, measurements are susceptible to incorrect calibration of measurement equipment. An even more common cause are errors originated in indirect measurements, estimates of physical quantity. For instance, measurements of total produced single oil and gas rate phases might be available, however total water rates might only be estimated by adding different measured water rates after separation and processing.

- A third reason may be the lack of informative data. Which could be in the sense that the data used to determine the parameters might be outdated, e.g., due to the costs and risks involved to perform some tests and measurements. Though, it could also be in the sense that the data may have insufficient excitation to uniquely determine the parameters of the model. The latter is a fairly common situation since, for example, to determine the rates of oil, gas and water produced from individual wells, the production of a single well is usually routed to a dedicated test separator where the rate of each separated component is measured. Since it is not always possible to perform multi-rate tests a well may operate in a condition different from the one which it was tested.

Bieker et al. [34] also pointed out two important issues with the models that are somehow related to these three later reasons for model inaccuracy. One of the issues is that models for production optimization are mostly pure steady-state, but, although reservoir changes caused by the drainage process occur in a time constant that can be ignored for this optimization, there are usually other transients happening during normal operation and their effects may have a significant impact on the capacity of the model to predict the production. The other issue is in line with the first. For using transient data to determine parameters for a steady-state model can result in erroneous parameter values. In practice all of those factors are usually present to some extent.

The standard approaches to production optimization (Section 2.1.4.1) simply disregard all the uncertainty mentioned above – which reflects in models inaccuracy, and considers nominal models, namely the production models that reflect the average observed behavior. However, for optimization problems, such approach may render the solution suboptimal and even infeasible when the actual models deviate significantly from the nominal conditions [16]. In fact, in a technological survey of 2007 [34] the authors state that the handling of model uncertainty is a key challenge for the success of production optimization. The challenge is actually twofold. The first relates to the need to identify and characterize the uncertainties in a production system, to have a model for them. The second one is the design of production optimization models that somehow can handle the uncertainties. Both issues have their importance and definitely the latter is highly dependent to the former.

In the context of the reservoir management layer (see Figure 2.4) modeling and incorporating uncertainty to the problem has received an increased attention in the last decade. The interested reader can refer to [70, 71, 72] and the references therein to learn more about the recent activities.

Yet, uncertainty in production optimization has received less attention. Here is a non-exhaustive list of works that to some extent try to model uncertainty or to consider uncertainty in the production optimization problem.

Elgsæter et al. [11] discussed the estimation of uncertainty in production optimization resulting from fitting models to production data with low information content. It also showed how system identification could be used to design models for production optimization based on production data. However they do not develop an approach to handle uncertainty in itself.

Camponogara and Nakashima [73] proposed dynamic programming algorithms to solve the problem of distributing a limited rate of lift-gas. Their formulation introduced precedence constraints on the activation of the wells and allows multiple WPC for one well to encompass uncertainty in these curves. It solves a max-min problem. An advantage of dynamic programming is the achievement of a family of solutions covering the whole range of lift-gas injection but it is difficult to add constraints to this problem.

Bieker et al. [13] dealt with uncertainty by formulating an optimization problem based on a priority list. They proposed to use information about the uncertainties of the gas or water oil ratios to find the order of opening and closing the wells to maximize the expected total oil production rate from the wells.

Elgsæter et al. [14] proposed an iterative approach to access the results of optimization in an uncertain scenario. Thus, uncertainty is not handled in the optimization problem, but in a post-optimization stage.

As a preliminary result of the work developed during this dissertation, Hülse and Camponogara [74] proposed an strategy of envelope curves to encapsulate the uncertainty of the production curves, and produce a robust solution for production optimization problem. Uncertainty is modeled as a range and as non-probabilistic. This strategy turned out to be in line with the results obtained by Soyster [57] for robust optimization.

There are not many works considering uncertainty for short-term production optimization, but model uncertainty is inherent to production systems. This leaves open for investigation the two challenges mentioned earlier: uncertainty identification/modeling and optimization models that can incorporate uncertain production models.

In this dissertation we focus on the latter issue and take for granted that the uncertainties can be identified and modeled. However, it is important to have a brief discussion on the former issue, since both are closely related. In this work, instead of considering that uncertainty has a stochastic model with a probabilistic description, which could be given or estimated, we assume that uncertainty has a simpler model, a set-based one (range model), with an undefined probabilistic distribution. Ben-Tal et al. [17] point out some advantages for this approach:

- More often than not, there are no reasons to assign a stochastic nature to the uncertainty.
- Even when perturbations can be considered as stochastic, it might be difficult to specify reliably data distribution.

We do agree that the first point might not always be true in the context of production optimization, but the second fits for most situations. Anyway, here it is important to make two notes. Although intuitively range modeling requires a least sophisticated uncertainty quantification system than stochastic modeling, which, by the way, could be a practical benefit for this approach, implying in an easier creation, maintenance, and usage of the uncertainty models, we understand that even coming up with meaningful ranges for the uncertain parameters may be a complex task. In the same line, it is natural that when a perturbation has a stochastic nature, and it is provided with a relevant stochastic definition, its much more informative description has the potential to produce a less conservative and more practical decision – if associated with a corresponding stochastic production optimization model.

As previously stated, our main objective is to design a production optimization methodology that accounts for uncertainties in the production model and thus generates a solution that has robustness against parameters perturbation. The present work follows the Robust Optimization paradigm to address parameter uncertainty in an optimization problem. This approach assumes that the uncertainty model is set-based instead of stochastic. Also, it constructs a solution that is optimal and immune to any realization of the uncertainty in a given set, but in a deterministic rather than a probabilistic sense [18].

In the next section we develop the idea behind this paradigm when applied to an oilfield production problem. More specifically, we choose typical continuous gas-lifted oilfield schemes to analyze the proposed methodology.

3.2 UNCERTAIN PRODUCTION OPTIMIZATION MODELS

Continuous gas-lift, or just gas-lift, is an artificial lift system where a controlled stream of compressed gas is injected in the well annulus. From there, it enters the production tubing through gas-lift valves, where it mixes with the reservoir fluids reducing the fluid density, thus enabling or enhancing the well production to reach the surface facilities (Figure 2.2). Different gas-lift producing system layouts exist, but all of them share most of these features:

- a lift gas compression system, which has a compression capacity.
- a gas-lift manifold, where gas-lift chokes are used to control the gas flow rate to individual wells;
- gas-lift manifolds and valves, which allow the gas enter the production tubing;
- a production choke in the wellhead to control the well head flowing pressure;
- flow lines to drive the production to the surface facilities. The pipelines may have fluid handling capacities;
- a well might have its production drained through an exclusive flow line (named satellite or platform well) or its production might pass first to subsea manifolds which are used to gather and route the production of different wells before it reaches the platform; manifolds pressure limits may apply;
- one or more separation facility to which production can be routed. They do not only separate the gas to be treated and sent for recompression, but they also separate the oil from the water and contaminants. Liquid handling limits apply to the equipment.

In order to introduce the proposed robust production optimization methodology and examine it step by step, we consider a simplified gas-lift scheme, a toy example, but one that is representative enough to our purpose. It consists of a single platform well that produce a stream of gas, oil and water to a single multiphase separator. Production is enhanced by a gas-lift system, where a choke valve can control the gas injection rate for this well. The production choke valve is considered to be always fully opened, thus not being a decision variable.

3.2.1 Nominal MINLP production problem

Assuming a nominal production model – e.g., one based on expected parameters values – the production optimization problem can be viewed as solving a conceptual MINLP problem, in the sense that well-production functions, a.k.a well performance curves (WPCs)s, are not explicitly available. Here, for simplicity, the objective function is considered the maximization of oil production of the well. The conceptual MINLP formulation follows:

$$
\begin{aligned}
P:\ \max\ \widehat{f} &= q_{\text{oil}} && \text{(3.1a)}\\
\text{s.t.}:\ q_{\text{inj}} &\leq Q_{\text{inj}}^{\max} && \text{(3.1b)}\\
q_r &= \widehat{q}_r(q_{\text{inj}}, G)\cdot y, \forall r \in \mathcal{R} && \text{(3.1c)}\\
q_{\text{inj}} &\geq y \cdot q_{\text{inj}}^{\min} && \text{(3.1d)}\\
q_{\text{inj}} &\leq y \cdot q_{\text{inj}}^{\max} && \text{(3.1e)}\\
q_r^{\text{S}} &\leq q_r^{\text{S},\max}, \forall r \in \mathcal{R} && \text{(3.1f)}\\
q_r^{\text{S}} &= q_r, \forall r \in \mathcal{R} - \{\text{gas}\} && \text{(3.1g)}\\
q_{\text{gas}}^{\text{S}} &= q_{\text{inj}} + q_{\text{gas}} && \text{(3.1h)}\\
y &\in \{0,1\} && \text{(3.1i)}
\end{aligned}
$$

The **decision variables** of the formulation are:

- q_{inj}^n is the rate of lift-gas injection into the well n;
- q_r is the rate of phase r produced according with the conceptual production function $\widehat{q}_r$ of the well;
- q_r^{S} is the total flow of phase r received by the separator;
- $y \in \{0,1\}$ is a well allocation variable. It takes 1 when well is active or else 0;

Notice that q_r and q_r^{S} are not actual decision variables. They are only used to define a given flow rate. Although the term is not precise to define this type of variable, heretofore we refer to them as state variables, in the sense that they are just used to describe the mathematical "state" of the system. The **parameters** are as follows:

- $\mathcal{R} = \{\text{oil}, \text{gas}, \text{water}\}$ has the mixed-flow phases: oil, gas, and water;
- $Q_{\text{inj}}^{\max}$ is the available lift-gas rate;

- $q_{\text{inj}}^{\min}$ is the lower bound for injection rate, established for example, by the production engineer for flow assurance.
- $q_{\text{inj}}^{\max}$ is the upper bound for lift gas injection. It could be set, e.g., to avoid slugging.
- $q_{\text{oil}}^{\max}$, $q_{\text{gas}}^{\max}$, $q_{\text{water}}^{\max}$ are limits on oil, gas, water handling of the separator;
- G is the set of parameters and correlations that characterize the well-production curves of the well. Since this is a satellite well, it includes the separator inlet pressure, the pipeline inside diameter (ID), vertical elevations, and horizontal length, the production tubing ID, length, and trajectory, the fluid data, as gas-liquid ratio (GLR), WC, specific densities, and bubble point pressure, the inflow performance relationship (IPR) and its parameters, the vertical and horizontal flow pattern correlations, to name a few.

and with the **functions**:

- $\widehat{f}$ is the total oil production from the well;
- $\widehat{q}_r(q_{\text{inj}}, G)$ is the conceptual function that relates the flow of component r produced by the well define by G with the rate of lift-gas injected to this well (q_{inj}).

Constraint (3.1b) represents the lift gas compression capacity for the gas lift compressor. The well production as a function of the lift-gas injection rate is modeled by Eq. (3.1c). Limits on lift-gas injection and separation capacities are given by constraints (3.1d)-(3.1e) and (3.1f), respectively. Eqs. (3.1g)-(3.1h) have the production arriving at the separation unit.

Note that for simplicity the production system constraints are only in the availability of lift gas, limits on gas injection, and fluid handling capacity in the separator. However, mathematical programming allows to describe other possible system constraints. For instance, through constraints such as,

$$q_{\text{oil}} + q_{\text{water}} \geq y \cdot q_{\text{liq}}^{\min} \tag{3.1j}$$

$$q_{\text{oil}} + q_{\text{water}} \leq y \cdot q_{\text{liq}}^{\max} \tag{3.1k}$$

it is also possible to specify, for example, a target flow, obtained from the reservoir management layer, or a liquid flow rate limit to avoid formation or perforations damage.

Uncertainty in Formulation (3.1) can be viewed as affecting the parameters defining the system limits and capacities – i.e., $Q_{\text{inj}}^{\max}$, and $q_r^{\text{S},\max}$, $\forall r \in \mathcal{R}$ – and the functions defining the multiphase production.

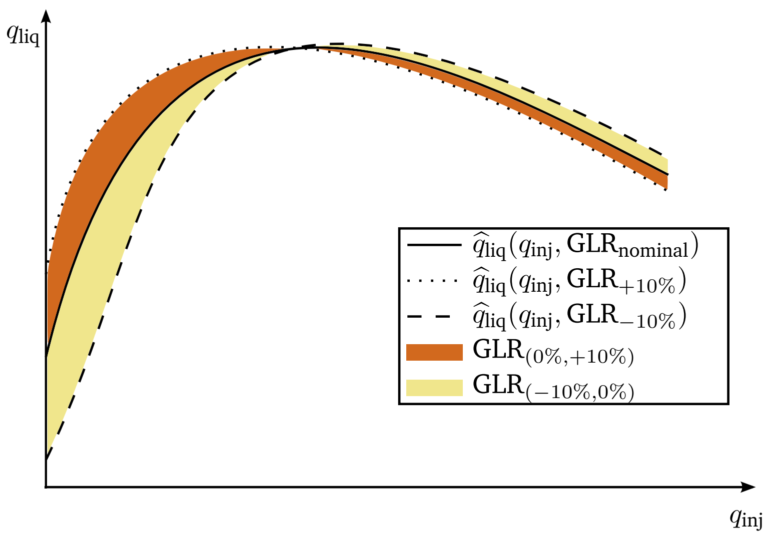
3.2.2 Uncertainties in the MINLP production problem

For uncertainty in limits and capacities, the uncertainty model associated to each of the uncertain parameters is explicitly present in the uncertain optimization problem. However, for uncertain values coming from the well-production functions this is not all clear. Let's for now focus on the latter uncertainty manifestation.

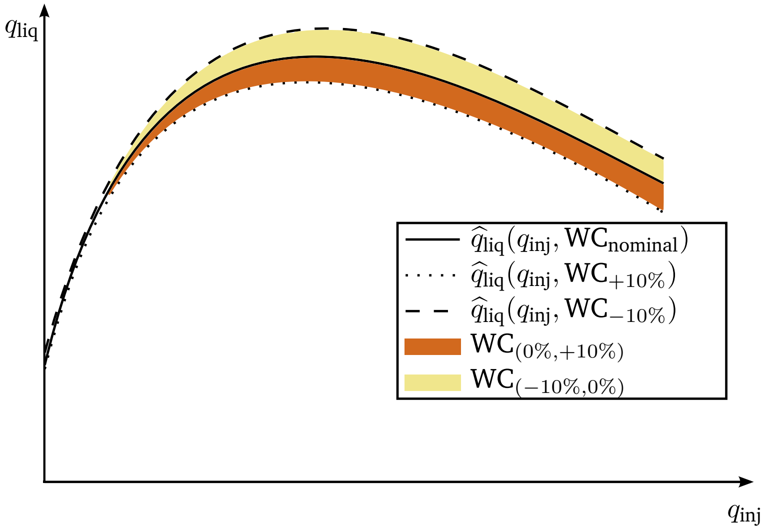
In fact, here we have to make a point. The well-production functions $\widehat{q}_r$ represents the real relationship between lift-gas injection, q_{inj}, and the production of component r, q_r. We assume in this work that given the "real" set of parameters and correlations (G), a multiphase flow simulator is capable of mirroring the steady-state behavior of the real well-production function. This assumption is not completely unjustified. First, as it was said, there is no precise and explicit formulation for the real well-production function due to the complex fluid dynamics involved. Therefore, the only alternative other than simulation to obtain useful data for the optimization problem, it is to dispose of enough and constantly updated well-test data (in number and in range of excitations) to then be fitted to a particular model. In practice, with few exceptions, this is an unrealistic scenario and with insufficient data, models based on system identification would also not be a perfect representation. Second, we must have a baseline production function formulation for comparison. Because real data is hard to obtain in quantities required for study, we use the simulated curves as our baseline. Third, there are many reliable multiphase flow simulators that are already widely used to aid engineering projects. So for now on, we treat the functions, $\widehat{q}_r(q_{\text{inj}}, G)$, that were referencing to the real well-production functions, as being the well-production curves generated by simulation, since we consider that both are equivalent.

By using the simulated curves, one is able to analyze the impact of uncertainty in the well-production function. Uncertainty in $\widehat{q}_r(q_{\text{inj}}, G)$ appears in the set of parameters and correlations G. Yet, not all parameters have the same susceptibility to perturbations. For instance, production tubing length and ID usually have accurate known values, and even if the ID of a tubing can decrease by wax deposition, this phenomenon happens in a time scale that can be disregard for the short-term production optimization problem. Empirical relations as the fluid flow correlations and the IPR that affect significantly the shape of the well-production curve, although uncertain in the beginning of the life of a well, over time, by calibration they tend to produce reliable representations that do not change greatly in the time frame of production optimization. However, there are key parameters in modeling well-production curves that are susceptible to the effects of uncertainty, as: the fluid data, such as the GLR and WC, and the reservoir data such as the productivity index (PI).

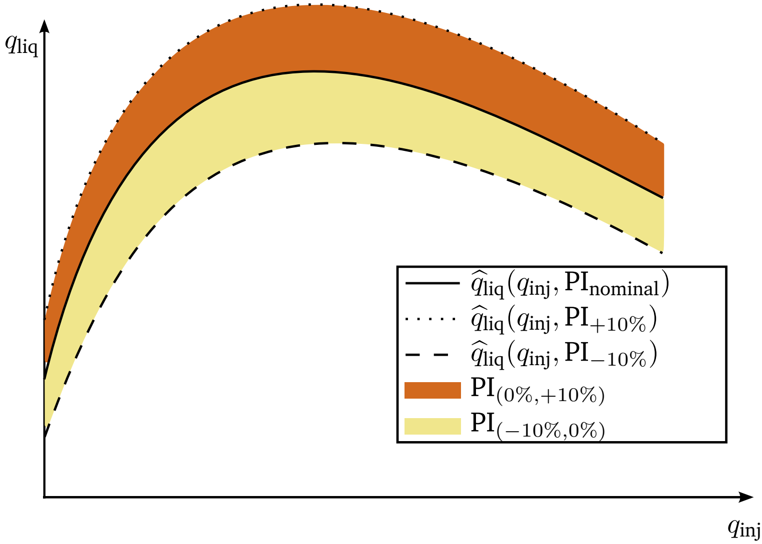
In order to get a feeling of exactly how the changes of different parameters affect the shape of the well-production curve, let us consider a hypothetical $\pm 10\%$ variation from a nominal value of GLR, WC, and PI, individually. The effects of each parameter in the liquid production curve ($\widehat{q}_{\text{liq}}(q_{\text{inj}}, G)$), which is the sum of $\widehat{q}_{\text{oil}}$ and $\widehat{q}_{\text{water}}$, are illustrated in Figure 3.2, where colored regions represent instances of this curve depending on the possible values assumed by GLR, WC and PI. In Figure 3.2(a) one can observe that there are two distinct regions, one before and one after the injection point that induces the maximum liquid production rate. Before this point,



(a)



(b)



(c)

Figure 3.2: Well production curves as functions of nominal, $+10\%$, and -10%, GLR (top), WC (center) and PI (bottom).

a higher GLR is a benefit for the production, helping to elevate the liquid stream. However, after the point of maximum production, a higher ratio of gas coming from the formation works against the flow of fluid, as an extra friction force. In Figure 3.1(b), for a same gas injection rate the production of liquid tend to be higher for lower water ratios, since in general water is denser than the oil. However, for some WC values this may be the other way around, since water is less viscous than oil. For Figure 3.1(c), as expected, a higher PI induces more liquid production for any lift gas rate.

Notice that those are qualitative illustrations of the effects of uncertainty in the well-production function, and quantifying it is much more difficult due to the nonlinearities and conceptual aspect of $\widehat{q}_r$.

3.2.3 Nominal MILP production problem

Now let's focus again on the nominal MINLP production problem (3.1). Both the simulated well-production curve and the real well-production function are conceptual, since one cannot have a explicit formulation for them. Thus, problem (3.1) continues to be a conceptual MINLP regardless the assumption we made earlier replacing the real function by the simulated one. Conceptual mathematical programming problems in general cannot be solved directly. Approximating these functions by fitting and validating simulated sample data – or even real measurements – to nonlinear models would yield to an approximated MINLP formulation, but this is in itself a challenging problem.

This motivates the use of PWL models to represent the aforementioned relations, which arise directly from the sample data – real or simulated. The PWL approach leads to an MILP approximation of the optimization problem, meaning that a solution close to the global optimum may be reached considering a sufficient number of sample points, and using, for example, off-the-shelf and specialized algorithms.

Figure 3.3 illustrates examples of a PWL approximation curves of a well-production curve. Notice that depending on the sample points used as the base for the linear model, the PWL curves differ. For instance, it is clear that, unless for the extremal sample points of gas injection (q_{inj}), the piecewise-linear curves PWL_1 and PWL_2 would predict a distinct oil production value (q_{oil}) for any lift gas rate, (see Figure 3.3). It is also clear that any piecewise-linear approximation assumes a different function value when compared to

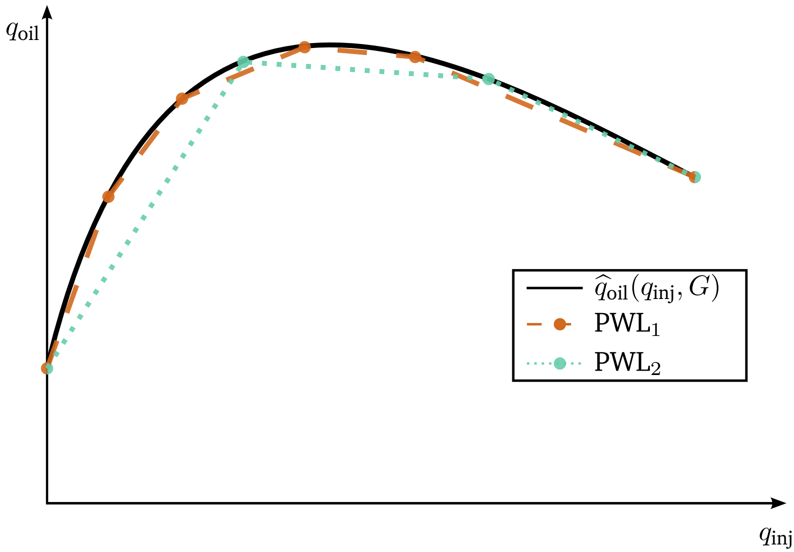


Figure 3.3: Well-production curve and two of its possible piecewise-linear approximation curves.

the original function value. This happens for any point in the domain that is not a sample point – given the original function has non zero concavity in its entire domain. This means that by using a PWL approximation for the well-production curve, model uncertainty would be introduced in the optimization problem. Well, this is not an exclusive issue of piecewise-linearization, since for fitting any proxy model to sample data would generate a model that to some extent has inaccuracy in predicting the original function value.

Many possible sources of uncertainty in the production models have already been presented and this is just another one to add to the count. Unlike the other sources of uncertainty this one is not related to the system, but it is introduced to make the optimization problem tractable. Since this is a perturbation inserted by the model builder it is intuitive to think that preventive actions that might at least diminish its effects might exist. In fact, there are a few possibilities.

One that first comes to mind would be to use as many as possible/available sample points to come up with a tight piecewise-linear representation. This approach has two issues. On one hand, depending on the simulation time and on the size of the problem, getting the sample data may become a time consuming pre-optimization step. On the other hand, for more complex optimization problems the size of the PWL models could become an issue (i.e. adding to many additional variables) yielding to harder to solve or even an unsolvable problem.

A procedure to remedy the latter issue is presented by Codas et al. [75]. It starts with the idea presented above, getting a high resolution piecewise-linear model, then it applies a greedy heuristic to reduce the number of sample points, while ensuring a maximum error with respect to the tight piecewise-linear representation. The problem with this approach is that it does not handle the issue of getting an overly detailed representation. Also it assumes that for any point in the domain, the difference between conceptual function value and the value predicted by the tight PWL function is not greater than the maximum tolerated error sought by the algorithm.

An approach to remedy both issues of a tight piecewise-linear representation is to design an iterative process for solving the optimization problem. Starting with a small number of sampling points the problem is solved, and solution is compared to a simulation of the complete production scenario running with the decision variable values calculated by the optimization problem. If all variables are in a given pre-specified maximum error the solution of the optimization problem is kept, otherwise, the production curves are oversampled according to some heuristic, and the process starts over until the tolerance is achieved. This approach has been used with success in [76], although being very sensitive since it relies greatly on the heuristics for being efficient. A drawback is that it compares the curves in a post-optimization stage.

Here we propose an alternative approach to those. This is only applicable to piecewise-linearization of conceptual well-production curves, or any conceptual function that has a convex or concave shape. It explores the fact that an error function, defined as the conceptual function minus its piecewise-linear approximation, is a convex function. Thus, it uses line-search without derivative methods [77] in a recursive algorithm to provide sampling points and generates a PWL approximation curve that guaranties a maximum error for all the domain. The algorithm does not ensure that the minimum number of sample points for a given tolerance would be obtained, but experimental results showed it is close enough. In Appendix A we go into the details of the algorithm.

By using one of the techniques mentioned above, and therefore having a PWL model that is of reasonable size and that has a negligible approximation error, it is possible to replace the simulated well-production function by its approximated piecewise-linear function, assuming no extra uncertainty is being added to the problem.

The general mathematical models for multidimensional piecewise-linear approximation include the convex combination (CC), the logarithmic version of CC (Log), the disaggregated convex combination (DCC), its logarithmic version DLog, the incremental (Inc), the multiple choice (MC) model, and special ordered sets of type 2 (SOS2) [40]. The SOS2 model is implemented in software by the branch-and-bound algorithm, whereas the other models are made explicit in the mathematical programming formulation using binary variables and additional constraints. Specific production optimization formulations for each of this piecewise-linear models are presented by [9]. A general MILP approximated formulation for Problem 3.1 (P) that is independent of the model chosen for piecewise linearization follows[1]:

$$
\begin{aligned}
\bar{P} :\ \max\ & \bar{q}_{\text{oil}} && (3.2a)\\
\text{s.t.} :\ & q_{\text{inj}} \leq Q_{\text{inj}}^{\max} && (3.2b)\\
& q_{\text{inj}} = \sum_{q_i \in \mathcal{K}} \lambda_{q_i} \cdot q_i, && (3.2c)\\
& \bar{q}_r = \sum_{q_i \in \mathcal{K}} \lambda_{q_i} \cdot \widehat{q}_r(q_i, G),\ \forall r \in \mathcal{R} && (3.2d)\\
& \left(\lambda_{q_i}\right)_{\mathcal{K}} \text{ induces a PWL function} && (3.2e)\\
& \bar{q}_r^{\text{S}} \leq q_r^{\text{S,max}},\ \forall r \in \mathcal{R} && (3.2f)\\
& \bar{q}_r^{\text{S}} = \bar{q}_r,\ \forall r \in \mathcal{R} - \{\text{gas}\} && (3.2g)\\
& \bar{q}_{\text{gas}}^{\text{S}} = q_{\text{inj}} + \bar{q}_{\text{gas}} && (3.2h)\\
& y \in \{0,1\} && (3.2i)
\end{aligned}
$$

where $\bar{P}$ is an MILP approximation of problem P. The additional **variables** for Problem (3.2) are:

- $\bar{q}_r$ is the rate of phase r produced according with the piecewise-linear function;

[1] actually it could be any of the aggregated, SOS2 or Incremental PWL models.

- $\bar{q}_r^{\mathrm{S}}$ is the approximated total flow of component r flowing to the separator;
- $\lambda_{q_i} \in [0, 1]$ is a weighting variable associated with the $(q_i)^{th}$ breakpoint of the piecewise-linear production function.

also the following **parameters** are introduced:

- $\mathcal{K}$ has the breakpoints of the production function; Breakpoints are of type (q_i), where q_i is a lift gas injection rate;
- $\widehat{q}_r(q_i, G)$ is the nonlinear production function evaluated in the breakpoint $q_i \in \mathcal{K}$ to form a sampling point.

The production of the well as a function of the lift-gas injection rate is modeled by Eqs. (3.2c)-(3.2d). Eq. (3.2e) induces a piecewise-linear function of the production curves. Notice that a full representation of the piecewise-linear model depends on the PWL model chosen. For instance, for a SOS2 model one has to add to the previous set of constraints of Formulation (3.2) the following equations,

$$\sum_{q_i \in \mathcal{K}} \lambda_{q_i} = y \tag{3.2j}$$

$$\lambda_{q_i} \geq 0, \ \forall q_i \in \mathcal{K} \tag{3.2k}$$

$$\left(\lambda_{q_i}\right)_{q_i \in \mathcal{K}} \text{ are SOS2.} \tag{3.2l}$$

Other models for piecewise-linearization may have impact on solving time [9], but the formulations will be equivalent regardless of the chosen PWL model. The equations not described here have similar or identical counterparts in Formulation (3.1). Also note that Eqs. (3.1d)-(3.1e) are eliminated in $\bar{P}$ by considering the extreme breakpoints are such that, $\min\{q_i : q_i \in \mathcal{K}\} = q_{\mathrm{inj}}^{\min}$ and $\max\{q_i : q_i \in \mathcal{K}\} = q_{\mathrm{inj}}^{\max}$

3.2.4 Uncertain MILP production problem

Formulation (3.2) is an approximation of Formulation (3.1) and uncertainty affects both likewise. However, the linear model $\bar{P}$ allows the use of more powerful tools, so we analyze the impact of uncertainty in the optimization problem through this framework. As previously mentioned uncertainty could arise in equipment capacities and in the well-production function, more precisely in the parameters that define the production curve. For our toy example, we identify the elements that could have uncertain values as being $Q_{\text{inj}}^{\max}$, $q_r^{\text{S},\max}$, $\forall r \in \mathcal{R}$ and $\widehat{q}_r(q_i, G)$, $\forall r \in \mathcal{R}$, $\forall q_i \in \mathcal{K}$. Notice that in the linear optimization problem these elements are all parameters. Even the functions $\widehat{q}_r$, since in fact, what is being used are their evaluations for specific injection rates, q_i. To make the notation clearer, let's denote that a parameter is uncertain by adding a tilde ($\sim$) over it.

It is possible to divide the uncertain parameters into two groups. One containing $\tilde{Q}_{\text{inj}}^{\max}$ and $\tilde{q}_r^{\text{S},\max}$, in which the uncertain element appears directly in the optimization problem, and one with $\widehat{q}_r(q_i, \tilde{G})$. For this second group the uncertainty is in reality present in some of the parameters of the set $\tilde{G}$, which of course, indirectly makes the function value for each injection point, also uncertain. Given the uncertain parameters above, an uncertain MILP version of $\bar{P}$ can be written as:

$$\tilde{P}: \max \ \tilde{q}_{\text{oil}} \tag{3.3a}$$

$$\text{s.t.}: q_{\text{inj}} \leq \tilde{Q}_{\text{inj}}^{\max}, \ \tilde{Q}_{\text{inj}}^{\max} \in \mathcal{U}_{Q_{\text{inj}}^{\max}} \tag{3.3b}$$

$$q_{\text{inj}} = \sum_{q_i \in \mathcal{K}} \lambda_{q_i} \cdot q_i, \tag{3.3c}$$

$$\tilde{q}_r = \sum_{q_i \in \mathcal{K}} \lambda_{q_i} \cdot \widehat{q}_r(q_i, \tilde{G}), \ \widehat{q}_r(q_i, \tilde{G}) \in \mathcal{U}_{\widehat{q}_r(q_i, G)}, \ldots \quad \ldots \forall r \in \mathcal{R} \tag{3.3d}$$

$$\tilde{q}_r^{\text{S}} \leq \tilde{q}_r^{\text{S},\max}, \ \tilde{q}_r^{\text{S},\max} \in \mathcal{U}_{q_r^{\text{S},\max}}, \ \forall r \in \mathcal{R} \tag{3.3e}$$

$$\tilde{q}_r^{\text{S}} = \tilde{q}_r, \ \forall r \in \mathcal{R} - \{\text{gas}\} \tag{3.3f}$$

$$\tilde{q}_{\text{gas}}^{\text{S}} = q_{\text{inj}} + \tilde{q}_{\text{gas}} \tag{3.3g}$$

$$y \in \{0, 1\} \tag{3.3h}$$

$$\sum_{q_i \in \mathcal{K}} \lambda_{q_i} = y \tag{3.3i}$$

$$\lambda_{q_i} \geq 0, \ \forall q_i \in \mathcal{K} \tag{3.3j}$$

$$\left(\lambda_{q_i}\right)_{q_i \in \mathcal{K}} \text{ are SOS2.} \tag{3.3k}$$

where $\tilde{q}_r$ and $\tilde{q}_r^{\text{S}}$ are state variables related respectively to $\bar{q}_r$ and $\bar{q}_r^{\text{S}}$. Yet, those are defined by equations containing uncertain parameters. All other parameters and variables are the same as in Formulation (3.2) except for the uncertain parameters $\tilde{Q}_{\text{inj}}^{\max}$, $\tilde{q}_r^{\text{S},\max}$, and

$\widehat{q}_r(q_i, \tilde{G})$, each one associated with a space $\mathcal{U}$ representing its possible states. Here we denote each $\mathcal{U}$ as being an uncertainty set, since an uncertain element is modeled as a range of possible values, with an undefined probabilistic distribution. However, in other uncertain optimization frameworks, where, for example, parameters are random variables, $\mathcal{U}$ could assume the role a of probability state space.

The uncertainty sets $\mathcal{U}_{Q_{\text{inj}}^{\max}}$ and $\mathcal{U}_{q_r^{\text{S},\max}}$ can be explicitly defined, provided a range model for the respective parameters are available. Defining $\mathcal{U}_{\widehat{q}_r(q_i,G)}$ is not as straightforward. What we assume is to have an uncertainty model for each uncertain parameter in $\tilde{G}$. But since $\widehat{q}_r$ are conceptual functions it is not possible to analytically arrive to an explicit model for $\mathcal{U}_{\widehat{q}_r(q_i,G)}$. So there is a need for using approximation techniques, such as, Monte Carlo methods. In general these methods have the following steps [78]:

(I) Define a domain of possible inputs;

(II) Generate inputs randomly over the domain;

(III) Perform a deterministic computation on the inputs;

(IV) Aggregate the results.

Let's exemplify this latter idea. Imagine that for all parameters in the set $\tilde{G}$, the only one that has an uncertain value is GLR. We then assume to have an explicit uncertainty model $\mathcal{U}_{\text{GLR}}$ for the parameter $\widetilde{\text{GLR}}$ as depicted in Figure 3.4. Now, for example, in order to define a range model for the uncertainty set $\mathcal{U}_{\widehat{q}_{\text{oil}}(q_i,G)}$ of the uncertain parameter $\widehat{q}_{\text{oil}}(q_i, \tilde{G})$, one has to apply a Monte Carlo method. The first step is already given because the range model $\mathcal{U}_{\text{GLR}}$ defines the domain of possible inputs. The second is to randomly, and conveniently, choose input points within the domain. The third step is to use the simulated function to calculate the oil flow rates for that specific injection rate q_i for each of the selected inputs. Finally, from the set of results one has to compute the maximum and the minimum values. Those are used to define a range model for $\mathcal{U}_{\widehat{q}_{\text{oil}}(q_i,G)}$. All these steps are illustrated in Figure 3.4.

The second step is the most critical one and requires experience with the nonlinearities of the production function to achieve a reliable final result. For instance, since what is being sought are the extremal values assumed by $\widehat{q}_{\text{oil}}(q_i, \tilde{G})$, which could be directly used to define a range model for the uncertainty set $\mathcal{U}_{\widehat{q}_{\text{oil}}(q_i,G)}$, it is intuitive to think that by choosing the extremal points of $\mathcal{U}_{\text{GLR}}$ would

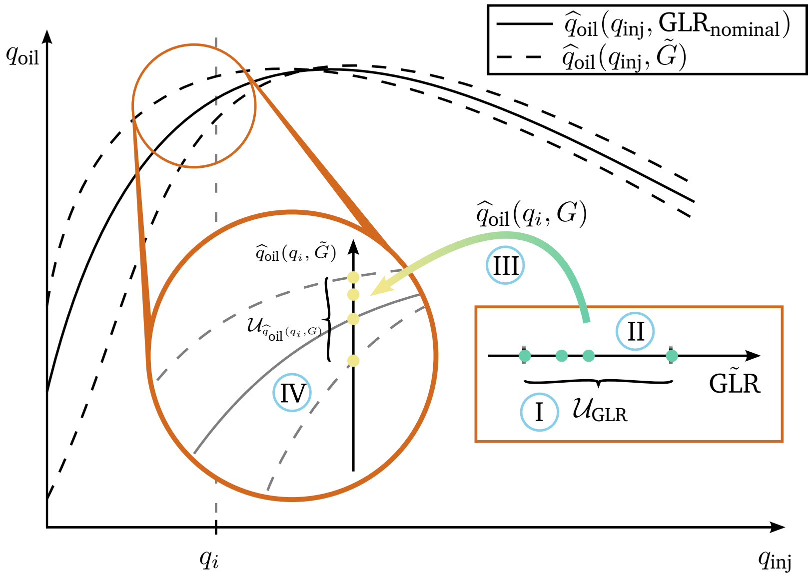


Figure 3.4: Steps of a Monte Carlo method for approximating $\mathcal{U}_{\widehat{q}_{\text{oil}}(q_i,G)}$ for an uncertain GLR.

be enough. For many situations this is true, but the nonlinearities leave no guaranty.

Notice that when more than one parameter of $\tilde{G}$ is uncertain, one has to use an uncertainty set $\mathcal{U}_G$ as the domain of possible inputs. This set defines the correlation between the uncertain parameters. While for uncorrelated data – or lack of knowledge about possible correlations, one would use an n-box model (a box in the $\mathbb{R}^n$) for $\mathcal{U}_G$, the method presented earlier allows the use of any shape for this uncertainty set. Figure 3.5 illustrates this for a well-production curve that has as uncertain parameters GLR and WC.

There are many possible mathematical representations for the uncertainty sets of Formulation (3.3). They can be, for example, represented by a central element and a symmetrical offset, or by the nominal value and asymmetrical offsets. For convenience, we chose a representation based on the limiting points, with tags $\inf$ and $\sup$. So, for instance, $\tilde{Q}^{\max}_{\text{inj}} \in \mathcal{U}_{Q^{\max}_{\text{inj}}}$ and $\mathcal{U}_{Q^{\max}_{\text{inj}}} = [Q^{\max}_{\text{inj},\inf}, Q^{\max}_{\text{inj},\sup}]$. In order to simplify the notation for $\widehat{q}_r(q_i, \tilde{G})$, $\mathcal{U}_{\widehat{q}_r(q_i,G)}$ is defined as $[\widehat{q}_{r,\inf}(q_i), \widehat{q}_{r,\sup}(q_i)]$

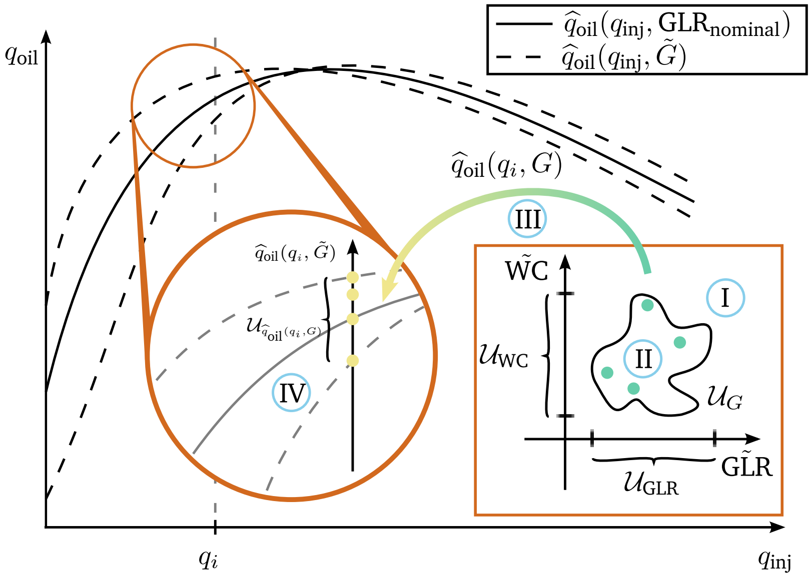


Figure 3.5: Steps of a Monte Carlo method for approximating $\mathcal{U}_{\widehat{q}_{\text{oil}}(q_i,G)}$ for uncertain GLR and WC.

Now that all uncertainty sets $\mathcal{U}_{Q_{\text{inj}}^{\max}}$, $\mathcal{U}_{q_r^{\text{S},\max}}$, and $\mathcal{U}_{\widehat{q}_r(q_i,G)}$ can be defined, the uncertain optimization problem $\tilde{P}$ has a complete description and we can move ahead to build a robust formulation of the production optimization problem.

3.3 ROBUST PRODUCTION OPTIMIZATION METHODOLOGY

Uncertain optimization problems cannot be solved directly since their structure is not the same as in standard optimization models. Various paradigms exist to transform an uncertain optimization problem into a corresponding one for which a solution can be computed. Each of these paradigms follows distinct modeling philosophies to perform the transformation of the uncertain problem, and for each one a subfield of optimization under uncertainty has emerged. Examples of these areas of study are stochastic programming, stochastic dynamic programing and robust optimization.

In the present work we explore the possibilities offered by the robust optimization paradigm. We are motivated by its simpler assumptions regarding the modeling of uncertainty and also for the possibility to generate simpler to compute robust counterpart (RC) problems when compared to other paradigms for optimizing under uncertainty.

The standard robust optimization paradigm follows four basic modeling philosophies that in short are [17]:

- the uncertain data in a problem is modeled in terms of ranges;
- decisions must be taken before any uncertain data become known;
- constraints of the problem are "hard";
- the objective is worst-case oriented.

In particular, it is used as support, the theory developed in a subarea called robust linear optimization, since all starting point uncertain models are linear in its variables.

Before starting building a RC it is necessary to identify how the uncertain parameters of problem $\tilde{P}$ relate to each other. For instance, in robust optimization there are two frameworks referred to as *column-wise* and *row-wise* (a.k.a. *constraint-wise*). In the column-wise framework, the uncertain parameters associated with the same variable are grouped under an uncertainty set that has a geometry reflecting the relationship between the individual uncertainty range models. For the row-wise framework, however, the parameters in a same constraint/objective are gathered in an uncertainty set. These sets that establish the correlated domain of possible value for each group of parameters are known respectively as column-wise uncertainty models and row-wise uncertainty models, depending on the framework chosen.

In order to analyze which of the frameworks is more suitable for problem $\tilde{P}$, let us reformulate it into a canonical form similar to:

$$\begin{aligned} \max \quad & \mathbf{c}^{\intercal}\boldsymbol{x} \\ \text{s.t.:}\ & \tilde{\mathbf{A}}\boldsymbol{x} \le \mathbf{b} \\ & \boldsymbol{x} \ge \mathbf{0}, \end{aligned}$$

for $\tilde{a}_{ij} \in \mathcal{U}_{a_{ij}}$ (for all uncertain coefficients of $\tilde{\mathbf{A}}$). Notice that all uncertainty is on the left hand side (LHS) of the constraints. Thus, starting with problem $\tilde{P}$ (Eqs. (3.3)), this is achieved by first replacing all the state variable occurrences by their corresponding equations. Then, by epigraph reformulation the uncertain objective equation is inserted in the constraints using an auxiliary variable t. In the next step, another auxiliary variable, v, is used to place the uncertain elements in the RHS into the LHS. Finally since all remaining equality constraints contain only certain parameters it is possible to simply transform them into two inequalities. Therefore, for a decision vector,

$$\boldsymbol{x}^{\intercal} = \begin{array}{c} (\forall q_i \in \mathcal{K}) \\ \begin{bmatrix} \lambda_{q_i} & \dots & y & t & v \end{bmatrix} \end{array}$$

$\mathbf{c}$, $\mathbf{b}$, and $\tilde{\mathbf{A}}$ would be,

$$\mathbf{c}^{\intercal} = \begin{array}{c} (\forall q_i \in \mathcal{K}) \\ \begin{bmatrix} 0 & \dots & 0 & 1 & 0 \end{bmatrix} \end{array},$$

$$\mathbf{b}^{\intercal} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix},$$

$$\tilde{\mathbf{A}} = \begin{array}{cc} & (\forall q_i \in \mathcal{K}) \\ \begin{array}{c} (3.3a) \\ (3.3b) \\ (3.3e) \\ (3.3e) \\ (3.3e) \\ (3.3i) \\ (3.3i) \\ \\ \\ \end{array} & \begin{bmatrix} -\widehat{q}_{\text{oil}}(q_i, \tilde{G}) & \dots & 0 & 1 & 0 \\ q_i & \dots & 0 & 0 & -\tilde{Q}_{\text{inj}}^{\max} \\ \widehat{q}_{\text{oil}}(q_i, \tilde{G}) & \dots & 0 & 0 & -\tilde{q}_{\text{oil}}^{\text{S},\max} \\ \widehat{q}_{\text{water}}(q_i, \tilde{G}) & \dots & 0 & 0 & -\tilde{q}_{\text{water}}^{\text{S},\max} \\ q_i + \widehat{q}_{\text{gas}}(q_i, \tilde{G}) & \dots & 0 & 0 & -\tilde{q}_{\text{gas}}^{\text{S},\max} \\ 1 & \dots & -1 & 0 & 0 \\ -1 & \dots & 1 & 0 & 0 \\ 0 & \dots & 0 & 0 & 1 \\ 0 & \dots & 0 & 0 & -1 \end{bmatrix} \end{array},$$

In analyzing matrix $\tilde{\mathbf{A}}$, one can see that:

- there are uncertain coefficients relating with each other in both fashions, column-wise and row-wise;
- row-wise correlations happen in rows 1, 3-5, between the coefficients $\widehat{q}_r(q_i, \tilde{G})$. For example, in the first row, the coefficients

$\widehat{q}_{\text{oil}}(q_i, \tilde{G})$, $\forall q_i \in \mathcal{K}$ relate with each other. However, the relationship that a value in $\mathcal{U}_{\widehat{q}_r(q_i,G)}$ has with all other breakpoints $q_i \in \mathcal{K}$ is given by the the production function $\widehat{q}_{\text{oil}}$. That means it could be difficult to establish these row-wise uncertainty set models, since they depend on the implicit nonlinearities of the production functions.

- column-wise correlations happen in columns 1-$|K|$, ($|K|$+3). For the last column, it is intuitive to consider that the coefficients $\tilde{q}_r^{\text{S,max}}$ are related by a straight-line linking the minimum possible capacities in each $\mathcal{U}_{q_r^{\text{S,max}}}$ for each $r \in \mathcal{R}$, to the maximum capacities. For the first $|K|$ columns the relation is between the production curves values for each phase, at a same breakpoint q_i. Clearly the relationships that arise are based on WC and GLR values. If those are certain the column-wise uncertainty models would also be a straight-line, otherwise, if there is a range of possible values of WC or GLR, the column-wise uncertainty models would be polyhedron.

- For uncorrelated coefficients, either in row-wise or column-wise uncertainty models, the coefficient can take any value within the its original range of possibilities.

Based on the considerations above, the column-wise framework is chosen. This means that the uncertain problem has in fact the following structure,

$$\begin{aligned} &\max\ \mathbf{c}^\intercal \boldsymbol{x} \\ &\text{s.t.:} \sum_{j=1}^{|K|+3} x_j \tilde{\mathbf{a}}_j \leq \mathbf{b},\ \forall (\tilde{\mathbf{a}}_j \in \mathcal{U}_j,\ j = 1, \ldots, |K| + 3) \\ &\qquad\qquad \boldsymbol{x} \geq \mathbf{0}. \end{aligned}$$

where $\mathcal{U}_j$ is the uncertainty set model relating the individual uncertainty models for the uncertain parameters in column j. As it was shown by Soyster, solving this formulation is equivalent to solve an ordinary LP in the form,

$$\begin{aligned} P_{\mathcal{U}} :\ &\text{Maximize}\ \ \mathbf{c}^\intercal \boldsymbol{x} \\ &\text{Subject to: } \mathbf{A}^* \boldsymbol{x} \leq \mathbf{b} \\ &\qquad\qquad\qquad \boldsymbol{x} \geq \mathbf{0}. \end{aligned}$$

where each entry $a_{ij}^* = \sup_{\tilde{\mathbf{a}}_j \in \mathcal{U}_j} (\tilde{\mathbf{a}}_j)_i$. Notice that this is the same as getting the supremum values of the projection of $\mathcal{U}_j$ over the uncertainty space of each individual uncertain parameter. Thus, this strategy disregards the relations imposed by $\mathcal{U}_j$ to the uncertain coefficients.

Therefore, for our problem the matrix $\mathbf{A}^*$ is,

$$\mathbf{A}^* = \begin{array}{cc} & \begin{matrix} \phantom{-\widehat{q}_{\text{oil,inf}}(q_i)} & (\forall q_i \in \mathcal{K}) & & & \end{matrix} \\ \begin{matrix} \text{(3.3a)} \\ \text{(3.3b)} \\ \text{(3.3e)} \\ \text{(3.3e)} \\ \text{(3.3e)} \\ \text{(3.3i)} \\ \text{(3.3i)} \\ \\ \\ \end{matrix} & \begin{bmatrix} -\widehat{q}_{\text{oil,inf}}(q_i) & \dots & 0 & 1 & 0 \\ q_i & \dots & 0 & 0 & -Q_{\text{inj,inf}}^{\max} \\ \widehat{q}_{\text{oil,sup}}(q_i) & \dots & 0 & 0 & -q_{\text{oil,inf}}^{\text{S,max}} \\ \widehat{q}_{\text{water,sup}}(q_i) & \dots & 0 & 0 & -q_{\text{water,inf}}^{\text{S,max}} \\ q_i + \widehat{q}_{\text{gas,sup}}(q_i) & \dots & 0 & 0 & -q_{\text{gas,inf}}^{\text{S,max}} \\ 1 & \dots & -1 & 0 & 0 \\ -1 & \dots & 1 & 0 & 0 \\ 0 & \dots & 0 & 0 & 1 \\ 0 & \dots & 0 & 0 & -1 \end{bmatrix} \end{array}$$

where the inf and sup elements come from the definitions of the uncertainty sets $\mathcal{U}_{Q_{\text{inj}}^{\max}}$, $\mathcal{U}_{q_r^{\text{S,max}}}$, and $\mathcal{U}_{\widehat{q}_r(q_i,G)}$.

From the formulation based on matrix $\mathbf{A}^*$ one can, by introducing convenient state variables and by suppressing variables t and v, arrive at a robust counterpart formulation in the form,

$$P_{\text{RC}} : \max \ \bar{q}_{\text{oil,inf}} \tag{3.4a}$$

$$\text{s.t.} : q_{\text{inj}} \leq Q_{\text{inj,inf}}^{\max} \tag{3.4b}$$

$$q_{\text{inj}} = \sum_{q_i \in \mathcal{K}} \lambda_{q_i} \cdot q_i, \tag{3.4c}$$

$$\bar{q}_{r,\text{inf}} = \sum_{q_i \in \mathcal{K}} \lambda_{q_i} \cdot \widehat{q}_{r,\text{inf}}(q_i), \ \forall r \in \mathcal{R} \tag{3.4d}$$

$$\bar{q}_{r,\text{sup}} = \sum_{q_i \in \mathcal{K}} \lambda_{q_i} \cdot \widehat{q}_{r,\text{sup}}(q_i), \ \forall r \in \mathcal{R} \tag{3.4e}$$

$$\bar{q}_{r,\text{sup}}^{\text{S}} \leq q_{r,\text{inf}}^{\text{S,max}}, \ \forall r \in \mathcal{R} \tag{3.4f}$$

$$\bar{q}_{r,\text{sup}}^{\text{S}} = \bar{q}_{r,\text{sup}}, \ \forall r \in \mathcal{R} - \{\text{gas}\} \tag{3.4g}$$

$$\bar{q}_{\text{gas,sup}}^{\text{S}} = q_{\text{inj}} + \bar{q}_{\text{gas,sup}} \tag{3.4h}$$

$$y \in \{0, 1\} \tag{3.4i}$$

$$\sum_{q_i \in \mathcal{K}} \lambda_{q_i} = y \tag{3.4j}$$

$$\lambda_{q_i} \geq 0, \forall q_i \in \mathcal{K} \tag{3.4k}$$

$$\left(\lambda_{q_i}\right)_{q_i \in \mathcal{K}} \text{ are SOS2.} \tag{3.4l}$$

where P_{RC} is a robust version of problem P. The additional state variables for the robust problem are:

- $\bar{q}_{r,\inf}$ is the infimum of the production phase r of the well;
- $\bar{q}_{h,\sup}$ is the supremum of the production phase r of the well;
- $\bar{q}^{\mathrm{S}}_{r,\inf}$ is the infimum flow of phase r flowing to the separator;
- $\bar{q}^{\mathrm{S}}_{r,\sup}$ is the supremum flow of phase r flowing to the separator;

also the following parameters are introduced:

- $Q^{\max}_{\mathrm{inj},\inf}$ is the lower bound value for an uncertain lift gas compression capacity;
- $q^{\mathrm{S},\max}_{r,\inf}$ is the lower bound value for an uncertain handling capacity of component r in the separator;
- $\widehat{q}_{r,\sup}(q_i)/\widehat{q}_{r,\inf}(q_i)$ is upper/lower bound value of production of phase r for an injection rate of q_i according with the production function $\widehat{q}_r$; Notice, that the existence of upper and lower bound values could be due to any uncertain parameters in $\tilde{G}$, which also define the production function.

Note that the set of Equations (3.2) is similar to the set of Equations (3.4). In fact, problem P_{RC} can be seen as having two piecewise-linear production curves enveloping all the uncertainty of the each well-production curve (see, Figure 3.6). The robust formulation uses sufficiently conservative objective and constraints in order to satisfy the system constraints for any realization of the uncertainties of the problem.

The methodology followed for the toy example can be easily extended to more complex production optimization problems. In Chapter 4 two distinct standard production models are investigated. In Section 4.1 a similar methodology is applied to an optimization problem with complexity similar to the one studied in [8], with

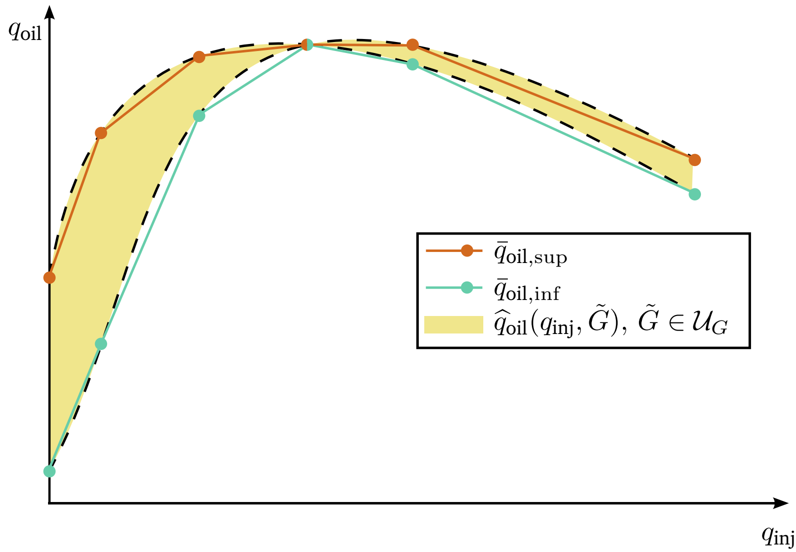


Figure 3.6: Well production envelope curves q_{lower} and q_{upper}, and their PWL functions $q_{\text{oil,inf}}$ and $q_{\text{oil,sup}}$.

satellite wells and routing decisions. Following, in Section 4.2, with slightly modifications this methodology is used for a production optimization scenario with subsea completion, similar to the structure of the problem presented in [9]. Next, in Chapter 5, the modeling results of the first problem, are subjected to a computational analysis, by applying the solutions of the robust and nominal formulations to a representative oil field available in simulation software.

3.4 SUMMARY

This chapter started discussing the causes of uncertainty in production models, and then provided a review of a few works that have mentioned or dealt with uncertainty in production systems. Following, uncertainty is examined in the perspective of production optimization problems. A toy example was used to exemplify the robust optimization methodology designed in this dissertation.

4 ROBUST PRODUCTION OPTIMIZATION MODELS

This chapter is divided in two sections. In each of them a different standard production optimization problem is presented. Then, the concepts developed in Chapter 3 are used to produce a robust counterpart of the problem. For the model in Section 4.2 the methodology is extended to account for uncertainty in equality constraints.

4.1 ROBUST PRODUCTION OPTIMIZATION – SATELLITE WELLS

This section presents a robust production optimization model for a production system of a group of satellite wells operated by gas lift which allows for the routing of production flows from wells to separators. The contents of this section were published in [74], but in a distinct format.

4.1.1 Problem Formulation

Let us consider the deterministic optimization problem described in [8] which consists in distributing a limited rate of pressurized gas, with the selection of production routing under multiple constraints. Such scenarios appear in multi-reservoir production systems, in which wells are satellite but can have their production routed to different separators. Figure 4.1 has the production network of the problem dealt in this section. Here, for simplicity, the objective function is considered the maximization of oil production of a group of wells. Considering nominal values for all parameters of the problem, the production optimization problem can be viewed as solving a conceptual MINLP problem, in the sense that well production functions and other relationships are not explicitly available. As shown in Section 3.2.3 these functions can be evaluated by simulation software approximated with piecewise-linear (PWL) models. Provided the breakpoints are well defined to approximate the well-production functions, it is possible to render the conceptual MINLP a MILP problem. The nominal production optimization problem suggested in this chapter extends the *compact formulation* from [8] to account for the distinct operating pressure of the separator to which each well may be connected:

$$\bar{P} : \max \sum_{n \in \mathcal{N}} \bar{q}^{n}_{\text{oil}}, \tag{4.1a}$$

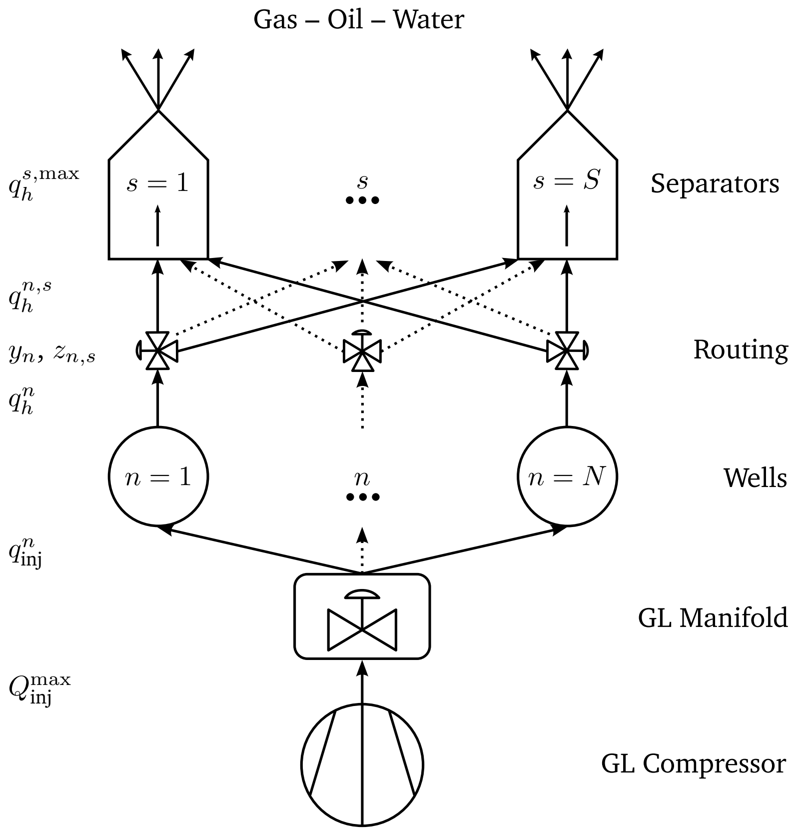


Figure 4.1: Production network – satellite wells.

subject to a limited compression capacity:

$$\text{s.t.}: \sum_{n \in \mathcal{N}} q_{\text{inj}}^n \le Q_{\text{inj}}^{\max}, \tag{4.1b}$$

and for each well $n \in \mathcal{N}$:

$$q_{\text{inj}}^{n,s} = \sum_{q_i \in \mathcal{K}^{n,s}} \lambda_{q_i}^{n,s} \cdot q_i,\ \forall s \in \mathcal{S}_n, \tag{4.1c}$$

$$\sum_{s \in \mathcal{S}_n} q_{\text{inj}}^{n,s} = q_{\text{inj}}^n, \tag{4.1d}$$

$$\bar{q}_h^{n,s} = \sum_{q_i \in \mathcal{K}^{n,s}} \lambda_{q_i}^{n,s} \cdot \widehat{q}_h(q_i, G^{n,s}),\ \forall s \in \mathcal{S}_n,\ \forall h \in \mathcal{H}, \tag{4.1e}$$

$$\sum_{s \in \mathcal{S}_n} \bar{q}_h^{n,s} = \bar{q}_h^n,\ \forall h \in \mathcal{H}, \tag{4.1f}$$

$$\sum_{s \in \mathcal{S}_n} z_{n,s} = y_n, \tag{4.1g}$$

with a SOS2 piecewise-linear model, for all well $n \in \mathcal{N}$ and separator $s \in \mathcal{S}_n$:

$$\sum_{q_i \in \mathcal{K}^{n,s}} \lambda_{q_i}^{n,s} = z_{n,s}, \tag{4.1h}$$

$$\lambda_{q_i}^{n,s} \geq 0,\ \forall q_i \in \mathcal{K}^{n,s}, \tag{4.1i}$$

$$\left(\lambda_{q_i}^{n,s}\right)_{q_i \in \mathcal{K}^{n,s}} \text{ are SOS2}, \tag{4.1j}$$

and for each separator $s \in \mathcal{S}$:

$$\bar{q}_h^s \leq q_h^{s,\max},\ \forall h \in \mathcal{H}, \tag{4.1k}$$

$$\sum_{n \in \mathcal{N}_s} \bar{q}_h^{n,s} = \bar{q}_h^s,\ \forall h \in \mathcal{H} - \{\text{gas}\}, \tag{4.1l}$$

$$\sum_{n \in \mathcal{N}_s} (\bar{q}_{\text{gas}}^{n,s} + q_{\text{inj}}^{n,s}) = \bar{q}_{\text{gas}}^s, \tag{4.1m}$$

with:

$$y_n \in \{0,1\},\ \forall n \in \mathcal{N}, \tag{4.1n}$$

$$z_{n,s} \in \{0,1\},\ \forall n \in \mathcal{N},\ \forall s \in \mathcal{S}_n. \tag{4.1o}$$

The **variables** of the formulation are:

- q_{inj}^n is the rate of lift-gas injection into well n;
- $q_{\text{inj}}^{n,s}$ is the lift-gas injection rate of well n when routed to separator s;
- $\bar{q}_h^n$ is the rate of component h produced according with the piecewise-linear function of well n;
- $\bar{q}_h^{n,s}$ is the component h of production rate from well n when directed to separator s;

- $\bar{q}_h^s$ is the approximated total flow of component h flowing to the separator s;
- $y_n \in \{0, 1\}$ is 1 when well n is active or else 0;
- $z_{n,s} \in \{0, 1\}$ assumes value 1 if well n is connected to separator s, and 0 otherwise;
- $\lambda_{q_i}^{n,s} \in [0, 1]$ is a weighting variable associated with the breakpoint $q_i \in \mathcal{K}^{n,s}$ of the piecewise-linear production function of well n to separator s;

The **parameters** are as follows:

- $\mathcal{N} = \{1, \ldots, N\}$ is the set of gas-lifted wells and $\mathcal{N}_s \subseteq \mathcal{N}$ is the subset of wells that can be connected to separator s;
- $\mathcal{S} = \{1, \ldots, S\}$ is the set of separators and $\mathcal{S}_n \subseteq \mathcal{S}$ is the subset of separators that can received production of a given well n;
- $\mathcal{H} = \{\text{oil}, \text{gas}, \text{water}, \text{liq}\}$ has the liquid-flow (liq) and also the mixed-flow components: oil, gas, and water; liquid-flow is the sum of the water and oil flow rates;
- $\mathcal{K}^{n,s}$ has the breakpoints of the production function of well n to separator s; Breakpoints are of type (q_i), where q_i is a lift gas injection rate;
- $G^{n,s}$ is the set of parameters and correlations that characterize the well-production curves of the well n when directed to separator s. Since this is a satellite well, it includes the separator inlet pressure p_{sep}^s, the pipeline ID, vertical elevations, and horizontal length, the production tubing ID, length, and trajectory, the fluid data, as gas-liquid ratio (GLR), WC, specific densities, and bubble point pressure, the inflow performance relationship (IPR) and its parameters, the vertical and horizontal flow pattern correlations, to name a few.
- $\widehat{q}_h(q_i, G^{n,s})$ is the nonlinear production function of component h of well n to separator s evaluated at breakpoint $q_i \in \mathcal{K}$ to define a sampling point;
- $Q_{\text{inj}}^{\max}$ is the available lift-gas rate;
- $q_{\text{oil}}^{s,\max}$, $q_{\text{gas}}^{s,\max}$, $q_{\text{water}}^{s,\max}$ and $q_{\text{liq}}^{s,\max}$ are limits on oil, gas, water and liquid handling of separator s.

The production of each well as a function of the lift-gas injection rate and routing is modeled by Eqs. (4.1c)-(4.1f). Constraints (4.1g)-(4.1h) dictate the routing from wells to separators. They ensure that well production is not split in the flow lines: the production of well n sent to separator s is null if $z_{n,s} = 0$, otherwise it is bounded by the capacity of the separator.

Eqs. (4.1h)-(4.1j) induces a piecewise-linear function of the production curves through a SOS2 model. Fluid handling capacities for the separators are given by constraints (4.1k). Eqs. (4.1l)-(4.1m) describe the flow rates arriving at the separators.

4.1.2 Robust Formulation

Formulation (4.1) considers nominal values for all parameters. As discussed in the previous chapter a few of these parameters may not be precisely known for one of the many reasons explained in Section 3.1. In this model, for simplicity, let us consider that all parameters have accurate values but some parameters of the set $G^{n,s}$, which define the production curves. We denote this set that have the uncertain parameters $\tilde{G}^{n,s}$. Thus, instead of nominal sampling point values, what appears are uncertain sampling points $\widehat{q}_h(q_i, \tilde{G}^{n,s})$ induced by the uncertain values of the parameters in $\tilde{G}^{n,s}$. Here we assume as given the uncertainty range models $\mathcal{U}_{G^{n,s}}$ that represent the uncertainty in the parameters of $\tilde{G}^{n,s}$ as given. By having these models and applying the procedure described in Section 3.2.4 – Monte Carlo method, it is possible to design the uncertainty range models $\mathcal{U}_{\widehat{q}_h(q_i,G^{n,s})}$ for the sampling points $\widehat{q}_h(q_i, \tilde{G}^{n,s})$, which are the uncertain parameters that would in fact appear in an uncertain formulation of $\bar{P}$. The mathematical representation of these models is $\mathcal{U}_{\widehat{q}_h(q_i,G^{n,s})} = [\widehat{q}^{\,n,s}_{h,\text{inf}}(q_i), \widehat{q}^{\,n,s}_{h,\text{sup}}(q_i)]$.

In line with the discussions of Section 3.3, we consider a column-wise relationship between the individual uncertain parameters $\widehat{q}_h(q_i, \tilde{G}^{n,s})$. Therefore, the robust production optimization problem can be viewed as:

$$P_{\text{RC}} : \max \sum_{n\in\mathcal{N}} \bar{q}^{n}_{\text{oil,inf}} \tag{4.2a}$$

subject to a limited compression capacity:

$$\text{s.t.} : \sum_{n\in\mathcal{N}} q^{n}_{\text{inj}} \leq Q^{\max}_{\text{inj}}, \tag{4.2b}$$

and for each well $n \in \mathcal{N}$:

$$q_{\text{inj}}^{n,s} = \sum_{q_i \in \mathcal{K}^{n,s}} \lambda_{q_i}^{n,s} \cdot q_i, \; \forall s \in \mathcal{S}_n \tag{4.2c}$$

$$\sum_{s \in \mathcal{S}_n} q_{\text{inj}}^{n,s} = q_{\text{inj}}^{n}, \tag{4.2d}$$

$$\bar{q}_{h,\text{inf}}^{n,s} = \sum_{q_i \in \mathcal{K}^{n,s}} \lambda_{q_i}^{n,s} \cdot \widehat{q}_{h,\text{inf}}^{n,s}(q_i), \; \forall s \in \mathcal{S}_n, \; \forall h \in \mathcal{H}, \tag{4.2e}$$

$$\bar{q}_{h,\text{sup}}^{n,s} = \sum_{q_i \in \mathcal{K}^{n,s}} \lambda_{q_i}^{n,s} \cdot \widehat{q}_{h,\text{sup}}^{n,s}(q_i), \; \forall s \in \mathcal{S}_n, \; \forall h \in \mathcal{H}, \tag{4.2f}$$

$$\sum_{s \in \mathcal{S}_n} \bar{q}_{h,\text{inf}}^{n,s} = \bar{q}_{h,\text{inf}}^{n}, \; \forall h \in \mathcal{H}, \tag{4.2g}$$

$$\sum_{s \in \mathcal{S}_n} z_{n,s} = y_n, \tag{4.2h}$$

with a SOS2 piecewise-linear model, for all well $n \in \mathcal{N}$ and separator $s \in \mathcal{S}_n$:

$$\sum_{q_i \in \mathcal{K}^{n,s}} \lambda_{q_i}^{n,s} = z_{n,s} \tag{4.2i}$$

$$\lambda_{q_i}^{n,s} \geq 0, \; \forall q_i \in \mathcal{K}^{n,s} \tag{4.2j}$$

$$\left(\lambda_{q_i}^{n,s}\right)_{q_i \in \mathcal{K}^{n,s}} \text{ are SOS2} \tag{4.2k}$$

and for each separator $s \in \mathcal{S}$:

$$\bar{q}_{h,\text{sup}}^{s} \leq q_h^{s,\max}, \; \forall h \in \mathcal{H}, \tag{4.2l}$$

$$\sum_{n \in \mathcal{N}_s} \bar{q}_{h,\text{sup}}^{n,s} = \bar{q}_{h,\text{sup}}^{s}, \; \forall h \in \mathcal{H} - \{\text{gas}\} \tag{4.2m}$$

$$\sum_{n \in \mathcal{N}_s} (\bar{q}_{\text{gas,sup}}^{n,s} + q_{\text{inj}}^{n,s}) = \bar{q}_{\text{gas,sup}}^{s} \tag{4.2n}$$

with:

$$y_n \in \{0,1\}, \; \forall n \in \mathcal{N}, \tag{4.2o}$$

$$z_{n,s} \in \{0,1\}, \; \forall n \in \mathcal{N}, \; \forall s \in \mathcal{S}_n. \tag{4.2p}$$

where P_{RC} is a robust version of problem $\bar{P}$. The additional **variables** for the robust problem are:

- $\bar{q}^{n,s}_{h,\inf}$ ($\bar{q}^{n,s}_{h,\sup}$) is the infimum (supremum) of the production component h from well n directed to separator s.

- $\bar{q}^{n}_{h,\inf}$ is the infimum production rate of component h of well n;

- $\bar{q}^{s}_{h,\sup}$ is the supremum production rate of component h flowing to separator s;

also the following **parameters** are introduced:

- $\widehat{q}^{n,s}_{h,\inf}(q_i)$ and $\widehat{q}^{n,s}_{h,\sup}(q_i)$ are respectively the lower and upper limits of the range model which defines the uncertainty set $\mathcal{U}_{\widehat{q}_h(q_i,G^{n,s})}$. This relates to the production of component h associated with the breakpoint q_i of well n when connected to separator s;

Note that the set of Eqs. (4.2) is similar to the set of Eqs. (4.1). Having introduced the notation above, the robust formulation of the lift-gas allocation problem is an MILP problem defined on the variables $q^{n,s}_{\text{inj}}$, q^{n}_{inj}, $\bar{q}^{n,s}_{h,\inf}$, $\bar{q}^{n,s}_{h,\sup}$, $\bar{q}^{n}_{h,\inf}$, $\bar{q}^{s}_{h,\sup}$, y_n, $z_{n,s}$ and $\lambda^{n,s}_{q_i}$. The solution it produces is robust optimal, in the sense that is the best solution within the set that satisfies any realization of the uncertain parameters that are within the uncertainty set.

4.2 ROBUST PRODUCTION OPTIMIZATION – SUBSEA MANIFOLD

This section describes a robust production optimization model of a more complex production system, which consists of a group of gas-lifted wells whose production travel through subsea manifolds and shared flow lines. Routing decisions also apply, but due to the subsea completion, pressure constraints exist.

4.2.1 Problem Formulation

Optimal production plans for multi-reservoir fields with subsea completion require models that consider pressure constraints besides the usual flow constraints, when operating conditions vary due to routing operations and equipment failure. Let us consider the daily production optimization problem with nominal parameters formulated in Silva and Camponogara [9] as a MINLP. It consists in distributing a limited rate of pressurized gas, with the selection of production routing under separation capacity, and flow and pressure constraints. Figure 4.2 has the production network of the problem dealt in this section. Silva and Camponogara propose several MILP reformulations for the nonlinear problem, each using a different PWL model. Below it is presented one approximation based on the SOS2 model. Here, for simplicity, the objective function is considered the maximization of oil production of a group of wells.

$$\bar{P}: \max \sum_{n\in\mathcal{N}} \bar{q}^{n}_{\text{oil}} \tag{4.3a}$$

subject to a limited compression capacity:

$$\text{s.t.}: \sum_{n\in\mathcal{N}} q^{n}_{\text{inj}} \leq Q^{\max}_{\text{inj}}, \tag{4.3b}$$

and for each well $n \in \mathcal{N}$:

$$q^{n,m}_{\text{inj}} = \sum_{(q_i,p)\in\mathcal{V}^{n,m}} \lambda^{n,m}_{q_i,p} \cdot q_i, \ \forall m \in \mathcal{M}_n, \tag{4.3c}$$

$$\sum_{m\in\mathcal{M}_n} q^{n,m}_{\text{inj}} = q^{n}_{\text{inj}}, \tag{4.3d}$$

$$\begin{aligned} \sum_{(q_i,p)\in\mathcal{V}^{n,m}} \lambda^{n,m}_{q_i,p} \cdot p \leq p^m \leq \sum_{(q_i,p)\in\mathcal{V}^{n,m}} \lambda^{n,m}_{q_i,p} \cdot p + \ldots \\ \ldots p^{m,\max}(1 - z_{n,m}), \ \forall m \in \mathcal{M}_n, \end{aligned} \tag{4.3e}$$

$$\bar{q}^{n,m}_{h} = \sum_{(q_i,p)\in\mathcal{V}^{n,m}} \lambda^{n,m}_{q_i,p} \cdot \widehat{q}_h(q_i, p, G^{n,m}), \ \forall m \in \mathcal{M}_n, \ \forall h \in \mathcal{H}, \tag{4.3f}$$

$$\sum_{m\in\mathcal{M}_n} \bar{q}^{n,m}_{h} = \bar{q}^{n}_{h}, \ \forall h \in \mathcal{H}, \tag{4.3g}$$

$$\sum_{m\in\mathcal{M}_n} z_{n,m} = y_n, \tag{4.3h}$$

$$\left(\lambda^{n,m}_{q_i,p}\right)_{\mathcal{V}^{n,m}} \text{ induces a PWL function}, \ \forall m \in \mathcal{M}_n, \tag{4.3i}$$

and for each manifold $m \in \mathcal{M}$:

$$\bar{q}^{m}_{h} \leq q^{m,\max}_{h}, \ \forall h \in \mathcal{H}, \tag{4.3j}$$

$$\bar{q}^{m}_{h} = \sum_{n\in\mathcal{N}_m} \bar{q}^{n,m}_{h}, \ \forall h \in \mathcal{H} - \{\text{gas}\}, \tag{4.3k}$$

$$\bar{q}^{m}_{\text{gas}} = \sum_{n\in\mathcal{N}_m} (\bar{q}^{n,m}_{\text{gas}} + q^{n,m}_{\text{inj}}), \tag{4.3l}$$

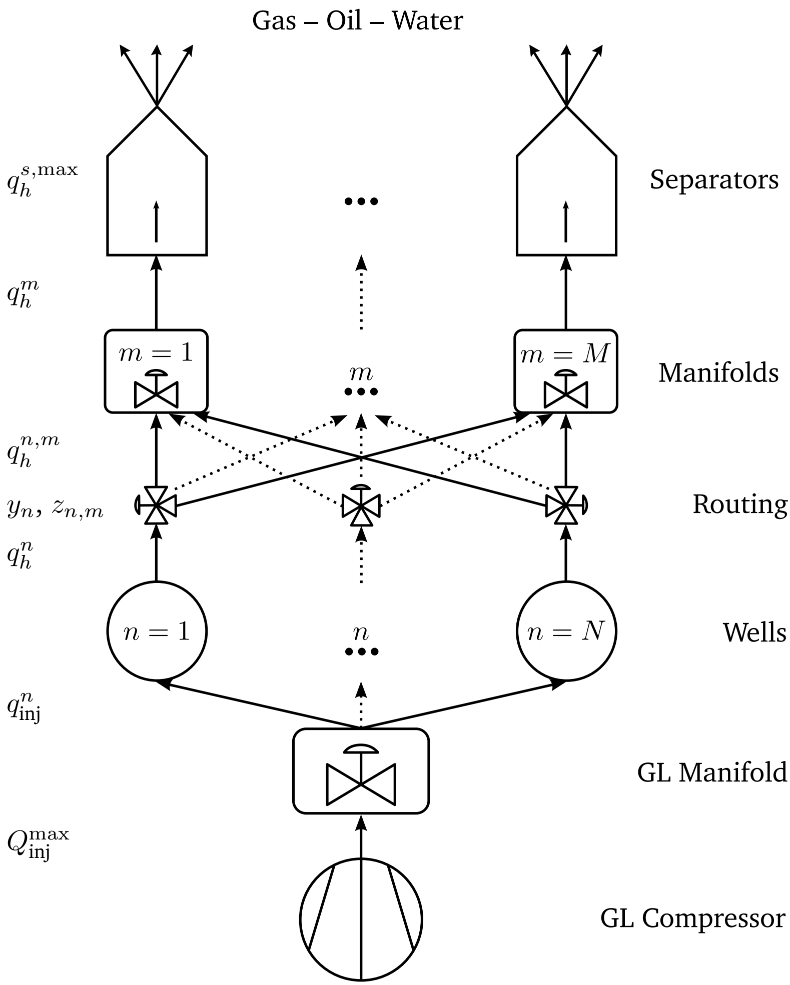


Figure 4.2: Production network – subsea completion.

$$\bar{q}^m_{\text{liq}} = \sum_{(q_{\text{liq}},glr,wc)\in\mathcal{T}^m} \gamma^m_{q_{\text{liq}},glr,wc} \cdot q_{\text{liq}}, \tag{4.3m}$$

$$\bar{q}^m_{\text{water}} = \sum_{(q_{\text{liq}},glr,wc)\in\mathcal{T}^m} \gamma^m_{q_{\text{liq}},glr,wc} \cdot wc \cdot q_{\text{liq}}, \tag{4.3n}$$

$$\bar{q}^m_{\text{gas}} = \sum_{(q_{\text{liq}},glr,wc)\in\mathcal{T}^m} \gamma^m_{q_{\text{liq}},glr,wc} \cdot glr \cdot q_{\text{liq}}, \tag{4.3o}$$

$$\bar{\Delta p}^m = \sum_{(q_{\text{liq}},glr,wc)\in\mathcal{T}^m} \gamma^m_{q_{\text{liq}},glr,wc} \cdot \widehat{\Delta p}(q_{\text{liq}}, glr, wc, I^m), \tag{4.3p}$$

$$p^m = p^m_{\text{sep}} + \bar{\Delta p}^m, \tag{4.3q}$$

$$p^{m,\min} \le p^m \le p^{m,\max}, \tag{4.3r}$$

$$z_{n,m} \le y_m,\ n \in \mathcal{N}_m, \tag{4.3s}$$

$$\sum_{n\in\mathcal{N}_m} z_{n,m} \ge y_m, \tag{4.3t}$$

$$\left(\gamma^m_{q_{\text{liq}},glr,wc}\right)_{\mathcal{T}^m} \text{ induces a PWL function}, \tag{4.3u}$$

with:

$$y_n \in \{0,1\},\ \forall n \in \mathcal{N}, \tag{4.3v}$$

$$y_m \in \{0,1\},\ \forall m \in \mathcal{M}, \tag{4.3w}$$

$$z_{n,m} \in \{0,1\},\ \forall n \in \mathcal{N},\ \forall m \in \mathcal{M}_n. \tag{4.3x}$$

For Eq. (4.3i) the SOS2 model the following equations apply for each well $n \in \mathcal{N}$ and manifold $m \in \mathcal{M}_n$:

$$\sum_{(q_i,p)\in\mathcal{V}^{n,m}} \lambda^{n,m}_{q_i,p} = z_{n,m}, \tag{4.4a}$$

$$\eta^{n,m}_p = \sum_{q_i\in\mathcal{V}^{n,m}_{\text{q}_\text{i}}} \lambda^{n,m}_{q_i,p}, \forall p \in \mathcal{V}^{n,m}_{\text{p}} \tag{4.4b}$$

$$\eta^{n,m}_{q_i} = \sum_{p\in\mathcal{V}^{n,m}_p} \lambda^{n,m}_{q_i,p}, \forall q_i \in \mathcal{V}^{n,m}_{\text{q}_\text{i}} \tag{4.4c}$$

$$\lambda^{n,m}_{q_i,p}, \in [0,1],\ \forall (q_i,p) \in \mathcal{V}^{n,m}, \tag{4.4d}$$

$$\left(\eta^{n,m}_p\right)_{p\in\mathcal{V}^{n,m}_{\text{p}}}, \left(\eta^{n,m}_{q_i}\right)_{q_i\in\mathcal{V}^{n,m}_{\text{q}_\text{i}}}, \text{ are SOS2} \tag{4.4e}$$

and for Eq. (4.3u) the SOS2 model for each manifold $m \in \mathcal{M}$:

$$\sum_{(q_{\text{liq}},glr,wc)\in\mathcal{T}^m} \gamma^m_{q_{\text{liq}},glr,wc} = y_m, \tag{4.4f}$$

$$\eta^m_{q_{\text{liq}}} = \sum_{glr\in\mathcal{T}^m_{\text{glr}}} \sum_{wc\in\mathcal{T}^m_{\text{wc}}} \gamma^m_{q_{\text{liq}},glr,wc}, \forall q_{\text{liq}} \in \mathcal{T}^m_{\mathtt{qliq}}, \tag{4.4g}$$

$$\eta^m_{glr} = \sum_{q_{\text{liq}}\in\mathcal{T}^m_{\mathtt{qliq}}} \sum_{wc\in\mathcal{T}^m_{\text{wc}}} \gamma^m_{q_{\text{liq}},glr,wc}, \forall glr \in \mathcal{T}^m_{\text{glr}}, \tag{4.4h}$$

$$\eta^m_{wc} = \sum_{q_{\text{liq}}\in\mathcal{T}^m_{\mathtt{qliq}}} \sum_{glr\in\mathcal{T}^m_{\text{glr}}} \gamma^m_{q_{\text{liq}},glr,wc}, \forall wc \in \mathcal{T}^m_{\text{wc}}, \tag{4.4i}$$

$$\gamma^m_{q_{\text{liq}},glr,wc}, \in [0,1], \forall (q_{\text{liq}}, glr, wc) \in \mathcal{T}^m, \tag{4.4j}$$

$$\left(\eta^m_{q_{\text{liq}}}\right)_{q_{\text{liq}}\in\mathcal{T}^m_{\mathtt{qliq}}}, \left(\eta^m_{glr}\right)_{glr\in\mathcal{T}^m_{\text{glr}}}, \left(\eta^m_{wc}\right)_{wc\in\mathcal{T}^m_{\text{wc}}}, \quad \text{are SOS2} \tag{4.4k}$$

The **decision variables** of the Formulation (4.3)-(4.4) are:

- q^n_{inj} is the rate of lift-gas injection into well n and $q^{n,m}_{\text{inj}}$ is the lift-gas injection rate of well n when routed to manifold m;

- $\bar{q}^n_h$ is the rate of component h produced according with the piecewise-linear function of well n;

- $\bar{q}^{n,m}_h$ is the component h of the approximated production rate from well n when directed to manifold m;

- $\bar{q}^m_h$ is the approximated total flow of component h flowing through manifold m;

- p^m is the pressure in manifold m;

- $y_n \in \{0,1\}$ is 1 when well n is active or else 0;

- $y_m \in \{0,1\}$ is 1 when manifold m is active or else 0;

- $z_{n,m} \in \{0,1\}$ assumes value 1 if well n is connected to manifold m, and 0 otherwise;

- $\bar{\Delta p}^m$ is the pressure drop in the flowline connecting manifold m to its separator according with the piecewise-linear function;

- $\lambda_{q_i,p}^{n,m} \in [0,1]$ is a weighting variable associated with the $(q_i, p)^{th}$ breakpoint of the piecewise-linear production function of well n to manifold m;
- $\gamma_{q_{\text{liq}},glr,wc}^{m} \in [0,1]$ is a weighting variable associated with the $(q_{\text{liq}}, glr, wc)^{th}$ breakpoint of the piecewise-linear pressure drop function in the pipeline connecting manifold m to its separator;
- $\eta_{q_i}^{n,m}$, $\eta_{p}^{n,m}$, η_{glr}^{m}, η_{wc}^{m}, $\eta_{q_{\text{liq}}}^{m}$ are auxiliary SOS2 variables.

the **parameters** are as follows:

- $\mathcal{N} = \{1, \ldots, N\}$ is the set of gas-lifted wells and $\mathcal{N}_m \subseteq \mathcal{N}$ is the subset of wells that can be connected to manifold m;
- $\mathcal{M} = \{1, \ldots, M\}$ is the set of manifolds and $\mathcal{M}_n \subseteq \mathcal{M}$ is the subset of manifolds that can received production of a given well n;
- $\mathcal{H} = \{\text{oil}, \text{gas}, \text{water}, \text{liq}\}$ has the liquid-flow (liq) and also the mixed-flow components: oil, gas, and water; liquid-flow is the sum of the water and oil flow rates;
- $Q_{\text{inj}}^{\max}$ is the available lift-gas rate;
- $q_{\text{oil}}^{m,\max}$, $q_{\text{gas}}^{m,\max}$, $q_{\text{water}}^{m,\max}$ and $q_{\text{liq}}^{m,\max}$ are limits on oil, gas, water and liquid handling of the separator connected to manifold m;
- $p^{m,\min}(p^{m,\max})$ is the minimum (maximum) operational pressure for manifold m;
- p_{sep}^{m} is the operational pressure of the separator connected to manifold m.
- $G^{n,m}$ is the set of parameters and correlations that characterize the well-production curves of the well n when directed to manifold m.
- I^{m} is the set of parameters and correlations that define the pressure drop in the pipeline connecting manifold m to its respective separator.
- $\widehat{q}_h(q_i, p, G^{n,m})$ is the nonlinear production function of component h of well n to manifold m evaluated at the breakpoint $(q_i, p) \in \mathcal{V}^{n,m}$ to define a sampling point;

- $\widehat{\Delta p}(\cdot)$ is the nonlinear pressure-drop curve between manifold and its separator as a function of the liquid rate, the gas-liquid ratio and the water-cut of the fluid passing through the flowing line. $\widehat{\Delta p}(q_{\mathrm{liq}}, glr, wc, I^m)$ is the evaluation of this function at the breakpoint $(q_{\mathrm{liq}}, glr, wc) \in \mathcal{T}^m$ to define a sampling point, for the flow line leaving manifold m;

- $\mathcal{V}^{n,m}$ has the breakpoints of the production function for well n when connected to manifold m; Breakpoints are of type (q_i, p), where q_i is a lift-gas injection rate and p is a pressure in the manifold;

- $\mathcal{T}^m$ has the breakpoints of the pressure-drop function for manifold m; Breakpoints are of type $(q_{\mathrm{liq}}, glr, wc)$, where q_{liq} is the liquid rate, glr is the gas-liquid ratio and wc is the water-cut of the fluid in manifold m;

In order to apply the SOS2 model we have to consider that $\mathcal{V}^{n,m}$ and $\mathcal{T}^m$ are full grids, thus $\mathcal{V}^{n,m} := \mathcal{V}^{n,m}_{\mathrm{qi}} \times \mathcal{V}^{n,m}_{\mathrm{p}}$ and $\mathcal{T}^m := \mathcal{T}^m_{\mathrm{q_{liq}}} \times \mathcal{T}^m_{\mathrm{glr}} \times \mathcal{T}^m_{\mathrm{wc}}$. Where $\mathcal{V}^{n,m}_{\mathrm{qi}}$ and $\mathcal{V}^{n,m}_{\mathrm{p}}$ are respectively the set of breakpoints for lift-gas injection rates and upstream pressure for each well n to manifold m, and $\mathcal{T}^m_{\mathrm{q_{liq}}}$, $\mathcal{T}^m_{\mathrm{glr}}$ and $\mathcal{T}^m_{\mathrm{wc}}$ are the set of breakpoints for liquid rate, gas-liquid ratio and water-cut for each manifold m.

The production of each well as a function of the lift-gas injection rate, manifold pressure and routing is modeled by Eqs. (4.3c)-(4.3g). In this formulation constraint (4.3h) dictates the routing from wells to manifolds/separators and ensure that there is no flow splitting. Separation capacities are given by constraint (4.3j). Eqs. (4.3k)-(4.3l) have the mass balance, while the pressure balance between manifolds and separators is given by constraints (4.3q). The pressure drop as a function of the resulting gas-liquid ratio, water-cut and total liquid rate in each manifold is modeled by Eqs. (4.3m)-(4.3p). Eq. (4.3r) has the pressure limits for the manifolds and Eqs. (4.3s)-(4.3t) ensure manifold that if a manifold is not active no well can route its production to it. Finally Eqs. (4.4) induce a piecewise-linear function of the production curves and the pressure drop curves.

4.2.2 Robust Formulation

Formulation (4.3)-(4.4) disregards any uncertainty in the values of the parameters and considers expected values for all of them. As in Section 4.1, let us assume that uncertainty is only in some of the parameters in the set $G^{n,m}$ that define the production curves. We denote this set of uncertain parameters $\tilde{G}^{n,m}$. All other parameters have known and accurate values.

Again, instead of nominal sampling point values, what comes out of the formulation described above are the uncertain sampling points $\widehat{q}_h(q_i, p, \tilde{G}^{n,m})$ induced by the uncertain values of the parameters in $\tilde{G}^{n,m}$. Considering the uncertainty range models $\mathcal{U}_{G^{n,s}}$ that represent the uncertainty in the parameters of $\tilde{G}^{n,s}$ are given, it is possible to obtain approximated uncertainty sets $\mathcal{U}_{\widehat{q}_h(q_i,p,G^{n,m})}$ for the sampling points $\widehat{q}_h(q_i, p, \tilde{G}^{n,m})$. This can be performed using, for example, Monte Carlo methods (see Section 3.2.4). The mathematical representation of these uncertainty sets can be $\mathcal{U}_{\widehat{q}_h(q_i,p,G^{n,m})} = [\widehat{q}^{\,n,m}_{h,\inf}(q_i, p), \widehat{q}^{\,n,m}_{h,\sup}(q_i, p)]$.

In line with the discussions of Section 3.3, we also consider a column-wise relationship between the individual uncertain parameters $\widehat{q}_h(q_i, \tilde{G}^{n,s})$. However, differently than in the example of Section 3.3 and the problem of Section 4.1, for the problem above, even after eliminating the state variables one is still left with uncertain equality constraints. Those are the mass balance equations (4.3m)-(4.3o).

As explained in the final remarks of Section 2.3.3, standard robust optimization theory does not consider problems with uncertainty in equality constraints. In order to handle uncertain equality constraints using the same framework detailed in the previous chapter, we have to modify formulation $\bar{P}$. That can be achieved relaxing the uncertain equality constraints by replacing them with two appropriate uncertain inequalities. The uncertain counterparts of Eqs. (4.3m)-(4.3o), eliminating the state variables $\bar{q}^m_{\text{liq}}$, $\bar{q}^m_{\text{water}}$, and $\bar{q}^m_{\text{gas}}$, would then become for all manifolds $m \in \mathcal{M}$:

$$\text{Eq. (4.3m)} \begin{cases} \tilde{q}^m_{\text{liq}} - \displaystyle\sum_{(q_{\text{liq}},glr,wc)\in\mathcal{T}^m} \gamma^m_{q_{\text{liq}},glr,wc} \cdot q_{\text{liq}} \geq 0, \\ -\,\tilde{q}^m_{\text{liq}} + \displaystyle\sum_{(q_{\text{liq}},glr,wc)\in\mathcal{T}^m} \gamma^m_{q_{\text{liq}},glr,wc} \cdot q_{\text{liq}} \geq 0 \end{cases} \tag{4.5a}$$

$$\text{Eq. (4.3n)} \begin{cases} \tilde{q}^m_{\text{water}} - \displaystyle\sum_{(q_{\text{liq}},glr,wc)\in\mathcal{T}^m} \gamma^m_{q_{\text{liq}},glr,wc} \cdot wc \cdot q_{\text{liq}} \geq 0, \\ -\,\tilde{q}^m_{\text{water}} + \displaystyle\sum_{(q_{\text{liq}},glr,wc)\in\mathcal{T}^m} \gamma^m_{q_{\text{liq}},glr,wc} \cdot wc \cdot q_{\text{liq}} \geq 0 \end{cases} \tag{4.5b}$$

$$\text{Eq. (4.3o)} \begin{cases} \tilde{q}^m_{\text{gas}} - \sum_{(q_{\text{liq}},glr,wc)\in\mathcal{T}^m} \gamma^m_{q_{\text{liq}},glr,wc} \cdot glr \cdot q_{\text{liq}} \geq 0, \\ -\tilde{q}^m_{\text{gas}} + \sum_{(q_{\text{liq}},glr,wc)\in\mathcal{T}^m} \gamma^m_{q_{\text{liq}},glr,wc} \cdot glr \cdot q_{\text{liq}} \geq 0 \end{cases} \tag{4.5c}$$

where $\tilde{q}^m_{\text{liq}}$, $\tilde{q}^m_{\text{water}}$, and $\tilde{q}^m_{\text{gas}}$ are defined as,

$$\tilde{q}^m_{\text{liq}} = \sum_{n\in\mathcal{N}_m} \sum_{(q_i,p)\in\mathcal{V}^{n,m}} \lambda^{n,m}_{q_i,p} \cdot \widehat{q}_{\text{liq}}(q_i, p, \tilde{G}^{n,m}), \ldots$$
$$\ldots \forall \widehat{q}_{\text{liq}}(q_i, p, \tilde{G}^{n,m}) \in \mathcal{U}_{\widehat{q}_{\text{liq}}(q_i,p,G^{n,m})}$$
$$\tilde{q}^m_{\text{water}} = \sum_{n\in\mathcal{N}_m} \sum_{(q_i,p)\in\mathcal{V}^{n,m}} \lambda^{n,m}_{q_i,p} \cdot \widehat{q}_{\text{water}}(q_i, p, \tilde{G}^{n,m}),$$
$$\forall \widehat{q}_{\text{water}}(q_i, p, \tilde{G}^{n,m}) \in \mathcal{U}_{\widehat{q}_{\text{water}}(q_i,p,G^{n,m})}$$
$$\tilde{q}^m_{\text{gas}} = \sum_{n\in\mathcal{N}_m} \sum_{(q_i,p)\in\mathcal{V}^{n,m}} \lambda^{n,m}_{q_i,p} \cdot \widehat{q}_{\text{gas}}(q_i, p, \tilde{G}^{n,m}),$$
$$\forall \widehat{q}_{\text{gas}}(q_i, p, \tilde{G}^{n,m}) \in \mathcal{U}_{\widehat{q}_{\text{gas}}(q_i,p,G^{n,m})}$$

With this substitution it is possible to use the same steps used in Section 3.3 to find the robust counterpart problem. Therefore, the robust production optimization problem can be viewed as:

$$P_{\text{RC}} : \max \sum_{n\in\mathcal{N}} \bar{q}^n_{\text{oil,inf}} \tag{4.6a}$$

subject to a limited compression capacity:

$$\text{s.t.} : \sum_{n\in\mathcal{N}} q^n_{\text{inj}} \leq Q^{\max}_{\text{inj}}, \tag{4.6b}$$

and for each well $n \in \mathcal{N}$:

$$q^{n,m}_{\text{inj}} = \sum_{(q_i,p)\in\mathcal{V}^{n,m}} \lambda^{n,m}_{q_i,p} \cdot q_i, \ \forall m \in \mathcal{M}_n, \tag{4.6c}$$

$$\sum_{m\in\mathcal{M}_n} q^{n,m}_{\text{inj}} = q^n_{\text{inj}}, \tag{4.6d}$$

$$\sum_{(q_i,p)\in\mathcal{V}^{n,m}} \lambda^{n,m}_{q_i,p} \cdot p \leq p^m \leq \sum_{(q_i,p)\in\mathcal{V}^{n,m}} \lambda^{n,m}_{q_i,p} \cdot p + \ldots$$
$$\ldots p^{m,\max}(1 - z_{n,m}), \ \forall m \in \mathcal{M}_n, \tag{4.6e}$$

$$\bar{q}^{n,m}_{h,\text{inf}} = \sum_{(q_i,p)\in\mathcal{V}^{n,m}} \lambda^{n,m}_{q_i,p} \cdot \widehat{q}^{\,n,m}_{h,\text{inf}}(q_i, p), \ \forall m \in \mathcal{M}_n, \ \forall h \in \mathcal{H}, \tag{4.6f}$$

$$\bar{q}_{h,\sup}^{n,m} = \sum_{(q_i,p)\in\mathcal{V}^{n,m}} \lambda_{q_i,p}^{n,m} \cdot \widehat{q}_{h,\sup}^{n,m}(q_i,p),\ \forall m \in \mathcal{M}_n,\ \forall h \in \mathcal{H}, \tag{4.6g}$$

$$\sum_{m\in\mathcal{M}_n} \bar{q}_{h,\inf}^{n,m} = \bar{q}_{h,\inf}^{n},\ \forall h \in \mathcal{H}, \tag{4.6h}$$

$$\sum_{m\in\mathcal{M}_n} \bar{q}_{h,\sup}^{n,m} = \bar{q}_{h,\sup}^{n},\ \forall h \in \mathcal{H}, \tag{4.6i}$$

$$\sum_{m\in\mathcal{M}_n} z_{n,m} = y_n, \tag{4.6j}$$

$$\left(\lambda_{q_i,p}^{n,m}\right)_{\mathcal{V}^n} \text{induces a PWL function},\ \forall m \in \mathcal{M}_n, \tag{4.6k}$$

and for each manifold $m \in \mathcal{M}$:

$$\bar{q}_{h,\sup}^{m} \leq q_h^{m,\max},\ \forall h \in \mathcal{H}, \tag{4.6l}$$

$$\bar{q}_{h,\sup}^{m} = \sum_{n\in\mathcal{N}_m} \bar{q}_{h,\sup}^{n,m},\ \forall h \in \mathcal{H} - \{\text{gas}\}, \tag{4.6m}$$

$$\bar{q}_{\text{gas},\sup}^{m} = \sum_{n\in\mathcal{N}_m} (\bar{q}_{\text{gas},\sup}^{n,m} + q_{\text{inj}}^{n,m}), \tag{4.6n}$$

$$\bar{q}_{h,\inf}^{m} = \sum_{n\in\mathcal{N}_m} \bar{q}_{h,\inf}^{n,m},\ \forall h \in \mathcal{H} - \{\text{gas}\}, \tag{4.6o}$$

$$\bar{q}_{\text{gas},\inf}^{m} = \sum_{n\in\mathcal{N}_m} (\bar{q}_{\text{gas},\inf}^{n,m} + q_{\text{inj}}^{n,m}), \tag{4.6p}$$

$$\bar{\Delta p}^m = \sum_{(q_{\text{liq}},glr,wc)\in\mathcal{T}^m} \gamma_{q_{\text{liq}},glr,wc}^{m} \cdot \widehat{\Delta p}(q_{\text{liq}},glr,wc,I^m), \tag{4.6q}$$

$$p^m = p_{\text{sep}}^m + \bar{\Delta p}^m, \tag{4.6r}$$

$$p^{m,\min} \leq p^m \leq p^{m,\max}, \tag{4.6s}$$

$$z_{n,m} \leq y_m,\ n \in \mathcal{N}_m, \tag{4.6t}$$

$$\sum_{n\in\mathcal{N}_m} z_{n,m} \geq y_m, \tag{4.6u}$$

$$\left(\gamma_{q_{\text{liq}},glr,wc}^{m}\right)_{\mathcal{T}^m} \text{ induces a PWL function}, \tag{4.6v}$$

with:

$$y_n \in \{0,1\},\ \forall n \in \mathcal{N}, \tag{4.6w}$$

$$y_m \in \{0,1\},\ \forall m \in \mathcal{M}, \tag{4.6x}$$

$$z_{n,m} \in \{0,1\},\ \forall n \in \mathcal{N},\ \forall m \in \mathcal{M}_n. \tag{4.6y}$$

and the equality relaxation approach – Eqs. Eq. 4.5a-Eq. 4.5c – result in the following constraints, for all $m \in \mathcal{M}$:

$$\bar{q}^{m}_{\text{liq},\text{inf}} \leq \sum_{(q_{\text{liq}},glr,wc)\in\mathcal{T}^m} \gamma^{m}_{q_{\text{liq}},glr,wc} \cdot q_{\text{liq}}, \tag{4.7a}$$

$$\bar{q}^{m}_{\text{liq},\text{sup}} \geq \sum_{(q_{\text{liq}},glr,wc)\in\mathcal{T}^m} \gamma^{m}_{q_{\text{liq}},glr,wc} \cdot q_{\text{liq}}, \tag{4.7b}$$

$$\bar{q}^{m}_{\text{water},\text{inf}} \leq \sum_{(q_{\text{liq}},glr,wc)\in\mathcal{T}^m} \gamma^{m}_{q_{\text{liq}},glr,wc} \cdot wc \cdot q_{\text{liq}}, \tag{4.7c}$$

$$\bar{q}^{m}_{\text{water},\text{sup}} \geq \sum_{(q_{\text{liq}},glr,wc)\in\mathcal{T}^m} \gamma^{m}_{q_{\text{liq}},glr,wc} \cdot wc \cdot q_{\text{liq}}, \tag{4.7d}$$

$$\bar{q}^{m}_{\text{gas},\text{inf}} \leq \sum_{(q_{\text{liq}},glr,wc)\in\mathcal{T}^m} \gamma^{m}_{q_{\text{liq}},glr,wc} \cdot glr \cdot q_{\text{liq}}, \tag{4.7e}$$

$$\bar{q}^{m}_{\text{gas},\text{sup}} \geq \sum_{(q_{\text{liq}},glr,wc)\in\mathcal{T}^m} \gamma^{m}_{q_{\text{liq}},glr,wc} \cdot glr \cdot q_{\text{liq}}, \tag{4.7f}$$

and the SOS2 models for $\lambda^{n,m}_{q_i,p}$ and $\gamma^{m}_{q_{\text{liq}},glr,wc}$ remain the same as in Eqs. (4.4).

Formulation (4.4)-(4.6)-(4.7) (P_{RC}) is a robust version of Formulation (4.3)-(4.4) ($\bar{P}$). The additional **variables** for the robust problem are:

- $\bar{q}^{n,m}_{h,\text{inf}}$ ($\bar{q}^{n,m}_{h,\text{sup}}$) is the infimum (supremum) of the production component h from well n directed to manifold m.
- $\bar{q}^{n}_{h,\text{inf}}$ ($\bar{q}^{n}_{h,\text{sup}}$) is the infimum (supremum) production rate of component h of well n;
- $\bar{q}^{m}_{h,\text{inf}}$ ($\bar{q}^{m}_{h,\text{sup}}$) is the infimum (supremum) production rate of component h flowing to manifold m;

also the following **parameters** are introduced:

- $\widehat{q}^{\,n,m}_{h,\text{inf}}(q_i,p)$ and $\widehat{q}^{\,n,m}_{h,\text{sup}}(q_i,p)$ are respectively the lower and upper limits of the range model which defines the uncertainty set $\mathcal{U}_{\widehat{q}_h(q_i,G^{n,m})}$. This relates to the production of component h associated with the breakpoint (q_i,p) of well n when connected to manifold m.

Note that the set of Eqs. (4.4)-(4.6)-(4.7) is similar to set of Eqs. (4.3)-(4.4). P_{RC} is also an MILP, but solving it produces a solution that satisfies any realization of the uncertain parameters that are within the uncertainty set and that is robust optimal.

4.3 SUMMARY

This chapter presented two standard production optimization problems. For both a MILP formulation was described considering nominal values for all parameters. For each problem, one containing satellite wells and one with wells and subsea completion, a robust reformulation of the problem was implemented considering uncertainty in the well-production curves. The methodology used was similar to the one detailed in Chapter 3, but for the second problem the methodology is extended to account for uncertain equality constraints.

5 EXPERIMENTS

In order to analyze the influence of uncertain parameters on daily production optimization, a synthetic but representative oil field was instantiated in a commercial multiphase flow simulator. The experimental scenario consists of the production plant layout, wells, processing facilities, and other equipment. The uncertainty arises in some of the parameters that characterize the wells of the oil field. The remainder of the chapter presents the oil field, the parameters that are assumed uncertain, and experimental results contrasting the performance of nominal and robust optimization of gas-lifted oil field with satellite wells.

5.1 OIL FIELD SCENARIO

The oil field has a structural configuration similar to the instances considered by Codas and Camponogara [8], which consists of one lift-gas compressor, two separation facilities, and five satellite wells, which were modified so that three wells have routing options and the remaining two have fixed routings to the separation units. Figure 5.1 illustrates this production network. For the equipment, only gas compression capacity and liquid separation limits are imposed. Two scenarios are created by combining different capacities for compressor and separators. Moreover, we consider distinct separation pressures for the separation unit. Table 5.1 gives the parameters and other data defining the production scenarios. The oil field is instantiated in the multiphase flow simulator *Pipesim*, which is widely used by the Oil and Gas Industry. The well production curves are also obtained from the same simulator. Field data used to represent the synthetic oil field is extracted from the five wells described in [79]. Some minor changes are made to adapt the length of the flow lines to suite the installation of routing valves and the connections to two separators. The simulator default values are used for

Table 5.1: Operating scenarios.

	Compressor	Separator 1	Separator 2
	—	$p^1_{\text{sep}} = 250$ psia	$p^2_{\text{sep}} = 150$ psia
Scenario	$Q^{\max}_{\text{inj}}$	$q^{1,\max}_{\text{liq}}$	$q^{2,\max}_{\text{liq}}$
1	7 mmscf/d	3500 STB/d	3700 STB/d
2	5 mmscf/d	3000 STB/d	3000 STB/d

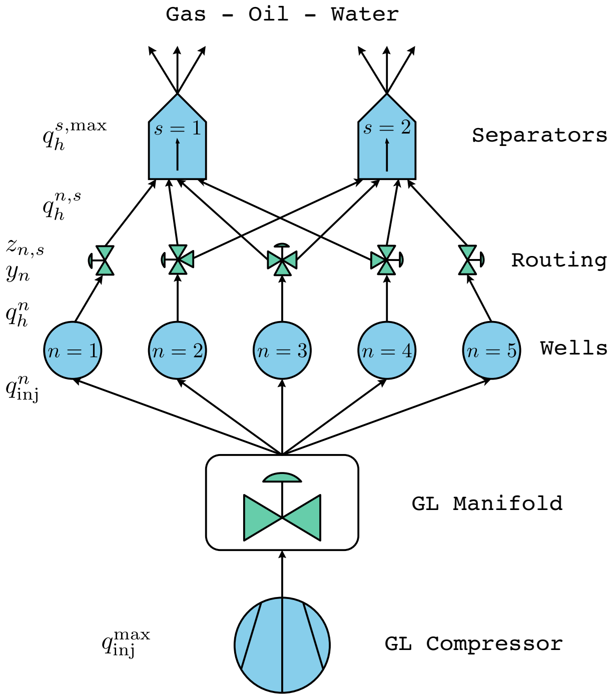


Figure 5.1: Production network.

the parameters and correlations that are not described in the paper. Table 5.2 presents part of the nominal values for the field data used in the simulated oil field.

Since the idea is to compare the results produced by standard production optimization using nominal data and robust production optimization, it is necessary to generate the piecewise-linear production curves for the standard problem. The well production curves are obtained by sampling the *Pipesim* simulator, using Algorithm 1 which is presented and detailed in Appendix A. The breakpoints for each production curve are defined allowing a maximum approximation error of 0.25% ($\varepsilon^{\%}$).

Table 5.2: Nominal field data.

Well # (n)	1	2	3	4	5
GLR [scf/STB]	**250**	**300**	**350**	**350**	**320**
WC [fraction]	**0.6**	**0.7**	**0.8**	**0.7**	**0.6**
SG Gas [fraction]*			0.8		
SG Oil [°API]*			35		
SG Water [fraction]*			1.07		
Reservoir pressure [psia]	1750	1850	1900	1850	1800
PI [STB/d/psi]	2	2.5	3	3.7	2.5
Reservoir temperature [°F]	180	190	190	190	190
Depth of perforations [ft]	4600	5000	5900	5900	5500
Tubing ID [inches]	2.875	3	2	2	2
Casing ID [inches]*			6.366		
Wellhead temperature [°F]*			110		
Flow lines ID [inches]*			4		
Flow line length[ft]	8000	5000	4000	5000	6000

* same values for all the wells

5.2 UNCERTAIN PARAMETERS

Although the robust formulation for production optimization presented in Section 4.1.2 allows selecting as uncertain any parameter of Table 5.2 – which characterize the production curves, the studies herein model uncertainties only in the gas-liquid ratio (GLR) and water cut (WC). This choice is not arbitrary because GLR and WC play a major part in characterizing a well behavior: these parameters tend to exhibit a higher short-term variability when compared, for example, with reservoir pressure and flow line ID. Another good candidate for uncertainty modeling would be the productivity indices, but for the present experiment their values are considered known and unique.

Table 5.3 has the percentage variation for GLR and WC for each well n from its nominal values presented in Table 5.2. This means that GLR and WC of each well does not have a unique and deterministic value. On the contrary, it is unknown but belongs to a given set centered at the nominal value, varying according with an unknown probability distribution. Notice that the percentage variations are merely illustrative, as they are not based on any study.

We define that for each well, the parameters composing this uncertainty set are not correlated, therefore the sets have the shape

Table 5.3: Percentage variation from nominal value.

Well # (n)	1	2	3	4	5
GLR	$\pm 20\%$	$\pm 25\%$	$\pm 20\%$	$\pm 6\%$	$\pm 15\%$
WC	$\pm 10\%$	$\pm 20\%$	$\pm 5\%$	$\pm 10\%$	$\pm 10\%$

of a box in $\mathbb{R}^2$. In order to generate the uncertainty set for the production value of each breakpoint, we developed a script in Python (programming language) that interacts with the simulator. It implements a Monte Carlo method to define upper ($\widehat{q}_{h,\mathrm{sup}}^{n,s}(q_i)$) and lower ($\widehat{q}_{h,\mathrm{inf}}^{n,s}(q_i)$) bound values for the breakpoints, which are used directly in the robust formulation (4.2).

5.3 COMPUTATIONAL RESULTS

All optimization problems are modeled in AMPL programming language and solved using IBM ILOG CPLEX 12.6 solver. The optimization models use data gathered in the steps described in Section 5.1 and 5.2. All those steps are depicted in the upper half of Figure 5.2, which is a diagram describing the entire experimental procedure. For each operating scenario (see Table 5.1) standard and robust production optimization models are solved. The calculated nominal-optimal and robust-optimal solutions – lift-gas injections and routing decisions – are stored.

The solutions are used as the operating settings for the synthetic oil field instantiated in the multiphase flow simulator. However, since our oil field has uncertain parameters, which have values modeled by ranges, we have to artificially create hypothetical instances of the oil field. This is achieved by randomly choosing values for the uncertain parameters within its range model. An instance, a random selection of all uncertain parameters, is here referred as a *case*. Figure 5.3 has the GLR and WC values selected for each well in each case. The hatched area corresponds to the bounds given in Table 5.3 for the respective uncertainty regions of each well. For each scenario of optimization, the total oil production obtained in every instance of the synthetic oil field using the robust setpoints is in the last row of Table 5.4 and 5.5. The values for the total oil production induced by nominal setpoints are in the middle rows of the same tables. In order to provide a benchmark for the oil productions mentioned above, we solved optimization problems for each oil field instance using production curves modeled with the exact

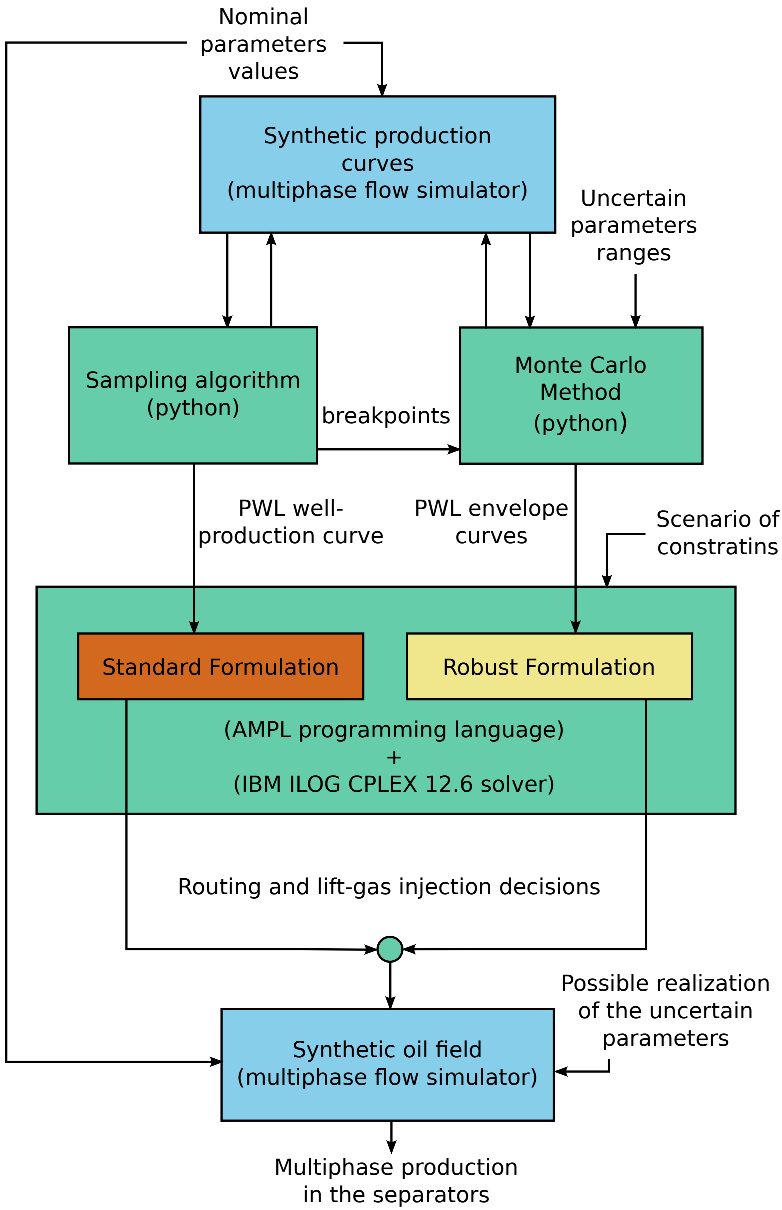


Figure 5.2: Experimental procedure diagram.

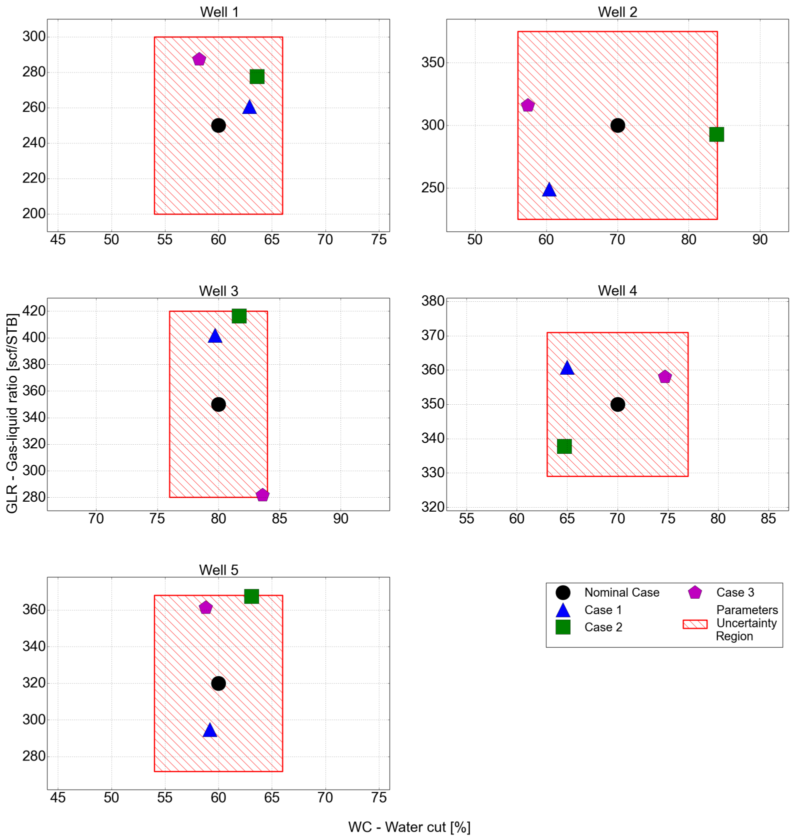


Figure 5.3: Instances of the uncertain data.

data for each respective case. The solutions of these optimization problems were than used in same particular oil field simulator instance. The total oil production achieved is referred as a theoretical, or let's say utopic, solution for the problem. Those values are also in Table 5.4 and 5.5.

Figure 5.4 condensates all the information on these tables. Nominal solutions are plotted next to the robust solution in each the corresponding scenario and case. The theoretical solution is also depicted in dashed bars. As expected, for each scenario, the theoretical total production differs depending on the case taken. For the first scenario, which is less restrictive, the nominal solution induces

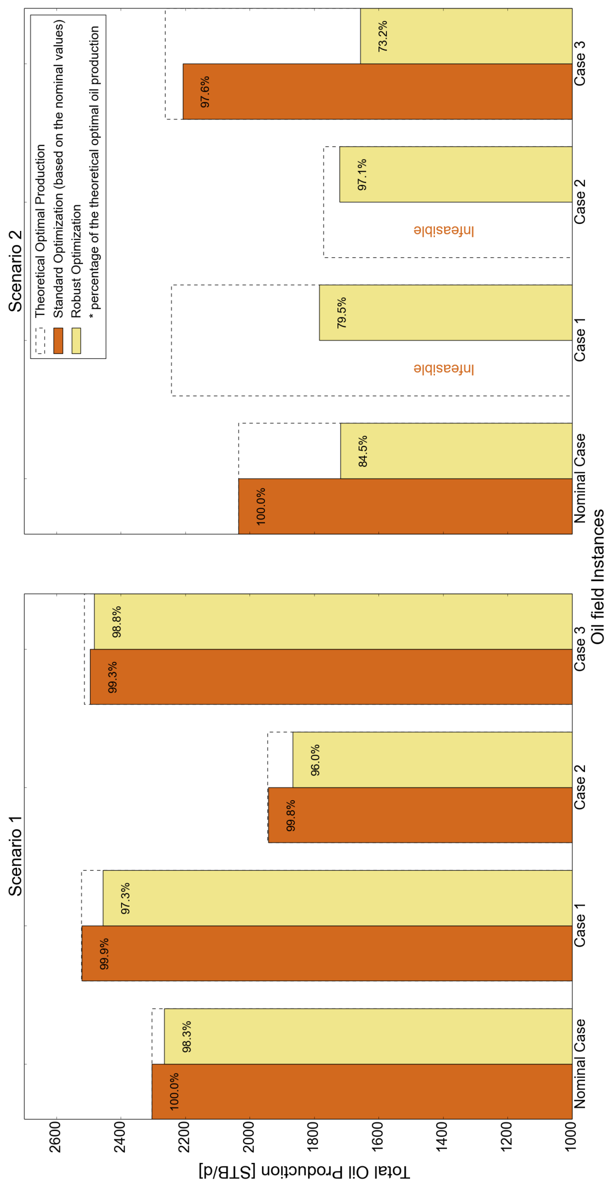


Figure 5.4: Total production comparison for each scenario of optimization.

Table 5.4: Scenario 1.

	Oil field Instances (Cases)			
	Nominal	1	2	3
Solution	q_{oil}	q_{oil}	q_{oil}	q_{oil}
Theoretical	2304	2523	1945	2513
Nominal	2304	2521	1942	2496
Robust	2265	2456	1867	2482

* oil rates in [STB/d]

Table 5.5: Scenario 2.

	Oil field Instances (Cases)			
	Nominal	1	2	3
Solution	q_{oil}	q_{oil}	q_{oil}	q_{oil}
Theoretical	2035	2244	1772	2263
Nominal	2035	Unfeasible	Unfeasible	2208
Robust	1719	1784	1721	1657

* oil rates in [STB/d]

an average oil production that is about 2% higher than the production obtained with robust optimization. In the second scenario, for the oil field instances of cases 1 and 2, separation capacities are violated by the nominal solution. This trend of a higher performance of the nominal solution over the robust solution in a optimization scenario where constraints are not so tight is expected. However, for this scenario when both solutions are feasible the robust approach produces 25% less oil than the nominal. It is important to notice that the difference between the optimal objective function (total oil production) calculated using piecewise-linear functions and obtained by simulation (using the optimized injections and routings) is, for all cases, less than 0.6% for both scenarios, which validates the use of the Algorithm 1 for sampling the production curves.

The results obtained suggest the use of robust solutions under critical situations, when operation is not allowed to suffer with constraint violations, or as a safe set of initial operational setpoints, that would then be further modified to enhance production.

5.4 SUMMARY

This chapter presented experimental results for the production optimization problem with satellite wells. Results of robust and standard solutions are compared in synthetic oil field instantiated in a commercial multiphase flow simulator. A Robust solution is indicated for critical operational scenarios where feasibility takes precedence over production maximization.

6 CONCLUSION

The daily operation of an oilfield in an optimal manner is a task of high complexity. In order to propose the optimal setpoints for operating an oil field, operators and petroleum engineers had count in the past, and still today, with their practical and theoretical knowledge, on the wells and equipment behavior, and with simulation tools, which mimic the oil fields, and enable sensitivity and scenario analysis. All this is not always sufficient to produce the actual optimal setpoints, specially under unconventional operating scenarios, e.g. an equipment is down due to failure or maintenance, a new well is put into operation. More recently, mathematical tools have being designed to help operators to take the appropriate decisions that would lead to an optimal operation. These tools are based on mathematical programming models which are used to solve a production optimization problem. Modeling is a key aspect of those approaches, since the strategies to solve the optimization problems are directly linked to the mathematical characteristics of the production models.

Specifically for oil fields operated by gas-lift systems, there are many different modeling approaches that were designed to solve distinct production problems. For example, some can only handle unconstrained problems, while others only satellite wells, others are very complete and can deal with subsea completion, routing decisions and more. However, the absolute majority of those modeling approaches do not account for uncertainties that might influence the quality of the models and consequently the results of the optimization problems. Uncertainty is inherent to production systems. It could arise from measurement errors, lack of informative data, and insufficiently detailed models, among others.

In this dissertation we proposed a methodology to incorporate uncertainty in the parameters of the wells and equipment into the production models used for optimization. Then, starting with the uncertain production models, we used the results from the robust linear optimization theory to develop robust MILP formulations for the production optimization problem. The robust formulations are independent of the parameters chosen to carry the uncertainties. They also do not require knowledge of probability distributions, only assuming that the uncertain parameters are confined within a known set.

For two standard production optimization of gas-lifted oil field problems the methodology is applied and the resulting robust production optimization models are presented. The first consists in distributing a limited rate of pressurized gas for a group of gas-lifted satellite wells, with the selection of production routing under multiple constraints. The second problem is similar, but the gas-lifted wells have their production traveling through subsea manifolds and shared flow lines. Thus, pressure constraints apply. For the latter problem uncertain equality constraints appear and we had to extend the classical theory for robust linear optimization to account for this type of structure.

For an oilfield scenario containing satellite wells, robust and standard (this based on nominal values) optimization problems are solved. A comparison between both solutions, when applied to different realizations within the uncertainty sets, shows that the standard solution may lead to an unfeasible operation, whereas the robust solution always gives a feasible operating condition. However, for some scenarios of the constraints the robust approach may provide a too conservative solution with underproduction. This seems to be a drawback for using the robust solution. In most operational situations, that may actually be true. Nevertheless, for critical situations, when even slight infeasibility is not an option, the robust solution can provide the best guideline for field operators.

6.1 FURTHER DEVELOPMENTS

Future work will consider the influence of other parameters and their variation, since the uncertain parameters may not vary as significantly as assumed in the experiments.

We also leave two suggestions with regards to the conservatism of the robust solution and its potential to be used in real world application.

One suggestion relates to the choice in this dissertation to seek for column-wise uncertainty models for the uncertain parameters in production optimization problems. This framework is notably more conservative. Under mild assumptions that were not made in our developments, it is possible to create a row-wise uncertainty model. This would lead to a different robust production optimization model and perhaps a less conservative solution. Another strategy in this same line would be to adapt the cardinality constrained uncertainty set developed by Bertsimas and Sim [64] to be used for production optimization problems.

A second suggestion is to use a distinct robust paradigm other than the classical one used here. Another approach idea would be to solve a series of robust problems in a path to the real optimal oil field operation, in a multistage setting. The first stage would provide solutions similar to the ones of this dissertation. However, for any forthcoming stage an evaluation of the impacts of the application of the previous solution would be used to refine the problem and reduce the level of uncertainty. The idea of this method is to create a safe path to optimality.

REFERENCES

1 ECONOMIDES, M. J. et al. *Petroleum production systems*. 2nd. ed. Upper Saddle River, NJ: Prentice Hall, 2012.

2 FOSS, B. A.; JENSEN, J. P. Performance Analysis for Closed-Loop Reservoir Management. *SPE Journal*, v. 16, n. 01, p. 183–190, apr. 2011.

3 REDDEN, J.; SHERMAN, T.; BLANN, J. Optimizing gas-lift systems. *Fall Meeting of the Society of Petroleum Engineers of AIME*, 1974.

4 BUITRAGO, S.; RODRIGUEZ, E.; ESPIN, D. Global optimization techniques in gas allocation for continuous flow gas lift systems. *SPE Gas Technology Symposium*, 1996.

5 KOSMIDIS, V. D.; PERKINS, J. D.; PISTIKOPOULOS, E. N. Optimization of Well Oil Rate Allocations in Petroleum Fields. *Industrial & Engineering Chemistry Research*, v. 43, p. 3513–3527, 2004.

6 CAMPOS, S. et al. Urucu Integrated Production Model. *Proceedings of SPE Intelligent Energy Conference and Exhibition*, n. 2005, p. 1–21, 2010.

7 GUNNERUD, V.; FOSS, B. Oil production optimization-A piecewise linear model, solved with two decomposition strategies. *Computers and Chemical Engineering*, v. 34, n. 11, p. 1803–1812, 2010.

8 CODAS, A.; CAMPONOGARA, E. Mixed-integer linear optimization for optimal lift-gas allocation with well-separator routing. *European Journal of Operational Research*, v. 217, n. 1, p. 222–231, feb. 2012.

9 SILVA, T. L.; CAMPONOGARA, E. A computational analysis of multidimensional piecewise-linear models with applications to oil production optimization. *European Journal of Operational Research*, Elsevier B.V., v. 232, n. 3, p. 630–642, feb. 2014.

10 LIMA, T. et al. Modeling of flow splitting for production optimization in offshore gas-lifted oil fields: Simulation validation and applications. *Journal of Petroleum Science and Engineering*, Elsevier, v. 128, p. 86–97, 2015.

11 ELGSÆTER, S.; SLUPPHAUG, O.; JOHANSEN, T. A. Production optimization; System Identification and Uncertainty Estimation. *Intelligent Energy Conference and Exhibition*, 2008.

12 NAKASHIMA, P.; CAMPONOGARA, E. Solving a gas-lift optimization problem by dynamic programming. *IEEE Transactions on Systems, Man, Cybernetics — Part A*, v. 36, n. 2, p. 407–414, 2006.

13 BIEKER, H. P.; SLUPPHAUG, O.; JOHANSEN, T. A. Well Management Under Uncertain Gas/ or Water/Oil Ratios. *Proceedings of Digital Energy Conference and Exhibition*, p. 1–6 – SPE106959, 2007.

14 ELGSÆTER, S. M.; SLUPPHAUG, O.; JOHANSEN, T. A. A structured approach to optimizing offshore oil and gas production with uncertain models. *Computers & Chemical Engineering*, v. 34, p. 163–176, 2010.

15 BEN-TAL, A.; NEMIROVSKI, A. Robust solutions of uncertain linear programs. *Operations Research Letters*, v. 25, n. 1, p. 1–13, aug. 1999.

16 BEN-TAL, A.; NEMIROVSKI, A. Robust solutions of linear programming problems contaminated with uncertain data. *Math. Programming*, v. 88, n. 3, p. 411–421, sep. 2000.

17 BEN-TAL, A.; GHAOUI, L. E.; NEMIROVSKI, A. *Robust Optimization*. New Jersey: Princeton University Press Requests, 2009.

18 BERTSIMAS, D.; BROWN, D. B.; CARAMANIS, C. Theory and Applications of Robust Optimization. *SIAM Review*, v. 53, n. 3, p. 464–501, jan. 2011.

19 DENEHY, D. *Oil & gas: the production story.* 2nd. ed. Austin, TX: The University of Texas at Austin, 2008.

20 BOMMER, P. *A primier of oilwell drilling.* 7th. ed. Austin, TX: The University of Texas at Austin, 2008.

21 THOMAS, J. E. *Fundamentos de engenharia do petróleo.* 2nd. ed. Rio de Janeiro, RJ: Editora Interciências, 2001.

22 DENEHY, D. *Fundamentals of Petroleum.* 5th. ed. Austin, TX: The University of Texas at Austin, 2011.

23 SAPUTELLI, L. et al. Promoting Real-Time Optimization of Hydrocarbon Producing Systems. *Proceedings of Offshore Europe*, p. 1–9, 2003.

24 BEGGS, H. *Production Optimization: using NODAL analysis*. 2nd. ed. Tulsa, Oklahoma: Oil & Gas Consultants Intl., 2003.

25 VOGEL, J. V. Inflow Performance Relationships for Solution-Gas Drive Wells. *Journal of Petroleum Technology*, v. 20, n. 01, p. 83–92, apr. 1968.

26 BRADLEY, H.; GIPSON, F. (Ed.). *Petroleum engineering handbook*. 3rd. ed. Richardson, TX: Society of Petroleum Engineers, 1987.

27 SANTOS, O. G.; BORDALO, S. N.; ALHANATI, F. J. S. Study of the dynamics, optimization and selection of intermittent gas-lift methods — a comprehensive model. *Journal of Petroleum Science and Engineering*, v. 32, p. 231–248, 2001.

28 BERTUZZI, A.; WELCHON, J.; POETTMAN, F. Description and Analysis of an Efficient Continuous-Flow Gas-Lift Installation. *Journal of Petroleum Technology*, v. 5, p. 271–278, 1953.

29 TRIBBE, C.; MÜLLER-STEINHAGEN, H. M. An evaluation of the performance of phenomenological models for predicting pressure gradient during gas-liquid flow in horizontal pipelines. *International Journal of Multiphase Flow*, v. 26, p. 1019–1036, 2000.

30 ULLMANN, a.; BRAUNER, N. Closure relations for two-fluid models for two-phase stratified smooth and stratified wavy flows. *International Journal of Multiphase Flow*, v. 32, p. 82–105, 2006.

31 MACH, J.; PROANO, E.; BROWN, K. E. A nodal approach for applying system analysis to the flowing and artificial lift oil or gas well. *SPE*, 1979.

32 SAPUTELLI, L.; NIKOLAOU, M.; ECONOMIDES, M. J. Self-Learning Reservoir Management. *SPE Reservoir Evaluation & Engineering*, v. 8, n. 06, p. 534–547, apr. 2005.

33 CAMPONOGARA, E. et al. An automation system for gas-lifted oil wells: Model identification, control, and optimization. *Journal of Petroleum Science and Engineering*, Elsevier B.V., v. 70, n. 3-4, p. 157–167, 2010.

34 BIEKER, H. P.; SLUPPHAUG, O.; JOHANSEN, T. A. Real-Time Production Optimization of Oil and Gas Production Systems: A Technology Survey. *SPE Production & Operations*, n. December 2006, p. 11–13, 2007.

35 SAPUTELLI, L. et al. A Critical Overview of Artificial Neural Network Applications in the Context of Continuous Oil Field

Optimization. *Proceedings of SPE Annual Technical Conference and Exhibition*, p. 1–2, 2002.

36 ALARCÓN, G. A.; TORRES, C. F.; GOMEZ, L. E. Global optimization of gas allocation to a group of wells in artificial lift using nonlinear constrained programming. *Journal of Energy Resources Technology*, v. 124, n. 4, p. 262–268, 2002.

37 CAMPONOGARA, E.; De Conto, A. M. Lift-Gas Allocation under Precedence Constraints: MILP Formulation and Computational Analysis. *IEEE Transactions on Automation Science and Engineering*, v. 6, n. 3, p. 544–551, 2009.

38 KOSMIDIS, V. D.; PERKINS, J. D.; PISTIKOPOULOS, E. N. A mixed integer optimization formulation for the well scheduling problem on petroleum fields. *Computers and Chemical Engineering*, v. 29, p. 1523–1541, 2005.

39 RASHID, K. Optimal allocation procedure for gas-lift optimization. *Industrial and Engineering Chemistry Research*, v. 49, p. 2286–2294, 2010.

40 VIELMA, J. P.; AHMED, S.; NEMHAUSER, G. Mixed-Integer Models for Nonseparable Piecewise-Linear Optimization: Unifying Framework and Extensions. *Operations Research*, v. 58, n. 2, p. 303–315, apr. 2010.

41 WILLIAMS, H. P. *Model building in mathematical programming*. 5th. ed. London: Wiley, 2013.

42 DANTZIG, G. B.; THAPA, M. N. *Linear Programming*. New York, NY: Springer-Verlag, 1997.

43 WOLSEY, L. *Integer Programming*. 1st. ed. New York, NY: Wiley-Interscience, 1998.

44 WOLSEY, L.; NEMHAUSER, G. L. *Integer and Combinatorial Optimization*. 1st. ed. New York, NY: Wiley-Interscience, 1999.

45 LI, D.; SUN, X. *Nonlinear Integer Programming*. New York, NY: Springer, 2006.

46 VANDERBEI, R. J. *Linear Programming - Foundations and Extensions*. New Jersey: Springer, 2008.

47 FOURER, R. *Algebraic Modeling Languages for Optimization*. Northwestern University, 2013.

48 FOURER, R.; GAY, D. M.; KERNINGHAN, B. W. *AMPL - A Modeling Language for Mathematical Programming*. New York, NY: Thomson Brooks/Cole, 2003.

49 RANGAVAJHALA, S.; MULLUR, A.; MESSAC, A. The challenge of equality constraints in robust design optimization: examination and new approach. *Structural and Multidisciplinary Optimization*, v. 34, n. 5, p. 381–401, mar. 2007.

50 DANTZIG, G. B. Linear Programming under Uncertainty. *Management Science*, v. 1, n. 3-4, p. 197–206, apr. 1955.

51 BELLMAN, R. *Dynamic Programming*. 1. ed. Princeton, NJ, USA: Princeton University Press, 1957.

52 CHARNES, A.; COOPER, W. W. Chance-constrained programming. *Management Science*, v. 6, n. 1, p. 73–79, 1959.

53 SAHINIDIS, N. V. Optimization under uncertainty: State-of-the-art and opportunities. *Computers and Chemical Engineering*, v. 28, p. 971–983, 2004.

54 ROCKAFELLAR, R. T. *Optimization under uncertainty*. Seattle, WA: University of Washington, 2001.

55 WALLACE, S. W. Decision making under uncertainty: Is sensitivity analysis of any use? . *Operations research*, v. 48, n. 1, p. 20–25, feb. 2000.

56 WALD, A. Statistical Decision Functions Which Minimize the Maximum Risk. *Annals of Mathematics*, v. 46, n. 2, p. 265–280, 1945.

57 SOYSTER, A. L. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations research*, v. 21, n. February 2015, p. 1154–1157, 1973.

58 BEN-TAL, A.; NEMIROVSKI, A. Robust Convex Optimization. *Mathematics of Operations Research*, v. 23, n. i, p. 769–805, 1998.

59 EL GHAOUI, L.; LEBRET, H. Robust Solutions to Least-Squares Problems with Uncertain Data. *SIAM Journal on Matrix Analysis and Applications*, v. 18, p. 1035–1064, 1997.

60 EL GHAOUI, L.; OUSTRY, F.; LEBRET, H. Robust Solutions to Uncertain Semidefinite Programs. *SIAM Journal on Optimization*, v. 9, n. 1, p. 33–52, 1998.

61 KOUVELIS, P.; YU, G. *Robust Discrete Optimization and its Applications*. 1. ed. Boston, MA: Kluwer Academic Publishers, 1997.

62 BERTSIMAS, D.; THIELE, A. Robust and data-driven optimization: Modern decision-making under uncertainty. *INFORMS Tutorials in . . .* , n. March, p. 1–39, 2006.

63 GUERRA VÁZQUEZ, F. et al. Generalized semi-infinite programming: A tutorial. *Journal of Computational and Applied Mathematics*, v. 217, p. 394–419, 2008.

64 BERTSIMAS, D.; SIM, M. The price of robustness. *Operations research*, v. 52, n. 1, p. 35–53, feb. 2004.

65 MINOUX, M. Robust Linear Programming with Right-Hand-Side Uncertainty, Duality and Applications. In: *Encyclopedia of Optimization, Second Edition*. New York, NY: Springer, 2009. p. 3317–3327.

66 RANGAVAJHALA, S.; MESSAC, A. Designer's preferences regarding equality constraints in robust design optimization. *Structural and Multidisciplinary Optimization*, v. 41, n. 1, p. 17–38, jul. 2009.

67 BERTSIMAS, D.; SIM, M. Robust discrete optimization and network flows. *Mathematical Programming*, v. 98, n. 1-3, p. 49–71, sep. 2003.

68 AVERBAKH, I. On the complexity of a class of combinatorial optimization problems with uncertainty. *Mathematical Programming*, v. 90, n. 2, p. 263–272, 2001.

69 ATAMTÜRK, A. Strong formulations of robust mixed 0-1 programming. *Mathematical Programming*, v. 108, n. 2-3, p. 235–250, 2006.

70 VAN ESSEN, G. M. et al. Robust Optimization of Oil Reservoir Flooding. In: *Conference on Control Applications*. [S.l.: s.n.], 2006. p. 699–704.

71 CHEN, C.; LI, G.; REYNOLDS, A. Robust constrained optimization of short and long-term NPV for closed-loop reservoir management. *SPE Reservoir Simulation Symposium*, n. February, p. 1–23, 2011.

72 CAPOLEI, A. et al. A mean–variance objective for robust production optimization in uncertain geological scenarios. *Journal of Petroleum Science and Engineering*, Elsevier, v. 125, n. 1, p. 23–37, 2015.

73 CAMPONOGARA, E.; NAKASHIMA, P. H. R. Solving a gas-lift optimization problem by dynamic programming. *European Journal of Operational Research*, v. 174, p. 1220–1246, 2006.

74 HÜLSE, E. O.; CAMPONOGARA, E. A mathematical programming formulation for robust production optimization of gas-lifted oil fields. In: RODRIGUES, H. et al. (Ed.). *Engineering Optimization 2014*. Lisbon: CRC Press, 2014. p. 891–896.

75 CODAS, A. et al. Integrated production optimization of oil fields with pressure and routing constraints: The Urucu field. *Computers & Chemical Engineering*, Elsevier Ltd, v. 46, p. 178–189, nov. 2012.

76 AGUIAR, M. A. S.; CAMPONOGARA, E.; SILVA, T. L. A mixed-integer convex formulation for production optimization of gaslifted oil fields with routing and pressure constraints. *Brazilian Journal of Chemical Engineering*, v. 31, n. 2, p. 439–455, 2014.

77 BAZARAA, M. S.; SHERALI, H. D.; SHETTY, C. *Nonlinear Programming: Theory and Algorithms*. 3rd. ed. New Jersey: Wiley-Interscience, 2006.

78 KALOS, M. H.; WHITLOCK, P. a. *Monte Carlo Methods*. 2nd. ed. Weinheim: Wiley, 2008. 203 p.

79 SAEPUDIN, D. et al. Optimization of Gas Injection Allocation in Multi Gas Lift Wells System. In: . [S.l.]: Proceedings of the International Conference on Engineering Optimization, 2008.

80 STONE, H. Approximation of curves by line segments. *Mathematics of Computation*, v. 15, n. 73, p. 40–40, 1961.

81 CANTONI, A. Optimal Curve Fitting With Piecewise Linear Functions. *IEEE Transactions on Computers*, C-20, n. 1, p. 59–67, 1971.

APPENDIX A — SAMPLING ALGORITHMS

Approximating a complicated function by piecewise functions is an alternative to reduce complexity. This is a common approach in optimization to create tractable versions of originally hard to solve problems. In this line, the most ordinary approach is to build piecewise-linear (PWL) functions, where in each interval a linear function is used to represent the original function. When the original function is well defined (with known derivatives), or at least has a mathematical description, many algorithms exist to find the appropriate breakpoints that define the intervals, and to designate their linear approximations, e.g. [80, 81].

However, if the original function is conceptual (i.e., has no mathematical description) those algorithms do not apply. Some algorithms have been proposed for this situation [75], but they rely on the fact that a very fine sampling step is performed a priori. Then unnecessary breakpoints are removed. The remaining points delimit the intervals and are used to define the linear functions.

Depending on the accuracy required, acquiring this overly sampled set of breakpoints may be time consuming. We notice, however, that a special set of conceptual functions can be seen as quasiconvex and that it is possible to explore this structure even if the function does not have a mathematical formulation. Following, two algorithms are suggested. For a given maximum tolerated error they determine breakpoints that define the intervals and the respective linear functions that approximate the original function. Both algorithms are based on the fact that, if a conceptual function is quasiconcave the error between this function and a linear approximation of it is also a quasiconcave function. Also, the two are only applicable to univariate functions.

ALGORITHMS

Let $\widehat{g}(x)$ be an unidimensional conceptual function with a concave shape in the interval $[l, u] \subset \mathbb{R}$. An algorithm receives as input a subset $[a, c] \subseteq [l, u]$ and a maximum tolerable approximation error (i.e. absolute error) ε. Then, it attempts to find a set of breakpoints $\mathcal{B} = \{b_1 = a, b_2, \ldots, b_{T-1}, b_T = c\}$, with $b_1 < b_2 < \cdots < b_T$, that define a piecewise-linear function,

$$\widehat{h}(x) = \begin{cases} \widehat{h}_{(b_k, b_{k+1})}(x) = \alpha_k \cdot x + \beta_k \quad \text{if } b_k \leq x \leq b_{k+1}, \ \ldots \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad \ldots \ k = \{1, \ldots, T-1\}. \end{cases}$$

where,

$$\alpha_k = \frac{\widehat{g}(b_{k+1}) - \widehat{g}(b_k)}{b_{k+1} - b_k}, \qquad \beta_k = \frac{\widehat{g}(b_k) \cdot b_{k+1} - \widehat{g}(b_{k+1}) \cdot b_k}{b_{k+1} - b_k}$$

which is an approximation of the function $\widehat{g}(x)$. α_k and β_k induce each individual linear segment to pass through the point $(b_k, \widehat{g}(b_k))$ and $(b_{k+1}, \widehat{g}(b_{k+1}))$. The algorithm select the breakpoints ensuring that $\widehat{e}(x) \leq \varepsilon$, $\forall x \in [a, c]$, where $\widehat{e}(x) = \widehat{g}(x) - \widehat{h}(x)$ is the absolute error of approximation. More specifically,

$$\widehat{e}(x) = \begin{cases} \widehat{e}_{(b_k, b_{k+1})}(x) = \widehat{g}(x) - \widehat{h}_{(b_k, b_{k+1})}(x) & \text{if } b_k \leq x \leq b_{k+1}, \ \ldots \\ & \ldots \ k = \{1, \ldots, T-1\}. \end{cases}$$

Following, we proposed two heuristics to achieve this result.

"Search method" algorithm

The heuristic used for Algorithm 1 is based on finding the maximum of $\widehat{e}(x)$ for a given interval. Then, if this value is larger than ε, three actions are taken: the breakpoint corresponding to the maximum point is added to set $\mathcal{B}$, the interval is split at this breakpoint, and the procedure is called recursively for the two new intervals. The process stops when all intervals covered have a $\widehat{e}(x)$ not greater than ε.

Because $\widehat{e}(x)$ is defined by $\widehat{g}(x)$, which has no mathematical description, thus no derivative explicitly available, one has to use derivative-free methods, as line search without using derivative methods [77], in order to compute the operators $\max$ and argmax in Algorithm 1. Four methods were studied: bisection method, dichotomous search, the golden section method, and the Fibonacci method. All of these methods have the same basic goal. For a function $\widehat{e}(x)$, $x \in [a, c]$, they start with the initial interval and after a series of iterations they return an interval $[a^\star, c^\star]$ of length 2δ. $x^\star \in [a^\star, c^\star]$, where $x^\star = \operatorname{argmax}_x\{\widehat{e}(x)\}$. The choice of the grid length δ affects directly the convergence of any of the methods. For a smaller grid length the number of iterations is higher

Theoretically, in analyzing the number of iterations required by the methods above, the dichotomous search method presents as the most efficient one. However, in experimental results using a quasiconcave function $\widehat{e}(x)$ with values obtained in a simulator, the golden section method presented similar efficiency but more ro-

Algorithm 1: "Search method" algorithm

input : $\widehat{g}(x)$, $[a, c]$, ε
$\mathcal{B} \longleftarrow \{a, c\}$
$\mathcal{B} \longleftarrow$ `search`$(a,\ c,\ \varepsilon,\ \mathcal{B})$
return $\mathcal{B}$

Procedure `search`$(a_t,\ c_t,\ \varepsilon,\ \mathcal{B})$
  compute $\widehat{g}(a_t)$ and $\widehat{g}(c_t)$
  define $\widehat{h}_{(a_t,c_t)}(x)$
  $e^{\max} \longleftarrow \max_x\{\widehat{e}_{(a_t,c_t)}(x) : a_t \leq x \leq c_t\}$
  $x^\star \longleftarrow \operatorname{argmax}_x\{\widehat{e}_{(a_t,c_t)}(x) : a_t \leq x \leq c_t\}$
  if $e^{\max} \geq \varepsilon$ **then**
    $\mathcal{B} \longleftarrow \mathcal{B} \cup \{x^\star\}$
    $\mathcal{A} \longleftarrow$ `search`$(a_t,\ x^\star,\ \varepsilon,\ \mathcal{B})$
    $\mathcal{B} \longleftarrow \mathcal{B} \cup \mathcal{A}$
    $\mathcal{A} \longleftarrow$ `search`$(x^\star,\ c_t,\ \varepsilon,\ \mathcal{B})$
    $\mathcal{B} \longleftarrow \mathcal{B} \cup \mathcal{A}$
  return $\mathcal{B}$

bust results. Therefore, this is the technique chosen for Algorithm 1. Algorithm 2 presents the golden section method.

"Cut method" algorithm

The heuristic used for Algorithm 3 is based on the golden section method for line search without derivatives. It starts sampling two points within the initial interval using the golden ratio ϕ as a rule. Then it evaluates the error function in the limiting points and in the two intermediary points. If one or more values are higher than ε, then it adds the breakpoints associated with the higher error to set $\mathcal{B}$ and recursively calls the the procedure for the intervals before and after this breakpoint. However, if all errors are lower than ε, then it still has to check if a higher error value does not exist for a value in the domain. Nevertheless it has to check it around the breakpoint with the higher error.

The entire procedure is performed recursively until the interval checked is smaller than a given grid length δ.

Algorithm 2: Golden section method

input : $\widehat{e}(x)$, $[a, c]$, δ
$\phi \longleftarrow 0.618$
$k \longleftarrow 1$
$a_k \longleftarrow a$
$c_k \longleftarrow b$
$\sigma_k \longleftarrow a_k + (1 - \phi)(c_k - a_k)$
$\mu_k \longleftarrow a_k + \phi(c_k - a_k)$
while $c_k - a_k \geq \delta$ **do**
  compute $\widehat{e}(\sigma_k)$ and $\widehat{e}(\mu_k)$
  if $\widehat{e}(\sigma_k) > \widehat{e}(\mu_k)$ **then**
    $a_{k+1} \longleftarrow \sigma_k$
    $c_{k+1} \longleftarrow c_k$
    $\sigma_{k+1} \longleftarrow \mu_k$
    $\mu_{k+1} \longleftarrow a_{k+1} + \phi(c_{k+1} - a_{k+1})$
  else
    $a_{k+1} \longleftarrow a_k$
    $c_{k+1} \longleftarrow \mu_k$
    $\sigma_{k+1} \longleftarrow a_{k+1} + (1 - \phi)(c_{k+1} - a_{k+1})$
    $\mu_{k+1} \longleftarrow \sigma_k$
  $k \longleftarrow k + 1$
$x^\star \longleftarrow \frac{a_k + c_k}{2}$
return $\widehat{e}(x^\star)$, $x^\star$

Important remarks

Absolute error vs. relative error: Algorithm 1 and Algorithm 3 rely on the fact that given a concave function $\widehat{g}(x)$ its corresponding absolute error function will also be concave. However, convexity is not maintained for the relative error function. In practice, sometimes it is more convenient to require a piecewise linearization that has a maximum relative error $\varepsilon^{\%}$ than a maximum absolute error ε. Note that it is possible to modify both algorithms to account for this. The idea is that in each interval that is being tested (each recourse) a corresponding absolute error has to be calculated. For example, before any of the procedures – search() or cut() – is called, one has to compute $\varepsilon = \varepsilon^{\%} \cdot \min\{\widehat{e}(a), \widehat{e}(b)\}$, where a and b are the limits of the interval being checked in a recourse. With this approach we guarantee an upper bound for the absolute errors being checked in

Algorithm 3: "Cut method" algorithm

input: $\widehat{g}(x)$, $[a, c]$, ε, δ
$\mathcal{B} \longleftarrow \{a, c\}$
$\mathcal{B} \longleftarrow$ `cut`$(a, c, a, c, \varepsilon, \delta, \mathcal{B})$
return $\mathcal{B}$
Procedure `cut`$(a_\mathrm{t}, c_\mathrm{t}, a_\mathrm{f}, c_\mathrm{f}, \varepsilon, \delta, \mathcal{B})$
  if $c_\mathrm{t} - a_\mathrm{t} \leq \delta$ **then**
    return $\mathcal{B}$
  $\sigma \longleftarrow a_\mathrm{t} + (1 - \phi)(c_\mathrm{t} - a_\mathrm{t})$
  $\mu \longleftarrow a_\mathrm{t} + \phi(c_\mathrm{t} - a_\mathrm{t})$
  compute $\widehat{g}(a_\mathrm{f}), \widehat{g}(a_\mathrm{t}), \widehat{g}(\sigma), \widehat{g}(\mu), \widehat{g}(c_\mathrm{t})$ and $\widehat{g}(c_\mathrm{f})$
  define $\widehat{h}_{(a_\mathrm{f}, c_\mathrm{f})}(x)$
  $e^{\max} \longleftarrow \max\{\widehat{e}_{(a_\mathrm{f},c_\mathrm{f})}(a_\mathrm{t}), \widehat{e}_{(a_\mathrm{f},c_\mathrm{f})}(\sigma), \widehat{e}_{(a_\mathrm{f},c_\mathrm{f})}(\mu), \widehat{e}_{(a_\mathrm{f},c_\mathrm{f})}(c_\mathrm{t})\}$
  case $\widehat{e}_{(a_\mathrm{f},c_\mathrm{f})}(\sigma) > \varepsilon$ *and* $\widehat{e}_{(a_\mathrm{f},c_\mathrm{f})}(\mu) > \varepsilon$
    if $\widehat{e}_{(a_\mathrm{f},c_\mathrm{f})}(\sigma) > \widehat{e}_{(a_\mathrm{f},c_\mathrm{f})}(\mu)$ **then**
      $\mathcal{B} \longleftarrow \mathcal{B} \cup \{\sigma\}$
      $\mathcal{B} \longleftarrow \mathcal{B} \cup$ `cut`$(a_\mathrm{t}, \sigma, a_\mathrm{t}, \sigma, \varepsilon, \delta, \mathcal{B})$
      $\mathcal{B} \longleftarrow \mathcal{B} \cup$ `cut`$(\sigma, b_\mathrm{t}, \sigma, b_\mathrm{t}, \varepsilon, \delta, \mathcal{B})$
    else
      $\mathcal{B} \longleftarrow \mathcal{B} \cup \{\mu\}$
      $\mathcal{B} \longleftarrow \mathcal{B} \cup$ `cut`$(a_\mathrm{t}, \mu, a_\mathrm{t}, \mu, \varepsilon, \delta, \mathcal{B})$
      $\mathcal{B} \longleftarrow \mathcal{B} \cup$ `cut`$(\mu, b_\mathrm{t}, \mu, b_\mathrm{t}, \varepsilon, \delta, \mathcal{B})$
  case $\widehat{e}_{(a_\mathrm{f},c_\mathrm{f})}(\sigma) > \varepsilon$
    $\mathcal{B} \longleftarrow \mathcal{B} \cup \{\sigma\}$
    $\mathcal{B} \longleftarrow \mathcal{B} \cup$ `cut`$(a_\mathrm{t}, \sigma, a_\mathrm{t}, \sigma, \varepsilon, \delta, \mathcal{B})$
    $\mathcal{B} \longleftarrow \mathcal{B} \cup$ `cut`$(\sigma, \mu, \sigma, \mu, \varepsilon, \delta, \mathcal{B})$
  case $\widehat{e}_{(a_\mathrm{f},c_\mathrm{f})}(\mu) > \varepsilon$
    $\mathcal{B} \longleftarrow \mathcal{B} \cup \{\mu\}$
    $\mathcal{B} \longleftarrow \mathcal{B} \cup$ `cut`$(\sigma, \mu, \sigma, \mu, \varepsilon, \delta, \mathcal{B})$
    $\mathcal{B} \longleftarrow \mathcal{B} \cup$ `cut`$(\mu, c_\mathrm{t}, \mu, c_\mathrm{t}, \varepsilon, \delta, \mathcal{B})$
  otherwise
    case $\widehat{e}(a_\mathrm{t}) = e^{\max}$
      $\mathcal{B} \longleftarrow$ `cut`$(a_\mathrm{t}, \sigma, a_\mathrm{f}, c_\mathrm{f}, \varepsilon, \delta, \mathcal{B})$
    case $\widehat{e}(\sigma) = e^{\max}$
      $\mathcal{B} \longleftarrow$ `cut`$(a_\mathrm{t}, \mu, a_\mathrm{f}, c_\mathrm{f}, \varepsilon, \delta, \mathcal{B})$
    case $\widehat{e}(\mu) = e^{\max}$
      $\mathcal{B} \longleftarrow$ `cut`$(\sigma, c_\mathrm{t}, a_\mathrm{f}, c_\mathrm{f}, \varepsilon, \delta, \mathcal{B})$
    case $\widehat{e}(c_\mathrm{t}) = e^{\max}$
      $\mathcal{B} \longleftarrow$ `cut`$(\mu, c_\mathrm{t}, a_\mathrm{f}, c_\mathrm{f}, \varepsilon, \delta, \mathcal{B})$
  return $\mathcal{B}$

an interval. This is possible because the function is concave thus the minimum in any interval will be a limit point of this interval.

Ideal Effectiveness: Ideally, it is desired that the algorithms do not only provide an ordinary set of breakpoints that produce a piecewise-linear approximation for a given maximum error, but actually the smaller set of breakpoints for that required error. Both algorithms do not achieve this goal. However, using an algorithm that is not presented here but that was designed with this objective, experimental results showed similar results to Algorithm 1.

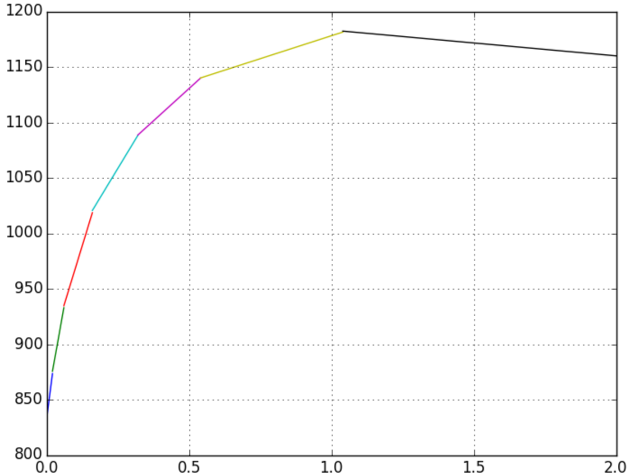
Extension to higher dimensions: The use of the algorithms presented here are limited to univariate functions. This reduce the number of situation where the algorithms may be applied. Once multidimensional search methods without using derivatives exit [77], it is intuitive to think that, for example, Algorithm 1 could be easily extended. However, this may not be trivial, since higher dimensions usually introduce more complexity for problems. Thus, this is left for future investigation.

EXPERIMENTAL RESULTS

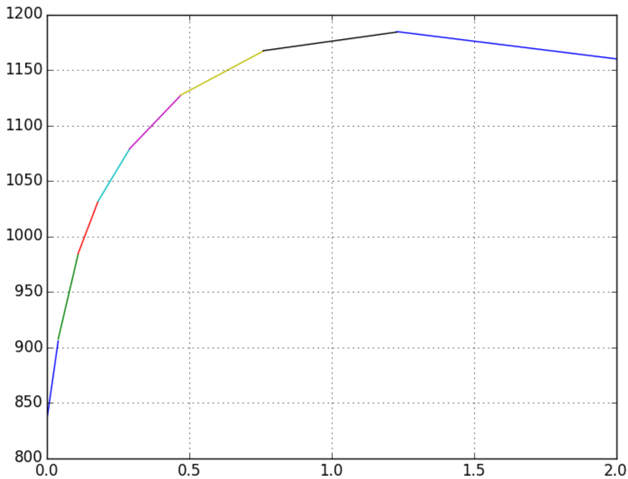
Algorithm 1 and Algorithm 3 were tested using a multiphase flow simulator to produced the quasiconcave functions $\widehat{g}(x)$ and the algorithms themselves were implemented in *Python* (programming language). For both algorithms the experimental results presented below are using the same initial interval $[0, 2]$, function $\widehat{g}(x)$ and minimum grid length $\delta = 1 \times 10^{-5}$. The only parameter that changes in each of the five tests is the desired maximum relative error $\varepsilon^{\%}$. The results for each test include:

- **Calls:** is the number of queries made to the simulator to evaluate function $\widehat{g}(x)$ in a point of its domain;
- **Breakpoints:** is the number of points defining the piecewise-linear curve, which guarantee a required maximum the relative error $\varepsilon^{\%}$.

The experimental results using Algorithm 1 are in Table 1 and for Algorithm 3 are in Table 2. Figure 1 illustrates the piecewise-linear curve formed by the breakpoints obtained by both algorithms.



(a) “Search method” - Test #1



(b) “Cut method” - Test #1

Figure 1: Piecewise-linear curves for Test # 1.

Table 1: "Search method" Algorithm 1.

Test	$\varepsilon^{\%}$	Calls	Breakpoints
#1	1%	132	6
#2	0.5%	189	9
#3	0.25%	241	12
#4	0.1%	345	20
#5	0.01%	813	65

Table 2: “Cut method” Algorithm 3.

Test	$\varepsilon^{\%}$	Calls	Breakpoints
#1	1%	57	7
#2	0.5%	64	10
#3	0.25%	95	16
#4	0.1%	145	44
#5	0.01%	279	112

Notice that in Algorithm 1 tended, in similar tests, to produce a smaller number of breakpoints compared to Algorithm 3. However, the latter used less function evaluations to achieve its results. Therefore, one has to choose whether to pay a higher “price” earlier, when producing the piecewise-linearization or later when using the piecewise-linear function in a optimization problem.

An advantage of both methods, when compared to the aforementioned sampling algorithm presented by Codas et al. [75], is in the reduced number of function evaluations performed. In fact, the algorithm in [75] does not sample the curve directly, but it reduces the number of points of an input oversampled curve. This input sample set has to be sufficiently large for the algorithm to be effective and guarantee a maximum required error. In the example above the value of δ was chosen to match the maximum accuracy of the simulator. Thus, to provide a safe input sampling set for the reduction algorithm, since the interval the function is being approximated is $[0, 2]$, a fine sampling could mean up to 2×10^5 evaluations of the function. Note that this number is much larger than the number of evaluations required by the algorithms presented above for any of the errors considered (see Tables 1 and 2).