Short-Term Supply Chain Management in Upstream Natural Gas Systems

by

Ajay Selot

Submitted to the Department of Chemical Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Chemical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2009

Author . . . .
Department of Chemical Engineering
November 25, 2008

Certified by. . . . .
Paul I. Barton
Lammot du Pont Professor of Chemical Engineering
Thesis Supervisor

Accepted by . . . .
William M. Deen
Carbon P. Dubbs Professor of Chemical Engineering
Chairman, Committee for Graduate Students

Short-Term Supply Chain Management in Upstream Natural Gas Systems

by

Ajay Selot

Abstract

Natural gas supply chain planning and optimization is important to ensure security and reliability of natural gas supply. However, it is challenging due to the distinctive features of natural gas supply chains. These features arise from the low volumetric energy density of natural gas and the significance of gas quality and pressure in supply chain operations. Contracts play a central role in the entire supply chain due to high capital cost, specificity and investment risks associated with gas infrastructure.

An upstream production planning framework is crucial for supply-side optimization and scenario evaluation in the natural gas supply chain. The technical features of upstream systems imply that the most efficient mode of operation is by single entity central control of the system, while their economics favor involvement of multiple parties in ownership. To resolve this conflict, upstream systems are generally operated by a single operator on the basis of *governing rules* that stem from agreements between the upstream operator, multiple stakeholders and consumer facilities. These agreements govern production sharing, operational strategy and gas sales in the upstream system.

A short-term operational planning framework (with a 2-12 weeks planning horizon) for upstream natural gas systems is presented that can help to maximize production infrastructure utilization and aid in its management, minimize costs and meet production targets while simultaneously satisfying governing rules. Its requirements are inspired by the Sarawak Gas Production System (SGPS), an offshore gas production system in the South China Sea, which supplies the liquefied natural gas (LNG) plant complex at Bintulu in East Malaysia. This is the first attempt to formulate a comprehensive modeling framework for an upstream gas production system that includes a production infrastructure model and a methodology to incorporate governing rules. The model has two components: *the infrastructure model* is a model of the physical system, i.e., of wells, trunkline network and facilities while *the contractual model* is a mathematical representation of *the governing rules*, e.g., production-sharing contracts (PSC), customer specifications and operational rules. The model formulation and objectives are from the perspective of the upstream operator.

The infrastructure model incorporates the capability to track multiple qualities of gas throughout the network and determine the optimal routing and blending of gas such that the quality specifications are satisfied at the demand nodes. Nonlinear pressure-flowrate relationships in wells and the network are included for predicting a sufficiently accurate pressure-flowrate profile thereby facilitating implementation of the production strategy on the network. Modeling of complex platform configurations with reversible lines, lines that can be shut-off in normal operation and compression facilities, further improve the realistic representation of the network. A simplified prediction of natural gas liquids (NGL) production is included to maximize NGL revenue.

The contractual model represents the framework for modeling the governing rules that are central to the operation of upstream systems. Modeling of production-sharing contracts is a two-fold challenge: accounting for gas volumes and converting the logical rules as stated in the system operations manual to binary constraints. A PSC network representation is proposed to account for gas volumes as well as interactions between different PSC. PSC rules are expressed as logical expressions in terms of availability, priority and transfer Boolean-states, and converted to binary constraints. Additional logical constraints are required to model the inference and intent of the rules. Operational rules can be modeled within the same framework.

The resulting mathematical program is a mixed-integer nonlinear program (MINLP) with nonconvex functions and can be solved with the current state-of-the-art global optimization approaches, provided careful attention is paid to the model formulation. A hierarchical multi-objective approach is proposed to address multiple objectives when operating upstream systems, by optimizing a lower priority objective over the multiple optimal solutions of a program with a higher priority objective to obtain a win-win scenario. A reproducible case study that captures all the features of natural gas upstream systems is constructed to facilitate future work in algorithm development for such problems. A preliminary comparison with the existing approach indicates that substantial benefits may be possible by using the proposed approach for short-term planning.

The application of a reduced-space global optimization approach to planning in upstream gas networks has also been demonstrated, which can significantly lower the number of variables in the branch-and-bound algorithm. The lower bounding problem is implemented using McCormick (convex) relaxations of computer evaluated functions and solved by implementing a nonsmooth bundle solver as a linearization tool to obtain a linear programming relaxation. The upper bounding problem is implemented using automatic differentiation and a local NLP solver. Branch-and-bound with reduction heuristics and linearization propagation is used for global optimization. This approach has been found to be competitive with current state-of-the-art global optimization algorithms for upstream planning problems.

Thesis Supervisor: Paul I. Barton
Title: Lammot du Pont Professor of Chemical Engineering

Acknowledgments

I firmly believe that we, as individuals, are not just products of our own efforts, but also of hard work, time and energy that people surrounding us invest in us and sacrifices they make for us. To this end, I would like to acknowledge people who have made possible my journey to this milestone.

First and foremost, I would like to thank Paul for his technical and intellectual input on the work presented in this thesis. I also appreciate his almost infinite patience in the face of several obstacles encountered during the course of this project. I am grateful to him for taking the time, in spite of his very busy schedule, to travel half way around the globe to Miri and Bintulu in Sarawak to participate in technical meetings with Shell engineers and, for his keen wisdom in discerning key technical issues during those meetings. Finally, I also thank him for giving me ample freedom to chart out my own way.

I would also like to thank the members of my thesis committee: Professor Gordon M. Kaufman, Professor Gregory J. McRae and Professor Jefferson W. Tester. They have always been enthusiastic about the subject matter, have provided excellent suggestions throughout the course of the project and have forced me to open my eyes to see the problem in wider contexts than I would have done myself.

This work would not have been possible without substantial help from several people from the Shell group of companies. I would like to express my sincere gratitude and thanks to all of them. Dr. Thomas L. Mason, (then at Shell International Exploration and Production (SIEP) in Rijswijk, The Netherlands) initiated the collaboration with Sarawak Shell Berhad (SSB), provided overall direction to the project and served as the liaison and an enthusiastic supporter for this project at SIEP. This project would have achieved much less without his involvement. I would like to express my gratitude to Loi Kwong-Kuok (SSB) for providing extensive information about the Sarawak Gas Production System (SGPS), answering my incessant queries promptly and serving as a primary contact for this project at SSB. I appreciate help from Andrew Hooks (SSB) in providing further information on the SGPS and the existing models. A sincere thanks to K. L. Tan (SSB) for hosting myself (and Paul) at the SGPS operations center in Bintulu and providing me with eye-opening firsthand insights into the SGPS operations. I am grateful to Mark Robinson (SSB) for providing additional technical information on the SGPS as well as for lending his whole-hearted support to the project and clearing out several administrative hurdles. I thank Nithi Bunchatheravat (SSB) for providing necessary files and help for the existing model and answering multiple queries about it. I am also grateful to Dr. Richard Sears for his help in resolving several administrative issues in his capacity as the liaison for all Shell projects at MIT. I would like to express my sincere gratitude to SSB Decision Review Board for permitting me to use the SGPS as a case study in this work. Finally, my sincere thanks to *Smart Fields, Shell International Exploration and Production, Rijswijk, The Netherlands* for funding this project.

To Geeta & Prem

Contents

List of Figures

List of Tables

Chapter 1

Overview

There are signs of a global energy crisis in the making in recent years. World energy demand is estimated to grow by 50% in the period from 2005 to 2030 [1, 2] without significant changes in government policies (the reference scenario). The unprecedented growth in energy usage is being primarily driven by rapid economic growth in emerging markets and their integration into the global economy. The International Energy Agency (IEA) estimates that the developing world will contribute 74% of the increase in energy usage in the reference scenario. The share of China and India alone is expected to be around 45% of the projected increase in the IEA reference scenario [1].

The rise in energy demand in conjunction with supply-side problems have resulted in a rapid rise in energy prices for the past several years. Underinvestment in energy during a period of low energy prices in the last decade has led to supply-side bottlenecks in terms of equipment, technology and human resources. A rise of resource nationalism (i.e., the desire of national governments or national oil companies to exercise tight control on resources) is preventing the flow of investments and technology into the most productive resources. This is resulting in low production rates and total recovery in existing oil and gas fields as well as hindering the development of new fields. Furthermore, the long-term investment climate in energy has been adversely impacted by the continuing policy uncertainty in the regulation of carbon emissions and alternative energy resources. This discourages investors from investing

in large-scale projects with long-term returns. Finally, geopolitical tensions and political instability in producing regions and transportation passages continue to have a negative impact on oil and gas markets.

For the foreseeable future, fossil fuels will continue to form the mainstay of global energy supply. They contributed 81% of global energy demand in 2005 [1]. Their share is estimated to almost stay the same at around 82% of the global energy demand in the IEA reference scenario (without any major policy changes) and drop to 76% in the IEA alternative policy scenario (i.e., with government policies to address energy security and climate change) by 2030. Hence, over the next two decades or more, ensuring reliable supplies of fossil fuels will be instrumental for global energy security and therefore for maintaining the high growth rate of the global economy that is crucial for breaking the cycle of poverty in the developing and poor economies.

1.1 Natural Gas

Natural gas contributed around a fifth of global energy demand in 2005 (Figure 1-1). Natural gas is primarily methane (CH_4) (usually in the 70-90 mole percent range for well-head gas). It contains varying amounts of ethane, propane, butane and other higher chain hydrocarbons. It can be formed by various processes including high pressure and temperature decomposition of organic matter, bio degradation of organic matter and abiogenic processes. Once gas is formed, it will rise to the surface through porous rocks unless it is trapped by a geological trap. A typical trap has a porous rock formation that holds gas with an umbrella shaped dome on top with dense impermeable rock that prevents it from escaping further.

Natural gas can be produced from gas only reservoirs in which case the gas contains none to a small fraction of heavier (than ethane) hydrocarbons. A fraction of the heavier hydrocarbons (especially if they are present in appreciable amount) condense when the pressure of the reservoir fluid is reduced at the well-head and an organic liquid phase separates out. Additionally, raw gas can also contain water which needs to be separated at the well-head to avoid formation of gas hydrates that can choke
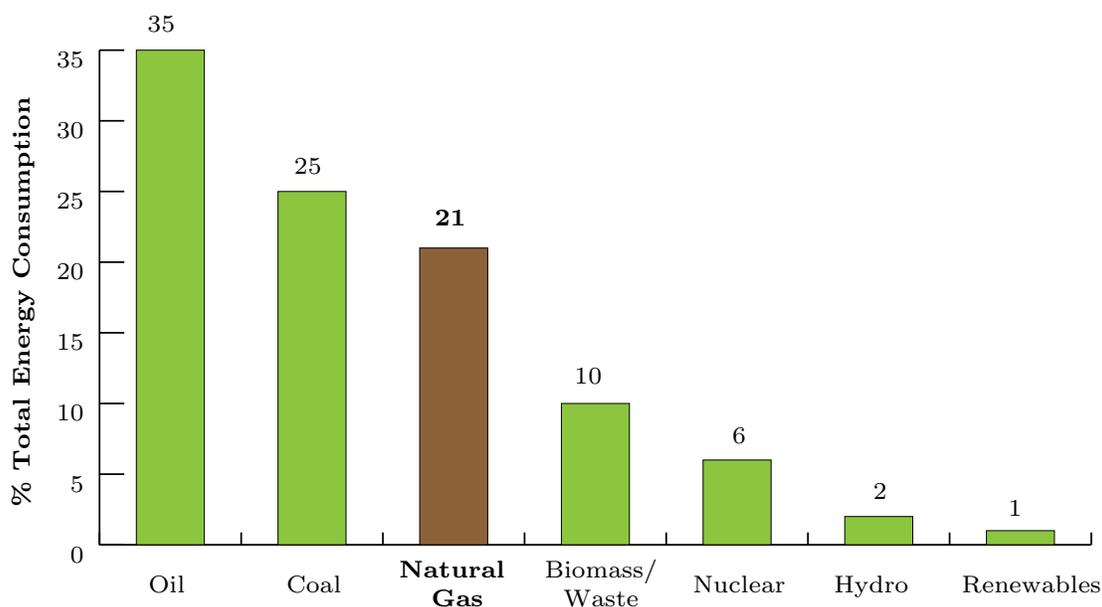
Figure 1-1: World energy usage by source in 2005 (based on the figures from *World Energy Outlook 2007, IEA*)

pipelines. Therefore, in general, a three-phase separation occurs at the well-head. The liquid organic phase is similar to light crude oil and is called *Natural Gas Liquids* (*NGL*) or simply *condensates*. When gas has a significant amount of condensate, it is sometimes referred to as *wet gas* as opposed to *dry gas*. Gas produced from such reservoirs is described as *non-associated gas*. Occasionally dry gas reservoirs may be referred to as gas only reservoirs to distinguished them from reservoirs producing wet gas also called condensate reservoirs.

Natural gas can also be produced from oil fields as a by-product, in which case it is termed *associated gas*. There are also unconventional sources of natural gas. These include tight gas (gas trapped in low permeability formations), shale, coalbed methane (gas adsorbed on coal, dissolved in water or stored in fractures and voids in coalbeds), natural gas hydrates and deep gas (gas found in very deep reservoirs). Natural gas can be classified as *sweet* and *sour*. Sour gas contains a relatively large amount of H_2S and CO_2 (though in a strict context, *sour* may exclusively refer to high levels of H_2S) while gas from *sweet* fields is mostly free of these contaminants.

Table 1.1: Natural gas: Proven reserves[a,b]

Country	tcf	tcm	% share
Russian Federation	1576.75	44.65	25.2
Iran	981.75	27.80	15.7
Qatar	904.06	25.60	14.4
Saudi Arabia	253.03	7.17	4.0
United Arab Emirates	215.07	6.09	3.4
US	211.08	5.98	3.4
Rest of the World	2121.60	60.08	33.9
World total	6263.34	177.36	100.0
Top 20 countries	5593.97	158.40	89.3

[a] BP Statistical Review of World Energy, BP, June 2008.
[b] End 2007.

1.1.1 Reserves

Proven natural gas reserves in the world currently stand at around 170-180 trillion cubic meters (tcm) i.e., 6,000-6,300 trillion cubic feet (tcf) [3, 4]. Overall reserves to production ratio for the world is 60.3 years [4]. The U.S. Geological Survey estimates that around 4,136 tcf of natural gas remains undiscovered [3]. Within the total resource base, 3,000 tcf is in reserves that are too far away from markets, also known as *stranded reserves* [3].

The distribution of proven reserves of natural gas is quite skewed as shown in Table 1.1 and Figure 1-2. The top 3 countries, Russia, Iran and Qatar, own close to 55% of the global proven reserves of natural gas. Moreover, the top 20 countries own around 90% of the total reserves. The biggest present and future consumers of gas, big economies in North America, Europe and Asia, do not have much proven reserves. The projected production from different regions is shown in Figure 1-3 and it is evident that much of the growth in production will take place outside the OECD (Organization for Economic Development, a group of mostly developed economies) countries. The major consumers of gas will have to rely on gas imports increasingly. The stark inequity of distribution raises concerns about the security of natural gas supply due to increasing intervention of national governments in the development and operation of gas projects.
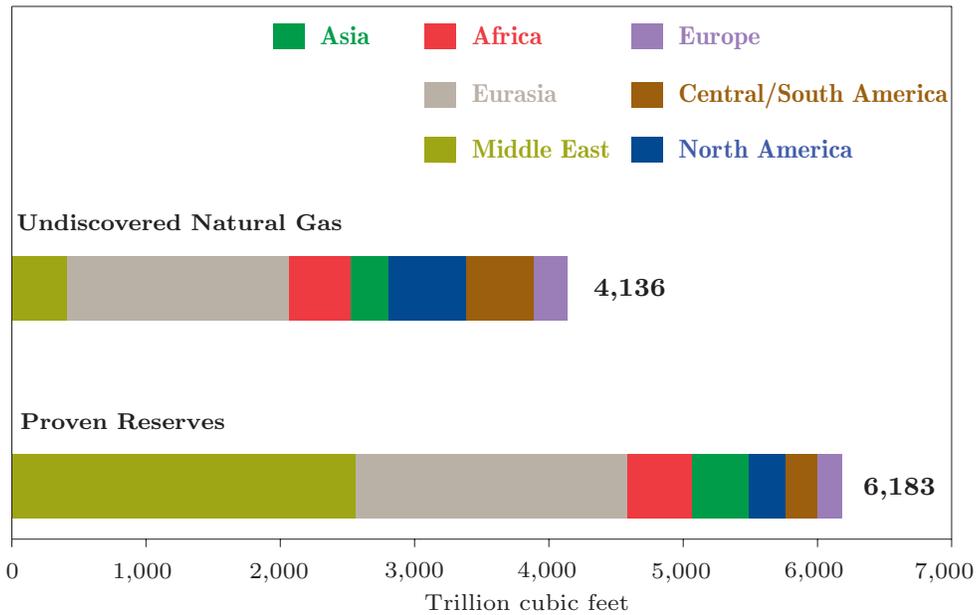
Figure 1-2: Natural gas reserves: Distribution by geographic region 2006-2025 (modified and adapted from the original in *International Energy Outlook 2007, EIA [3]*)
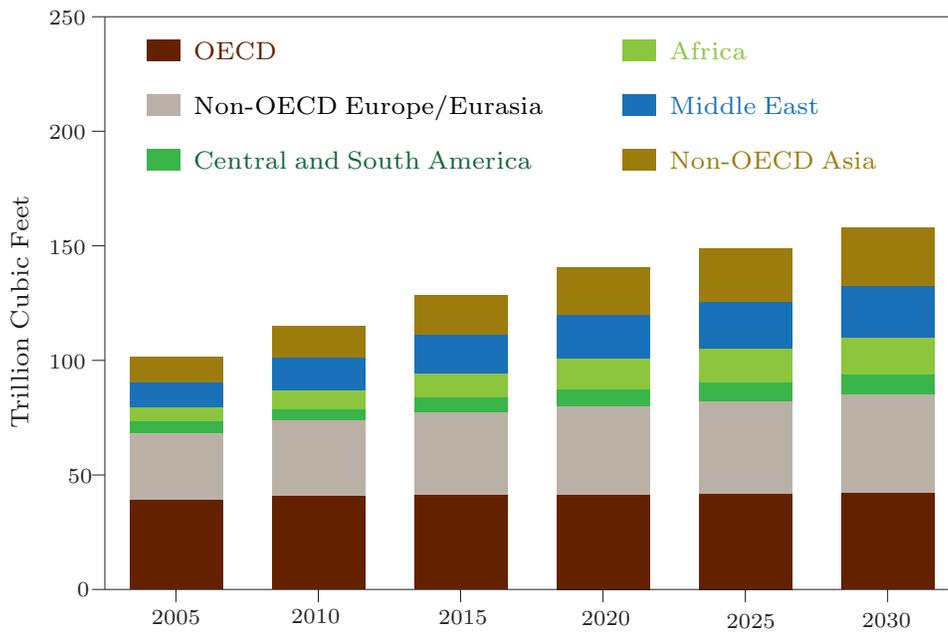
Figure 1-3: Natural gas production by region 2006-2030 (modified and adapted from the original in *International Energy Outlook 2008 (Highlights), 2008, EIA [2]*)

1.1.2 Demand

Natural gas is a less carbon-intense fuel than either oil or coal, i.e., its combustion produces less greenhouse emissions per unit energy produced. Moreover, it produces relatively lower sulfur, NOx and particulates emissions on combustion compared to other fossil fuels. It is therefore the cleanest fossil fuel and is expected to play an important role in the transition to alternative clean energy resources. Currently, industrial uses and power generation are the major consumers of natural gas. A rough division for global use of natural gas is shown in Figure 1-4.

Global natural gas demand in 2007 was 2,921.9 billion cubic meters (bcm) [4]. The global natural gas demand is expected to rise from 2,854 bcm in 2005 to 4,779 bcm in 2030 in the IEA 2007 reference scenario [1], rising at a rate of 2.1% annually.

There are three main uses of natural gas based on the sectors:

1. Residential and commercial users use natural gas for heating space and water, and cooking. Residential and commercial demand is strongly dependent on weather. Also these users are mostly *captive*, i.e., they cannot change their usage patterns easily and therefore, residential demands take time to adjust in face of a price change or a supply shock.

2. Industrial users may use gas as feedstock, for in-house power generation on a small-scale and as a heat source. Their demand is fairly predicable and stable. Industrial users can easily substitute another fuel for natural gas during periods of price increase and shortages.

3. Natural gas use for power generation is rising rapidly in the OECD countries with a doubling of gas use for power in the past fifteen years [5]. Gas-fired power increasingly meets peak summer demand in several countries. In Europe, almost two-thirds, and in North America, half, of new electricity plants are based on natural gas [5]. Most new plants use the combined-cycle gas turbine (CCGT) technology. These are favored economically because they are highly efficient and require less capital investment. The efficiency of CCGT plants can be as high
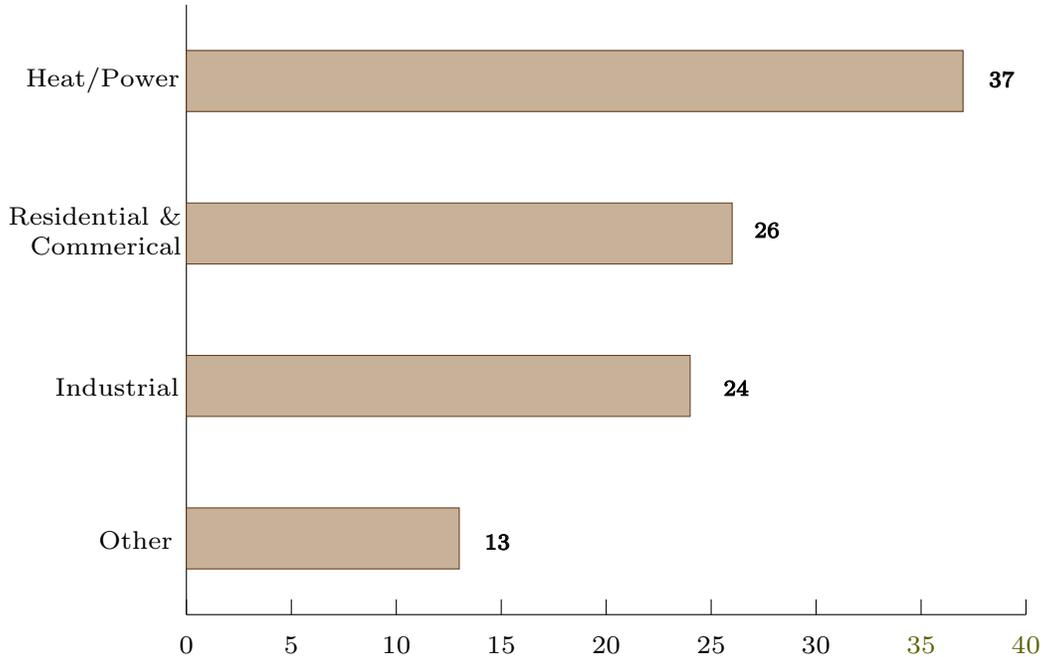
Figure 1-4: Global natural gas demand by sector (based on figures from *IEA Natural Gas Market Review, 2006 [6]*)

as 60%, the highest for any thermal power plant [5]. Increasing environmental concerns favor natural gas as the fuel of choice for power generation due to gas being the cleanest of all fossil fuels. Gas-fired plants are also being preferred in the developed world because coal-fired generation is being held up by policy uncertainly on carbon emissions. Finally, some renewable technologies such as wind-powered generation may actually favor intermittent gas-fired generation that can be quickly ramped up to fill in the supply-demand gap when the primary renewable source is not able to fulfil the demand [5].

Gas-to-liquids (GTL) is another potentially promising area for exploiting stranded reserves. GTL involves converting natural gas into liquid transportation fuels at source. However, there has been a rapid rise in GTL project costs that have dampened the development of projects [5]. Also, GTL projects are in competition with Liquefied Natural Gas (LNG) projects for the feed gas. LNG projects also draw investments away from GTL projects. Therefore, with rising feed gas prices and increasing trade in LNG, the viability of GTL projects is uncertain at this stage [5].

1.1.3 Natural Gas Supply Chain

Natural gas is conventionally produced from gas-only fields (non-associated gas) or as a by-product from oil fields (associated gas). A three-phase separation at wellheads removes water and natural gas liquids (NGL). Removal of water is essential to avoid the formation of hydrates (usually methane hydrates) that can potentially choke trunklines. NGL can be transported in the same trunklines as natural gas. If gas is sweet, it requires no further processing (it may be still blended to match heating value specifications). If gas has a high H_2S and (or) CO_2 content, it needs additional processing to remove them.

If there are regional or national markets close to the fields, gas is fed into relatively short-distance transportation networks (usually on the order of few hundred kilometers) that carry it to markets or big consumers. In the case, fields are faraway from markets, inter-regional transportation (over several thousand kilometers) is required to deliver gas to markets. Transportation as liquefied natural gas (LNG) involves specialized infrastructure to liquefy, transport and regasify natural gas. LNG may also be preferred as a mode of transportation over short distances under special circumstances. Inter-regional transport is also possible using long-distance trunklines. These have additional complications over the national networks such as crossing of multiple international borders. Large natural gas consumer facilities whose products are easier to transport, e.g., petrochemicals or gas-to-liquids (GTL) plants, may also choose to locate near remote fields.

Regional networks are linked to local distribution networks that operate on much lower pressure and supply residential and small commercial customers. Industrial consumers (e.g., power plants, petrochemicals plants, fertilizer plants) may be directly supplied by national or regional networks. Due to the fluctuations in natural gas demand right from an hourly basis to seasonal variations, regional networks are generally tied into short-, medium- and long-term storage facilities to match supply and demand.

An abstraction of the natural gas supply chain is presented in Figure 1-5.
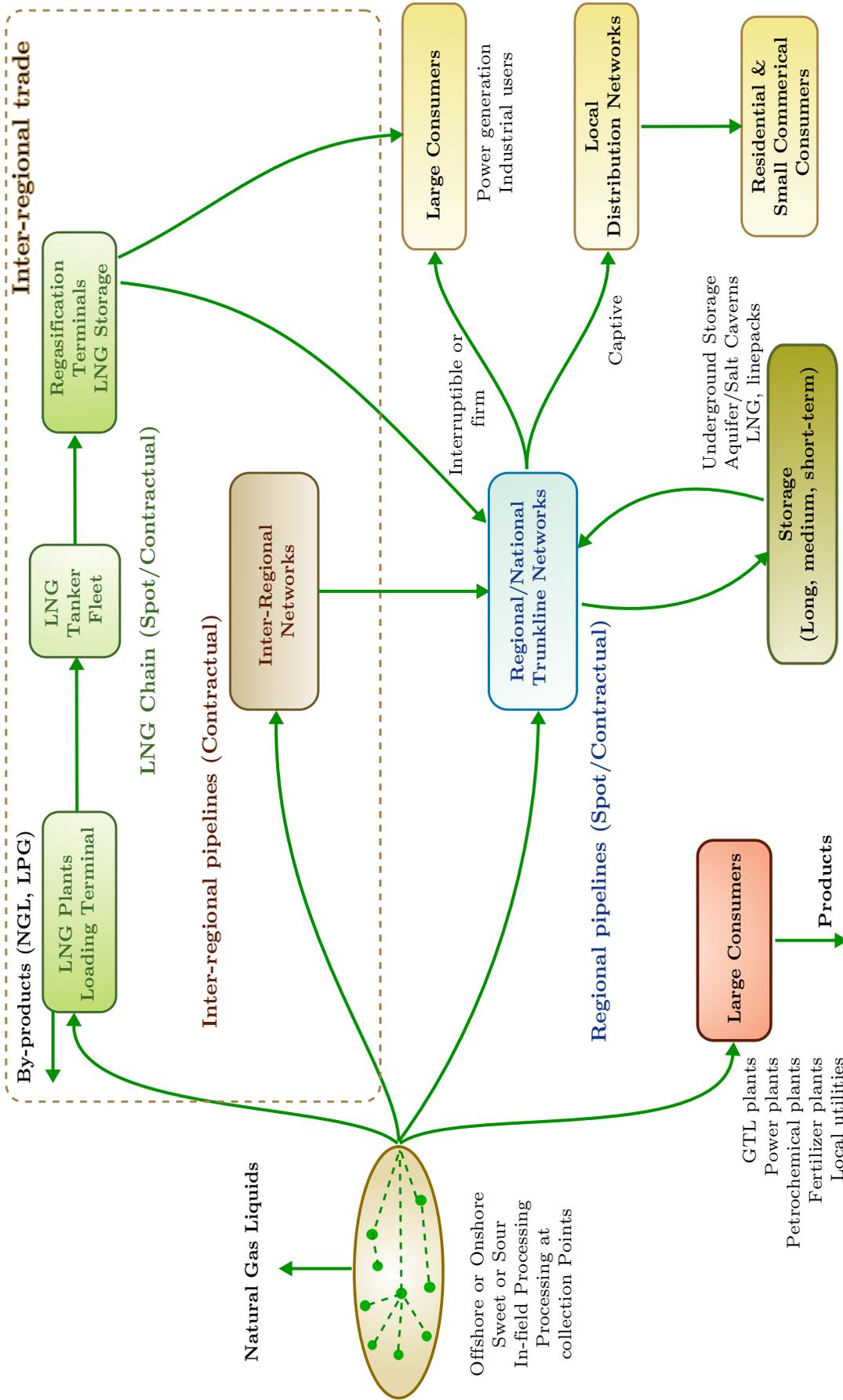
Figure 1-5: Gas supply chain overview

Inter-regional trade

LNG Chain (Spot/Contractual)

Inter-regional pipelines (Contractual)

Regional pipelines (Spot/Contractual)

By-products (NGL, LPG)

LNG Plants
Loading Terminal

LNG
Tanker
Fleet

Regasification
Terminals
LNG Storage

Large Consumers

Power generation
Industrial users

Local
Distribution Networks

Captive

Residential &
Small Commerical
Consumers

Interruptible or
firm

Inter-Regional
Networks

Regional/National
Trunkline Networks

Underground Storage
Aquifer/Salt Caverns
LNG, linepacks

Storage
(Long, medium, short-term)

Large Consumers

Products

GTL plants
Power plants
Petrochemical plants
Fertilizer plants
Local utilities

Natural Gas Liquids

Offshore or Onshore
Sweet or Sour
In-field Processing
Processing at
collection Points

29

1.1.4 Liquefied Natural Gas (LNG)

Liquefied natural gas (LNG) is natural gas (usually processed to around 90+% methane) that has been liquefied at near atmospheric pressure and cryogenic temperature (around 110 K). The liquefaction reduces the volume by roughly 600 times. This liquid is then carried by special ships expressly designed for the purpose, usually called *LNG tankers* to the markets. LNG tankers unload the liquid at *LNG terminals* (also called *regasification terminals*). LNG is converted back to gas at these terminals. Gas is fed into downstream transportation networks to take it to markets or supplied directly to bulk consumers. LNG regasification terminals may also have storage facilities for LNG.

The entire LNG chain is arguably more complicated than normal pipeline transmission and consists of the following:

1. The upstream natural gas production system.

2. Liquefaction plants that process raw gas and liquefy it.

3. The LNG shipping terminals.

4. The LNG tanker fleet.

5. Receiving terminals (often with storage) that are connected to pipeline networks.

Traditionally, gas producers have sold most of their production to regional consumers through pipeline networks. Most gas is still distributed in this fashion. The share of gas supply that is traded within major regions of the world was only 13% of the total gas supply in 2005 compared with roughly 48% of supply (2006 figures) for inter-regional oil trade [1]. Figure 1-6 shows that gas trade as well as LNG trade has been rapidly rising during the last decade.

LNG based natural gas transportation usually has a large fixed cost at liquefaction and regasification facilities but the dependence on distance is weak. On the other hand, the fixed cost of pipeline transportation is strongly dependent on distance.
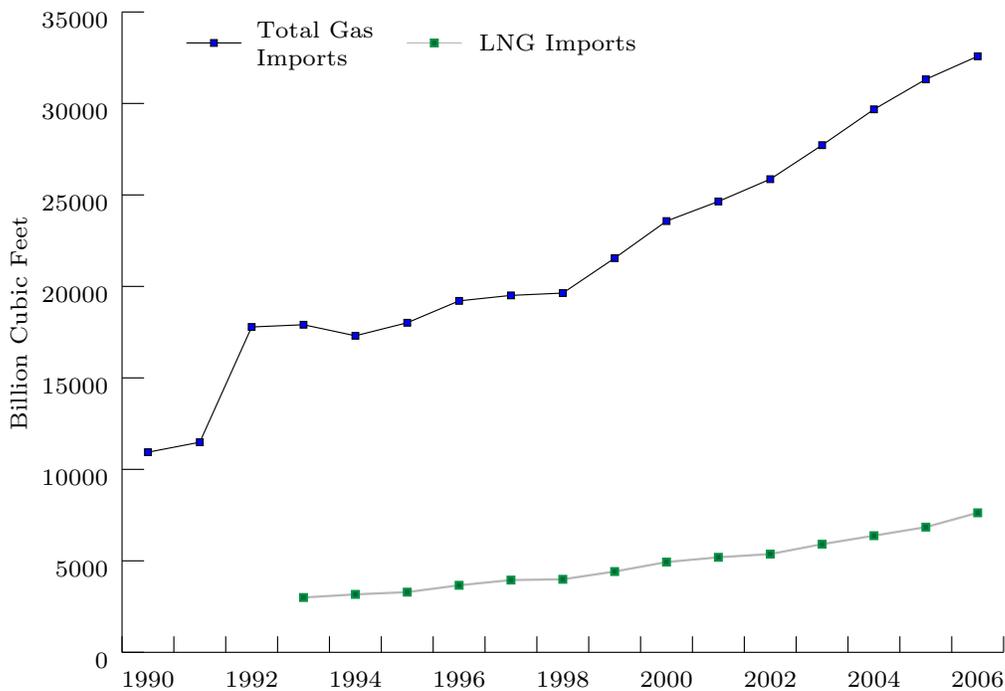
Figure 1-6: Global gas and LNG imports trend 1990-2006 (*Energy Information Administration [7]*)

Usually, a single train LNG project breaks even with a 42 inch pipeline for a 4,000 km onshore pipeline and just 2,000 km for an offshore pipeline [8]. LNG is also favored by several other factors. LNG does not suffer from the right of way issues such as the pipeline crossing multiple international frontiers, and therefore may be favored even if the economics are slightly against it. Since right of way is less of an issue for LNG, the geopolitical risk for LNG is much less than pipelines. Small reserves and fields are more favorable for exploitation as LNG because pipelines require a certain scale of transportation to be viable. LNG offers much more flexibility to producers than pipelines because pipelines (especially inter-regional pipelines) tend to lock producers into markets while LNG can be diverted easily to other markets. Gas-fired plants are also expected to increase LNG demand as it is convenient to build (and supply) these plants near a LNG terminal in densely-populated coastal areas (and therefore, near electricity markets). In certain geographical settings such as North America, LNG is almost the only option for gas imports. LNG is also favored by inadequacy of the

pipeline network in a country when it is easier to build a LNG terminal feeding part of the deficient network, instead of building an interconnection to transfer gas from another part of the network. A further discussion of some of these issues can be found in the 2004 IEA report on LNG [8].

Recent years have seen the rise of a global gas market with rapid construction of LNG plants, shipping facilities and regasification terminals. The IEA estimates that global liquefaction and shipping capacity will double by the next decade. The global LNG market continues to grow with a projected increase from 189 bcm in 2005 to 393 bcm in 2015 and 758 bcm in 2030 in the IEA reference scenario [1]. More optimistic estimates predict that global LNG capacity will rise from 240 bcm (2005) to 360 bcm in 2010 and 500-600 bcm by 2015. By 2015, LNG is expected to fulfil 14-16% of global gas demand [5] and a quarter of gas demand in OECD countries [8].

A global LNG market still continues to evolve. Innovative contractual agreements for LNG are still being worked out and experimented with in various parts of the world. A further discussion can be found in Section 1.2.6. Eventually these developments will lead to markets that are flexible and resilient with supply backups from LNG spot markets.

1.2   Issues in the Natural Gas Supply Chain

Natural gas, being a gas, suffers from low volumetric energy density (energy per unit volume). As a result, it is inherently difficult to store and transport compared to oil for a fixed amount of energy. The issues outlined here are all inter-related and related in some way to the fact that natural gas is a gas and, therefore, is difficult to store and transport. Oil and natural gas are occasionally referred together in contexts such as discussions about fossil fuels or energy economics, but they are quite different from all other standpoints: production, transportation, storage and consumption. Again, the main difference stems from the obvious fact that crude oil and its end-products (e.g., gasoline and diesel) are liquids while natural gas is handled and consumed as a gas in most uses.

1.2.1 Single Entity Operation of Upstream Systems

Pressures in reservoirs, wells and upstream networks play an important role in natural gas production. A substantial amount of investment over the lifetime of upstream networks is devoted to maintaining flow in the network. Due to this feature, an upstream natural gas system needs to be managed by a single operator so that all the fields are managed in coordination with each other. This coordinated operation and management makes sure that the entire collection (gathering) network (either surface or subsea) keeps flowing smoothly. Collection networks (and therefore entire systems) can be adversely impacted by inefficient operation in one part of the system. For example, failure to remove all water from the gas produced by a field (e.g., due to a malfunction) can lead to choking of pipelines due to gas hydrates somewhere downstream. Similarly, uncontrolled production from high pressure wells can choke low pressure wells. Investments are required over time to install compressors to produce from low pressure reservoirs and build new interconnections in the network to maintain flow. As new reservoirs are developed, investments may be needed to install new platforms and pipelines, and connect them to the existing network at appropriate locations such that they have minimal effect on existing production. All this requires central management of the entire upstream network.

The central operation of upstream gas networks is a technical necessity (and not just for reasons of economies of scale as may be true for oil networks) because the upstream pipeline network couples the entire system (i.e., multiple reservoirs and consumer facilities) and is able to transmit disturbances from one part to the other. Therefore, to have effective control over the system, an operator needs to have control over the entire network and all elements connected to it. This means that even when multiple stakeholders are involved, efficient practices will gravitate towards centralizing the control and operation of an upstream system. Recently, with sophisticated online sensors and actuators, and real-time data acquisition and processing, it has become even more beneficial, efficient and easier to operate large upstream systems centrally. Finally, gas production is technically intensive and, therefore, needs sophis-

33

tication and experience which are only available to large operators (mostly international oil companies) that naturally leads to centrally-controlled systems even when operators do not entirely own the system.

Some of the above arguments may also apply to downstream gas networks. However, there is an important difference which distinguishes them from upstream networks. Downstream networks can be upgraded by installing additional inter-connections, bypasses and compression facilities to decouple them from each other if it makes economic sense to do so. On the other hand, upstream networks are connected to gas wells, the pressures in which are governed by reservoir dynamics. Moreover, upstream networks are by nature, collection networks, that deliver gas to certain demand nodes and hence have a (converging to demand) topology that favors integration and not decoupling. The network is expected to become cheaper (per unit of gas transported) downstream as economies of scale for pipelines come into play. Besides, upstream networks are likely to be offshore or in other difficult locations where the capital cost of any additional facility may be too high. All these factors make decoupling parts of the upstream system less favorable economically and make central operation a more attractive proposition.

1.2.2 Issues in Production and Transportation Infrastructure

Gas production and transportation has some unique issues associated with it that distinguish it from oil infrastructure.

Capital-intensity

Production and transportation infrastructure for natural gas requires a large capital investment. Upstream investments in gas account for 56% of total gas sector expenditure [5]. Most upstream investment goes towards the development of new and existing fields.

Pipeline transportation of gas requires an extensive pipeline network and compression stations at regular intervals. Gas pipelines are significantly more expensive than

their oil counterparts because they have a larger diameter and operate at much higher pressure (e.g., 70-100 bar for cross-country pipelines). Long-distance gas pipelines can cost up to US $1-2 billion for every 1,000 km [5] depending on the routing and terrain. With the rising price of steel, they are expected to become even more expensive in the near future.

Investments in liquefaction facilities, LNG tankers and regasification terminals are required for transportation as LNG. A typical LNG train produces 3-4 million ton per annum (mtpa)[1] of LNG and can cost up to US $1 billion as per 2004 estimates [8]. LNG has enormous economies of scale and therefore, the size of trains continues to increase. New trains being set up in Qatar are in 5-7.5 mtpa range. It is estimated that costs for a new (so called greenfield) LNG project are in the range US $3-5 billion for liquefaction, US $2 billion for shipping and US $0.8-1 billion for the regasification terminal [5]. For the same amount of energy carried, a large LNG carrier costs up to four to five times more than a large oil tanker [5]. 2003-2004 estimates for a 135,000-140,000 cubic meter LNG carrier are around US $160-170 million.

Overall some estimates claim that gas costs up to 10 times more to store and transport than oil [6].

Specificity

Natural gas infrastructure is specific and designed to strict capacity, pressure and composition specifications for processing and transporting natural gas from a particular set of fields. It is therefore specific to the natural gas it can transport. This means that it cannot be put to any other use (i.e., for carrying natural gas of a different specification) without extensive modification of facilities. In the case of downstream natural gas networks, even the direction can be specific and the capacity of the network may be asymmetrical with respect to the designated (during design) direction of flow and may require additional investment in compressors, inter-connectors and by-passes, to make it symmetrical. The specificity of natural gas infrastructure raises the risk of natural gas projects to investors. If the economics of a project are adversely

---

[1]A metric ton of LNG is approximately 2.47 cubic meters

affected, almost none of the associated infrastructure can be put to another use to at least partially recover capital costs. For example, an inter-regional pipeline, linking a source to a market, once constructed, has the express purpose to transport gas at a specified quality from the source to the market. Inter-regional gas pipelines result in a long-term tie-up between the gas producer and a certain market. If, due to some external factors, demand in the market collapses or gas at source is not available, the entire rationale of the project collapses.

This specificity is in contrast to other infrastructure such as oil product pipelines, tankers and storage tanks that can handle a much wider range of products. Oil can be easily transported using conventional infrastructure such as roads or railways with relatively modest additional capital costs (in road or rail tankers). This is usually not possible without large investments in additional facilities (e.g., compression, liquefaction, insulated or high pressure tankers) for natural gas. The specificity of natural gas infrastructure can be seen at its extreme when comparing a LNG plant with a crude oil refinery. Both involve a similar level of investment and impact. However, a LNG plant is exclusively designed to process gas from a particular set of fields and its returns can be significantly impacted if the supply of gas cannot be maintained to required level or gas quality is significantly different from the plant design or enough LNG shipping capacity is not available. On the other hand, a crude oil refinery can handle a much wider range of crude qualities, can rely on the global oil market to keep functioning and can transport products using conventional infrastructure.

Large Footprint

Natural gas facilities have large footprints; inter-regional gas pipelines occupy a strip of land for several hundreds or thousands kilometers, production systems may be spread over hundreds of square kilometers with several hundred kilometers of upstream collection network, liquefaction and regasification facilities require large pieces of land in densely-populated coastal areas and so on. The large footprint of natural gas facilities in turn requires administrative and environmental clearances from various entities. Inter-regional natural gas pipelines may cross multiple international

or ethnic frontiers and suffer from right of way permissions and transit fee issues. There are significant geopolitical risks associated with international pipeline projects with involvement of multiple national governments with different strategic interests. Finally, large liquefaction and regasification facilities as well as pipelines can have significant environmental impact.

1.2.3 Role of Gas Quality

Oil is refined to convert it into final products. Therefore, refining and blending stages for oil products can usually take care of specifications. Crude oil quality only affects refineries' operations, and they are designed to accept a range of crude qualities and can adapt to produce a more or less consistent product quality. On the other hand, natural gas is only minimally processed before being ready for use. This means that quality specifications on final consumer-grade natural gas tend to impact the operation of the upstream system. Raw gas quality can vary over extreme ranges. Some of the difficult reservoirs being considered for production currently can have $20+\%$ CO_2 content that only a specialized infrastructure can handle. The same holds for high H_2S reservoirs.

Gas Quality and Global Gas Trade

Quality also potentially hinders global gas trade [6]. Although, gas sourced from different sources have different compositions, the consumer equipment is designed and optimized to accept a narrow range of gas quality for safety and efficiency reasons. This quality may have historically been chosen based on the gas quality available in the region. However, with global LNG trade, the quality delivered may not match with consumer equipment and can create serious safety, environmental and economic problems. For example [8], the bulk of LNG traded in the world has a higher heating value and is richer in heavier hydrocarbons than specified by some North American natural gas pipeline operators that require lean gas for transportation. Similarly, California has strict limits on composition as a result of which the bulk of traded

LNG does not qualify. In the UK [8], several LNG imports have a Wobbe Index (ratio of higher heating value of gas to square root of its specific gravity, roughly the amount of heat released by a gas burner at constant pressure) that exceeds the current regulations. Higher Wobbe Index gas requires more oxygen than lower index gas to burn. Gas appliances designed for low index gas may not be able to supply sufficient oxygen for complete combustion of high index gas. This creates safety and environmental issues such as flame lifting, back firing, excess CO and NOx emission and increased sooting.

The solution to this lies in the processing or blending of natural gas or the modification of consumer equipment, all of which have an economic impact [6]. For example, ethane can be stripped from rich LNG; however, ethane does not have direct uses [8]. At least one US terminal has an ethane stripping plant that can handle rich LNG. Another interesting proposal is to combine LNG terminals with LPG and power generation facilities that consume stripped heavier hydrocarbons. Blending with leaner gas at a central location is another option. Finally, equipment modification is another solution, though it suffers from the logistical difficulty of carrying out an upgrade of all consumer equipment, besides its economic costs.

The quality of LNG is also a hindrance to spot trading of LNG for exactly the same set of reasons.

1.2.4 Seasonal Nature of Demand

A big share of natural gas is used for residential and commercial heating purposes, as well as for electricity generation. The demand is therefore strongly coupled with seasonal patterns, especially in the developed world. In cold regions, demand is low in summer months and high in winters when heat is required. Even during a single winter, demands can spike within a region when a cold spell hits. In regions with gas-fired electricity generation and hot summers, the demand can peak during summer months when electricity demand for cooling is high.

The industry relies on long-term storage to offset seasonal fluctuations in demand. However, even if gas is available in the storage, it must be able to be transported to

the consumer. During periods of peak demand, the required transportation capacity may not be available as discussed later in Section 1.2.5. Seasonal demand fluctuations also have a regulatory dimension for local distribution companies (LDC) and utilities. Especially in places which have colder winters, the loss of residential heating may be life-threatening and regulators demand strict guarantees from LDCs that such a situation does not occur. LDCs in turn have contractual arrangements (briefly discussed later in Section 1.2.6) with bulk consumers to make sure supplies will be available.

Again, these issues arise because any excess amount above the rated capacity of gas storage and transportation infrastructure is expensive and not even technically feasible at times. Gas infrastructure is inherently inflexible to handle spikes in demand. Moreover, it is uneconomical to over-design the infrastructure for the worst case scenarios (e.g., once in fifty years winter) and it will be done only if there is a regulatory requirement to do so.

1.2.5   Issues in Storage

Natural gas cannot be easily stored due to its low volumetric energy density. Hence to store a given amount of energy, a large volume is required unless the gas is stored at a high pressure or is liquefied. The nature of storage for natural gas differs depending on the time-scales of the storage as follows:

1. Short-term storage of natural gas is on the order of hours to satisfy peak demand during the day. *Linepack storage* is achieved by raising the pressure in the pipeline system to enable it to store extra gas that can relieve peak demand on the order of hours. LNG has also been used on a small-scale for satisfying peak demand in some natural gas networks where liquefaction is done during off-peak hours and LNG so produced is regasified during peak hours.

2. Medium-term storage is on the order of several days to several weeks. Salt caverns are underground cavities where salt has been dissolved by water and gas can be stored in these cavities if compressed. Use of rock caverns for storage has

39

also been investigated in recent years. Since these are essentially like compressed gas storage tanks, they provide very fast injection and withdrawal profiles. LNG has become a promising and viable option for storage with the rise in global LNG trade and storage facilities are being made available at regasification terminals. Since LNG terminals are usually near markets (i.e., near populated coastal areas), they have additional advantages from a transportation standpoint as transportation bottlenecks and disruptions over a short distance are less likely.

3. Long-term storage facilities serve to smooth out the seasonal variations of natural gas demand. These therefore operate on an yearly cycle. These are mostly geological formations such as depleted fields and large aquifers. Gas is compressed and injected into formations during a low demand period. The formations can be produced as a gas field during peak demand months. However, suitable formations may not be available near all markets or may not have the requisite capacity. The other issue with these facilities is that they have specific injection and withdrawal profiles that can put rate limits on drawing from storage.

It is important to note again that each of the storage options above involve compression or liquefaction and incurs a cost. Underground storage options also suffer from a cost in terms of lost gas. Hence, gas storage incurs an operational cost per unit of energy stored as opposed to oil storage which incurs little or no such costs. As indicated earlier, some estimates give this cost to be 10 times as much as oil. Not only that, except for short-term storage, most storage development is much more capital-intensive than oil storage. For example, developing underground storage involves not just investment in compression and transportation facilities, but also requires an expenditure in *cushion gas* to bring up the storage to working pressure which can be significant for large storage capacity. LNG storage can contribute up to 40-50% of costs in a regasification terminal [8].

Availability of Transportation

Delivering gas supplies during periods of shortages or peak demands usually involves two steps:

1. Ascertaining the availability and maximum rate of withdrawal from the storage.

2. Ascertaining whether gas can actually be delivered to the market.

The second step may not be possible during peak demands for a particular market, mainly because gas transportation relies heavily on pipeline networks for delivery and pipelines have strict capacity limits. This is in contrast with the situation in oil markets where (unless there is a nationwide crisis) additional road and rail oil tankers may be pressed into service to reinforce supplies. Additionally, as opposed to rail or road transportation, a lot of natural gas pipeline infrastructure favors unidirectional flow.

The difficulty with storage and transportation of natural gas makes it challenging to deal with supply shocks and market disruptions for natural gas. Oil storage pools, e.g., the U.S. strategic reserve, can help to absorb temporary supply shocks (such as weather or political events) in oil markets. Based on the difficulties outlined above, a comparable (in terms of energy content) strategic reserve for gas would be prohibitively expensive to commission and operate, and would not be as effective as an oil storage pool because of withdrawal and transportation difficulties.

1.2.6   Contractual Framework

Contractual frameworks exist almost in the entire gas supply chain from production to the final consumer. It is important to note here that contractual frameworks are an integral part of the gas supply chain resulting from the characteristics of natural gas and the associated infrastructure. They do not exist as add-ons in the gas business, i.e., gas markets would not function without them. In fact, it can be argued that contracts are responsible for the rise of natural gas as a viable fuel. Gas infrastructure would not have developed without the supply and demand guarantees

that contractual frameworks provide. Contractual frameworks also play a central role in the operation of gas infrastructure because production, transportation and storage of natural gas have to be in line with the contractual rules.

Contractual frameworks in natural gas systems mostly result from the factors that have been discussed so far:

1. Capital-intensity and specificity of gas infrastructure for each part of the chain pose a significant risk for investors, producers, transportation operators, consumers and so on. They reduce their risk by obtaining guarantees from each other. These guarantees are formally instated as contractual agreements.

2. High capital-intensity and risk also imply that there are multiple stakeholders in the entire value chain and hence there is a need for formal agreements.

3. The absence of a liquid gas market increases the risk for producers and investors by making the sales of gas uncertain. Similarly, there is a risk for buyers as to whether they will be able to source the required amount from the markets. Therefore, there is both supply-security (from the buyers' perspective) and demand-security (from the producers' perspective) rationale for contracts.

4. Imperfect markets also make discovering the market price of gas difficult. Therefore, formal arrangements are required to set the gas price. This imperfect and distorted market is a direct result of difficulties with transportation and storage.

Contracts for exploration and production exist as licenses granted by the state (or the owner of the land, depending upon the laws in a particular country) either for exploration or production or both. The license may be provided to producers for a specified period of production. These contracts are usually long-term. A producer will pay a royalty to the licensing-granting entity in exchange for the right to produce. The royalty may be flat (a fixed share of production) or may be determined in a complex way.

In some upstream systems, multiple firms have stakes in the production system and a single operator (which is usually a firm that holds a stake in the upstream

system) is responsible for day to day operations. In such systems, there is a need for formal agreements (that govern sharing of products or revenue and the agreed operational strategy) between stakeholders and the operator. These are usually called *production-sharing contracts*. Upstream producers (or in the case described here, upstream operators and stakeholders) in turn may be involved in contractual arrangements with big consumers such as LNG and GTL plants or pipeline exporters for a guaranteed offtake of gas.

International gas sales have traditionally been governed by long-term contracts between producers and consumers with periods ranging from 15-30 years. The length of contracts depends on the economics of the project, with a tighter and more uncertain economics leading to longer-term contracts. Contracts may usually define not just the amount, but also gas quality. Most imports to the developed economies take place through *Take-or-Pay* (ToP) contracts between buyers and seller [8]. These contracts have played a major role in development and investment in the industry. Under ToP contracts, producers are obligated to supply a predefined volume at a mutually agreed price subject to revision and buyers are obligated to buy the predefined volume. Under this arrangement, producers take a price risk while buyers take a marketing risk. Pricing arrangements in such long-term contracts are based on the *replacement value principle* [8], i.e., the gas price is tied to price (value) of a substitute (replacement) fuel (e.g., heavy fuel oil for gas-fired power stations and industrial uses). The replacement value principle is not simple when capital costs are taken into account. For example [8], if the price of gas for power generation is inferred from coal as an alternative, one comes up with an artificially low price for gas due to the fact that long-run marginal costs are not considered which are lower for coal. If the fact that gas-fired plants are efficient and cheaper to build (per unit generation capacity) is taken into account, the price inferred for gas is high and the marginal cost of generation from gas-fired plants becomes uncompetitive with coal-fired generation. The price of gas in long-term contracts may also be linked with prices in a domestic liquid gas market if one exists, such as in the US or the UK. A further discussion of contractual arrangements in LNG can be found in [8].

Contracts also exist in the transportation chain. Investors in a large-scale pipeline project may need guarantees that enough transportation capacity can be sold in the market. LNG producers may have agreements with LNG shipping operators. Similarly, owners of a large LNG receiving terminal may need guarantees that gas can be transported to markets from the terminal. Again the length of a contract depends on economics and the risk of a project.

Finally, contracts also exist between downstream suppliers and bulk consumers. These contracts may be *interruptible* (as opposed to *firm* or *uninterruptible* contracts when the supply is guaranteed), i.e., under peak demand the supply can be disrupted. This is especially true for utilities where there may be a regulatory requirement to supply residential and small commercial customers.

Although necessary, the contracts may further distort gas markets, that are already imperfect and make them inflexible and unresponsive to external signals (e.g., changes in technology, resource scarcity and so on) due to long-term lock-ins.

Evolution of Contracts

With the rise of LNG trade and infrastructure, and the continued trend of liberalization of markets in the developed world, the contractual framework continues to evolve rapidly. Specifically, the maximum length of contracts is being shortened from 20-30 years to 15-20 years [8]. Also, contracts are increasingly becoming flexible. For example, some long-term contracts now contain *swing provisions* that allow buyers to vary the volume up to a certain maximum limit if they desire. Some LNG contracts also allow the transfer of gas to spot markets at certain price thresholds. The volumes involved in the contracts are getting smaller. The rise of spot markets has further removed some inflexibility in the gas markets. Excess gas left over can be sold in the spot market that is profitable for sellers. It is now possible to link contract prices to open market spot prices. Therefore, price-indexation of LNG is moving away from the prices of oil products. Additionally, upstream operators are increasing acquiring interests in LNG terminals to ensure offtake of LNG.

As one of the first examples of the changes sweeping through the LNG contracts,

MLNG Tiga negotiated a contract with three Japanese utilities (Tokyo Gas, Toho Gas and Osaka Gas) in February 2002 with a buyer option to reduce contracted volumes and a portion of the contract to be supplied *Free-on-board* (i.e., the buyer is responsible for shipping from the exporting terminal as opposed to *ex-ship* arrangements) [8].

Nevertheless, it is expected that long-term contracts will continue to coexist with short-term arrangements in gas supply chains. Most experts agree that the LNG short-term trade (both physical and paper) will never reach the scale of the oil market [8]. Most optimistic estimates claim that the spot LNG trade will ultimately be around 15-30% of global LNG trade [8]. Even in a liberalized market, the bulk of the gas is still delivered under long-term contracts as of 2003-2004, e.g., in the US half of the wholesale gas and in the UK, 85% of gas was delivered under long-term contracts [8] (although, long-term in the US and the UK refers to a shorter period than other parts of the world, 8-10 years).

1.2.7   Issues in the Overall Supply Chain

All previously discussed, unique features of gas production, transportation and storage have major implications for the entire supply chain. It is clear that gas supply chains operation involves tightly coupled subsystems and they must be closely coordinated and aligned with each other for the proper functioning of the entire chain. This means that the weakest link can disrupt the entire supply chain severely. This is in contrast to the oil chain where it is possible for the system to recover from a disruption of a single link, e.g., a disruption in crude supply can be circumvented by drawing from a reserve or buying from oil spot markets, a disruption in refining can be substituted by importing oil products, a disruption in a pipeline can be alleviated by using road or rail and so on.

The fragility of gas chains has serious implications for the security of gas supply. With the decline of gas reserves in the developed world and the growth of emerging markets, the supply chains to markets will be stretched further in the near future and therefore will be more prone to disruptions due to natural or geopolitical events.

The rise of resource nationalism is a retrogressive development for gas supply chains as national oil companies do not have the resources or will to invest in infrastructure and technology that are crucial to a robust gas supply chain. The central role of technology, advanced project management practices and capital in gas chains also calls for expertise from international oil companies that have not been allowed to operate (freely) by several national governments.

1.3 Planning Overview

In general, decisions in the petroleum industry involve a multitude of technical, economic, regulatory, geopolitical and environmental factors and can be quite complex. Petroleum infrastructure involves large investments and the industry has large turnovers and volumes. Hence, even small fractional performance gains made in the design and operation of petroleum infrastructure can translate into significant increases in profits. It is therefore not surprising that the petroleum industry has been a pioneer in the use of systematic mathematical programming methodology for decision-making. For example, the earliest optimization models for blending problems in refineries date back to at least 1952 [9].

The domain of planning in the industry encompasses the entire supply chain: exploration and production, transportation, processing and distribution. Traditionally, the part of the industry involved with exploration and production is termed *upstream* while processing and distribution are termed *downstream*. Based on this classification, planning problems can be categorized as being upstream or downstream problems depending on their domain:

1. *Upstream problems* are concerned with exploration and production decisions in the fields. These problems may concern technical and economic decisions regarding field development, reservoir management, production infrastructure development and expansion, and production operations.

2. *Downstream problems* are concerned with decision-making in the transporta-

tion, processing and distribution of oil and natural gas. Examples of such problems include crude and end-products supply chains, refinery design and operation and natural gas transmission and distribution.

The decisions involved in a planning problem can vary with the time horizon over which the plan is made. Based on this, planning problems may be roughly distributed into three categories:

1. *Long-term planning problems* tend to have a time horizon of roughly over five years to over a decade (or even several decades). The key decisions are major investment decisions over the planning period. For example, long-term planning problems in natural gas systems may include decisions about the development of fields, offshore production platforms, pipeline infrastructure, processing facilities, compression stations, liquefaction plants and so on. The level of uncertainty is quite high in these problems. The sources of uncertainty can be technical, e.g., uncertainties in amount and quality of recoverable resources, or they can be economic, such as demands and market prices, or a combination of both, for example, total project costs.

2. *Medium-term plans* generally run from several months to several years. Examples of medium-term decisions may include reservoir management, major debottlenecking and expansion, and maintenance of facilities.

3. *Short-term planning problems* involve decisions over several weeks to several months. Due to the timescales involved they are usually concerned with operational decision-making. Decisions may be concerned with the state of the individual wells and fields, transportation of feed and final products, processing facilities operating states and delivery to consumers.

There is often a substantial overlap and interaction between different planning models, both between upstream and downstream models, and between short, medium and long-term models.

1.3.1 Planning in Natural Gas Systems

Planning in natural gas systems has additional issues due to the factors described in Section 1.2. In most cases, planning models for natural gas must include a technical model of the infrastructure and cannot be purely economic or commercial (an exception may be when infrastructure has excess capacity). This is because of the specificity and capacity limitations of production, transportation and storage infrastructure. Without a technical representation, the model will likely generate impossible scenarios. Accurate prediction of pressures and gas quality specifications usually play an important role in natural gas planning problems. Natural gas models need additional constraints to take contractual, regulatory and operational rules into account. Finally, due to the strong coupling of subsystems within gas chains, it is beneficial to model the entire chain with a single planning model and generate a coordinated management scenario if the mode of operation and control of the system permits to do so. Such planning models can reduce systemic risks intrinsic to gas chains by enabling operators to evaluate scenarios, identify weak links in the chain and evaluate potential remedies.

The next chapter explores in detail the specific issues and requirements arising from short-term planning in upstream natural gas systems.

1.4 Literature Review

As described earlier, the oil and natural gas industry has been a pioneer in the application of mathematical programming with work going back to at least 1952 [9]. Therefore, the body of literature on this topic is vast. There are multiple disciplines that have contributed to the area. These include operations research and management sciences, petroleum engineering, energy economics and chemical engineering. However, there has been little interaction and collaboration between these disciplines. Each community has formulated and solved models in its own area of interest and according to its own perspective. Petroleum engineers have focused on decision-making in upstream engineering, e.g., gas-lift calculations, flow calculations in wells and fa-

cilities, equipment scheduling and so on. Logistics, economic decision-making and market modeling in the downstream business have been within the realm of the operations research community and energy economists. The chemical engineering community has mainly focused on technical models (especially nonlinear models) for design and operations of upstream and downstream infrastructure. Due to the involvement of multiple disciplines, even the terminology across the field is not consistent. An exhaustive survey of the field is therefore very difficult.

The subset of works described here are in some way related to natural gas production, processing, transportation and distribution or have some relationship to the modeling approach described later. Oil production and downstream system modeling are not discussed here because most of the problems encountered and models used in oil systems do not directly apply to natural gas systems. The main exception to this is the class of refinery blending problems, so-called *pooling problems* [10, 11], the constraints corresponding to which are found in gas networks when quality is tracked across the entire network. There is a substantial literature on gas-lift optimization in oil fields that is not been not discussed here since the modeling approach and objectives of those problems differ widely from the actual gas problems.

Basic information about natural gas production can be found in the works by Katz and Lee [12], Ikoku [13] and Lyons and Plisga [14]. A discussion of underground storage can be found in Tek [15].

Dougherty [16] presents a review of works until 1970 from a petroleum engineering perspective that covers both oil and natural gas applications. Another survey of the work prior to 1977 can be found in Durrer and Slater [17]. Broadly, the work relevant to natural gas can be divided into the following topics:

1. Planning in natural gas systems, both from the supply chain perspective and the subsystem perspective.

2. Design, simulation and optimization of gas transmission systems.

3. Decision support models for local distribution companies (LDC) or utility companies to plan purchase, storage and transportation of natural gas.

4. Infrastructure development planning, both in oil and natural gas fields.

5. Some relevant models for oil fields.

1.4.1 Natural Gas Systems

There are only a few relevant and detailed works in this area due to the difficulty and size of optimization problems that result from modeling of entire subsystems. One of the earliest discussions of the factors involved in production planning in natural gas systems from a technical and economic perspective appears in Van Dam [18]. A qualitative discussion of the long-term planning system for gas production and processing operations owned by Santos in Australia is presented in Dougherty et al. [19]. A simple operational planning model of the gas network in a commercial field planning and optimization tool is optimized using SQP in Dutta-Roy et al. [20]. A superficial presentation of the tools used for the North Sea gas fields can be found in Mortimer [21]. Bitsindou and Kelkar [22] present a model for gas well production optimization by solving the network sequentially and fitting historical data. A detailed discussion of issues involved in the natural gas supply chain planning appears in Tomasgard et al. [23]. The Energy Information Administration of the US Department of Energy has developed over the years a demand, supply and transportation matching model for the North American natural gas market [24, 25]. Midthun [26] discusses issues in the optimization models along the entire natural gas value chain. Mason et al. [27] formulate a production planning problem for a natural gas system with nonlinear pressure-flowrate relationships and solve it with a derivative-free and a nonsmooth method.

1.4.2 Gas Transportation Networks

A substantial body of work exists on design and operational optimization of gas pipeline transmission systems, the main objective being to minimize the capital and operational costs for the network. Osiadacz [28] and Kraálik et al. [29] contain a discussions of methods for simulation and analysis of gas networks. Il'kaev et al. [30]

present detailed fluid dynamic simulation approach for gas pipelines. A dynamic programming based model for optimizing gas pipeline operations is presented in Peretti and Toth [31]. An optimal control perspective on the problem can be found in Marqués and Morari [32] and Osiadacz and Bell [33]. Furey [34] presents a modified successive quadratic programming (SQP) algorithm for optimizing natural gas pipeline networks. A bundle method is used to solve the pipeline design problem in De Wolf and Smeers [35]. In a later work by the same authors [36], an iterative method using piecewise linearizations of nonlinear functions and LP simplex is employed to solve the network problem. An optimal routing problem for natural gas transportation is presented in Dahl et al. [37]. A simulation model for natural gas pipeline systems is presented in Nimmanonda et al. [38]. A reduction method for networks is presented in Ríos-Mercado et al. [39]. Techniques for constructing piecewise linear approximations of the nonlinear functions involved in the gas transmission network and the properties and solution of the resulting mixed-integer linear program (MILP) have been discussed in Martin et al. [40]. A methodology based on dynamic programming for minimizing fuel consumption in gas networks that contain cycles is presented in Ríos-Mercado et al. [41]. Various numerical and mathematical aspects of cost minimization in gas transmission networks are discussed in [42–46]. Arsegianto et al. [47] present a simulation-based design of a gas transmission network. A discussion of capacity allocation in pipelines appears in Cremer et al. [48]. A model for planning investment in residential gas network is presented in Davidson et al. [49]. Issues in the maintenance of infrastructure networks is explored in Papadakis and Kleindorfer [50]. Abbaspour et al. [51] develop an optimization model for linepack operation of compressor stations and solve it with a sequential unconstrained minimization technique. Kabirian and Hemmati [52] present a nonlinear programming (NLP) model for design of natural gas transmission networks. A nonconvex NLP model for the design of a natural gas distribution network is presented and solved with Floudas' GOP based global optimization approach in Wu et al. [53].

1.4.3 Decision-making for Local Utilities

Substantial effort has been put into decision support models for making purchasing and storage decisions for local distribution companies (LDC) and gas transportation companies. An optimal schedule for withdrawal from storage reservoirs appears in Wattenbarger [54]. An allocation problem within a statewide trunkline network with users having different priorities is solved in O'Neill et al. [55] that introduces the idea of a pseudonetwork to model swaps between different pipeline systems. Levary and Dean [56] present a model for gas procurement by a natural gas utility. A linear programming (LP) framework to evaluate supply scenarios for planners on a statewide or national level is presented in Brooks [57]. A model that considers storage deliverability for managing natural gas purchases for a LDC is developed in Bopp et al. [58]. A chance constrained approach to making purchasing and storage decisions for a utility is presented in Guldmann [59]. In a work by the same author [60] a marginal-cost pricing model that includes gas supply, storage and transmission, for a gas utility is discussed. A LP model for determining utility decisions appears in Avery et al. [61]. A decision support model for natural gas dispatch is developed in Chin and Vollmann [62]. Guldmann and Wang [63] present a model for choosing the optimal mix of natural gas supply contracts for a LDC. A similar contract selection approach for a North American gas producer is presented in Haurie et al. [64]. Butler and Dyer [65] develop a multi-period LP model for natural gas purchase by an electric utility that considers purchasing, storage and usage. A model for a Chilean LDC with contracts is presented in Contesse et al. [66]. Recently, Gabriel et al. [67] present a mixed nonlinear complementarity model of natural gas markets. In another work, the same authors present a stochastic equilibrium model for deregulated natural gas markets [68] A complementarity model for the European gas market appears in Egging et al. [69]. A combined upstream and downstream market model for Europe is presented in Holz et al. [70]. Chen and Baldick [71] discuss a model for optimizing the short-term natural gas supply portfolio for natural gas based power generation for an electric utility. Attempts have been made to estimate residential and commercial demand

([72]) as well as industrial demand ([73]).

1.4.4 Infrastructure Development

Offshore infrastructure development and long-term planning is the earliest and possibly the most widespread application of mathematical programming in this area. This is not surprising given the enormous capital cost and risk associated with offshore field development. However, because this thesis is concerned with operational planning, only some representative works and trends in the area are outlined and one should refer to these to explore the field further. A combination of the infrastructure and operational problems from a petroleum engineering perspective appears in Huppler [74], Flanigan [75] and O'Dell et al. [76]. A well location problem with a simplified reservoir model is solved in Murray III and Edgar [77]. McFarland et al. [78] use a simple tank model for reservoir dynamics to formulate an optimal control problem and solve it using a generalized reduced gradient method. Beale [79] describes a long-term offshore field development problem with options to install compressors that is solved as a MILP approximation. An approach for making exploration decisions in oil and gas fields is presented in Beale [80]. Haugland et al. [81] present a long-term multi-period MILP model for making decision in field and infrastructure development. A long-term multiperiod MILP model in use by the Norwegian regulator to aid in field development decision making is presented in Nygreen et al. [82]. van den Heever et al. [83] present a model for long-term infrastructure planning with complex economic objectives. A model of the production system in Saudi Arabia and issues associated with it are presented by Gao et al. [84]. An oil well spacing and production control problem is solved in Ayda-zade and Bagirov [85] by first formulating a two dimensional partial differential equation (PDE) model of the reservoir and then converting it to a conventional optimization problem. Recently, works have started focusing on handling the uncertainty involved in planning using stochastic programming formulations [86, 87]. Jonsbråten [88] presents a stochastic programming model for oil field development and operations under price uncertainty. A simple analysis of the profitability of development projects in the presence of production-sharing contracts

appear in Yusgiantoro and Hsiao [89]. The analysis does not involve mathematical programming and is based on simply evaluating several scenarios.

1.4.5   Relevant Models in Oil Production

There have been works on short-term production planning in oil fields. Dawson and Fuller [90] present a multi-period mixed-integer nonlinear programming (MINLP) model for production planning in an oil field, however optimize it heuristically using Generalized Benders Decomposition. Kosmidis et al. [91] present a model for well rate allocation that comprises naturally flowing and gas lift wells. The model is solved by linearizing well models to formulate an approximate MINLP and then solving a sequence of MILPs. Khezzar and Seibi [92] discuss a NLP model for optimizing upstream oil production systems with gas-lifted wells and solve it with SQP. Ortíz-Gómez et al. [93] present MILP and MINLP short-term multiperiod oil production models, however nonconvex MINLP models are not solved to global optimality and they only address oil production systems where the gathering system is not strongly coupled with the wells. Queipo et al. [94] present an integrated model of reservoir and surface facilities for an upstream oil production system and solve it with SQP and derivative-free methods. Barragán-Hernández et al. [95] solve a dynamic model with an interior point solver for an oil field.

1.4.6   Miscellaneous Works

A hybrid systems model that includes reservoir dynamics and economics for gas wells is presented in Chermak et al. [96]. Teisberg and Teisberg [97] discuss a contract valuation methodology for natural gas. An introduction to the application of options theory to oil and gas is discussed in Paddock et al. [98] and Smith and McCardle [99]. Murphy et al. [100] discuss analysis of natural gas regulatory proposals in the US. Chen and Forsyth [101] develop an approach for valuation of a natural gas storage facility as a stochastic control problem and solve the resulting Hamilton-Jacobi-Bellman equation using a semi-Lagrangian approach. An exploration of regulatory issues in

the gas markets based on a model of gas transportation pricing model appears in Cremer and Laffont [102]. A generalized network model of the US integrated energy system is presented in Quelhas et al. [103] and Quelhas and McCalley [104]. A model for LNG terminal design with a supply-chain perspective is presented in Özelkan et al. [105].

1.4.7 Relationship with Electricity Planning

Gas production and transportation planning correspond respectively to power generation and transmission planning. This is because electricity shares several characteristics with natural gas: the storage of electricity is difficult and expensive, the expansion and development of electric generation capacity requires that a big share of demand be guaranteed and peak power demand can result in non-availability of sufficient power generation capacity or deliverability problems due to inadequacy of the power grid. However, there are also important differences from a modeling and planning perspective. Natural gas planning is complicated by nonlinear pressure-flowrate relationships in wells and transportation networks, gas quality issues, complex storage deliverability, involved transportation issues due to pressures and gas quality, and complex contractual rules. Also, natural gas faces uncertainty not only from demand (as electricity) but also from the supply-side infrastructure in term of production expansion in the long-term (exploration and recoverable reserve uncertainty) and transportation availability. Finally, the capital intensity of the overall natural gas chain is higher than electricity infrastructure. Electricity power planning problems traditionally focus on generation expansion, scheduling and planning and competitive market issues. Gas planning problems tend to focus on modeling of gas production, transmission, storage and contractual issues. In general, issues in production (generation), transportation (transmission), and distribution differ for both and hence the modeling approach in one is not directly applicable to the other. Nevertheless, some aspects of the modeling and solution methods can be similar. A further discussion is out of scope here and more information can be found in Kagiannas et al. [106] and Hobbs [107]. An interesting (and worrying) development is the increasing use of

natural gas in electricity generation using high-efficiency CCGT plants which is increasingly intertwining both markets [5], thereby enabling supply shocks from natural gas markets to travel to electricity markets.

Chapter 2

Short-term Planning in Upstream Natural Gas Systems

A short-term planning framework for upstream natural gas systems can streamline production operations and enable operators to realize the full system potential in terms of their preferred metrics while simultaneously satisfying governing rules (as defined later). However, there has been little work in this area, partly because computing power and optimization theory have been insufficient in the past to handle the kind of moderate to large-scale nonconvex mixed-integer nonlinear programs (MINLP) that result from these planning problems. However, with the advances in global optimization theory and algorithms made in the last decade, this problem is now tractable, provided a reasonably good optimization modeling approach is used in conjunction with exploitation of problem structure.

2.1   General System Definition

The upstream natural gas system in this context is a set of natural gas fields that produce into a common collection (gathering) network to transport gas to a large consumer facility. The system includes reservoirs, wells, processing facilities connected to wells and the surface or subsea pipeline network that supplies the consumer. The consumer is usually a large facility that consumes all (or most of) the natural gas

produced by the upstream system. Examples of such consumer facilities may be a LNG plant, a GTL plant, a petrochemicals complex, a power plant, a big pipeline exporter, or a combination of several such facilities. The consumer facility may or may not be included in the planning scope depending on the mode of operation and control of the upstream system. If the facility and the upstream production system are operated in an integrated fashion, it is highly desirable that the operational planning frameworks include the consumer facilities. The planning horizon is short-term, typically ranging from a few weeks to several months.

2.2 General Features

Most upstream systems have multiple stakeholders because they are too large to be owned by a single entity. Fields, facilities and network may be split between different parties due to the high capital costs of developing upstream infrastructure. The upstream system may be operated by a single operator which is usually one of the stakeholders in the system. The above arrangement is progressively becoming more widespread due to national oil companies owning large gas projects where they need to bring in technology and investments from international oil companies. The consumer facility itself may be a different entity from the group of producers. As a result, such systems are operated on the basis of comprehensive contractual agreements that may involve multiple stakeholders in the upstream infrastructure and consumer facilities as well as the upstream operator. The contractual agreements define the operational strategy for the system that facilitates sharing of products or revenue from the system. They may also specify gas quality, delivery amount and pressures as dictated by the operation of consumer facilities. Usually consumer facilities are designed to accept a narrow band of quality and out of specification feed gas can disrupt the facility operations. Some systems may also impose economic penalties on producers if there is a failure to deliver the requisite amounts in contracts. Similarly, consumer facilities, in turn, have sales agreements with buyers that may define amounts, quality specifications and penalties on non-delivery. System operations must also comply

with governmental regulations, e.g., safety and environmental codes. Finally, there is decades of human experience in operation of such upstream systems that result in operational heuristics to work around design deficiencies.

These contractual and commercial agreements, regulatory rules and operational heuristics govern the upstream operations and are referred to later as the *governing rules* for the upstream system. As a result of these governing rules, an optimization of the production infrastructure model alone is not sufficient to generate operational policies that are consistent with the governing rules and therefore permissible. On the other hand, a simplistic accounting of volumes based on the contractual rules and gas qualities may not be consistent with the physical behavior of the production infrastructure and therefore, may not be useful as an operational policy. A model that integrates a production infrastructure model with the governing rules is therefore the preferred approach for upstream systems.

Short-term planning problems are intrinsically operational planning problems because most decisions in short periods (a few weeks to several months) are concerned with determining an operational state of the infrastructure that meets certain production goals while satisfying system constraints. There are two required components of such models:

1. A reasonable representation of the production infrastructure that is able to predict pressures and gas quality throughout the upstream network so that the resulting policy is physically realizable on the network.

2. A representation of the governing rules (whether contractual, regulatory, commercial or operational) is required for the policy to be acceptable within the framework agreed by all stakeholders.

2.3 Technical and Business Benefits

The upstream planning framework can yield multiple business and technical benefits for production operations and by extension, overall gas supply chains. These tools

generate an operational strategy that help system operators to fulfill their primary production targets. As long as the models include the governing rules for the upstream system, policies generated from these models are guaranteed to comply with them. They can help to fulfill the gas quality specifications frequently mandated by contracts for smooth functioning of consumer facilities. Additionally, several upstream production systems and associated processing facilities produce by-products such as Natural Gas Liquids (NGL) and Liquefied Petroleum Gas (LPG). By-product optimization can further improve returns from the upstream system. Finally, it is also possible to integrate complex commercial and economic clauses of contracts into these model to take economic factors and penalties into account and create a techno-economic model of the upstream system that maximizes the economic value generated by the system.

These models can facilitate management of assets in the upstream system in the short term. Maintenance scheduling using these tools to evaluate scenarios and using intelligent objectives and constraints can help to minimize supply disruptions. The majority of newly discovered reservoirs are expected to have high levels of contaminants, especially CO_2 and H_2S. These models can help to intelligently route gas through the network so as to blend it with sweet gas to satisfy gas quality specifications. In this way, they aid in the distributed depletion of fields in the presence of quality constraints. A post optimality sensitivity analysis of a planning model can also precisely point out the bottlenecks in the system, i.e., which part of the system must be upgraded to obtain an increase in the particular production objective, so that capital investment decisions for capacity expansion can be targeted precisely and the return obtained can be maximized. For example, an inability to sustain a particular level of production from a sour field in the presence of quality constraints indicates that a separation facility must be installed to further increase the production without violating quality constraints or alternatively a quality constraint must be relaxed, if feasible.

Short-term planning frameworks can interact beneficially with medium and long-term upstream planning models. A short-term planning model can help to follow trajectories specified by the medium and long-term planning more accurately while

simultaneously satisfying production or revenue targets and governing rules in the short term. For example, short-term models can enforce rate limits and profiles computed by reservoir management models for maximum recovery. The information can also flow the other way. Capacity expansion information and other operating information can be factored into the medium and long-term models to correct them and hence apply mid-course corrections to the plans, or reevaluate them entirely if required.

Such models can also improve the resilience of upstream systems to failures. It is possible to run various operational scenarios and evaluate the availability and plan in advance. These model can also help to evaluate the delivery potential of the systems in face of unexpected emergency demands (e.g., an unexpected bout of cold weather) or unexpected supply-side problems (e.g., the failure of a facility). With increasing LNG spot trade, the same approach can be applied to query the model as to whether spot gas is available, and because spot markets fetch higher prices, this increases overall revenue from the system. In this way, these models provide real-time decision support. If a particular scenario makes violation of at least some governing rules inevitable, these model can used to ascertain which set of rule violations involve the least economic or operational penalty. These models can also aid in redesigning of contractual agreements between parties by evaluating the impact of governing rules on operations and identifying the most unfavorable rules.

With a move towards more automated upstream systems with online sensors and actuators, and advanced control systems, these planning models can be modified to use data from sensors in the field for effective and automatic calibration and to provide target profiles for control systems. Consumer facilities can be integrated into the planning framework if it is permitted by their operational mode. This can help to determine the optimal response of facilities to fluctuations in the upstream network and/or demand.

As discussed earlier, the entire gas supply chain has to be coordinated to ensure reliable supply. An entire source-to-market model can help to make the supply chain resilient and robust to disruption. These source-to-market models are expected to

be short-term models (except when they are models for designing supply chains) since supply chain availability planning is only relevant over short term. A short-term upstream system model is an important component of the supply side of such a model.

2.4 Mathematical Characteristics

Mass, volume or specieswise molar balances and pressure-flowrate relationships in wells, pipelines and facilities account for a majority of the constraints in an upstream production infrastructure model. Similarly, if a planning model incorporates contracts, contractual volume accounting constraints are required. All these constraints are equality constraints that are either linear or nonlinear. There are only a few "real" inequality constraints in the problem that represent physical or contractual bounds on volumes, composition and pressure variables (although, the actual number of inequalities is larger due to the addition of redundant constraints). A rough measure of the number of *degrees of freedom (DOF)* in such a model is the difference between the numbers of variables and equalities. If the DOF are fixed, the rest of the model can be solved as a system of nonlinear equations, provided it is well-posed. Therefore, if the feasible region is nonempty and the problem involves only *smooth functions* with continuous variables, any feasible solution of the system lies on a *smooth manifold* (which is the solution of the nonlinear system parameterized by the DOF) of a dimension that is equal to the number of degrees of freedom of the optimization problem. Hence, the feasible sets of such problems are nonconvex. Furthermore, if governing rules are incorporated in the upstream planning framework, either binary variables and constraints or disjunctive constraints are required to represent these logical conditions and model these rules. Hence, upstream planning problems are expected to be mixed-integer nonlinear programs with nonconvex functions.

A general mixed-integer nonlinear programming problem (MINLP) can be repre-

sented as:

$$\min_{\mathbf{x},\mathbf{y}} \quad f(\mathbf{x},\mathbf{y})$$

$$\text{subject to} \quad \mathbf{g}(\mathbf{x},\mathbf{y}) \leq \mathbf{0}$$

$$\mathbf{h}(\mathbf{x},\mathbf{y}) = \mathbf{0}$$

$$\mathbf{x} \in \mathbb{X}, \quad \mathbf{y} \in \{0,1\}^{n_y}$$

where $\mathbb{X} \subset \mathbb{R}^{n_x}$ is a nonempty convex set, usually an interval defined by variable bounds. $f : \mathbb{X} \times [0,1]^{n_y} \to \mathbb{R}$, $\mathbf{g} : \mathbb{X} \times [0,1]^{n_y} \to \mathbb{R}^{n_g}$ and $\mathbf{h} : \mathbb{X} \times [0,1]^{n_y} \to \mathbb{R}^{n_h}$ are continuous functions, some or all of which are nonconvex. If the binary variable vector $\mathbf{y}$ is relaxed to be in the interval $[0,1]^{n_y}$, a nonconvex nonlinear program is obtained. It is worth noting here that nonlinearity of equality constraints (even though they may involve convex functions) is sufficient to make the nonlinear program nonconvex, in almost all cases. MINLPs in which some or all the participating function are nonconvex will be referred to as *nonconvex MINLPs*.

Deterministic optimization algorithms can be characterized on the basis of the nature of their solution set as being *global optimization algorithms* and *local optimization algorithms*. Local optimization algorithms are methods that use local information (e.g., function values, gradients or Hessians at the points) to generate the next iterate and therefore, define their solution set as points satisfying some local optimality conditions. Since they rely on local information, they are only guaranteed to converge to a local minimum (more precisely, to a point satisfying some necessary condition or stationarity condition for a local minimum, e.g., the Karush-Kuhn-Tucker (KKT) conditions [108, 109]). Almost all nonlinear programming algorithms (e.g., Sequential Quadratic Programming (SQP)) fall into this category. As is demonstrated later, local algorithms perform poorly on nonconvex programs. On the other hand, global optimization algorithms are methods for which the solution set is defined as the set of global minima of the problem. Most global algorithms use information from the entire feasible set. They guarantee (by definition) convergence to a global minimum for continuous nonconvex programs and by similar extension of logic to a global minimum

for nonconvex MINLPs.

Most global algorithms are not specific algorithms, but are algorithmic frameworks within which details should be filled in to arrive at a specific algorithm. In general, these algorithms work by solving a series of subproblems using local solvers to generate upper and lower bounds on the optimal solution value. Assuming that the problem is a minimization problem, upper bounds are usually generated by solving a *restriction* of the original problems by fixing variables or adding additional constraints that may make it easy to solve. The feasible set of a restriction is a subset of the original feasible set. Lower bounds are obtained by constructing and solving a *relaxation* such that the feasible set of the relaxation contains the original feasible set. There are details that are specific to an algorithm and implementation that define how restrictions and relaxations are constructed, as well as how the next iterate is determined. These upper and local bounds will converge to within the specified accuracy after a finite number of iterations under mild assumptions on the nature of the subproblems. The robustness of global optimization algorithms stems from their exhaustive search over the feasible set of the problem, infeasibility decisions relying on overestimation (i.e., on relaxation) of the feasible set, and their globally convergent nature (convergence is independent of the initial guess for starting the algorithm). However, global optimization algorithms suffer from worst-case exponential runtime and hence solving relatively moderate- to large-scale nonconvex MINLPs (e.g., several hundred continuous variables and tens of binary variables) to global optimality can be quite challenging. Therefore, in general, global optimization algorithms need to be customized, i.e., the general frameworks need to be tailored to a specific problem class using features of the class to accelerate convergence for larger problems.

More information on general nonlinear programming theory can be found in Bertsekas [108] and Bazaraa et al. [109]. Further information on MINLP solution methods can be found in the review by Grossmann [110].

It is a common misconception that the tradeoff between using global and local optimization algorithms for upstream planning problems (and in general, for any nonconvex NLP) is that when using global optimization, a global minimum is found

while when using local optimization algorithms a locally optimal solution will be found. Skeptics argue that real-world models are inherently uncertain and imprecise, and therefore, that any potential benefits promised by using global optimization algorithms may not be realizable and are not worth the complexity and computational burden of these algorithms. Also, they claim that a robust behavior can be obtained by using multistart approaches in conjunction with local solvers and that global algorithms are incapable of handling complicated models. There are several arguments against these objections:

1. The arguments above assume that a locally optimal solution can be found for such problem using local solution methods. In general, there is no guarantee that a feasible solution will be found for such problems using local solver. There are numerous theoretical and numerical difficulties associated with applying local solution methods to nonconvex programs. As is demonstrated in the example later, local method can be misleading even in determining feasibility for such problems.

2. A strong parametric sensitivity of the solution is certainly observed in many nonconvex programs. Therefore, a parametric uncertainty can potentially mislead about a solution. However, this is a feature of the nonlinearity of such problems that results in drastic changes in the feasible set topology with parameter variations, and applies to both local and global solution methods. A local optimum (as well as a global optimum) can be equally sensitive to parameter changes for nonconvex NLP. Moreover, this does not preclude the optimization model from being useful as long as care is taken to obtain a reasonably accurate estimate of the parameter and an eye is kept on the structure of the solution with variation of the parameters. Also sensitive parameters can be deduced from physical arguments, problem structure or numerical experiments. Efforts can then be made to either estimate the parameters more accurately or even comprehensively review the model formulation to reduce sensitivity. Often such instability of solutions can indeed be a feature of the system itself and cannot

65

be avoided. Finally, the solution effort for nonconvex programs is also strongly dependent on parameter values and bounds due to similar arguments.

3. Integer variables cannot be handled (robustly) in continuous local solver frameworks. This means that any logical condition based constraint cannot be incorporated in the planning frameworks, which severely limits the usefulness of such frameworks for gas supply chains where governing rules are of paramount importance for operations.

4. A multistart approach (i.e., performing solution attempts from multiple initial points) can indeed alleviate some problems associated with local solution methods. Multistart approaches rely on the fact that one of the initial points will be close enough to the actual feasible region and therefore will succeed in finding a global or local optimum. However, there is no guarantee and success is very much dependent on the structure of the problem and distribution of the set of initial points. Moreover, some inferior minima may have large radii of convergence while some that are of interest may have a small convergence radii and therefore, are hard to find. Besides, it can be argued that a global optimization framework already incorporates the multistart approach in a smart and systematic fashion (i.e., the upper bound initial points are chosen based on lower bounding information) and therefore, computational effort and time is better spent in systematic global approaches than the brute-force multistart solution methods.

5. A global optimization framework is indeed complex from a theoretical and implementation perspective but need not be complex from a user perspective. Global optimization can almost always perform at least as good as a local algorithm (simply because local solution methods are embedded in them for subproblem solutions). Therefore, to argue complexity from a user perspective is incorrect. A properly designed implementation can gracefully fall back to local methods (or heuristic based approaches for a mixed-integer problem) if the problem is too computationally intensive to solve or can even provide a

tentative local solution immediately for the user to analyze while it continues on computation of global solution. Finally, implementations can be designed to replicate the current iterative (or heuristic-based) approaches (an example is presented in later in Section 2.5.1) to recover from failures.

6. MINLP modeling is complicated and the model formulation is important for efficient solution as is discussed later. However, once a modeling formulation and methodology for a specific upstream system has been decided, adding additional constraints to the model is not complicated from a user perspective. This combined with a properly designed implementation means that the complexity of this approach from a user perspective is no more than conventional local solver based approaches. The only other option to MINLP models is the iterative approach described later in Section 2.5.1 which is only defensible if one is interested in a feasible solution of the problem and not in a optimal solution.

7. Most global optimization algorithms require explicit functional representation of the governing equations (e.g., flowrate-pressure relationships). A black-box function evaluator (e.g., a simulator to calculate a quantity such as a reservoir pressure) cannot be used with them. This has been raised as an objection to these frameworks in the past. However, this is really not a handicap at all. A functional representation can be chosen to represent a particular governing equation, preferably derived from a mechanistic approach. Parameters in the functional representation can be then be regressed in the expected range from data generated from a detailed simulation model (e.g., a detail fluid dynamics model for flow in the reservoir) or even from real-time data obtained from online sensors. Not only that, this entire calibration procedure can be easily automated. These regressed functional relationships can be used in MINLP models.

8. Global algorithms provide guaranteed bounds on system performance which is a huge benefit. If the performance required by a plausible scenario is outside the performance bounds provided by the global algorithm, that scenario can be

immediately deemed unattainable.

2.4.1 Local Solver Behavior: A Case Study

This example outlines the downside of using local solvers to solve upstream planning problems that are expected to be strongly nonlinear and nonconvex problems. Consider the following example problem:

$$\text{minimize} \quad x_3$$
$$\text{subject to} \quad x_1^2 - x_2^2 - x_3^2 = 0$$
$$5x_1 - 4x_2 - 3x_3 - 6 = 0$$
$$1 \leq x_1 \leq 20$$
$$1 \leq x_2 \leq 20$$
$$1 \leq x_3 \leq 20.$$

The first constraint is similar to the standard relationship for modeling pipeline pressure drops in gas pipelines and wells (e.g., Equation (1), page 85). The region feasible in the second linear constraint may result from the combination of other constraints in the original problem, e.g., molar balances that are linear.

The feasible set of the problem is shown in Figure 2-1. Note that the feasible set is the intersection of a nonlinear surface represented by the first constraint and the hyperplane represented by the second. It is a one-dimensional curve with an "upper" and a "lower" branch that are disjoint. The dimension of the feasible set is one consistent with the fact that the problem has just one degree of freedom. Therefore the feasible set is *nonconvex* as well as *not connected*. The optimal solution to the problem is $(x_1, x_2, x_3) = (8.7712, 8.7140, 1.0000)$ with an objective value of 1.0000 (in the lower branch) obtained with a *Branch-and-Reduce* global optimization solver. There is a local minimum for the problem $(x_1, x_2, x_3) = (4.8435, 1.0000, 4.7391)$ with an objective value 4.7391 in the upper branch.

Table 2.1 present the results from various solvers with the example problem. It
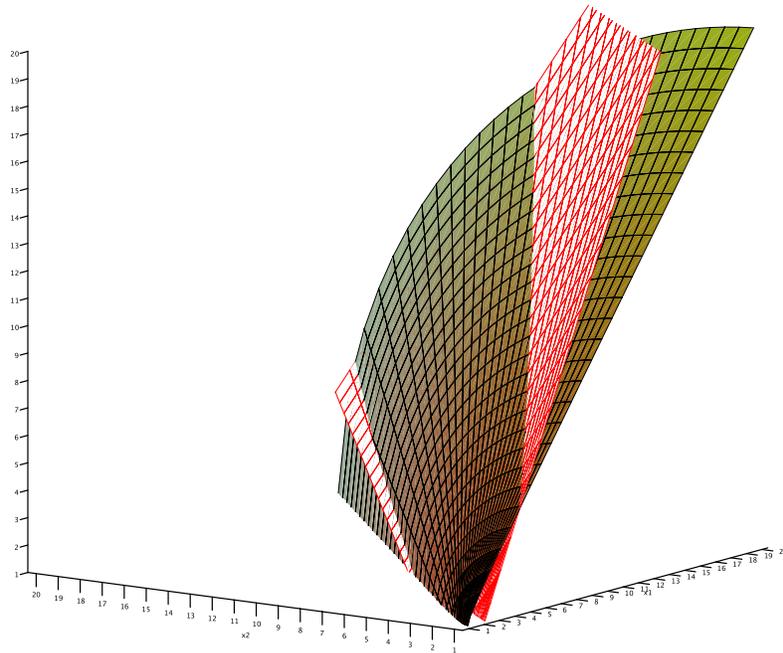
Figure 2-1: Feasible region of the example is the intersection of the surface with the hyperplane

shows that the solutions obtained by local solvers depend on the initial point used to start the algorithms. In the worst cases, as can be seen from the table, a local method may report the problem as infeasible (sometimes after a single iteration). Even worse, one can find several points, for which such behavior is observed. For example, there are several starting points for which the SQP method (SNOPT) will report the problem as infeasible: $(5, 4, 3)$, $(20, 2, 1)$, $(10, 4, 1)$ and possibly even more. In cases where the local solvers do find feasible points, nonconvexity in the problem results in convergence to the inferior local minimum and not finding the best possible objective value. Only for certain starting points do the local solvers converge to the global minimum.

The failure of the SQP method is of particular concern since it is one of the most

Table 2.1: Local solvers – A case study: Objective value[a]

Solver	Type	Algorithm	Initial Point					
			Midpoint (10.5,10.5,10.5)	A[b] (5,4,3)	B[b] (10,4,1)	Vertex (1,1,1)	Vertex (20,20,20)	Vertex (20,20,1)
CONOPT	Local	Generalized Reduced Gradient	4.7391	Infeasible	1.0000	Infeasible	4.7391	1.0000
MINOS	Local	Projected Lagrangian	4.7391	Infeasible	1.0000	Infeasible	4.7391	1.0000
SNOPT	Local	SQP	4.7391	Infeasible	Infeasible	Infeasible	4.7391	1.0000
KNITRO	Local	Interior Point	4.7391	1.0000	4.7391	4.7391	4.7391	1.0000
IPOPT	Local	Interior Point	4.7391	1.0000	1.0000	1.0000	4.7391	1.0000
BARON	Global	Branch-and-Reduce	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

[a] GAMS 22.8 with default options

[b] Example points
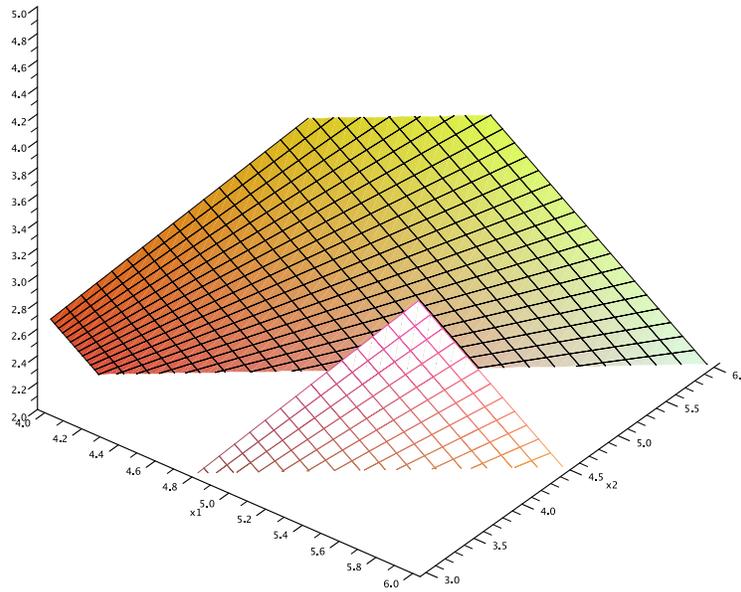
Figure 2-2: The feasible region of QP subproblem at (5,4,3) is intersection of the two parallel hyperplanes resulting in an empty feasible set

popular methods in process and petroleum engineering software suites. Advantages of the SQP method include its speed, ability to handle large NLPs and the need for first order information only.

SQP Failure

This analysis serves to demonstrate an example of the underlying causes of local solver failures. An example point (5,4,3) is chosen to demonstrate the problems associated with SQP. In a *major iteration*, the SQP method generates linearizations of the constraints and a quadratic approximation of the objective to formulate a *Quadratic Programming* (QP) subproblem and solves it for a descent direction. It then performs a *line search* in the resulting direction. The linearizations at $(5, 4, 3)$ are depicted in Figure 2-2. These yield a QP subproblem whose feasible set is composed of the intersection of two *strictly separated* parallel hyperplanes and is therefore empty. The QP subproblem at $(5, 4, 3)$ is therefore infeasible and cannot be solved. The convergence theory of the SQP method usually assumes that the QP problem can be

71

solved at every iteration to obtain a descent direction. This certainly holds true for convex programs with non-empty feasible sets and therefore, SQP perform quite well for such problems.

Generally, SQP solvers use certain heuristics to overcome QP subproblem infeasibility. For example, the solver used here (SNOPT) switches to a weighted minimization of local constraint infeasibilities when the QP subproblem is infeasible, but this heuristic breaks down for several initial points as demonstrated above and the problem is deemed as infeasible. This can happen even if a large weight is assigned to the sum of infeasibilities. Some solvers also relax the infeasible constraints and attempt to generate a feasible QP, but again such procedures can be shown to fail.

2.4.2   Implications for Upstream Planning Problems

It is demonstrated here that local algorithms are unsuitable for nonconvex programming. There are both theoretical and numerical difficulties associated with erratic behavior of local solvers when applied to nonconvex programs. Nonconvex programs usually have multiple local minima some of which can be suboptimal. KKT conditions are not sufficient to characterize a (global or local) minimum for nonconvex programs. Additionally, several assumptions specific to the algorithms and their implementation break down for these problems. Finally, due to nonlinearity of nonconvex programs, subproblems in iterations (e.g., the QP direction finding problem is SQP) can become numerically ill-conditioned or unstable and fail to solve. In the worst case, local solvers can be misleading, reporting a problem whose feasible set is non-empty as being infeasible.

It is therefore clear that for upstream planning problems, local solvers may not only converge to a suboptimal local minimum, they may also fail to locate even a feasible point. It can be potentially hard to ascertain if the feasible region of the problem is empty and the problem is indeed infeasible or if the structure of the feasible region combined with the fact that only a very small part of the search space is feasible is making it hard for the solver to locate a feasible point. This can be especially true if the system is operating close to its maximum potential and a very small region of the

hyperrectangle defined by bound constraints is feasible and therefore, it is unlikely that an initial guess will be close to the feasible region. Under such conditions, feasible scenarios may be declared infeasible resulting in lost spot sales or contract violations. It is also expected that there are multiple suboptimal local minima in the problem and the point to which a local solver converges is strongly dependent on the starting point of the solution method. Global optimization methods are therefore required for a reliable solution of upstream planning problems.

The previous case study also serves to illustrate why the "black-box optimization of simulation models" approach, i.e., an optimization solver coupled with a complicated upstream simulation model to evaluate functions, is dangerous and can be misleading. It is hard to even ascertain the basic properties of functions evaluated by complicated models with embedded iterative solution methods in unit operation models and/or a sequential simulation approach with logical conditions on the output of unit operation models. The functions involved can be nonconvex, nonsmooth, discontinuous or even "undefined" on the required "domain". There is no way to guarantee that core mathematical assumptions necessary for the application and convergence of an optimization algorithm are being satisfied. So the solutions obtained by these approaches (if they can be solved at all) are entirely unreliable. Frequently, they will simply fail to locate any solution. At best, one can say that the solution points obtained from such approaches are feasible. In the real-world, this unreliable behavior breeds distrust of planning models among users and operators because it is easy to best the model using simple heuristics and trial-and-error.

Finally, due to the network structure of the problems, there are multiple production profiles that can achieve similar production goals. The upstream problem with operational objectives can therefore have multiple globally optimal solutions as well. It is beneficial to add constraints reflecting operational aspects and preferences to the problems, even though these might not represent any actual costs. This will certainly help to reduce the number of multiple solutions. Also in such a scenario, even if there are multiple solutions, they will be operationally equivalent, so that one need not be concerned that any other solution may be more promising.

Importance of Model Formulation

A problem can have several MINLP representations that are equivalent, i.e., the optimal solution set and optimal solution value are the same. However, these representations may not be equivalent in terms of the effort required to solve them. This is because different equivalent MINLP representations result in different subproblems that must be solved at an iteration of a global algorithm, thus strongly affecting its convergence. To judge whether a formulation is better or worse for use in conjunction with a particular optimization algorithm, a good understanding of the algorithm is required. Hence, attention to modeling is of paramount importance for nonconvex MINLP and is as crucial as the algorithms employed to solve the model. Finally, the value of adding extra redundant constraints (to tighten relaxations) is well known for integer programming and is even more important for MINLP. In particular, extra constraints help to tighten the relaxation.

2.5   The Sarawak Gas Production System

The requirements and scope of the general modeling framework presented in the following chapters has been inspired by the Sarawak Gas Production System (SGPS) in East Malaysia. It is therefore instructive to discuss the SGPS features to put in context the model requirements and overview presented later. The Sarawak Gas Production System (SGPS) is located in the South China Sea off the coast of the state of Sarawak in East Malaysia (Figure 2-3). There are 12 offshore gas fields in the system. Additionally, *associated gas* from 3 oil fields is fed into the system. The daily production rate of dry gas from the SGPS is around 4,000 million standard cubic feet per day (MMscfd). Additionally the system also produces 90,000 barrels per day (bpd) of natural gas liquids. The annual revenue from the SGPS is around US $5 billion, that is approximately 4% of Malaysia's GDP as of 2005[1].

The gas from the system is fed to the Petronas LNG complex in Bintulu, Sarawak. The complex is one of the largest LNG production facilities at a single location in the

---

[1]All figures from 2005.

Figure 2-3: The Sarawak Gas Production System
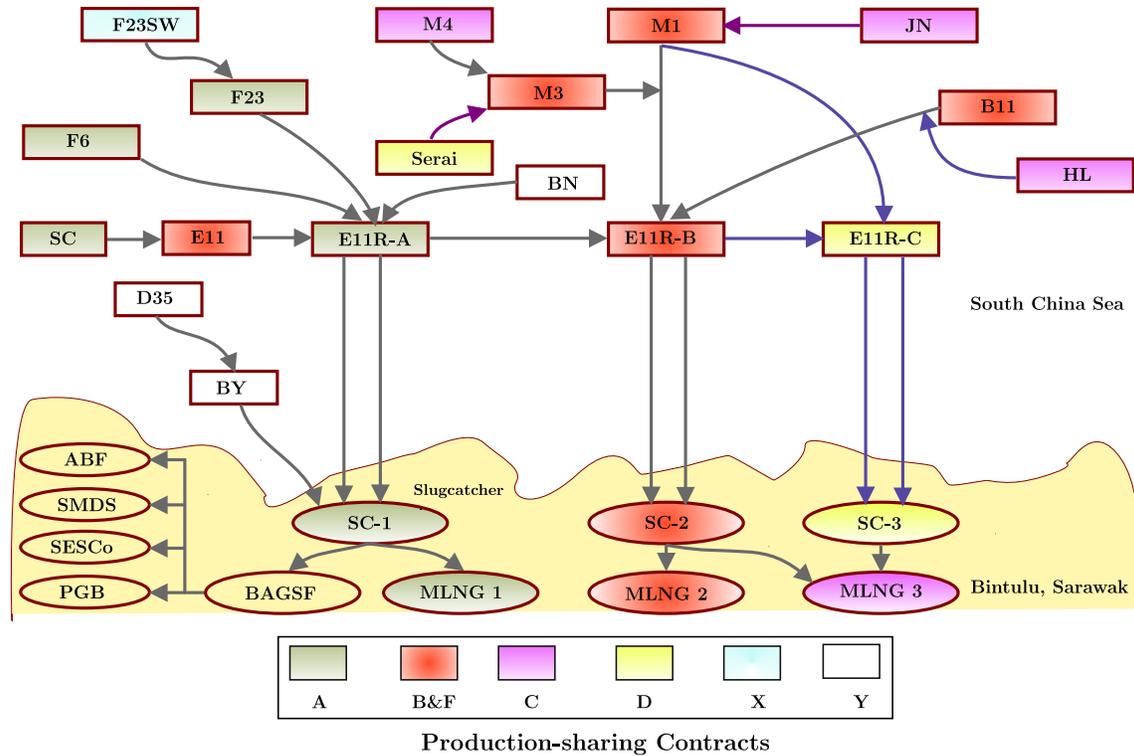
Figure 2-4: The Sarawak Gas Production System: Network overview (not to scale)

world. It has three plants and produces around 21 million tonnes of LNG. The main customers of this complex are Japan (58%), South Korea (25%) and Taiwan (17%)[1].

From a modeling standpoint, the system comprises wells in the fields, well platforms, the pipeline network, riser platforms and the facilities onshore. Gas from the wells belonging to a particular field is collected at a well platform. A well platform may serve more than one field. Well platforms have dehydration facilities that perform a three phase separation of gas, natural gas liquids (NGL) and water. They may also have compression facilities, in case the field pressure is insufficient to drive the flow. Once dehydrated, dry gas and natural gas liquids are remixed (after compression if a platform has compression facilities) before injection into the pipeline. A subsea pipeline network (referred to also as the trunkline network later) connects the well platforms to the facilities onshore. The flow in the network is two phase (gas and liquid).

---

[1]Petronas Annual Report, 2007.

Table 2.2: Bintulu Petronas LNG complex: Stakeholders[a]

Plant	Ownership	Capacity (MMt/y)	Start-up
MLNG	Petronas (65%), Shell (15%), Mitsubishi (15%), Sarawak local government (5%)	8.1	1983
MLNG Dua	Petronas (60%), Shell (15%), Mitsubishi (15%), Sarawak local government (10%)	7.8	1996
MLNG Tiga	Petronas (60%), Shell (15%), Nippon Oil (10%), Sarawak local government (10%), Diamond Gas (5%)	6.8	2003
Total Liquefaction Capacity		22.7	

[a] EIA, U.S. DOE, (http://www.eia.doe.gov/cabs/Malaysia/Full.html).
Original source: Petronas

The subsea trunklines end at one of the three *slugcatchers* corresponding to the three LNG plants at the complex in Bintulu, Sarawak. The slugcatchers are units that remove NGL from the two phase flow coming out of the trunklines. The liquids are sent to stabilizers to remove volatiles and the dry gas is fed into the LNG plants. There are also small customers: a power generation company, a fertilizer plant, a local utility and a petrochemical plant. However, these users consume close to 5% of the total production and moreover their demand is mostly fixed and hence can be represented by adding a small constant factor to the minimum demand rate of the first LNG plant. Hence these need not be considered explicitly for planning. The complex also contains a liquefied petroleum gas (LPG) plant. Hence, there are two by-products, NGL and LPG, from the system.

2.5.1 Operational Aspects

The fields, facilities and plants in the system are not owned by a single entity but instead several parties either have stakes in them or may fully own some of them. The stakeholders for Bintulu LNG complex are shown in Table 2.2. However, almost the entire system (excluding the customer plants) is operated by a single upstream

operator.

As a consequence of ownership issues, a particular field cannot arbitrarily supply any customer in the system. There are *Production-Sharing Contracts* (*PSC*) that determine how the products are shared between different parties. PSC (referred to as simply "contracts" later) define the field to plant assignment, i.e., which field should supply a particular LNG plant. They also define the course of action when a demand cannot be fulfilled as per the contractual assignment, e.g., which fields should step in to fill in the deficit if a particular set of fields cannot meet its mandated demand. Moreover, the contracts also contain operational details (if necessary) to implement such a flow redistribution on the network. Additionally, due to ownership issues, contracts often also dictate the use of facilities. Finally, contracts specify the customer requirements, particularly the amount, the delivery pressure, heating value of dry gas and composition specifications. This extensive set of contractual and operational rules govern the operation of the SGPS and must be satisfied at all times. Making routine operational decisions about production and routing in the network is therefore very difficult and cumbersome.

Traditionally this has been done in two stages: first solving a production planning problem using the production system model, i.e., a model of production infrastructure in a commercial software suite with a local solver (e.g., SQP) to determine feasible values for the well production rates and trunkline flowrate-pressure distribution, and then manually ascertaining if the contractual (including gas quality specifications) and operational rules are satisfied. If not, then another scenario can be evaluated by enforcing different constraints and bounds on the production system model and checking the contract rules again. Iteratively, a feasible solution that satisfies all rules may be found. However, this approach suffers from several problems:

1. Due to the nonlinear pressure-flowrate relationships in the wells and trunklines, the problem is nonconvex. Hence, solution of the first stage production planning problem is liable to fail. As was demonstrated in the numerical experiment presented earlier, some instances of such problems may be reported infeasible by local solvers when they are in fact feasible.

2. Even if this iterative procedure does converge, there is no guarantee concerning the quality of the solution, i.e., there is no information if a far superior solution exists.

3. For a large system containing tens of fields, such a scheme is too tedious and error prone to devise a consistent operating strategy. With an expansion in the upstream system, this approach is only going to get more difficult to apply.

4. It is also possible that this procedure may not generate any solution point at all that satisfies all the rules.

5. Finally, such an approach requires too much human intervention and intuition, and is not maintainable in long-term. Depending on the person who is running the problem solution procedure, the results will differ and will not be consistent with someone else.

2.6 The Short-term Upstream Planning Model: Overview

This work focuses on short-term production allocation in the upstream system. By short term, planning on the order of a few days to a few weeks is implied. The upstream system is defined as from the bottom of the well bore to the LNG plants, however excluding the LNG plants. The operating state of the system is determined by the following decision variables:

1. Production share of dry gas from each well (and therefore each field).

2. Associated pressures at the well bore and well head for each well.

3. Pressure and flowrate distribution in the trunkline network.

4. The state of inter contract transfers and operational rules.

5. Amount and quality of gas delivered to the LNG plants.

2.6.1 Requirements

The objectives of the problem are with respect to the perspective of the upstream operator managing the production system. Hence, the goal is to formulate a production policy such that the operator can optimize its objectives while simultaneously satisfying the contractual rules and customer requirements. Only operational objectives have been considered since the focus of this work is to assist operational decision making. The model is supposed to serve as a decision support tool for operators controlling the system and to plan a steady-state operation between disruptions or planned events. These events can be disruptive, e.g., a field needs to be temporarily shutdown due to a breakdown or a facility needs emergency repairs, or they can be planned, e.g., a scheduled maintenance shutdown. Due to this a multiperiod formulation is unnecessary for this problem.

Following are the requirements for the production planning model and discussion of some model features resulting from these requirements:

1. The entire network is controlled by regulating the pressure at slugcatchers. Hence, it is essential to model accurately the pressure-flowrate relationships in the trunkline network and in the wells. However, even the simplest possible expressions for pressure-flowrate relationships in gas pipelines and wells are nonlinear equalities and therefore nonconvex.

2. There are different qualities of gas (i.e., gas with different composition) in the network and hence species flowrates need to be tracked throughout the network. These introduce additional nonconvexities in the model due to bilinearities in formulating species balances since the network contains both splitters and mixers.

3. The model needs to include customer specifications: the maximum and minimum amounts to be delivered, the maximum and minimum delivery pressures, the gross heating value of the dry gas and the composition specifications.

4. The contractual rules and operational heuristics need to be included in the
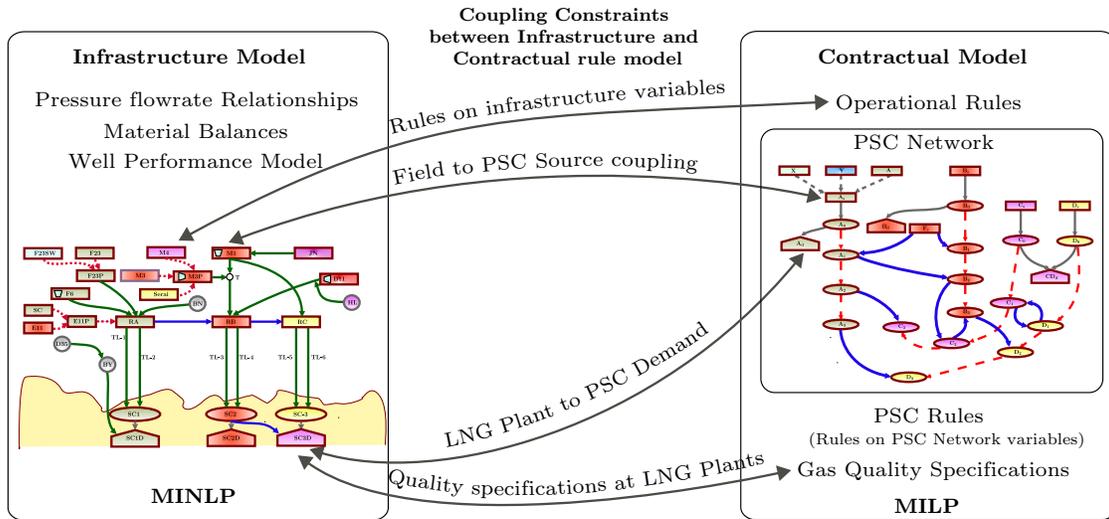
Figure 2-5: Upstream planning model: Overview of components

model. These involve logical conditions and therefore, representation of these rules requires binary variables and constraints.

5. A requirement for the model is that it be extensible so that more detailed models for facilities can be added later.

From these requirements, it is clear that the final model will be a relatively large-scale nonconvex MINLP.

2.6.2 Components

An overview of the model is presented in Figure 2-5. It is instructive to view the overall model as being the two following sub-models that are coupled:

1. **The Infrastructure Model:** This is the model of the physical system that includes wells, pipeline network and processing facilities.

2. **The Contract Model:** This is the model of the production sharing contracts and customer requirements.

Both sub-models are network models with additional constraints and hence the overall model can also be viewed as two networks whose sources and sinks are coupled as

<div align="center">**Table 2.3: Common symbols**</div>

Symbol	Description	Units	Type
$P_{(.)}$	Pressures	bar	Variable
$Q_{(.)}^{(.)}$	Dry gas volumetric rates in infrastructure model	hm^3/day^a	Variable
$Q_{L(.)}^{(.)}$	Condensate volumetric rates in infrastructure model	m^3/day	Variable
$q_{(.)}^{(.)}$	Volumetric rates in contract network	hm^3/day	Variable
$F_{(.)}^{(.)}$	Molar rates	Mmoles/day	Variable
$y_{(.)}^{(.)}$	Binary variables		Variable
$\pi_{(.)}$	Pressures	bar	Constant
$\theta_{(.)}$	Temperatures	K	Constant
$\mathcal{A}, \mathcal{B}\dots$	Sets		Set symbol
(i,j)	A directed arc from node i to j		Index
i, j, k	Indices		Index

^{a} 1 $\text{hm}^3 = 10^6 \text{ m}^3$ (since 1 hectometer $= 10^2$ m)

shown in Figure 2-5.

2.6.3 Notation

The following conventions are used in the model description:

1. The constraints are numbered continuously throughout the work (except for Section 3.6 on alternative infrastructure model formulation which has a separate numbering scheme to avoid confusion). Any non-numbered expression is not a constraint in the model and is an intermediate equality/inequality or the value of a constant or the bounds on a variable.

2. All Greek letters denote parameters in the model with the exception of the universal gas constant R.

3. Lower and upper case Roman alphabets denote decision variables.

4. Superscripts U and L to variables imply upper and lower bounds respectively.

A description of symbols that appear often in the model is presented in Table 2.3.

Chapter 3

The Infrastructure Model

The infrastructure model is a model of the physical production infrastructure. In particular, from a modeling perspective, it can be broken down into:

1. The trunkline network model: this is a model of the flow network that includes pipelines and subsea connections.

2. The well performance model: this represents pressure-flowrate relationships in the wells.

3. The compression model: the compression model is a calculation of the power consumed by compressors.

3.1   Assumptions

Following are the primary assumptions in the infrastructure model:

1. The gas is assumed to be ideal at standard conditions. Standard conditions (in the natural gas industry) are defined as the conditions at which volumetric flowrate is metered. The standard conditions in this work is taken as 15ºC and 1 atmosphere, which is close to the industry standard.

2. The reservoir pressure is assumed to be constant over the planning period. This is justified by the planning period length of a few days, over which the reservoir

pressure is not expected to change substantially.

3. The composition of the reservoir fluid for a field is assumed invariant over the planning period. This is justified by the same argument as above. This assumption implies that the composition of gas from fields and the condensate gas ratio (CGR) stays the same over the planning period.

4. Perfect mixing is assumed at junctions, since the SGPS network is operated without any preferential routing. For other systems, this assumption is easy to relax provided the exact configuration of junctions is known.

3.2   Trunkline Network Model

The trunkline network is modeled as a *directed graph*. The nodes in this graph are fields, well platforms, riser platforms and LNG plants. The trunklines and subsea connections (for platforms serving several fields) are modeled as arcs of this graph.

Let $(\mathcal{N}, \mathcal{A})$ be the directed graph representation of the trunkline network where $\mathcal{A}$ is the set containing all arcs and $\mathcal{N}$ is the set containing all nodes.

3.2.1   Flow Model

Pressure-flowrate relationships are quite important in the system because the network is controlled by regulating the pressure at certain nodes in the network. Hence, a reasonable prediction of pressure is essential for an operational planning model to be useful.

As pointed out earlier, the flow in pipelines is a two-phase mixture of gas and natural gas liquids. A full model of multiphase flow is not possible for use in conjunction with current state-of-the-art global optimization algorithms that require an explicit functional representation of constraints. Hence, the standard gas flow equation [12] (described later) is used as a flow model. The pressure drop constant in this equation can be estimated from historical operating data. It has been observed that this relationship works well for long trunklines (> 20 km) under steady-state operation.
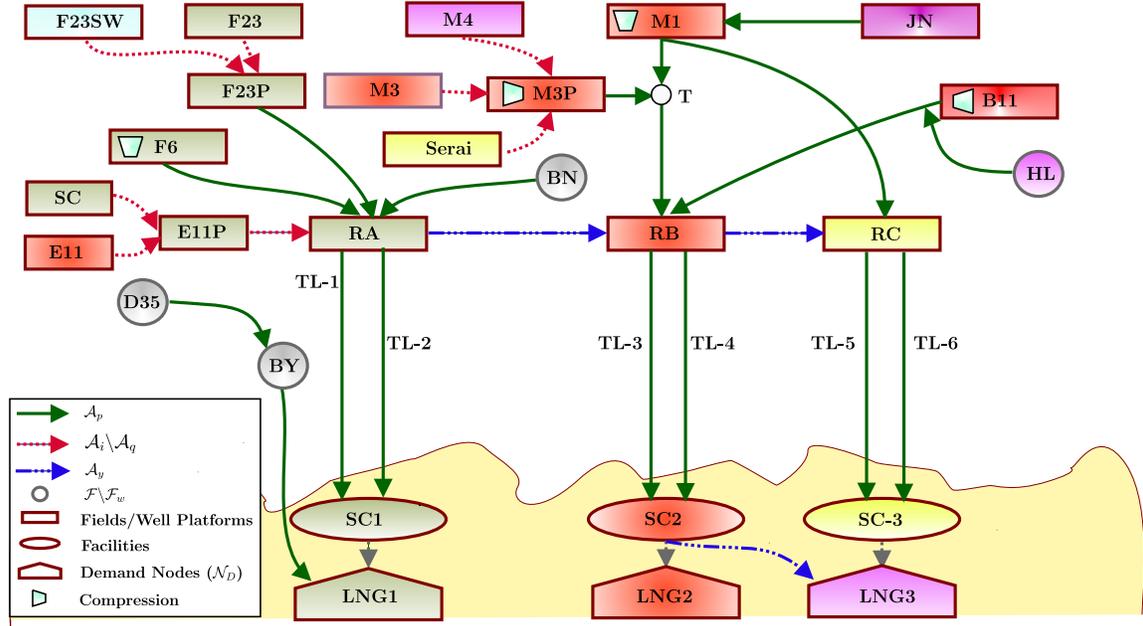
Figure 3-1: The SGPS trunkline network model: Directed graph representation
.

However, for short pipeline sections or for pipelines where the flow fluctuates a lot, the predictions are not satisfactory and hence these are not modeled using this equation.

A representation of the network as in the model is presented in Figure 3-1. One should note the differences in topology of the networks represented in Figure 2-4 and Figure 3-1. These differences facilitate easy representation of the pressure and flowrates constraints in a directed graph framework. The set of arcs $\mathcal{A}$ is partitioned into four subsets for the purposes of modeling the flow. Set $\mathcal{A}_q \subset \mathcal{A}$ denotes the set of arcs over which a volumetric flowrate variable $Q_{a,(i,j)}$ is defined.

1. For most trunklines, the flow is described by the standard gas pressure-flowrate relationship [12]. This set is denoted by $\mathcal{A}_p \subset \mathcal{A}_q$. Therefore for this set

$$P_i^2 - P_j^2 = \kappa_{(i,j)}\, Q_{a,(i,j)}^2, \quad \forall (i,j) \in \mathcal{A}_p, \tag{1}$$

where P_i and P_j are pressures at the inlet and outlet, respectively, and $Q_{a,(i,j)}$ is the volumetric flowrate at standard conditions. This equation is one of the major sources of nonconvexity in the model as it is a nonlinear equality. The

coefficient $\kappa_{(i,j)}$ was estimated from the operating data for the SGPS, however it has been changed in the case study presented later as it is business sensitive information.

2. The second set $\mathcal{A}_y \subset \mathcal{A}_q$ involves pipelines that can be shut off during normal operation. These pipelines are only a few kilometers in length and have a complicated configuration (e.g., multiple valves). Hence predictions from the regressed standard gas flow equation (Equation (1)) do not match well with operating data and therefore any equation of similar form is clearly not suitable to model these lines.

These are modeled in the following way: when these lines are in the open state, a pressure inequality between the inlet and outlet is enforced and any flowrate up to the capacity of the lines is allowed. This is justified since the pressure drop across these lines is quite small (less than 1 bar) at typical flowrates in the network. When the lines are in the closed state, the pressure inequality need not be enforced and the flowrate is pinned to zero.

To represent the above mathematically, a single binary variable $y^l_{(i,j)}$ per line is introduced such that $y^l_{(i,j)} = 1$, if the line is open and 0 otherwise. The resulting constraints can be reformulated as per Glover [111] that require the introduction of two extra variables $w_{u,(i,j)}$ and $w_{d,(i,j)}$ to represent the upstream and downstream pressure respectively.

The following four constraints force $w_{u,(i,j)}$ to the upstream pressure P_i if $y^l_{(i,j)} = 1$ and to 0 if $y^l_{(i,j)} = 0$.

$$P_i - (1 - y^l_{(i,j)})P_i^U - w_{u,(i,j)} \leq 0, \quad \forall (i,j) \in \mathcal{A}_y, \tag{2}$$

$$w_{u,(i,j)} - P_i + (1 - y^l_{(i,j)})P_i^L \leq 0, \quad \forall (i,j) \in \mathcal{A}_y, \tag{3}$$

$$y^l_{(i,j)}P_i^L - w_{u,(i,j)} \leq 0, \quad \forall (i,j) \in \mathcal{A}_y, \tag{4}$$

$$w_{u,(i,j)} - y^l_{(i,j)}P_i^U \leq 0, \quad \forall (i,j) \in \mathcal{A}_y. \tag{5}$$

Similarly, the following four constraints force $w_{d,(i,j)}$ to the downstream pressure

P_j if $y^l_{(i,j)} = 1$ and to 0 if $y^l_{(i,j)} = 0$.

$$P_j - (1 - y^l_{(i,j)})P^U_j - w_{d,(i,j)} \leq 0, \quad \forall(i,j) \in \mathcal{A}_y, \tag{6}$$

$$w_{d,(i,j)} - P_j + (1 - y^l_{(i,j)})P^L_j \leq 0, \quad \forall(i,j) \in \mathcal{A}_y, \tag{7}$$

$$y^l_{(i,j)}P^L_j - w_{d,(i,j)} \leq 0, \quad \forall(i,j) \in \mathcal{A}_y, \tag{8}$$

$$w_{d,(i,j)} - y^l_{(i,j)}P^U_j \leq 0, \quad \forall(i,j) \in \mathcal{A}_y. \tag{9}$$

Finally, the following inequality represents the actual constraint relating the upstream and downstream pressure if $y^l_{(i,j)} = 1$. If $y^l_{(i,j)} = 0$, this constraint evaluates to 0 and hence is irrelevant

$$w_{d,(i,j)} - w_{u,(i,j)} \leq 0, \quad \forall(i,j) \in \mathcal{A}_y. \tag{10}$$

Additionally, the following constraint forces the flowrate $Q_{a,(i,j)}$ to zero if $y^l_{(i,j)} = 0$ and otherwise keeps it within its bounds.

$$y^l_{(i,j)}Q^L_{a,(i,j)} \leq Q_{a,(i,j)} \leq y^l_{(i,j)}Q^U_{a,(i,j)}, \quad \forall(i,j) \in \mathcal{A}_y. \tag{11}$$

3. There is a third set of arcs $\mathcal{A}_{sc} \subset \mathcal{A}_q$ for which a constant pressure drop is assumed. This is because these arcs actually represent slugcatchers in LNG plants. The operational data suggest that the pressure drop across slugcatchers is constant. Then the pressure relationship between inlet and outlet pressures is simply

$$P_i - P_j = \Delta\pi_{(i,j)}, \quad \forall(i,j) \in \mathcal{A}_{sc}, \tag{12}$$

where $\Delta\pi_{(i,j)}$ is the pressure drop associated with the facility.

4. There is a set of trunklines which can be adequately represented by molar balances and pressure inequalities. These trunklines are subsea connections from fields to the well platforms (that serve multiple fields). This is justified since all these lines have chokes that reduce the pressure to the common header

level. Hence any pressure drop modeling for these lines is not important. These features imply that a flowrate variable need not be defined on these arcs since the sources at their origins (subsea fields) and ends (well platforms modeled as sources) can be directly related. Also there is no need to represent explicitly this subset since the constraints are already represented in the common molar balances and pressure inequalities.

Finally, a pressure inequality constraint is enforced over $\mathcal{A}_i \subset \mathcal{A}$, the set of all trunklines for which pressure at the inlet should always be greater than the pressure at the outlet. This is redundant for arcs in the set $\mathcal{A}_p \subset \mathcal{A}_i \cap \mathcal{A}_q$ (since it follows directly from Equation (1)), but may be useful for strengthening relaxations.

$$P_j - P_i \leq 0, \quad \forall (i,j) \in \mathcal{A}_i. \tag{13}$$

3.3 Material Balances

The material balances are formulated in terms of molar flowrates of chemical species. This facilitates modeling of multiple qualities of gas (i.e., gas with different compositions) in the network. Eight species are modeled. The set of species is denoted by $\mathcal{S}$:

$$\mathcal{S} = \{\mathrm{CO_2}, \mathrm{N_2}, \mathrm{H_2S}, \mathrm{C_1}, \mathrm{C_2}, \mathrm{C_3}, \mathrm{C_4}, \mathrm{C_5+}\}.$$

The balances need to be formulated separately for junctions and for nodes that are sources or sinks. Let $\mathcal{N}_J$ be the set of nodes that are junctions. The set of nodes that are sources and sinks, i.e., that have volumetric production rate $Q_{s,i}$ and componentwise molar production rate $F_{s,i,k}$ associated with them is denoted as $\mathcal{N}_s$. However not all nodes in this set form the origin or destination of arcs in set $\mathcal{A}_q$, the reason being that production from some of the sources is transferred to other sources directly and therefore the former nodes need not be on the sub-network defined by $\mathcal{A}_q$. The sources that are directly connected to this sub-network are denoted as set $\mathcal{N}_q$. With these definitions, the species molar balances at nodes can be easily formulated

as:

$$F_{s,i,k} + \sum_{v:(v,i)\in\mathcal{A}} F_{a,(v,i),k} - \sum_{v:(i,v)\in\mathcal{A}} F_{a,(i,v),k} = 0, \quad \forall(i,k) \in \mathcal{N}_q \times \mathcal{S}, \qquad (14)$$

$$\sum_{v:(v,i)\in\mathcal{A}} F_{a,(v,i),k} - \sum_{v:(i,v)\in\mathcal{A}} F_{a,(i,v),k} = 0, \quad \forall(i,k) \in \mathcal{N}_J \times \mathcal{S}, \qquad (15)$$

where $F_{a,(i,v),k}$ denotes componentwise molar flowrate on arc $(i,v) \in \mathcal{A}_q$ and $F_{s,i,k}$ denote the componentwise production rate from source $i \in \mathcal{N}_q$.

Let $\mathcal{F} \subset \mathcal{N}_s$ be the set of all fields and $\mathcal{N}_{wp} \subset \mathcal{N}_s$ be the set of well platforms. A node corresponding to a well platform that serves only one field and the node corresponding to that field are the same node and lie in $\mathcal{N}_{wp} \cap \mathcal{F}$. Denote the set of well platforms that serve multiple fields as $\mathcal{N}_{wp,m}$. Then production at well platforms in the set $\mathcal{N}_{wp,m}$ is given as

$$\sum_{j\in\mathcal{F}_i} F_{s,j,k} - F_{s,i,k} = 0, \quad \forall(i,k) \in \mathcal{N}_{wp,m} \times \mathcal{S}, \qquad (16)$$

where set $\mathcal{F}_i$ is the set of fields connected to the platform $i \in \mathcal{N}_{wp,m}$.

3.3.1 Relationship with Volumetric Flowrate

The relationship between molar flowrate in arcs and volumetric flowrate is formulated using the ideal gas assumption. The total molar flowrate in an arc is proportional to the volumetric flowrate:

$$\sum_{k\in\mathcal{S}} F_{a,(i,j),k} - \phi\, Q_{a,(i,j)} = 0, \quad \forall(i,j) \in \mathcal{A}_q, \qquad (17)$$

where ϕ is given by ideal gas equation of state:

$$\phi = 10^5\, \frac{\pi_{sc}}{R\theta_{sc}},$$

where π_{sc} and θ_{sc} are the pressure and temperature, respectively, at standard conditions. At fields, a relationship between the molar production rate and volumetric

flow rate is written:

$$F_{s,i,k} - \chi_{i,k}\phi\, Q_{s,i} = 0, \quad \forall (i,k) \in \mathcal{F} \times \mathcal{S}, \tag{18}$$

where $\chi_{i,k}$ is the mole fraction of species k in gas from field i. For well platforms that serve multiple fields, i.e., set $\mathcal{N}_{wp,m}$, and for demand nodes, i.e., set $\mathcal{N}_D$, the total molar rate should match the volumetric rates:

$$\sum_{k \in \mathcal{S}} F_{s,i,k} - \phi\, Q_{s,i} = 0, \quad \forall i \in \mathcal{N}_{wp,m}, \tag{19}$$

$$\sum_{k \in \mathcal{S}} F_{s,i,k} - \phi\, Q_{s,i} = 0, \quad \forall i \in \mathcal{N}_D. \tag{20}$$

3.3.2   Splitters and Mixers

There are both mixers and splitters of streams with unknown composition in the network and hence bilinearities associated with models of splitting or mixing cannot be avoided. There are at least two approaches to modeling a combination of splitters and mixers. The first is to model the balances in terms of species-wise molar flowrates at all nodes and split fractions at splitters. In this case, nodes that are mixers do not require any special treatment (and are linear) while splitters require nonconvex bilinear constraints to model them. The second approach is to model the balances in terms of total flowrates and species-wise compositions in which case splitters become linear while mixers now require bilinear constraints to calculate composition. Since upstream networks are collection networks that converge as the demand node gets closer, they are expected to contain more mixers than splitters. Most splitters in upstream networks exist to provide interconnections and bypasses between parts of network. The first approach is the preferred one that makes balances at mixers linear and therefore, minimizes the number of nonconvex (bilinear) terms associated with network balances.

Let $\mathcal{N}_x$ be the set of splitters. Define

$$\mathcal{N}_{x,J} = \mathcal{N}_x \cap \mathcal{N}_J$$
$$\mathcal{N}_{x,q} = \mathcal{N}_x \cap \mathcal{N}_q.$$

Define a subset of arcs that are immediately downstream of splitters, i.e., $\forall (i,j) \in \mathcal{A}_q$ such that $i \in \mathcal{N}_x$. $s_{(i,j)}$ is the split fraction defined over a subset $\mathcal{A}_x$ of this set. $\mathcal{A}_x$ is defined by excluding exactly one arc corresponding to each splitter from the set defined above. $s_{(i,j)}$ varies between zero and one and represents the fraction that goes into arc (i,j) of the total flow coming into the splitter. It is not defined over one of the arcs downstream of a particular splitter since flow in that arc is implied by the molar balance constraints (equations (14) and (15)). At splitters that are junctions:

$$F_{a,(i,j),k} - s_{(i,j)} \sum_{v:(v,i)\in\mathcal{A}} F_{a,(v,i),k} = 0, \quad \forall (i,j,k) \in \{\mathcal{A}_x : i \in \mathcal{N}_{x,J}\} \times \mathcal{S}. \qquad (21)$$

For splitters with a source term,

$$F_{a,(i,j),k} - s_{(i,j)} \left(\sum_{v:(v,i)\in\mathcal{A}} F_{a,(v,i),k} + F_{s,i,k} \right) = 0, \quad \forall (i,j,k) \in \{\mathcal{A}_x : i \in \mathcal{N}_{x,q}\} \times \mathcal{S}. \qquad (22)$$

The above equalities are valid only with a perfect mixing assumption. However, this framework can be easily extended to represent preferential routing and blending in the network by defining mixing fractions for a particular outgoing arc over incoming arcs (instead of splitting fraction on outgoing arc), i.e., each outgoing arc can choose a fraction of different qualities of incoming gas. However, such a formulation is not presented as it is not the case for the SGPS.

The constraints arising from the models of splitters and mixers are the same as the classical pooling problem ([10, 11]). Therefore, some of the customized solution strategies that have been developed for the pooling problem may also be applied to this problem.

Figure 3-2: The well-performance model

.

3.4 Well-performance Model

Let $\mathcal{W}$ denote the set of all wells. The well-performance model comprises:

1. The well flow model.

2. The well material balances.

An overview of the well-performance model is shown in Figure 3-2.

3.4.1 Well Flow Model

Variation of the reservoir pressure $\pi_{r,w}$ occurs in the vicinity of wells even if they belong to the same field. This variation corresponds to a pressure distribution in the

reservoir that can be assumed invariant over a period of a few days. There are two pressures associated with each producing well. The *bottom-hole pressure* $P_{b,w}$ is the pressure at the bottom of the well bore. The *flowing tubing-head pressure* $P_{t,w}$ is the pressure at the well head.

Following are the two relations that relate the well dry gas production rate $Q_{w,w}$ to the pressures $P_{b,w}$, $P_{t,w}$ and $\pi_{r,w}$:

1. **In-Flow Performance (IFP):** This models the flow from the reservoir bulk to the well bore [12]:

$$\alpha_w Q_{w,w} + \beta_w Q_{w,w}^2 = \pi_{r,w}^2 - P_{b,w}^2, \quad \forall w \in \mathcal{W}. \tag{23}$$

Here α_w is Darcy's constant and β_w is the non-Darcy correction factor for modeling gas flow.

2. **Vertical-lift Performance (VLP):** This models flow in the well bore itself:

$$\vartheta_w Q_{w,w}^2 = P_{b,w}^2 - \lambda_w P_{t,w}^2, \quad \forall w \in \mathcal{W}. \tag{24}$$

Additionally, constraints can be forced on the pressures from physical considerations. The bottom-hole pressure should be less than the reservoir pressure for all wells:

$$P_{b,w} - \pi_{r,w} \leq 0, \quad \forall w \in \mathcal{W}. \tag{25}$$

The tubing-head pressure must be less than the bottom-hole pressure for all wells:

$$P_{t,w} - P_{b,w} \leq 0, \quad \forall w \in \mathcal{W}. \tag{26}$$

Implicit-choke Assumption: This assumption on the well head implies that the pressure at the common header (into which all wells produce) must be less than the flowing tubing-head pressure for all wells connected to that header. In reality, this is achieved by a choke valve at each well head, however an explicit model of the choke valve is not considered here. The implicit choke provides the required pressure drop.

This constraint needs to be formulated differently for platforms that have compression and those that produce directly into the trunkline network.

Let set $\mathcal{N}_{wp,c} \subset \mathcal{N}_{wp}$ denote the platforms that have compression. Note that $\mathcal{N}_{wp,c} \subset \mathcal{F}_w \cup \mathcal{N}_{wp,m}$, i.e., the nodes that represent platforms with compression are either well platforms serving a single field and therefore are the same as the node representing that particular field (as indicated earlier) or they are well platforms serving multiple fields which require compression. For these platforms, the compression inlet pressure $P_{c,i}$ is the common header pressure and this pressure must be less than or equal to the well-head pressures of the wells producing to that platform:

$$P_{c,i} - P_{t,w} \leq 0, \quad \forall w \in \mathcal{W}_i, \quad i \in \mathcal{N}_{wp,c}. \tag{27}$$

For the set of fields for which well performance is modeled and that have no compression, i.e., the set $\mathcal{F}_{w,nc} \subset \mathcal{F}_w$, the common header pressure is the same as the pressure P_i of the node corresponding to the field,

$$P_i - P_{t,w} \leq 0, \quad \forall w \in \mathcal{W}_i, \quad i \in \mathcal{F}_{w,nc}. \tag{28}$$

3.4.2   Well Material Balances

Wells produce a mixture of gas, natural gas liquids (also termed *condensates*) and water. The NGL volume produced from a well is directly proportional to the volume of dry gas produced from that well and the condensate gas ratio σ_w is assumed constant. This can be justified partially by the assumption on the constant composition of the reservoir fluid. Then the NGL production rate $Q_{Lw,w}$ from a well w is given as

$$Q_{Lw,w} = \sigma_w Q_{w,w}, \quad \forall w \in \mathcal{W}. \tag{29}$$

Finally total dry gas production from a field (for which well performance is modeled, i.e., it is in set $\mathcal{F}_w$), is the sum of productions from all wells that belong to that field

(set $\mathcal{W}_i$):

$$Q_{s,i} = \sum_{w \in \mathcal{W}_i} Q_{w,w}, \quad \forall i \in \mathcal{F}_w. \tag{30}$$

The same is true for the total NGL production:

$$Q_{Ls,i} = \sum_{w \in \mathcal{W}_i} Q_{Lw,w}, \quad \forall i \in \mathcal{F}_w. \tag{31}$$

However, the transport of NGL through trunklines is not modeled due to impracticality of modeling multiphase flow through trunklines in a global optimization framework as discussed earlier. It is assumed that NGL produced can be transported to the demand nodes and that the transport of NGL does not limit the transfer of dry gas.

3.5   Compression Model

It is assumed that the compression equation is given by the polytropic work of compression. The outlet pressure corresponds to the pressure of the node. Then the power of compression in MW is given by

$$W_i = \omega_i \, Q_{s,i} \left[\left(\frac{P_i}{P_{c,i}} \right)^{\nu} - 1 \right], \quad \forall i \in \mathcal{N}_{wp,c}, \tag{32}$$

where the constant ω_i is given by

$$\omega_i = \frac{1}{\eta_i} \frac{\zeta}{\zeta - 1} \frac{\pi_{sc}}{R\theta_{sc}} R\theta_{m,i} \frac{1}{\tau_{sec}} = \frac{1}{\eta_i} \frac{\pi_{sc}}{\theta_{sc}} \theta_{m,i} \frac{1}{\nu t_{sec}}, \quad \forall i \in \mathcal{N}_{wp,c},$$

and ν is given as

$$\nu = \frac{\zeta - 1}{\zeta}.$$

Here π_{sc} and θ_{sc} are the pressure and temperature respectively at standard conditions, R is the universal gas constant, η_i is the compression efficiency, $\theta_{m,i}$ is the mean operating temperature of compression and ζ is the polytropic constant for the process.

The power is constrained by the maximum rated power of a compressor as follows

$$\Psi_i^L \leq W_i \leq \Psi_i^U, \quad \forall i \in \mathcal{N}_{wp,c}.$$

Note that this is not treated as a constraint but instead as a bound. When the value of W_i at the optimal solution or at an intermediate point in the solution procedure is zero, the compression constraint (32) will be an equilibrium constraint. To avoid this, the lower bound Ψ_i^L is set to a strictly positive small nonzero value. Physically, this means that compression stations should never be shut down and therefore the lowest production flowrates from these fields cannot go to zero, but only to a small value that ensures that the compressors are operating at their minimum power.

3.6   Alternative Formulation

The model formulation presented in this section is an alternative formulation for the production infrastructure model. The results and the case study presented later does not correspond to this infrastructure representation because this formulation was implemented only for the actual SGPS parameters and therefore is business sensitive. The main advantage of this formulation is that it is simpler to follow and maintain, however it suffers from a slightly larger number of variables along with a larger number of equalities. This formulation also incorporates complex platform configurations and flow reversals in certain trunklines in the network. There are also minor variations on the relationships used for pressure-flowrate relationship.

The infrastructure network for this alternative formulation is different from Figure 3-1. It is larger than the network shown in Figure 3-1 with more fields. Also, it has additional complicating features such as complex platform configurations and flow reversals. Hence, this formulation also adds additional constructs to the infrastructure modeling framework.

3.6.1 Notation

The notation in this section is not referred to outside this discussion (not even in the nomenclature to avoid conflict of notation). Hence, any references to symbols outside this section (including appendices) refer exclusively to the formulation of infrastructure model presented earlier (unless stated explicitly to refer to this section). Set definitions in this section are independent of the previous formulation. Nevertheless, the variable, set and parameter naming conventions are in line with Table 2.3. Also, most variable and parameter symbols are consistent with and retain the same or a similar meaning as the earlier formulation. The constraints in this formulation are numbered independently of the previous numbering scheme.

3.6.2 The Network Model

The entire network is represented in this formulation as a directed graph $(\mathcal{N}, \mathcal{A})$ as earlier. However, the definition of arcs and nodes differs from the earlier formulation. Arcs in this formulation consists of not only trunklines, but also facilities and subsea connections. The nodes consist of common field headers (to which a well produces), well platforms, facility inlets and outlets, and junctions.

Arc Operating Equations

Each arc in this formulation is associated with one or more operating equations depending on whether it is a trunkline, a facility or a subsea link. These operating equations establish relationships between inlet and outlet pressures, and volumetric flowrate associated with a particular arc.

1. Modeling of pressure-flowrate relationship in trunklines corresponding to set $\mathcal{A}_p$ in the earlier formulation is more or less similar. Equation (1), the standard flowrate-pressure relationship for gas flow, is used to model flow in most trunklines (set $\mathcal{A}_{sp}$) as below:

$$P_i^2 - P_j^2 = \kappa_{(i,j)} \, Q_{a,(i,j)}^2, \quad \forall (i,j) \in \mathcal{A}_{sp}, \tag{i}$$

97

However, for some trunklines (subset $\mathcal{A}_{tp}$), a two parameter relationship provides a better fit to the data.

$$\kappa_{u,(i,j)} P_i^2 - \kappa_{d,(i,j)} P_j^2 = Q_{a,(i,j)}^2, \quad \forall (i,j) \in \mathcal{A}_{tp}, \tag{ii}$$

2. Lines that can be closed or opened in normal operation (set $\mathcal{A}_y$) are modeled by a binary variable as earlier. However, Glover's reformulation has been dropped in this formulation in favor of a direct bilinear formulation. There is at least some evidence that a direct formulation gives a stronger relaxation and a faster convergence within the branch-and-reduce algorithm.

$$y_{(i,j)}^l (P_j - P_i) \leq 0, \quad \forall (i,j) \in \mathcal{A}_y. \tag{iii}$$

Note that this formulation may be sensitive to how the constraint is input for some algorithms, i.e., depending on the implementation, $y_{(i,j)}^l (P_j - P_i)$ may be treated differently from $y_{(i,j)}^l P_j - y_{(i,j)}^l P_i$. The former should give rise only to a single bilinear term in a good implementation, while the latter may be relaxed as two bilinear terms. Also, constraints on flowrate are formulated using species-wise molar flowrates instead of volumetric flowrate in the earlier formulation. This gives rise to a higher number of constraints per binary variable involved that may potentially result in a tighter relaxation.

$$y_{(i,j)}^l F_{a,(i,j),k}^L \leq F_{a,(i,j),k} \leq y_{(i,j)}^l F_{a,(i,j),k}^U, \quad \forall (i,j,k) \in \mathcal{A}_y \times \mathcal{S}. \tag{iv}$$

3. There are trunklines in this formulation that can be reversed, the set $\mathcal{A}_r$. It is assumed that (i,j) is the normal direction of the flow corresponding to $y_{(i,j)}^r = 0$ and $y_{(i,j)}^r = 1$ implies a reversal of flow, i.e., flow in direction (j,i). The flow reversal is represented as negative flow in the (i,j) direction and the standard network balance formulation is therefore valid for the reversal.

The pressure inequality must switch to allow for reversal.

$$P_j - P_i + 2\,y^r_{(i,j)}\,(P_i - P_j) \le 0, \quad \forall (i,j) \in \mathcal{A}_r. \tag{v}$$

This can be reformulated as Glover's formulation, but as described above is kept as a bilinear term in the model. The constraints on the flowrates are enforced in terms of species-wise molar flowrates similar to the case for lines that can be open or closed in normal operation.

$$y^r_{(i,j)} F^L_{a,(i,j),k} \le F_{a,(i,j),k} \le (1 - y^r_{(i,j)}) F^U_{a,(i,j),k}, \quad \forall (i,j,k) \in \mathcal{A}_r \times \mathcal{S}. \tag{vi}$$

Here the lower bound is the negative of the upper bound on reversed flow, i.e., $F^L_{a,(i,j),k} = -|F^U_{a,(j,i),k}|$ and therefore $F^L_{a,(i,j),k} < 0$. This the only constraint where a flowrate bound is allowed to go negative in the infrastructure model.

4. Compressors are represented as the set $\mathcal{A}_c \subset \mathcal{A}$ with the similar equation for the calculation of power as Equation (32) in the earlier framework.

$$W_{(i,j)} = \omega_{(i,j)}\, Q_{a,(i,j)} \left[\left(\frac{P_j}{P_i} \right)^\nu - 1 \right], \quad \forall (i,j) \in \mathcal{A}_c. \tag{vii}$$

The premultiplying factor $\omega_{(i,j)}$ for the compression equation is defined similarly as in Section 3.5:

$$\omega_{(i,j)} = \frac{1}{\eta_{(i,j)}} \frac{\pi_{sc}}{\theta_{sc}} \theta_{m,(i,j)} \frac{1}{\nu t_{sec}}, \quad \forall (i,j) \in \mathcal{A}_c.$$

5. The slugcatchers are modeled with fixed pressure drop as earlier.

$$P_i - P_j = \Delta \pi_{(i,j)}, \quad \forall (i,j) \in \mathcal{A}_{sc}, \tag{viii}$$

Except for compressors, reversible trunklines and lines that can be closed or open, all other trunklines and connector lines must have a pressure inequality enforced
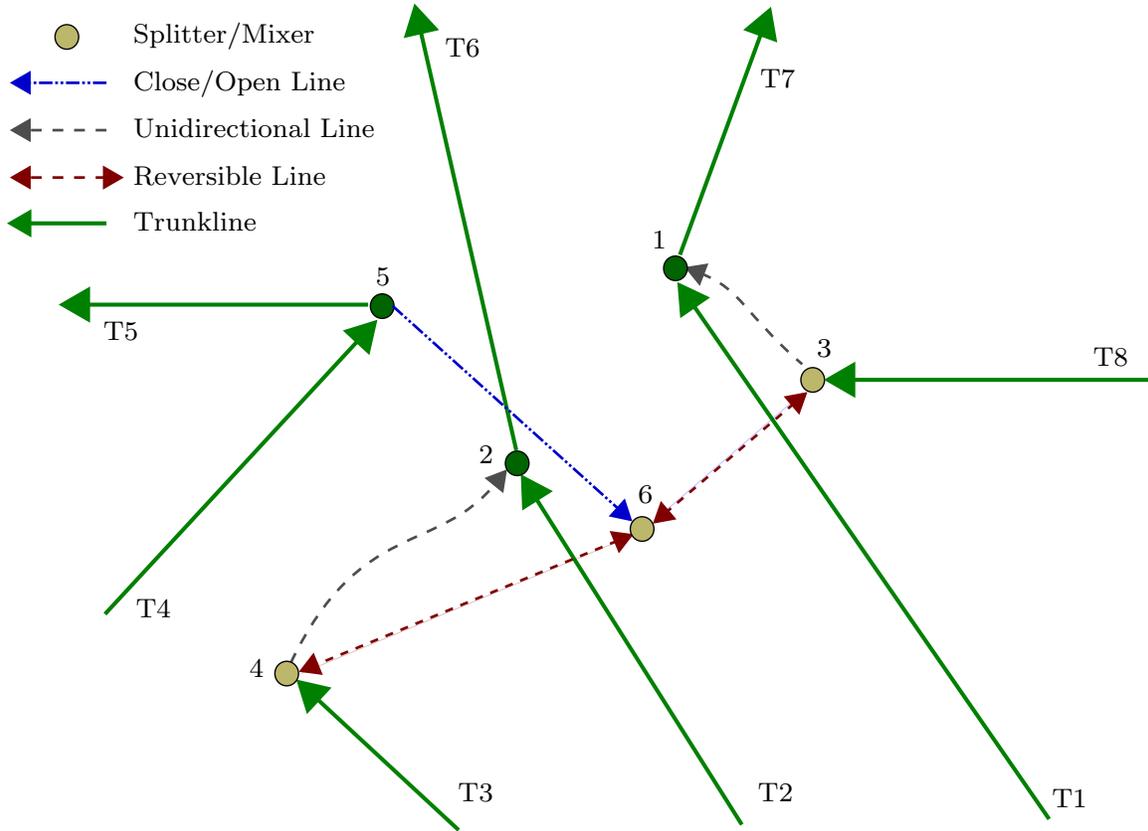
Figure 3-3: Complex hub configuration: An example
.

between the inlet and the outlet.

$$P_j - P_i \leq 0, \quad \forall (i,j) \in \mathcal{A}_i. \tag{ix}$$

3.6.3 Complex Platform Configuration: An Example

Complex hubs like the one shown in Figure 3-3 collect gas from several fields and can play an important role in blending sour gas with sweet gas. Hubs of this kind provide flexibility to network operations by routing gas across different parts of the network. A realistic representation of these hubs in the upstream planning problem is important so that the model accounts for the complexity and flexibility of the system and the solution represents an implementable routing in the network.

The modeling approach for the complex hub shown in Figure 3-3 is outlined here.

There are four types of lines in the hub shown. There are normal long-distance trunklines that terminate (or originate) at the hub and are modeled by the pressure-flowrate relationships. There are unidirectional links that connect junctions on the hub that are always open (i.e., have no controls on them) and can be adequately represented by pressure inequalities and molar balances. There are links that carry unidirectional flow and can be closed in normal operation. They can be modeled using Equations (iii) and (iv). Finally, there are lines that can carry reverse flow in normal operation if it is desirable and possible (i.e., pressures at the end-points are favorable to reversal) to do so. Pressure constraints and flowrate constraints on the reversible links are the same as Equations (v) and (vi). Except for the trunklines ending or originating at the hub, all the other links are short inter-connectors that tie major trunkline junctions and therefore any pressure modeling on them is not important.

As shown, lines (3,6) and (6,4) can be reversed. Therefore, depending on the direction of flow in lines (3,6) and (6,4), nodes 4, 6 and 3 can be either splitters or mixers. The splitters' constraints modeling is based on the same arguments as Section 3.3.2 with species-wise molar flowrate balances at nodes and the definition of split fractions at splitters. The splitter constraints are switched on and off as follows:

1. Node 3 is a splitter when (3,6) is flowing in the arc direction (i.e., direction (3,6)) which implies that $y^r_{(3,6)} = 0$ should force the splitter constraint:

$$F_{a.(3,1),k} - s_{(3,1)}\, F_{a,(\text{T8}),k} - y^r_{3,6}\, S^U_{(3,1),k} \leq 0, \quad \forall k \in \mathcal{S} \tag{x}$$

$$y^r_{(3,6)}\, S^L_{(3,1),k} - F_{a.(3,1),k} + s_{(3,1)} F_{a,(\text{T8}),k} \leq 0, \quad \forall k \in \mathcal{S} \tag{xi}$$

where $S^L_{(3,1),k}$ and $S^U_{(3,1),k}$ are lower and upper bounds on the $F_{a.(3,1),k} - s_{(3,1)}\, F_{a,(\text{T8}),k}$ that are derived from the corresponding flow bounds,

$$S^L_{(3,1),k} = F^L_{a,(3,1),k} - F^U_{a,(\text{T8}),k}, \quad \forall k \in \mathcal{S},$$

$$S^U_{(3,1),k} = F^U_{a,(3,1),k} - F^L_{a,(\text{T8}),k}, \quad \forall k \in \mathcal{S}.$$

The expression for $S^U_{(3,1),k}$ (and similar expressions later) is valid only for $F^L_{a,(\text{T8}),k} \geq$

0 and $F^U_{a,(T8),k} \geq 0$.

2. Modeling for node 4 is almost identical except for the fact that node 4 is a splitter when (4,6) reverses, i.e., splitter condition corresponds to $y^r_{(6,4)} = 1$,

$$F_{a,(4,2),k} - s_{(4,2)} F_{a,(T3),k} - \left(1 - y^r_{(6,4)}\right) S^U_{(4,2),k} \leq 0, \quad \forall k \in \mathcal{S}, \qquad \text{(xii)}$$

$$\left(1 - y^r_{(6,4)}\right) S^L_{(4,2),k} - F_{a,(4,2),k} + s_{(4,2)} F_{a,(T3),k} \leq 0, \quad \forall k \in \mathcal{S} \qquad \text{(xiii)}$$

with a similar definition of $S^L_{(4,2),k}$ and $S^U_{(4,2),k}$ as before,

$$S^L_{(4,2),k} = F^L_{a,(4,2),k} - F^U_{a,(T3),k}, \quad \forall k \in \mathcal{S},$$

$$S^U_{(4,2),k} = F^U_{a,(4,2),k} - F^L_{a,(T3),k}, \quad \forall k \in \mathcal{S}.$$

3. Modeling is more complicated for node 6. This is because the splitter mode of node 6 cannot be directed related to a single binary variable. An additionally binary variable is defined to indicate if this node is a splitter. $y^s_6 = 1$ if node 6 is a splitter and zero otherwise. Node 6 is a splitter if line (3,6) reverses and line (6,4) flows normally. This condition can be represented logically as $R_{36} \wedge \neg R_{64} \Rightarrow S_6$ where $R_{(.)}$ is an atomic proposition, true when the flow is reversed in the corresponding arc and S_6 is atomic proposition that is true when the node 6 is splitter. This can be converted to a binary constraint as explained in Section 4.3.10 to obtain the following binary constraints:

$$y^r_{(3,6)} - y^r_{(6,4)} - y^s_6 \leq 0. \qquad \text{(xiv)}$$

The formulation then is identical to the above:

$$F_{a,(6,4),k} - s_{(6,4)} F_{a,(5,6)} - y^s_6 S^U_{(6,4),k} \leq 0, \quad \forall k \in \mathcal{S}, \qquad \text{(xv)}$$

$$y^s_6 S^L_{(6,4),k} - F_{a,(6,4),k} + s_{(6,4)} F_{a.(5,6),k} \leq 0, \quad \forall k \in \mathcal{S}, \qquad \text{(xvi)}$$

$$\text{(xvii)}$$

with

$$S^L_{(6,4),k} = F^L_{a,(6,4),k} - F^U_{a,(5,6),k}, \quad \forall k \in \mathcal{S},$$

$$S^U_{(6,4),k} = F^U_{a,(6,4),k} - F^L_{a,(5,6),k}, \quad \forall k \in \mathcal{S}.$$

There are additional logical constraints that can be enforced for node 6. In particular, the negation of the splitter condition ($\neg(R_{36} \wedge \neg R_{64}) \Rightarrow \neg S_6$) to give the constraints:

$$y^s_6 - y^r_{(3,6)} \leq 0, \tag{xviii}$$

$$y^s_6 + y^r_{(6,4)} - 1 \leq 0. \tag{xix}$$

Finally, line (3,6) flowing in the normal direction while line (6,4) is reversed is impossible because there is no where for the incoming flow to leave node 6. The molar balances will prohibit the $\{1,0\}$ realization for $\{y^r_{(6,4)}, y^r_{(3,6)}\}$ (except for identically zero flowrate on all lines). However, a cut is added to cut off this realization explicitly,

$$y^r_{(6,4)} - y^r_{(3,6)} \leq 0. \tag{xx}$$

For the splitter molar balances, it is possible to use a direct trilinear formulation instead of converting to a bilinear formulation outlined above. For example, for node 3, the splitter constraint can be formulated as:

$$y^r_{3,6} \left(F_{a.(3,1),k} - s_{(3,1)} F_{a,(\text{T8}),k} \right) = 0, \quad \forall k \in \mathcal{S}$$

In cursory numerical experiments, this formulation does not perform satisfactorily and therefore was not considered. The performance is susceptible to how the algorithm constructs the convex relaxations of the term and a different algorithm with a better relaxation for trilinear terms may perform better with this formulation.

3.6.4 Network Balances

Simplified molar balances in the upstream network is a major advantage of this formulation. The definition of multiple sets to distinguish well platforms, field headers, compressor inlets and so on is no longer required. This is achieved at the expense of an increase in the number of variables in the molar balance formulation (and the corresponding increase in equality constraints so that the degrees of freedom stay the same). The balances are now written over the entire set of sources/sinks (set $\mathcal{N}_s$) and junctions (set $\mathcal{N}_J$) as opposed to a subset only in Equations (14) and (15):

$$F_{s,i,k} + \sum_{v:(v,i)\in\mathcal{A}} F_{a,(v,i),k} - \sum_{v:(i,v)\in\mathcal{A}} F_{a,(i,v),k} = 0, \quad \forall(i,k) \in \mathcal{N}_s \times \mathcal{S}, \tag{xxi}$$

$$\sum_{v:(v,i)\in\mathcal{A}} F_{a,(v,i),k} - \sum_{v:(i,v)\in\mathcal{A}} F_{a,(i,v),k} = 0, \quad \forall(i,k) \in \mathcal{N}_J \times \mathcal{S}. \tag{xxii}$$

The balances on the splitters (i.e., normal splitters, excluding ones encountered in complex platform configurations with flow reversals that were described in the earlier section) are the same as earlier.

$$F_{a,(i,j),k} - s_{(i,j)} \sum_{v:(v,i)\in\mathcal{A}} F_{a,(v,i),k} = 0, \quad \forall(i,j,k) \in \{\mathcal{A}_x : i \in \mathcal{N}_{x,J}\} \times \mathcal{S},$$

$$\tag{xxiii}$$

$$F_{a,(i,j),k} - s_{(i,j)} \left(\sum_{v:(v,i)\in\mathcal{A}} F_{a,(v,i),k} + F_{s,i,k} \right) = 0, \quad \forall(i,j,k) \in \{\mathcal{A}_x : i \in \mathcal{N}_{x,s}\} \times \mathcal{S}.$$

$$\tag{xxiv}$$

The definitions of sets are similar to as defined earlier. $\mathcal{N}_x$ is the set of splitters, $\mathcal{A}_x$ is the set of arcs downstream of a splitter, $\mathcal{N}_{x,s} = \mathcal{N}_x \cap \mathcal{N}_s$ and $\mathcal{N}_{x,J} = \mathcal{N}_x \cap \mathcal{N}_J$. Also this equation is only written over all except one arcs downstream of a splitter.

The relationship between molar flowrate in arcs and volumetric flowrate is formulated using the ideal gas assumption as in Section 3.3.1, however, there is no longer a need to distinguish arcs on which flowrate variables are defined as earlier and the

constraint can be formulated for all arcs without qualification:

$$\sum_{k \in \mathcal{S}} F_{a,(i,j),k} - \phi Q_{a,(i,j)} = 0, \quad \forall (i,j) \in \mathcal{A}. \tag{xxv}$$

Finally, relationships at between the molar production rate and volumetric flow rate is written for fields and demands:

$$F_{s,i,k} - \chi_{i,k} \phi Q_{s,i} = 0, \quad \forall (i,k) \in \mathcal{F} \times \mathcal{S}, \tag{xxvi}$$

$$\sum_{k \in \mathcal{S}} F_{s,i,k} - \phi Q_{s,i} = 0, \quad \forall i \in \mathcal{N}_D. \tag{xxvii}$$

Definition of ϕ is identical to Section 3.3.1.

3.6.5 Well-performance Model

The well-performance model in this formulation is similar to the one described earlier.

1. **In-Flow Performance (IFP):** This models the flow from the reservoir bulk to the well bore [12]:

$$\alpha_w Q_{w,w} + \beta_w Q_{w,w}^2 = \pi_{r,w}^2 - P_{b,w}^2, \quad \forall w \in \mathcal{W}. \tag{xxviii}$$

Here α_w is Darcy's constant and β_w is the non-Darcy correction factor for modeling gas flow.

For some high-pressure wells (with reservoir pressure > 200 bar), a linear single parameter relationship seems to provide a more consistent relationship with the data obtained from more detailed simulations:

$$\kappa_w Q_{w,w} = \pi_{r,w} - P_{b,w}, \quad \forall w \in \mathcal{W}. \tag{xxix}$$

2. **Vertical-lift Performance (VLP):** This models flow in the well bore itself:

$$\vartheta_w Q_{w,w}^2 = P_{b,w}^2 - \lambda_w P_{t,w}^2, \quad \forall w \in \mathcal{W}. \tag{xxx}$$

Additionally constraints can be forced on the pressures from physical considerations as in the earlier formulation.

$$P_{b,w} - \pi_{r,w} \leq 0, \quad \forall w \in \mathcal{W} \tag{xxxi}$$

$$P_{t,w} - P_{b,w} \leq 0, \quad \forall w \in \mathcal{W}. \tag{xxxii}$$

The *Implicit-choke assumption* is simplified since all field headers are on the network representation of the system and no special treatment is required to treat fields with compression separately.

$$P_i - P_{t,w} \leq 0, \quad \forall w \in \mathcal{W}_i, \quad i \in \mathcal{F}_w. \tag{xxxiii}$$

Finally the mass balances are the same as before.

$$Q_{Lw,w} = \sigma_w Q_{w,w}, \quad \forall w \in \mathcal{W}, \tag{xxxiv}$$

$$Q_{s,i} = \sum_{w \in \mathcal{W}_i} Q_{w,w}, \quad \forall i \in \mathcal{F}_w, \tag{xxxv}$$

$$Q_{Ls,i} = \sum_{w \in \mathcal{W}_i} Q_{Lw,w}, \quad \forall i \in \mathcal{F}_w. \tag{xxxvi}$$

Chapter 4

The Contract Modeling Framework

Contracts are central to the operation of the upstream production system. Any supply chain planning tool needs to take these rules into account to be implementable on the real system. A framework for incorporating these rules is presented here. It is useful to view contractual rules as comprising the following four subcategories:

1. *Demand rates and delivery pressures*: these are maximum and minimum supply rates and delivery pressures to the LNG plants.

2. *Gas quality specifications*: these are customer requirements, e.g., the heating value and the composition of the feed gas to the LNG plants.

3. *Production-Sharing Contracts (PSC)*: these define field to plant assignment rules that designate certain fields to supply a specific plant under certain conditions.

4. *Operational rules*: operational rules enforce conditions on the network to ensure proper operation of the system and also, to aid implementation of production-sharing rules on the network.

Strictly, operational rules are not part of the contractual framework. However, they are included here because the approach for modeling them is the same as the PSC rules and, moreover, several PSC rules also invoke operational constraints in order to implement sharing and transfer on the physical network.

To summarize, the contract modeling framework encompasses all constraints that do not originate from modeling of the physics of the production infrastructure.

4.1 Demand Rates and Delivery Pressure

Contracts enforce a maximum demand rate at each LNG plant. This reflects the maximum intake of the LNG plant and the supply should not exceed the amount specified by the contracts. Additionally, the demands are bounded from below by a minimum amount that should be supplied. Together, this defines a narrow window of operation as defined by the LNG plant operators. Since the demand nodes are sinks with a negative production rate, the bounds are reversed as follows

$$-\Lambda_{d,i}^{U} \leq Q_{s,i} \leq -\Lambda_{d,i}^{L}, \quad \forall i \in \mathcal{N}_D,$$

where $\Lambda_{d,i}^{U}$ and $\Lambda_{d,i}^{L}$ are maximum and minimum demand rates respectively. These requirements are enforced as bounds and not as constraints.

The primary means of control for the network is to regulate the pressure at slugcatchers. The LNG plants require the feed gas to be at a certain pressure for proper operation. If the pressure is too low, the gas flowrate has to be cut back to maintain the pressure. Since there is two phase flow in the pipelines, pressure regulation is also quite important to make sure that the liquids properly separate out in slugcatchers. If the pressure is too high, a lot of liquid will go through to the LNG plants which is undesirable since in normal circumstances NGL belong to the upstream operator, but in this case, LNG plants will take a share and hence a loss for the upstream operator:

$$\pi_{d,i}^{L} \leq P_i \leq \pi_{d,i}^{U}, \quad \forall i \in \mathcal{N}_{sc},$$

where $\pi_{d,i}^{L}$ and $\pi_{d,i}^{U}$ are lower and upper limits of the operating pressure at slugcatchers. This constraint is also represented as a bound in the model.

4.2 Gas Quality Specifications

The specifications on the feed gas to the LNG plants are enforced by the sales agreement between the upstream operators and the LNG plant operators. These comprise the constraints on the gross heating value of the feed gas and mole percentages of species in the feed gas. Feed gas quality directly affects LNG quality. As explained in Section 1.2.3, gas quality plays a major role in gas trade. LNG quality is usually specified by LNG customers in their long-term contracts with suppliers. Customer equipment and facilities are designed to these specifications and off-specification LNG can seriously impact and even disrupt customer operations. Therefore, off-specification LNG may have implications for reputation of LNG suppliers and may even carry an economic cost for them either as penalties or as additional costs in further processing to conform to the specifications.

An example of feed gas specifications as outlined in the contracts (similar to the SGPS in structure but completely changed otherwise to preserve sensitive information) is presented in Table 4.1 for the demand nodes in the network in Figure 3-1.

4.2.1 Gross Heating Value

The most important gas quality specification is the *Gross Heating Value* (GHV) of the feed gas to the LNG plants, since it has a direct effect on the calorific value of the LNG produced. The gross heating value of the feed gas is measured in terms of the heat of combustion per unit mass of gas (excluding CO_2 since it is separated before liquefaction of the gas). The GHV is specified in a range.

For this work, the GHV is constrained to be *higher* than the lower limit of the range. This is because a low heating value for the LNG is the major concern here and a high heating value is not considered a problem. However, in general, GHV should be constrained from both above and below as a high GHV may not be preferable for some markets (as explained in Section 1.2.3). The energy content of the gas is calculated using the superior calorific values of C_1 through C_5 at a prespecified temperature and pressure. In this work, it is assumed to be calculated at 15^oC and

Table 4.1: Customer requirements on gas quality

Specification	Constraint	$i \in \mathcal{N}_D$		
		LNG1	LNG2	LNG3
Γ_i^s (MJ/kg excl. CO_2)	Range	53.0-56.0	53.0-56.0	53.0-56.0
χ_{i,CO_2}^s	less than	5.8 mol %	5.8 mol %	6.0 mol %
χ_{i,N_2}^s	less than	1.2 mol %	1.0 mol %	1.2 mol %
χ_{i,H_2S}^s	less than	250 ppmV	270 ppmV	30 mg/m^3
$\chi_{i,S}^s$	less than	29.0 mg/m^3	22.0 mg/m^3	25.0 mg/m^3
χ_{i,C_2}^s (Excl. CO_2)	greater than	3.2 mol %	3.4 mol %	4.0 mol %
χ_{i,C_3}^s (Excl. CO_2)	greater than	2.1 mol %	2.7 mol %	3.0 mol %
χ_{i,C_4}^s (Excl. CO_2)	less than	2.4 mol %	2.7 mol %	2.7 mol %
$\chi_{i,C5+}^s$ (Excl. CO_2)	less than	1.0 mol %	1.7 mol %	1.7 mol %

1 atmosphere. The superior calorific values at 15°C and 1 atmosphere were obtained from the ISO standard for calculation of calorific values of natural gas [112, page 12, table 4]. For C_1 through C_3, the superior calorific values for methane, ethane and propane respectively are used. For C_4, the mean of superior calorific value of n-butane and 2-methylpropane is used. For heavier components, the heating value and molecular weight of n-hexane is used. The inequality representing the GHV constraint is given as:

$$\sum_{k \in \mathcal{S}_h} \gamma_k \, \mu_k F_{s,\{i,k\}} - \Gamma_i^s \sum_{k \in \mathcal{S} \backslash CO_2} \mu_i F_{s,\{i,k\}} \leq 0, \quad \forall i \in \mathcal{N}_D, \tag{33}$$

where γ_k is the superior calorific value of component k on per unit mass basis, μ_k is the molecular weight of the component, Γ_i^s is the lower limit of GHV specification at demand i and $\mathcal{S}_h$ is the set of species that are used to calculate heating value. It should be noted that the inequality is the opposite to what may seem intuitive because the demand nodes are sinks and all molar rates are negative. The heating values and molecular weights used for the GHV calculation are presented in Table B.9 in Appendix B.

4.2.2 Composition Specifications

Additional quality specifications are related to satisfying composition specifications by customers as well as maintaining the capacity and ensuring proper operation of LNG plants. For example, if the CO_2 mole fraction in the feed gas goes above a certain threshold, the plants have to cut back the total amount of feed gas being processed because the capacity of CO_2 extraction units is exceeded.

It should be noted that the inequalities are the opposite to what may seem intuitive because the molar rates are negative at demand nodes:

1. The CO_2 and N_2 mole fractions should be less than the threshold:

$$10^{-2} \chi_{i,k}^s \sum_{j \in \mathcal{S}} F_{s,i,j} - F_{s,i,k} \leq 0, \quad \forall (i,k) \in \mathcal{N}_D \times \{CO_2, N_2\}. \tag{34}$$

2. The amount of sulfur (mass per unit volume) should be below specified limits:

$$10^{-6} \chi_{i,S}^s Q_{s,i} - \mu_S F_{s,i,H_2S} \leq 0, \quad \forall i \in \mathcal{N}_D. \tag{35}$$

3. As is evident from Table 4.1, different demand nodes have different units in which the hydrogen sulfide concentration specification is expressed. The hydrogen sulfide (H_2S) concentration should be less the specified ppmV (parts per million by volume) for the set $\mathcal{N}_{D,HSP} \subset \mathcal{N}_D$:

$$10^{-6} \chi_{i,H_2S}^s \phi\, Q_{s,i} - F_{s,i,H_2S} \leq 0, \quad \forall i \in \mathcal{N}_{D,HSP}. \tag{36}$$

The hydrogen sulfide (H_2S) concentration should be less the specified mg/m^3 by volume for the set $\mathcal{N}_{D,HSM} \subset \mathcal{N}_D$:

$$10^{-6} \chi_{i,H_2S}^s Q_{p,i} - \mu_{H_2S} F_{s,i,H_2S} \leq 0, \quad \forall i \in \mathcal{N}_{D,HSM}. \tag{37}$$

4. The C_2, C_3 mole fractions excluding CO_2 should be greater than the threshold:

$$F_{s,i,k} - 10^{-2} \chi_{i,k}^s \sum_{j \in \mathcal{S} \backslash CO_2} F_{s,i,j} \leq 0, \quad \forall (i,k) \in \mathcal{N}_D \times \{C_2, C_3\}. \qquad (38)$$

5. The C_4, C_{5+} mole fraction excluding CO_2 should be less than the threshold:

$$10^{-2} \chi_{i,k}^s \sum_{j \in \mathcal{S} \backslash CO_2} F_{s,i,j} - F_{s,i,k} \leq 0, \quad \forall (i,k) \in \mathcal{N}_D \times \{C_4, C_{5+}\}. \qquad (39)$$

4.3 Production-Sharing Contracts (PSC): Modeling Framework

As was pointed out in Section 2.5.1, several parties have stakes in different parts of the system and therefore a complicated framework of *production-sharing contracts* (referred to as simply *contracts* or *PSC* in this section) exists. A plant cannot receive supply from arbitrary fields. It can only receive supply from fields that are "produced" under the "contract" authorized to supply it. On the other hand, gas from different fields is blended in the network, so the gas molecules originating from a field under a particular contract do not all end up at the LNG plant associated with this contract. Instead, only the gas volume produced by this field must be supplied to the LNG plant associated with its contract, the actual gas may come from a different field. One of the main reasons for this is that a field may be physically connected to the network in such a way that it is easier to supply a different LNG plant than the LNG plant corresponding to the producing contract, which must be then compensated from some other field. Therefore, there is a need for accounting of volumes.

A framework for modeling these contracts is presented here.

4.3.1 Terminology

Every field in the system is associated with a *PSC* (*contract*). In the industry terminology, "the field is produced under a contract." A contract can contain several

fields. The sum of productions from the fields associated with a particular contract is the *supply* of that contract. A contract is mandated to supply to a particular demand (a LNG plant). This is termed the *primary demand* of the contract. *Inter-contract transfers* are exchanges of gas volumes between different contracts. They are required because the primary demand and supply of a contract may not match. A contract is in *excess* if its supply exceeds its primary demand and it is in *deficit* otherwise. This is the *state of the contract* (*contract state*). A deficit can happen due to production network constraints or the customer specifications and many other reasons. The *inter-contract transfer rules* (*transfer rules*) are the set of rules that govern the inter-contract transfers. The transfer rules can be viewed as *deficit rules* associated with each contract that define, "which contracts and the order in which they should supply", the contract in question if it is in deficit. Alternatively, they can be interpreted as *excess rules* that dictate the order in which a contract in excess should supply other contracts in deficit. In other words, a deficit rule at the contract that is borrowing is the same as an excess rule at the contract that is in excess and supplying. Additionally, these rules may also invoke operational rules to implement the transfers on the pipeline network. When a transfer takes place as per a rule, the rule is said to have been *activated* (also termed the *state* of an inter-contract transfer (rule) is *active*).

4.3.2 Issues in Mathematical Representation

The primary challenges in representing these contracts in a mathematical programming framework are as follows.

The inter-contract transfer rules are activated by a particular contract being in deficit or excess and, in certain cases, based also on several additional conditions that represent priorities and operational concerns. Hence, they are based on logical conditions and therefore require binary variables and constraints to model them.

The rules also interact with each other. This is because to establish whether a supply of gas is available from a contract, it is not sufficient to know the primary demand and supply of the contract, but also whether other rules have activated and

borrowed from the contract under consideration.

Modeling the inference of a rule

The way a rule is stated in the operational documents for the real system may not cover all logical possibilities. This is because for a human operator, the logic and inference of a rule are obvious. Hence, the statement of a rule may not dictate the outcomes that can be easily deduced by a human.

A simple of example of this problem is as follows. Transfer rules as stated only define *sufficient* conditions for transfer activations, i.e., they are of the form, "if some condition holds, then transfer should take place." However, they may or may not be *necessary* for transfer activation and this missing information is not stated explicitly since a human operator can deduce it. Hence an exact mathematical representation of transfer rules will only model the *sufficiency* clause. This will lead to a problem when the clause was indeed *necessary*, generating feasible operations where the clause is false, but the transfer activates. This operation would be deemed infeasible by a human operator. For example, the deficit rules state that when a contract is in deficit, what other contracts must supply it. However, they do not explicitly prohibit transfers to contracts in excess since this is obvious to a human operator. A straightforward modeling of rules as stated in operational documents will result in transfers to contracts in excess. When transfer to a contract in excess is prohibited, the deficit condition is now both *sufficient* (deficit implies transfer) and *necessary* (transfer implies deficit).

Though the above example may seem obvious, for certain complex rules and scenarios, it may not be clear if an exact representation of the inference of a rule has been embedded in the model or if more logical constraints must be added to avoid infeasible scenarios at the solution. This is because activation conditions for complex transfer rules may have (subtle) logical dependencies on other rules or states and this is not always obvious from their operational statements. Additionally, for complex rules, it is not at all clear whether their activation conditions are necessary. In either case, it is usually easy to pinpoint violations in the solution, though it may be still be

114

hard to deduce the exact rule (rules) causing it. Therefore, converting a rule written in human language to an exact and equivalent set of logical clauses is challenging.

A rigorous verification as to whether a precise representation of rules has been achieved in the model may require evaluating all feasible integer realizations of the problem which is clearly impractical. The modeling approach in this work has been to add as many logical constraints as was obvious or had become apparent during the development of the model at that stage and then the solution was evaluated for any infeasibility and if it was found that solution will be deemed infeasible by a human, additional logical constraints were added to avoid such a scenario and iterate until the solution was satisfactory.

4.3.3   Contracts: Network Representation

Contractual rules interact with each other and therefore there are *levels* of *excesses* or *deficits* in a contract. *Excess* (or *deficit*) from a particular contract at the nth *level* is defined as the excess (or deficit) after n excess (or deficit) rules for that particular contract have been considered. In other word, it is the excess (or deficit) once the decision on the transfers as per n excess (or deficit) rules have taken place. It is not necessary that the transfer as per one of these rules must have taken place, i.e., the rules need not have been *activated*.

This formulation can be represented using a *directed graph* $(\mathcal{L}_l, \mathcal{E}_l)$ with nodes of the graph (set $\mathcal{L}_l$) indicating the levels in each contract and flows in arcs (set $\mathcal{E}_l$) indicating excesses or deficits between the levels. The flow in an arc is permitted to be either negative or positive as opposed to traditional network theory. It is positive if the direction of the flow is the same as the arc direction and negative otherwise. A positive flow incident out of a particular node (i.e., the flow in the direction of the arc originating at that node) indicates excess in the corresponding level. On the other hand, a negative flow indicates deficit corresponding to that level and this therefore means that the flow is incident into the node (i.e., the flow direction is opposite to the arc originating at that node) corresponding to that particular level.

Each contract has a single source node corresponding to it in the contract graph
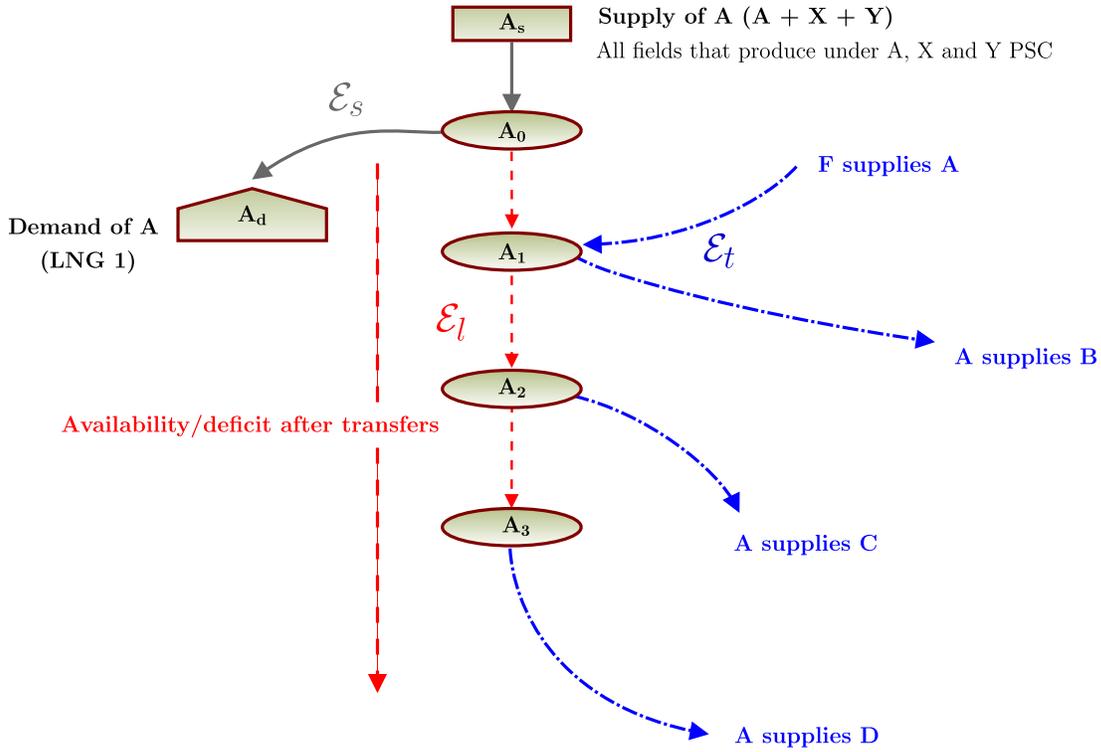
Figure 4-1: Network representation of a contract

whose production rate is equal to the sum of production rates of fields producing under that contract. Sink nodes in the infrastructure network have counterparts in the contract graph representing LNG plants. Together sources and sinks form the set $\mathcal{L}_s$. The supply and demand arcs connect the source nodes and the demand nodes respectively to the contract levels sub-graph $(\mathcal{L}_l, \mathcal{E}_l)$. Together the supply and demand arcs form the set $\mathcal{E}_s$.

All nodes that correspond to one particular contract can be grouped to indicate the levels of excess or deficit in this contract. The nodes are labeled with index of the contract and level, so that the node corresponding to contract p at level i is p_i ($\in \mathcal{L}_l$). The supply nodes and demand nodes are termed p_s and p_d ($\in \mathcal{L}_s$), respectively.

The inter-contract transfer rules can now be represented as arcs between the nodes from different contracts. For example, a transfer from contract p to q may be represented as a unique arc (p_i, q_j) on this graph where i and j are determined by the priority of that transfer over other transfers. The set of these arcs is represented by

$\mathcal{E}_t$. The subnetwork for a contract is shown in Figure 4-1.

It should be noted that the originating node p_i for a transfer (p_i, q_j) from contract p to q can be a supply node (i.e., $i = s$) under special conditions. This is because there are fields in the system that can supply to multiple production-sharing contracts and hence need to be treated as special cases. These are modeled by treating them as a separate contract that has no explicit mandated demand. In such a case, there is no need for supply or demand arcs and the states corresponding to these fields supplying one or the other contract are represented by transfer arcs originating at the corresponding supply node.

The contract network representation is given by the directed graph $(\mathcal{L}, \mathcal{E}) = (\mathcal{L}_s \cup \mathcal{L}_l, \mathcal{E}_s \cup \mathcal{E}_l \cup \mathcal{E}_t)$. The excess and deficit calculations at different levels (including inter-contract transfers) are equivalent to material balances on this network. Figure 4-2 represents an example of the contract network representation which is similar in complexity to the real system, though not identical for confidentiality reasons. It contains four contracts A, B, C and D and three demands. F corresponds to a field that can supply multiple contracts. This network is used in the case study presented in Chapter 5. Also, note the correspondence between Figure 2-4 and Figure 4-2 with fields of a particular color in Figure 2-4 belonging to the contract of the same color in Figure 4-2.

4.3.4 Excess (Deficit) Policy

The actual constraints for activating the excess/deficit rules can be formulated within the contract network framework presented in the previous section. An atomic proposition representing the contract state is formulated which is true when the contract is in excess and false otherwise.

Once a contract is in excess (or deficit) at a particular level, it should maintain that state for all further levels following it. If this is not so, there is a possibility of transfers that should not be normally permitted.

Example: Consider the following two rules that state:

117
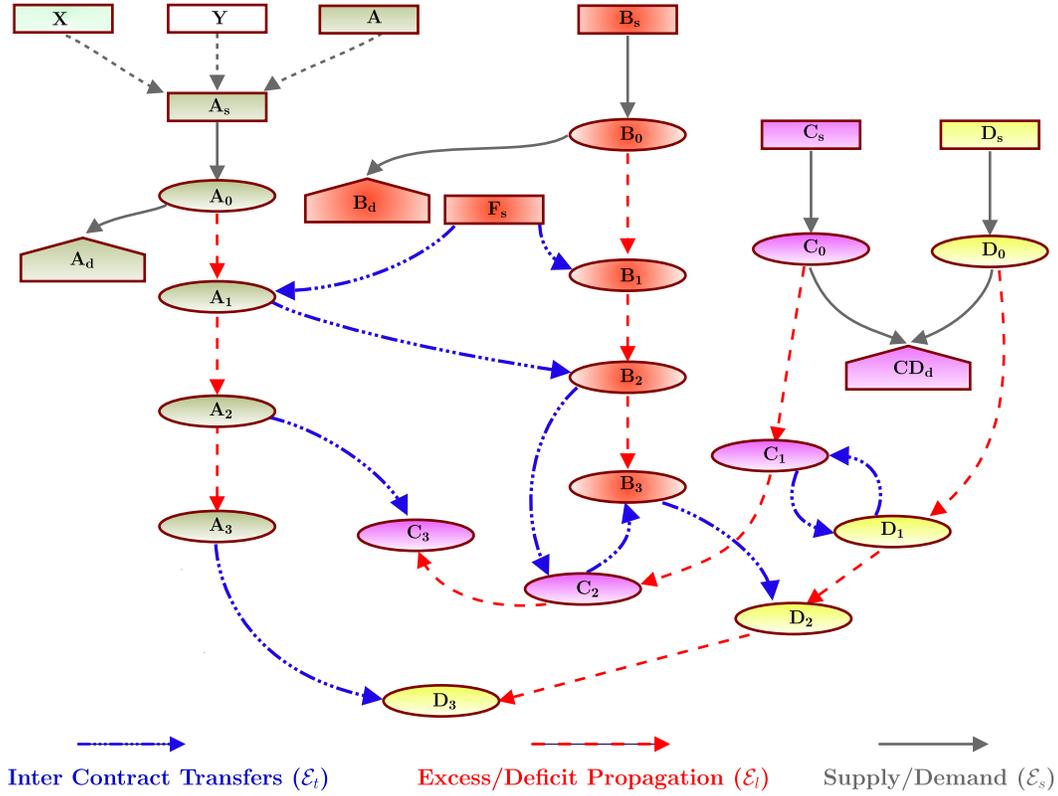
Figure 4-2: Contracts in the case study: Network representation

1. *If A is in excess and B is in deficit, then A should supply B.*

2. *If B is in excess and C is in deficit, then B should supply C*

It is possible to get a scenario, when A is in excess and both B and C are in deficit. A can supply B to make it in excess and B now can supply C. Hence A indirectly supplies C, although this was never intended by the two rules as stated. The policy stated ensures that this cannot happen. This example also demonstrates the difficulty outlined in Section 4.3.2 that a literal modeling of rules may not be sufficient to represent the inference of those rules.

The above policy implies that a single atomic proposition is needed for representing the contract state for most contracts, i.e., the flow from only a single arc corresponding to a contract in the contract network representation needs to be examined to ascertain if that contract is in excess (or deficit). Notationally, atomic proposition $E_{(p_i, p_{i+1})}$ is true if the flow in arc (p_i, p_{i+1}) is nonnegative in the contract network representation

118

and this is equivalent to contract p being in excess. The level i used for defining the contract state is 0 for most contracts, i.e, the excess flag is set before any transfer rules have activated (therefore, the excess flag for a contract p is given by $E_{(p_0,p_1)}$ for most cases).

However, for some contracts, there is a need to define multiple atomic propositions to represent excess. This is because these contracts contain fields that can supply to other contracts under special conditions. These fields feed downstream of the node corresponding to 0th level in the contract network and hence flow in arc (p_0, p_1) is not sufficient to resolve the state. This arrangement exists because a field in one contract system may be physically connected to the other contract system and hence might require special rules. In such a case, there are several atomic propositions $E_{(p_i,p_{i+1})}$ each corresponding to a different arc $(p_i, p_{i+1}) \in \mathcal{E}_{l,a}$.

4.3.5 Transfer State

Atomic proposition $T_{p,q}$ is used to indicate the state of an inter-contract transfer of gas from contract p to contract q. $T_{p,q}$ is true implies that the transfer takes place (or the transfer is activated). It is equivalent to flowrate on an inter-contract transfer arc (p_i, q_j) being non-negative, (p_i, q_j) being the unique arc in the contract network representing transfer flowrate from contracts p to q.

4.3.6 Transfer Priorities

When a contract p is in excess and there are several contracts (say q, r) that are in deficit and should receive supply from contract p (as per the rules), there is a need to define priorities of transfer because the excess may not be enough to fulfill all the transfer demands being placed on contract p. The rules indeed provide these priorities. If these transfer priorities are not modeled, a solution may violate these priorities (though the balances will still be closed and all deficits will be fulfilled but not as per the priorities dictated by the rules).

Hence, atomic propositions $S_{(q_j,q_{j+1})}$ need to be defined corresponding to transfer

from contract p to contract q (arc (p_i, q_j) or atomic proposition $T_{p,q}$). $S_{(q_j, q_{j+1})}$ flags the state of the receiving contract once a transfer has been made. It being true implies that arc (q_j, q_{j+1}) has a nonnegative flowrate and hence the transfer has fulfilled the deficit. Another transfer $T_{p,r}$ from the supply contract p can only go ahead if the proposition $S_{(q_j, q_{j+1})}$ is true.

However, it is important to note that a $S_{(r_j, r_{j+1})}$ need not be defined corresponding to every transfer $T_{p,r}$. Specifically, it is not required for transfers that are at the lowest priority on the supplying side, i.e., for transfer arc (p_i, r_j), p_i is the *terminal level node* (the last node representing level) corresponding to contract p in the contract network, because there is no transfer rule that follows it and needs to depend on it. Of course this lowest priority transfer rule still needs to honor all the previously defined priorities. Another condition when the priority flag is not required is when a receiving contract r has the transfer $T_{p,r}$ at the lowest priority, i.e., r tries to borrow from p, only when everything else fails. In this case r_j is a terminal level node and as per material balance on the network, the demand should be satisfied. To summarize, when either p_i or r_j is a terminal level node, it is not required to define a priority corresponding to the transfer arc connecting them.

The set of arcs over which priority atomic propositions are defined is denoted by $\mathcal{E}_{l,S} \subset \mathcal{E}_l$.

4.3.7 Coupling Constraints between Infrastructure and Contract Networks

Equations (40), (41) and (42) represent the coupling constraints between the infrastructure and the contract networks. These are essentially constraints linking the sources and sinks of both networks.

If multiple contracts have no special rules distinguishing them from one of the others in the group (i.e., they either have no specific rules or have the same rules as the others), they are collapsed to represent a single source and are not represented separately on the contract network.

Let $\mathcal{C}^S$ be the superset of all contracts, some of which may not be represented on the contract network. The set of contracts that are actually represented is $\mathcal{C} \subset \mathcal{C}^S$. These sets for the example shown in Figure 4-2 are defined as follows:

$$\mathcal{C}^S = \{A, B, C, D, F, X, Y\},$$

$$\mathcal{C} = \{A, B, C, D, F\}.$$

Set $\mathcal{C}^S$ is maintained in the model for future flexibility so as to add additional rules, at which point the contract network can be expanded to incorporate those contracts. Denote the set of contracts in the set $\mathcal{C}^S$ that are collectively represented as a single source $p_s \in \mathcal{C}$ in the contract network as $\mathcal{C}_p^S$.

The supply rate of a contract $q_{c,i}$ is given by

$$q_{c,i} = \sum_{j \in \mathcal{F}_i} Q_{s,j}, \quad \forall i \in \mathcal{C}^S, \tag{40}$$

where $\mathcal{F}_i$ is the set of fields producing under contract i. The source rates for the contract network is given by

$$q_{s,p_s} = \sum_{j \in \mathcal{C}_p^s} q_{c,j}, \quad \forall p_s \in \mathcal{C}. \tag{41}$$

The sink nodes in the contract network are mapped to the demand nodes in the infrastructure network.

$$q_{s,u_i} = Q_{s,i}, \quad \forall i \in \mathcal{N}_D \backslash \{i\}, \tag{42}$$

where u_i is the demand node in the contract network corresponding to the demand node i in the infrastructure network. This equation should not be formulated for exactly one demand node in the infrastructure network. The reason is that the last equality is implied by the combined material balances in the infrastructure and contract networks. If this is included, it violates the linear independence constraint qualification and has been observed to create severe problems for local solver convergence in subproblem solutions.

4.3.8 Volumetric Balances in the Contract Network Representation

Volumetric balances in the contract network represent the actual constraints for supply and demand balances as well as for the excess/deficit calculations for each contract. Volumetric balances are easier to formulate over the entire contract network than referring to individual contracts. Hence the convention of referring to a node as p_i and it being associated with contract p has been dropped in this section.

Let $e_{(u,v)}$ be the excess volumetric flowrate in arc $(u,v) \in \mathcal{E}_l$. It is positive when a contract is in excess and negative otherwise. Let $t_{(u,v)}$ represent the transfer rates in arcs $(u,v) \in \mathcal{E}_t$. Furthermore $q_{a,(u,v)}$ represents flowrates in arcs $(u,v) \in \mathcal{E}_s$, i.e., supply and demand arcs. At the nodes that are junctions in the contract network (i.e., all nodes except the supply and demand nodes, same as set $\mathcal{L}_l$), the balance can be represented as:

$$\sum_{v:(v,u)\in\mathcal{E}_l} e_{(v,u)} - \sum_{v:(u,v)\in\mathcal{E}_l} e_{(u,v)} + \sum_{v:(v,u)\in\mathcal{E}_t} t_{(v,u)}$$
$$- \sum_{v:(u,v)\in\mathcal{E}_t} t_{(u,v)} + \sum_{v:(v,u)\in\mathcal{E}_s} q_{a,(v,u)} - \sum_{v:(u,v)\in\mathcal{E}_s} q_{a,(u,v)} = 0, \qquad \forall u \in \mathcal{L}_l. \quad (43)$$

Nodes that are supply or demand nodes to the network (i.e., that are not junctions) have a production term. It should be noted that these nodes do not have arcs representing levels terminating at or originating from them.

$$\sum_{v:(v,u)\in\mathcal{E}_s} q_{a,(v,u)} - \sum_{v:(u,v)\in\mathcal{E}_s} q_{a,(u,v)}$$
$$+ \sum_{v:(v,u)\in\mathcal{E}_t} t_{(v,u)} - \sum_{v:(u,v)\in\mathcal{E}_t} t_{(u,v)} + q_{s,u} = 0, \quad \forall u \in \mathcal{L}_s. \quad (44)$$

4.3.9 Relationship between Atomic Propositions and Flowrates in Contract Network

A binary variable is introduced in the model corresponding to each atomic proposition defined previously. The atomic proposition being true is equivalent to the binary variable being equal to 1.

Binary variable $y^e_{(p_i,p_{i+1})}$ corresponds to the atomic proposition $E_{p,(p_i,p_{i+1})}$ and is the contract state binary variable. It should be noted again that this variable is not defined over every arc representing excess or deficit in the system. The set of arcs over which this binary variable is defined is termed $\mathcal{E}_{l,a} \subset \mathcal{E}_l$.

Let $e_{(p_i,p_{i+1})}$ denote the excess rate for contract p at level i. $e_{(p_i,p_{i+1})}$ is negative if the supply of contract is less than the primary demand of contract p and the contract is in deficit. The binary variables $y^e_{(p_i,p_{i+1})}$ are related to the flowrates in the excess arcs $e_{(p_i,p_{i+1})}$ over the set of all contracts $\mathcal{C}$ as follows:

$$e^L_{(p_i,p_{i+1})} - y^e_{(p_i,p_{i+1})} \, e^L_{(p_i,p_{i+1})} - e_{(p_i,p_{i+1})} \leq 0, \quad \forall (p_i,p_{i+1}) \in \mathcal{E}_{l,a}, \tag{45}$$

$$e_{(p_i,p_{i+1})} - y^e_{(p_i,p_{i+1})} \, e^U_{(p_i,p_{i+1})} \leq 0, \quad \forall (p_i,p_{i+1}) \in \mathcal{E}_{l,a}, \tag{46}$$

where $e^L_{(p_i,p_{i+1})}$ and $e^U_{(p_i,p_{i+1})}$ are lower and upper bounds respectively to $e_{(p_i,p_{i+1})}$. These constraints ensure that $y^e_{(p_i,p_{i+1})} = 1$ is equivalent to $0 \leq e_{(p_i,p_{i+1})} \leq e^U_{(p_i,p_{i+1})}$ and $y^e_{(p_i,p_{i+1})} = 0$ is equivalent to $e^L_{(p_i,p_{i+1})} \leq e_{(p_i,p_{i+1})} \leq 0$. These constraints therefore couple the contract state binary variables with the actual flowrates in the contract network representation.

$e^L_{(p_i,p_{i+1})}$ is strictly negative and is set by making the assumption that the supply of contract p is zero and the primary demand of contract p is fulfilled exclusively by inter-contract transfers and is therefore the maximum deficit rate that is possible for a contract. Hence, it is set to the lower bound of the rate at the sink node corresponding to contract p (physically this is the maximum demand rate at the LNG plant corresponding to this sink). On the other hand, $e^U_{(p_i,p_{i+1})}$ is the maximum excess rate possible and is set assuming that the primary demand of the contract is

zero and all the supply is available for inter-contract transfers. It is therefore equal to the upper bound of the supply of the contract and is strictly positive. A further discussion can be found in Appendix C.1.2.

An exactly analogous constraint exists for priority atomic propositions. Let $y^s_{(q_j,q_{j+1})}$ correspond to atomic proposition $S_{(q_j,q_{j+1})}$. Then

$$e^L_{(q_j,q_{j+1})} - y^s_{(q_j,q_{j+1})} e^L_{(q_i,q_{j+1})} - e_{(q_j,q_{j+1})} \leq 0, \quad \forall (q_j, q_{j+1}) \in \mathcal{E}_{l,S}, \tag{47}$$

$$e_{(q_i,q_{j+1})} - y^s_{(q_i,q_{j+1})} e^U_{(q_j,q_{j+1})} \leq 0, \quad \forall (q_j, q_{j+1}) \in \mathcal{E}_{l,S}. \tag{48}$$

This constraint performs exactly the same function as equations (45) and (46) and bounds on $e_{(q_j,q_{j+1})}$ are calculated as above and they satisfy the same properties.

Let $t_{(p_i,q_j)}$ denote the transfer rate between contract p and q (i.e., the flowrate in arc $(p_i, q_j) \in \mathcal{E}_t$). Also, let $y^t_{p,q}$ be the binary variable corresponding to the atomic proposition $T_{p,q}$ representing the state of the (p, q) transfer such that $y^t_{p,q} = 1$ when $T_{p,q}$ is true. Then this binary variable $y^t_{p,q}$ can be coupled with the actual transfer flowrate $t_{(p_i,q_j)}$ by the following relationships

$$-t_{(p_i,q_j)} \leq 0, \quad \forall (p_i, q_j) \in \mathcal{E}_t, \tag{49}$$

$$t_{(p_i,q_j)} - y^t_{p,q} t^U_{(p_i,q_j)} \leq 0, \quad \forall (p_i, q_j) \in \mathcal{E}_t. \tag{50}$$

It should be noted that the upper bound for $t_{(p_i,q_j)}$ is set to $e^U_{(p_{i-1},p_i)}$ since the maximum amount of transfer that is possible is the maximum flowrate possible in the excess arc upstream of node p_i in the transferring contract (except for transfer arcs originating at supply nodes in which case it is set to the upper bound on the production rate from that supply node). An explicit representation of these bounds can be found in Appendix C.1.2. These constraints ensure that $t_{(p_i,q_j)} = 0$, i.e., no transfer takes place when $y^t_{p,q} = 0$ and $t_{(p_i,q_j)} \geq 0$ when $y^t_{p,q} = 1$.

4.3.10 Transfer-Activation Constraints

An inter-contract transfer rule can now be expressed in terms of a logical expression involving atomic propositions representing contract states, the transfers states, the transfer priorities flags and operational flags. This logical expression can be converted to its *conjunctive normal form* (CNF) and this CNF can be converted to constraints involving binary variables [113, 114]. All contract rules used in the case study can be found in Section 4.4. Following is an example from the case study.

Example (Transfer-Activation Constraint): Consider the following contractual rule:

A to B Transfer Rule: *When demand at LNG 2 (SC2D) cannot be fulfilled by contract B, "borrow gas" from contract A shall supply to LNG 2 (SC2D) to meet the demand and the RA to RB trunkline shall be open at this stage.*

The following atomic propositions are required:

- $E_{(A_0,A_1)}$: A is in excess (contract A state),

- $E_{(B_1,B_2)}$: B is in excess (contract B state),

- $T_{A,B}$: A supplies B (transfer state),

- $C_{(RA,RB)}$: RA to RB trunkline is open (additional operational state).

The logic in the above rule can be expressed as:

$$(E_{(A_0,A_1)} \land \neg E_{(B_1,B_2)}) \Rightarrow (T_{A,B} \land C_{(RA,RB)}).$$

Here $\land$, $\lor$, $\neg$ and $\Rightarrow$ are logical AND, OR, NOT and IMPLICATION operators respectively. The above is equivalent to

$$\neg (E_{(A_0,A_1)} \land \neg E_{(B_1,B_2)}) \lor (T_{A,B} \land C_{(RA,RB)}).$$

The CNF of the statement is:

$$(\neg E_{(A_0,A_1)} \lor E_{(B_1,B_2)} \lor T_{A,B}) \land (\neg E_{(A_1,A_2)} \lor E_{(B_1,B_2)} \lor C_{(RA,RB)}).$$

The CNF can be converted to binary constraints

$$1 - y^e_{(A_0,A_1)} + y^e_{(B_1,B_2)} + y^t_{A,B} \geq 1,$$
$$1 - y^e_{(A_0,A_1)} + y^e_{(B_1,B_2)} + y^l_{(RA,RB)} \geq 1,$$

so that the final form of the constraints representing the A to B transfer rule is as follows:

$$y^e_{(A_0,A_1)} - y^e_{(B_1,B_2)} - y^t_{A,B} \leq 0,$$
$$y^e_{(A_0,A_1)} - y^e_{(B_1,B_2)} - y^l_{(RA,RB)} \leq 0.$$

4.3.11 Additional Logical Constraints

Additional logical constraints must be added to formulate a satisfactory representation of the contracts as well as to strengthen relaxations. These are listed as follows:

1. Cuts should be added by analysing the negation of transfer-activation conditions. If the negation corresponds to a condition when the transfer should not activate (i.e., the original transfer-activation condition was both necessary and sufficient for the transfer to take place), an additional logical constraint which states that no transfer should take place when the negation is true must be added to the problem. It is important to note that for some transfers, a negation of the activation condition does not imply that the transfer cannot activate (e.g., they may have multiple activation conditions) and hence the above logical cut will be invalid. Therefore, such an analysis must be carried out on a case by case basis. This is important to avoid transfers when they should not take place. This is required because inter-contract transfer rules represent only *sufficient* conditions for transfer. The negation of transfer rules also introduces logical constraints that imply that no transfer should be made to a contract in excess and hence this need not be enforced explicitly.

2. If a contract q requiring a single excess flag is in excess, i.e, the unique excess

126

flag $E_{(q_i, q_{i+1})}$ corresponding to it is true, then any priority atomic propositions $S_{(q_j, q_{j+1})}$ defined for that contract are also true. Also, if a contract q has multiple excess flags, then a priority flag $S_{(q_j, q_{j+1})}$ is true if the closest excess flag to $S_{(q_j, q_{j+1})}$ is true (i.e., $E_{(q_{i_j}, q_{i_j+1})}$, $(q_{i_j}, q_{i_j+1}) \in \mathcal{E}_{l,a}$ is true, (q_{i_j}, q_{i_j+1}) is the closest element of set $\mathcal{E}_{l,a}$ to node q_j in the sub-graph of contract q). This can be represented logically as follows

$$E_{(q_{i_j}, q_{i_j+1})} \Rightarrow S_{(q_j, q_{j+1})}, \quad i_j < j, \quad \forall (q_{i_j}, q_{i_j+1}) \in \mathcal{E}_{l,a}, \quad (q_j, q_{j+1}) \in \mathcal{E}_{l,S},$$

where (q_{i_j}, q_{i_j+1}) is the closest arc in set $\mathcal{E}_{l,a}$ to node q_j. This can be converted to the following binary constraint

$$y^e_{(q_{i_j}, q_{i_j+1})} - y^s_{(q_j, q_{j+1})} \leq 0, \quad i_j < j, \quad \forall (q_{i_j}, q_{i_j+1}) \in \mathcal{E}_{l,a}, \quad (q_j, q_{j+1}) \in \mathcal{E}_{l,S}. \quad (51)$$

3. There are transfer rules that cannot activate simultaneously. For example, either $T_{p,q}$ is true or $T_{q,p}$ is true. Both cannot happen simultaneously.

4. For contracts that have multiple excess atomic propositions, once an atomic proposition is true, all excess atomic propositions on further levels must also be true.

4.4   PSC Framework: An Example

In this section, the production-sharing contracts (PSC) framework for the case study is described. It is of equal complexity and size as the original SGPS contractual model, however the contract field assignments and rules have been changed to preserve confidentiality.

4.4.1   Contract Definitions

There are five production-sharing contracts in the system. The field contract assignment is provided in Table 4.2. Following are the rules dictating contract plant

assignments:

1. LNG plant 1 must be supplied by contract A, X and Y. There are no special rules for separating them, so supplies from A, X and Y are combined and denoted by contract A.

2. LNG plant 2 must be supplied by contract B.

3. LNG plant 3 must be supplied in the ratio 600:350 by contract C and D. Hence the following constraint is enforced for flowrates in the contract graph:

$$350\, q_{a,(\mathrm{C_0,CD_d})} = 600\, q_{a,(\mathrm{D_0,CD_d})}. \tag{52}$$

4. Contract F should normally supply contract B (i.e., LNG plant 2), however under certain conditions, it may supply contract A (further details are provided in Section 4.4.1).

The methodology for modeling inter-contract transfer rules and operational rules has already been developed in Section 4.3. The definition of atomic propositions required for representation of the rules can be found in Table 4.3.

Table 4.2: Field contract assignments

$i \in \mathcal{C}$	$j \in \mathcal{C}_i^S$	Fields $\mathcal{F}_j$
A	A	SC, F6, F23
	X	F23SW
	Y	BN,BY,D35
B	B	M1, M3, B11
C	C	M4, HL, JN
D	D	SE
F	F	E11

4.4.2 Inter-Contract Transfer Rules

This section describes the inter-contract transfer rules and their representation as binary constraints. The priorities for supply when a contract is in excess are defined

Table 4.3: Atomic proposition definitions

Symbol	Definition	Equivalent Representation
$E_{(A_0,A_1)}$	A is in excess	$e_{(A_0,A_1)} \geq 0$
$E_{(B_0,B_1)}$	B is in excess without F supply	$e_{(B_0,B_1)} \geq 0$
$E_{(B_1,B_2)}$	B is in excess including F transfer	$e_{(B_1,B_2)} \geq 0$
$E_{(C_0,C_1)}$	C is in excess	$e_{(C_0,C_1)} \geq 0$
$E_{(D_0,D_1)}$	D is in excess	$e_{(D_0,D_1)} \geq 0$
$T_{A,B}$	A supplies B	$t_{(A_1,B_2)} \geq 0$
$T_{A,C}$	A supplies C	$t_{(A_2,C_3)} \geq 0$
$T_{A,D}$	A supplies D	$t_{(A_3,D_3)} \geq 0$
$T_{B,C}$	B supplies C	$t_{(B_2,C_2)} \geq 0$
$T_{B,D}$	B supplies D	$t_{(B_3,D_2)} \geq 0$
$T_{C,B}$	C supplies B	$t_{(C_2,B_3)} \geq 0$
$T_{C,D}$	C supplies D	$t_{(C_1,D_1)} \geq 0$
$T_{C,B}$	D supplies C	$t_{(D_1,C_1)} \geq 0$
$T_{F,A}$	F supplies A	$t_{(F_s,A_1)} \geq 0$
$T_{F,B}$	F supplies B	$t_{(F_s,B_1)} \geq 0$
$S_{(B_2,B_3)}$	B is *not* in deficit after considering A-B transfer	$e_{(B_2,B_3)} \geq 0$
$S_{(C_2,C_3)}$	C is *not* in deficit after considering B-C transfer	$e_{(C_2,C_3)} \geq 0$
$S_{(D_1,D_2)}$	D is *not* in deficit after considering C-D transfer	$e_{(D_1,D_2)} \geq 0$
$C_{(RA,RB)}$	(RA,RB) is open	$Q_{a,(RA,RB)} \geq 0,$ $P_{RA} \geq P_{RB}$
C_{M1}	M1 production is greater than 500 MMscfd	$Q_{s,(M1)} \geq \varrho_g 500$
$C_{JN,(M1,RC)}$	All JN production is diverted into (M1,RC)	-

in Table 4.4.

Following are the rules and binary constraints corresponding to them.

1. *When demand of LNG plant 2 cannot be fulfilled by contract B, then gas shall be borrowed from A to supply plant 2. RA to RB shall be open at this stage.*

   (a) The above statement can be expressed logically as

   $$(E_{(A_0,A_1)} \wedge \neg E_{(B_1,B_2)}) \Rightarrow (T_{A,B} \wedge C_{(RA,RB)}).$$

Table 4.4: Priority of supply

Supplying Contract	Priority of Receiving if in Deficit
A	B, C, D
B	C, D
C	D, B
D	C
F	B, A

Conjunctive normal form of the above is given by

$$\left(\neg\, E_{(A_0,A_1)} \;\vee\; E_{(B_1,B_2)} \;\vee\; T_{A,B} \right) \wedge \left(\neg\, E_{(A_0,A_1)} \;\vee\; E_{(B_1,B_2)} \;\vee\; C_{(RA,RB)} \right),$$

which can be converted to the following binary constraints:

$$y^e_{(A_0,A_1)} - y^e_{(B_1,B_2)} - y^t_{A,B} \leq 0, \tag{53}$$

$$y^e_{(A_0,A_1)} - y^e_{(B_1,B_2)} - y^l_{(RA,RB)} \leq 0. \tag{54}$$

(b) Additionally, the additional constraint that "if contract A supplies LNG 2, then (RA,RB) should be open" is enforced:

$$T_{A,B} \;\Rightarrow\; C_{(RA,RB)}$$

that is equivalent to

$$\neg T_{A,B} \;\vee\; C_{(RA,RB)}$$

and therefore the constraint

$$y^t_{A,B} - y^l_{(RA,RB)} \leq 0. \tag{55}$$

(c) A negation of the original rule is also added to strengthen the relaxations.

In case of negation, the state of (RA,RB) is immaterial:

$$\neg\,(E_{(A_0,A_1)} \,\wedge\, \neg E_{(B_1,B_2)}) \,\Rightarrow\, \neg\,T_{A,B},$$

or the following CNF

$$(E_{(A_0,A_1)} \,\vee\, \neg T_{A,B}) \,\wedge\, (\neg E_{(B_1,B_2)} \,\vee\, \neg T_{A,B}).$$

Final binary constraints are given by

$$y_{A,B}^{t} - y_{(A_0,A_1)}^{e} \leq 0, \tag{56}$$

$$y_{(B_1,B_2)}^{e} + y_{A,B}^{t} - 1 \leq 0. \tag{57}$$

2. *If demand of LNG plant 2 cannot be fulfilled by contract B and A is unable to fulfill this deficit, then gas shall be borrowed from contract C to supply plant 2.* This clause requires a priority clause because the first priority of C is to supply D and hence it needs to be tested if demand from D has been fulfilled:

$$(\neg\,E_{(B_1,B_2)} \,\wedge\, E_{(C_0,C_1)} \,\wedge\, S_{(D_1,D_2)}) \,\Rightarrow\, T_{C,B}.$$

This can be written as a series of disjunctions:

$$E_{(B_1,B_2)} \,\vee\, \neg E_{(C_0,C_1)} \,\vee\, \neg S_{(D_1,D_2} \,\vee\, T_{C,B},$$

which can be directly converted to a binary constraint

$$y_{(C_0,C_1)}^{e} + y_{(D_1,D_2)}^{s} - y_{(B_1,B_2)}^{e} - y_{C,B}^{t} - 1 \leq 0. \tag{58}$$

Negation

The negation of this constraint states the transfer should not take place if the

left-hand side (LHS) is negated.

$$\neg\,(\neg\,E_{(B_1,B_2)} \ \wedge \ E_{(C_1,C_2)} \ \wedge \ S_{(D_1,D_2)}) \ \Rightarrow \ \neg\,T_{C,B}.$$

Its CNF is given as

$$(\neg\,E_{(B_1,B_2)} \ \vee \ \neg\,T_{C,B}) \ \wedge \ (E_{(C_0,C_1)} \ \vee \ \neg\,T_{C,B}) \ \wedge \ (S_{(D_1,D_2)} \ \vee \ \neg\,T_{C,B}),$$

or in terms of binary variables

$$y^e_{(B_1,B_2)} + y^t_{C,B} - 1 \le 0, \tag{59}$$

$$y^t_{C,B} - y^e_{(C_0,C_1)} \le 0, \tag{60}$$

$$y^t_{C,B} - y^s_{(D_1,D_2)} \le 0. \tag{61}$$

3. *If contract D cannot meet its allocated production share, C shall have the first priority to fulfill this deficit, followed by contract B, then A.*

   This rule actually contains three rules defining whether three transfers $T_{C,D}$, $T_{B,D}$ and $T_{A,D}$ should take place. Additionally it also defines the priorities. It should also be noted that B needs to first fulfill any deficit from C (hence the priority flag $S_{(C_2,C_3)}$) and A needs to supply B and C (no priority required for $T_{A,C}$ since it is a terminal level transfer):

$$(\neg E_{(D_0,D_1)} \ \wedge \ E_{(C_0,C_1)}) \ \Rightarrow \ T_{C,D},$$

$$(\neg E_{(D_0,D_1)} \ \wedge \ E_{(B_1,B_2)} \ \wedge \ S_{(C_2,C_3)}) \ \Rightarrow \ T_{B,D},$$

$$(\neg E_{(D_0,D_1)} \ \wedge \ E_{(A_0,A_1)} \ \wedge \ S_{(B_2,B_3)}) \ \Rightarrow \ T_{A,D}.$$

Converting implication to disjunction

$$E_{(D_0,D_1)} \lor \neg E_{(C_0,C_1)} \lor T_{C,D},$$

$$E_{(D_0,D_1)} \lor \neg E_{(B_1,B_2)} \lor \neg S_{(C_2,C_3)} \lor T_{B,D},$$

$$E_{(D_0,D_1)} \lor \neg E_{(A_0,A_1)} \lor \neg S_{(B_2,B_3)} \lor T_{A,D}.$$

These disjunctions can be directly formulated into binary constraints:

$$y^e_{(C_0,C_1)} - y^e_{(D_0,D_1)} - y^t_{C,D} \le 0, \tag{62}$$

$$y^e_{(B_1,B_2)} + y^s_{(C_2,C_3)} - y^e_{(D_0,D_1)} - y^t_{B,D} - 1 \le 0, \tag{63}$$

$$y^e_{(A_0,A_1)} + y^s_{(B_2,B_3)} - y^e_{(D_0,D_1)} - y^t_{A,D} - 1 \le 0. \tag{64}$$

Negation

The negation of these conditions states that these transfers should not take place if there is a violation of LHS.

$$\neg\left(\neg E_{(D_0,D_1)} \land E_{(C_0,C_1)}\right) \Rightarrow \neg T_{C,D},$$

$$\neg\left(\neg E_{(D_0,D_1)} \land E_{(B_1,B_2)} \land S_{(C_2,C_3)}\right) \Rightarrow \neg T_{B,D},$$

$$\neg\left(\neg E_{(D_0,D_1)} \land E_{(A_0,A_1)} \land S_{(B_2,B_3)}\right) \Rightarrow \neg T_{A,D}.$$

The corresponding CNF is given by

$$\left(\neg E_{(D_0,D_1)} \lor \neg T_{C,D}\right) \land \left(E_{(C_0,C_1)} \lor \neg T_{C,D}\right),$$

$$\left(\neg E_{(D_0,D_1)} \lor \neg T_{B,D}\right) \land \left(E_{(B_1,B_2)} \lor \neg T_{B,D}\right) \land \left(S_{(C_2,C_3)} \lor \neg T_{B,D}\right),$$

$$\left(\neg E_{(D_0,D_1)} \lor \neg T_{A,D}\right) \land \left(E_{(A_0,A_1)} \lor \neg T_{A,D}\right) \land \left(S_{(B_2,B_3)} \lor \neg T_{A,D}\right).$$

Converting to binary constraints:

$$y^e_{(D_0,D_1)} + y^t_{C,D} - 1 \le 0, \tag{65}$$

$$y^t_{C,D} - y^e_{(C_0,C_1)} \le 0, \tag{66}$$

$$y^e_{(D_0,D_1)} + y^t_{B,D} - 1 \le 0, \tag{67}$$

$$y^t_{B,D} - y^e_{(B_1,B_2)} \le 0, \tag{68}$$

$$y^t_{B,D} - y^s_{(C_2,C_3)} \le 0, \tag{69}$$

$$y^e_{(D_0,D_1)} + y^t_{A,D} - 1 \le 0, \tag{70}$$

$$y^t_{A,D} - y^e_{(A_0,A_1)} \le 0, \tag{71}$$

$$y^t_{A,D} - y^s_{(B_2,B_3)} \le 0. \tag{72}$$

4. *If contract C cannot meet its allocated production share, D shall have the first priority to fulfill this deficit, followed by contract B, then A.* This rule is similar to the previous one and governs three transfers:

$$(\neg E_{(C_0,C_1)} \ \wedge \ E_{(D_0,D_1)}) \ \Rightarrow \ T_{D,C},$$

$$(\neg E_{(C_0,C_1)} \ \wedge \ E_{(B_1,B_2)}) \ \Rightarrow \ T_{B,C},$$

$$(\neg E_{(C_0,C_1)} \ \wedge \ E_{(A_0,A_1)} \ \wedge \ S_{(B_2,B_3)}) \ \Rightarrow \ T_{A,C}.$$

Converting implication to disjunction:

$$E_{(C_0,C_1)} \ \vee \ \neg E_{(D_0,D_1)} \ \vee \ T_{D,C},$$

$$E_{(C_0,C_1)} \ \vee \ \neg E_{(B_1,B_2)} \ \vee \ T_{B,C},$$

$$E_{(C_0,C_1)} \ \vee \ \neg E_{(A_0,A_1)} \ \vee \ \neg S_{(B_2,B_3)} \ \vee \ T_{A,C}.$$

Binary constraints corresponding to the above are given as

$$y^e_{(D_0,D_1)} - y^e_{(C_0,C_1)} - y^t_{D,C} \leq 0, \tag{73}$$

$$y^e_{(B_1,B_2)} - y^e_{(C_0,C_1)} - y^t_{B,C} \leq 0, \tag{74}$$

$$y^e_{(A_0,A_1)} + y^s_{(B_2,B_3)} - y^e_{(C_0,C_1)} - y^t_{A,C} - 1 \leq 0. \tag{75}$$

Negation

The negation of the transfer-activation conditions imply that no transfer should take place:

$$\neg \left(\neg E_{(C_0,C_1)} \wedge E_{(D_0,D_1)} \right) \Rightarrow \neg T_{D,C},$$

$$\neg \left(\neg E_{(C_0,C_1)} \wedge E_{(B_1,B_2)} \right) \Rightarrow \neg T_{B,C},$$

$$\neg \left(\neg E_{(C_0,C_1)} \wedge E_{(A_0,A_1)} \right) \wedge S_{(B_2,B_3)} \Rightarrow \neg T_{A,C}.$$

The CNF of the above logical conditions is

$$\left(\neg E_{(C_0,C_1)} \vee \neg T_{D,C} \right) \wedge \left(E_{(D_0,D_1)} \vee \neg T_{D,C} \right),$$

$$\left(\neg E_{(C_0,C_1)} \vee \neg T_{B,C} \right) \wedge \left(E_{(B_1,B_2)} \vee \neg T_{B,C} \right),$$

$$\left(\neg E_{(C_0,C_1)} \vee \neg T_{A,C} \right) \wedge \left(E_{(A_0,A_1)} \vee \neg T_{A,C} \right) \wedge \left(S_{(B_2,B_3)} \vee \neg T_{A,C} \right).$$

The binary constraints can now be obtained from CNF:

$$y^e_{(C_0,C_1)} + y^t_{D,C} - 1 \leq 0, \tag{76}$$

$$y^t_{D,C} - y^e_{(D_0,D_1)} \leq 0, \tag{77}$$

$$y^e_{(C_0,C_1)} + y^t_{B,C} - 1 \leq 0, \tag{78}$$

$$y^t_{B,C} - y^e_{(B_1,B_2)} \leq 0, \tag{79}$$

$$y^e_{(C_0,C_1)} + y^t_{A,C} - 1 \leq 0, \tag{80}$$

$$y^t_{A,C} - y^e_{(A_0,A_1)} \leq 0, \tag{81}$$

$$y^t_{A,C} - y^s_{(B_2,B_3)} \leq 0. \tag{82}$$

5. The supply clause for Contract F is complicated. The rule only states that *contract F supply normally belongs to contract B. However, if B is in excess and A is in deficit, F may supply A.* Following is the representation of this rule with the first statement representing the rule and rest of the statement being additional inferences of the rule.

(a) F can supply A when B is in excess either at level 0 or 1, and A cannot fulfill demand of LNG plant 1

$$(\neg E_{(A_0,A_1)} \wedge (E_{(B_0,B_1)} \vee E_{(B_1,B_2)})) \Rightarrow T_{F,A}. \tag{83}$$

(b) When B is in deficit at level 0, F should supply B

$$\neg E_{(B_0,B_1)} \Rightarrow T_{F,B}. \tag{84}$$

(c) F is forbidden to supply A, when B is in deficit at level 1

$$\neg E_{(B_1,B_2)} \Rightarrow \neg T_{F,A}. \tag{85}$$

(d) F is forbidden to supply A, when A is in excess

$$E_{(A_0,A_1)} \Rightarrow \neg T_{F,A}. \tag{86}$$

(e) When F supplies B, (RA,RB) shall be open

$$T_{F,B} \Rightarrow C_{(RA,RB)}. \tag{87}$$

The conjunctive normal forms for these logical expressions are as follows:

$$(E_{(A_0,A_1)} \lor T_{F,A} \lor \neg E_{(B_0,B_1)}) \land (E_{(A_0,A_1)} \lor T_{F,A} \lor \neg E_{(B_1,B_2)}),$$

$$E_{(B_0,B_1)} \lor T_{F,B},$$

$$E_{(B_1,B_2)} \lor \neg T_{F,A},$$

$$\neg E_{(A_0,A_1)} \lor \neg T_{F,A},$$

$$\neg T_{F,B} \lor C_{(RA,RB)}.$$

The following binary constraints are required:

$$y^e_{(B_0,B_1)} - y^e_{(A_0,A_1)} - y^t_{F,A} \leq 0, \tag{88}$$

$$y^e_{(B_1,B_2)} - y^e_{(A_0,A_1)} - y^t_{F,A} \leq 0, \tag{89}$$

$$1 - y^e_{(B_0,B_1)} - y^t_{F,B} \leq 0, \tag{90}$$

$$y^t_{F,A} - y^e_{(B_1,B_2)} \leq 0, \tag{91}$$

$$y^e_{(A_0,A_1)} + y^t_{F,A} - 1 \leq 0, \tag{92}$$

$$y^t_{F,B} - y^l_{(RA,RB)} \leq 0. \tag{93}$$

Negation

The negation of only statement (a) needs to be considered since negations of the other rules do not provide any new constraints because the right-hand side (RHS) of the respective logical implications may or may not be true:

$$\neg (\neg E_{(A_0,A_1)} \land (E_{(B_0,B_1)} \lor E_{(B_1,B_2)})) \Rightarrow \neg T_{F,A}.$$

The CNF of the above is given by

$$(\neg E_{(A_0,A_1)} \lor \neg T_{F,A}) \land (E_{(B_0,B_1)} \lor E_{(B_1,B_2)} \lor \neg T_{F,A}).$$

The first operand of $\land$ is same as statement (d), hence only the second operand

needs to be added as a constraint and is as follows

$$y_{F,A}^t - y_{(B_0,B_1)}^e - y_{(B_1,B_2)}^e \leq 0. \tag{94}$$

6. Only one of the following two transfers should activate: $T_{C,D}$ and $T_{D,C}$,

$$\neg (T_{D,C} \wedge T_{C,D})$$

or

$$\neg T_{D,C} \vee \neg T_{C,D}.$$

The final binary constraint is given by

$$y_{D,C}^t + y_{C,D}^t - 1 \leq 0. \tag{95}$$

7. Contract B has two excess atomic propositions $E_{(B_0,B_1)}$ and $E_{(B_1,B_2)}$. If the contract is already in excess at level 0, it must be excess at further levels. Logically this can be represented as

$$E_{(B_0,B_1)} \Rightarrow E_{(B_1,B_2)},$$

and in the CNF as

$$\neg E_{(B_0,B_1)} \vee E_{(B_1,B_2)}.$$

The binary constraint representation of the above is as follows

$$y_{(B_0,B_1)}^e - y_{(B_1,B_2)}^e \leq 0. \tag{96}$$

4.5   Operational Rules

Operational rules can be modeled in the same framework as outlined for production-sharing contracts, i.e., by defining atomic propositions representing the states of var-

ious operating lines and facilities and then forming logical expressions representing the rules. Finally, the resulting logical expression rules can be converted into binary constraints in exactly the same way as described above.

4.5.1 An Example

Consider the following operational rules:

1. *A minimum of 500 MMscfd shall be maintained in the (M1,T) line. In the event that M1 production is less than 500 MMscfd, part of JN production shall be diverted into this pipeline section.*

    Define the following atomic propositions:

    (a) C_{M1}: M1 production is greater than 500 MMscfd,

    (b) $C_{JN,(M1,RC)}$: *All of JN production* is diverted into (M1,RC).

    The following states that if M1 production is greater than 500 MMscfd, all of JN production is carried by (M1,RC)

    $$C_{M1} \Rightarrow C_{JN,(M1,RC)}.$$

    The above can be converted to a disjunction

    $$\neg C_{M1} \lor C_{JN,(M1,RC)},$$

    and finally to binary constraint

    $$y^c_{M1} - y^c_{JN,(M1,RC)} \leq 0. \tag{97}$$

    If M1 production is less than 500 MMscfd, the (M1,T) flowrate is pinned to 500 MMscfd (an extra binary variable to do so is not required, since C_{M1} can also force this). This implies that only enough production from JN is diverted so as to make the shortfall and the bulk is still carried by (M1,RC). The following

constraints relate M1 production to the binary variable y^c_{M1}

$$(1 - y^c_{M1})\, Q^L_{s,M1} + y^c_{M1}\, \varrho_g\, 500 - Q_{s,M1} \leq 0, \tag{98}$$

$$Q_{s,M1} - y^c_{M1} Q^U_{s,M1} - \varrho_g\, 500 \leq 0. \tag{99}$$

The following constraints force (M1,RC) flowrate to be equal to JN production rate when binary variable $y^c_{JN,(M1,RC)}$

$$(Q^L_{a,(M1,RC)} - Q^U_{s,JN})\, (1 - y^c_{JN,(M1,RC)}) - Q_{a,(M1,RC)} + Q_{s,JN} \leq 0, \tag{100}$$

$$Q_{a,(M1,RC)} - Q_{s,JN} - (1 - y^c_{JN,(M1,RC)})\, (Q^U_{a,(M1,RC)} - Q^L_{s,JN}) \leq 0. \tag{101}$$

The (M1,T) flowrate is bounded below by 500 MMscfd

$$\varrho_g 500 - Q_{a,(M1,T)} \leq 0. \tag{102}$$

The following constraint combined with the above will pin (M1,T) flowrate to 500 MMscfd when $y^c_{M1} = 0$:

$$Q_{a,(M1,T)} - y^c_{M1}\, Q^U_{a,(M1,T)} - \varrho_g\, 500 \leq 0. \tag{103}$$

2. *Processing capacity at M1 platform is 1300 MMscfd of which 750 MMscfd belongs to JN.*

Processing capacity at M1 platform is 1300 MMscfd

$$Q_{s,M1} + Q_{s,JN} - \varrho_g\, 1300 \leq 0. \tag{104}$$

750 MMscfd capacity belongs to JN production

$$Q_{s,JN} - \varrho_g 750 \leq 0. \tag{105}$$

Chapter 5

The Case Study

A case study is presented in this chapter to demonstrate the application of the modeling approach described so far. This case study has been carefully constructed so that it captures all the features of the SGPS model and hence, is a faithful and accurate representation of the application to a real-world production system. However, it is not a model of the Sarawak gas production system for the reasons outlined below. This has been done to preserve the confidentiality of original system parameters that are business sensitive.

1. The parameters in the system including reservoir pressures, compositions, water and condensate ratios, well performance parameters, field production bounds, constants in flowrate-pressure relationships and maximum demand rates have been altered from their values in the SGPS model. Hence the flowrate-pressure distribution in the infrastructure model is totally different from the SGPS model and does not relate to it in any way.

2. Although the trunkline network used in the case study is the same as the SGPS, the facilities have been moved around due to changes in the parameters.

3. The contractual model has been altered. However, it is of the same complexity as the SGPS contract model. The production-sharing contracts are as in Figure 4-2 and the set of rules is as presented in Section 4.4. The customer requirements

have been altered and are as in Table 4.1. Finally, operational rules included are the ones presented in Section 4.5.1.

Information on subsets of nodes and arcs in the infrastructure and the contract model can be found in Appendix A. Appendix B lists additional model parameter values for the case study. The results in this section are presented in industry units since the study is based on the SGPS and hence these are the most natural set of units for the purpose of comparison and analysis. The volumetric flowrates are in million standard cubic feet per day (MMscfd) and the NGL volume rates are in barrels per day (bpd). However the detailed results for the base case are presented in the model units, i.e., in SI units.

5.1   Planning Objectives

The objective functions considered in the proposed framework are operational objectives. The planning objectives are from the perspective of the single upstream operator operating the production system. This operator has the obligation to operate the system in such a way that all contractual rules and customer requirements are met. These form the constraints of the model and have been delineated earlier. Within these constraints, the operator may want to meet its own production targets. The model is a production-allocation model, therefore the optimal solution point (representing the production rates, flowrates and pressure distribution that satisfies all requirements) is of more interest than the optimal solution value.

Following are the three objectives that are of interest from an operational perspective:

1. The upstream operator is interested in maximizing the delivery of dry gas to the LNG plants. This is because the more gas sold, the higher the revenue for the operator. Furthermore, the gas supply is stipulated by the gas sales agreement and therefore, is the primary target for the operator. The gas flowrates cannot go above the maximum demand rates set by gas sales agreements. Mathematically this is a minimization since demand nodes are sinks and hence $Q_{s,i}$ is

142

nonpositive. This objective function is therefore bounded above by 0 and below by the negative of the sum of maximum demand rates $\sum\limits_{i \in \mathcal{N}_D} -\Lambda^U_{d,i}$.

$$z_g = \sum_{i \in \mathcal{N}_D} Q_{s,i} \tag{106}$$

2. NGL sharing is not governed by production-sharing contracts. Instead NGL are shared according to the ownership of the fields. Hence, if two production strategies produce the same amount of dry gas but different amounts of condensates, the one with higher condensate production is preferred since the upstream operator receives more condensate and hence a higher revenue.

$$z_L = -\sum_{i \in \mathcal{F}_w} Q_{Ls,i} \tag{107}$$

3. It may be of interest to the upstream operator to prioritize production from certain fields. The logic for doing so may come from long range production-planning models or reservoir-management models, that may dictate the long term production profile for a particular field. Over a short term, the interpretation of these profiles may be to prioritize certain fields. The simplest model for prioritizing production is to simply maximize the production from these fields. Define a set $\mathcal{F}_{pr} \subset \mathcal{F}$ that is the set of fields that should have high priority. Then this objective is represented as:

$$z_{pr} = -\sum_{i \in \mathcal{F}_{pr}} Q_{s,i} \tag{108}$$

In this case study, the high priority fields are assumed to be sour fields (in this context defined as the fields with higher CO_2 and H_2S content). Quality constraints favor a higher production rate from sweet fields and therefore, can lead to a faster depletion of these fields. This can result in a situation in the future when there is insufficient sweet gas available in the system to satisfy quality requirements and additional investments must be made in sour gas processing

facilities. A more sensible strategy is to produce the maximum possible amounts from sour fields and then blend the sour gas with gas from sweet fields so that quality specifications are just met and therefore, delay the capital investments as far out as possible into the future. All over the world, several new fields currently under development have a high content of CO_2 and H_2S and therefore, this is going to be an important concern in the future.

It should be noted that equalities (106), (107) and (108) are represented as constraints in the model to facilitate calculation of all the three quantities when using one of them as objective. The actual objective function is given by:

$$\min z$$
$$z = z_o, \tag{109}$$

where subscript o is either g, L or pr depending on the objective for the particular instance of the model. The domain of the optimization is not indicated explicitly to simplify notation, but it is over all the decision variables defined earlier.

5.2   Estimation of Bounds

The importance of estimating the tightest possible bounds for the decision variables in a nonconvex optimization problem is well known. The most important variable bounds on the system are bounds on the demand rates at the LNG plants, the pressures at the slugcatchers and the dry gas production rate from fields.

This problem has been observed to be especially sensitive to the bounds on the production rate. The optimal solution point is strongly influenced by the bounds set for production rates and pressures. A completely different solution point with the same optimal solution value can be found by varying the bounds. Roughly, the flowrate and pressure distribution in the network is driven by the bounds set on the field variables while the optimal solution value is dependent on the bounds on the demand rates and delivery pressures. Bounds are therefore as important as model

parameters in this problem and the discussion of the case study is incomplete without a statement about setting the bounds.

Appendix C contains an exhaustive discussion of the bounds. However, here are some of the general features (excluding the field dry gas production rate bounds that are discussed in the following subsection):

1. Upper bounds for the pressures at nodes corresponding to fields are set using the maximum reservoir pressure $\pi_{r,i}^M$ among all wells in that field. The lower bound is set to atmospheric pressure since the pressure does not matter when the field is shut down.

2. All the lower bound on flowrates and productions rates are generally set to zero (except at demand points).

3. The trunkline flowrate bounds are set either using the actual design capacities of the lines (that have been changed in the case study to preserve confidentiality) or are set by propagating either the field production rate bounds or the demand rate bounds.

4. The contract model bounds are set by propagating the production rate bound in the infrastructure model through the contract network.

5.2.1   Field Production Estimate

The field production rate bounds are derived from the well performance model. The maximum theoretical rate of production from a well is estimated based on the well performance model. This can be calculated by assuming that the tubing head pressure is equal to the atmospheric pressure. The bottom-hole pressure can be eliminated out of the equations (23) and (24) (page 93) to yield an upper bound on the production rate from a well:

$$
Q_{w,w}^U = \frac{-\alpha_w + \sqrt{(\alpha_w^2 - 4(\lambda_w \pi_{atm}^2 - \pi_{r,w})(\beta_w + \vartheta_w))}}{2(\beta_w + \vartheta_w)}.
$$

Finally, the maximum possible production from a field can be calculated simply by summing the maximum productions from the wells corresponding to that field

$$Q_{s,i}^{U} = \sum_{w \in \mathcal{W}_i} Q_{w,w}^{U}, \quad \forall i \in \mathcal{F}_w.$$

However, these bounds are not the tightest bounds obtainable for several reasons; the assumption about the wells producing to atmospheric pressure and a very simplified well performance model. The historical production data on the system can be used to estimate tighter bounds estimates. That is in fact the case with the SGPS model. Also, detailed reservoir management and well performance models may enforce rate limits on wells that can be used to fix the corresponding well and field production rate bounds.

In this case study, the bounds are estimated from the relationship presented above due to confidentiality reasons. For fields with no well performance modeling, the bounds are presented in Table C.4 in the Appendix.

5.3 Solution Approach

The final model is a MINLP with nonconvex constraints. The model is formulated in GAMS [115]. It has a total of 827 variables, with 804 continuous variables and 23 binary variables (reported using GAMS CONVERT by converting the model from a set-based GAMS representation to a scalar GAMS representation). There are a total of 1,086 constraints with 702 equalities (of which 220 are nonlinear) and 384 inequalities (not including variable bounds).

The model is solved with a branch-and-reduce algorithm [116, 117] as implemented in BARON 7.5 [118] with GAMS 22.2 (64-bit version). The CPU times are as reported by BARON on a 3.2 GHz Xeon dual processor machine running Linux kernel. SNOPT [119] was used as the NLP solver and CPLEX [120] was used as an LP solver for BARON. The constraint satisfaction tolerance was 10^{-6}.

All variables in the model are in the same units as presented so far with the

exception of pressures (that are scaled by a factor of 10 from the units in the paper) and NGL rates (that are scaled by a factor of 100 from the units in the paper). It must be noted that this scaling introduces multiplying factors for the parameters in the constraints involving these variables. Also, all composition units in the model are in mole fractions although the field compositions in Table B.2 (Appendix B) and quality specifications in Table 4.1 are in mole percentages.

5.4 Dry Gas Maximization

In order to elucidate the characteristics of the model and the solution, it is instructive to solve the following three cases with dry gas maximization (i.e., minimization of z_g). The first case excludes the PSC model and the gas quality specifications and is a solution of the infrastructure model with the rate and pressure constraints at demand nodes. The second case is the infrastructure model with gas quality constraints but excluding the PSC and operational rules. Finally, the third case is the entire model as presented.

All three cases are a MINLP. They are solved with a termination criterion of 1% relative gap between upper and lower estimates on the solution value. The base case (case (3)) was initialized with the solution from case (2). The lower bound on z_g was set equal to the lower bound on the solution value $z_g^{L,2}$ at termination (and therefore satisfying the termination criterion) obtained by solving case (2).

$$z_g^L = z_g^{L,2} = -102.934959524[1]$$

This is quite important since in the absence of this bound, the solution procedure fails to converge even in 26+ hours.

It should be noted that the solution times are not directly comparable since all three cases have different numbers of variables, both binary and continuous, as well as different numbers of constraints. Objective values and solution times for the three

---

[1]The precision in this constraint is more than the input data precision, however, BARON is sensitive to precision and therefore, the constraint is represented exactly as it was inputted

Table 5.1: Dry gas maximization objective

Case	DryGas MMscfd	NGL bpd	Sour Gas MMscfd	Best Possible MMscfd	Time CPUs
(1) No gas quality, PSC and operational rules	3,865	143,290	1,681	3,873	237
(2) No PSC and operational rules	3,599	151,688	1,188	3,635	205
(3) Full model (Base Case)	3,333	134,854	1,073	3,367	19,237

cases are compared in Table 5.1. "Best possible" column in the table presents the lower bound on the objective value at termination (of the corresponding minimization problem, therefore the best-possible value bounds the actual production value from above). The detailed solution for case (3) (i.e., the base case) is presented in Appendix D. An overview of the base case solution is shown in Figures 5-1 and 5-2.

The following general features of the problem can be noted:

1. The binary relaxation of the base case (which is a nonconvex NLP) can be solved with SNOPT in less than 1 second and the relaxed solution value (3,435 MMscfd) is very tight with respect to the base case solution value (3,333 MMscfd). However, this does not mean that getting a contractually feasible solution (i.e., integer feasible solution) is easy. This is clear from the fact that the time required for the solution of the full model is two orders of magnitude greater than the first and second cases.

2. The branch-and-bound on the base case behaves contrary to the usual behavior observed in nonconvex NLP and MINLP. The conventional wisdom is that branch-and-bound usually locates the global solution in the first 10-20 % of solution time and rest of the time is spent verifying that it is indeed the solution. However, in this instance, for just under two hours (close to 30% of the total solution time), the upper and lower estimates on the solution are 3,635 and 2,100, respectively, with the actual solution being 3,333 MMscfd. This behavior can be even worse in certain cases when weak estimates can persist until the
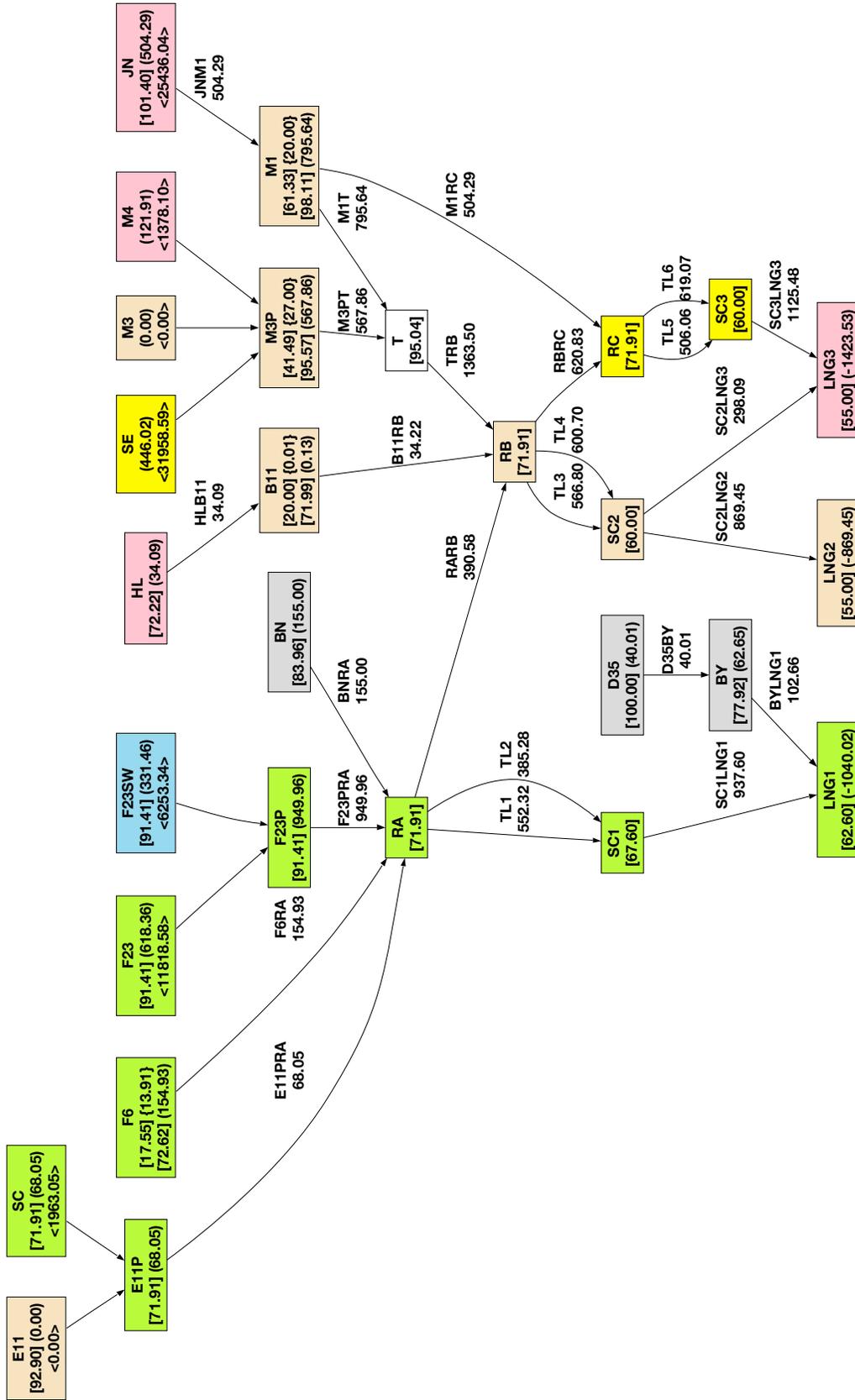
Figure 5-1: Dry gas maximization *base case solution*: Trunkline network (*Legend: [Pressure in bar], (production/delivery rate in MMscfd) <NGL in bpd>, {compression power in MW}, number next to a line: flowrate in MMscfd*)
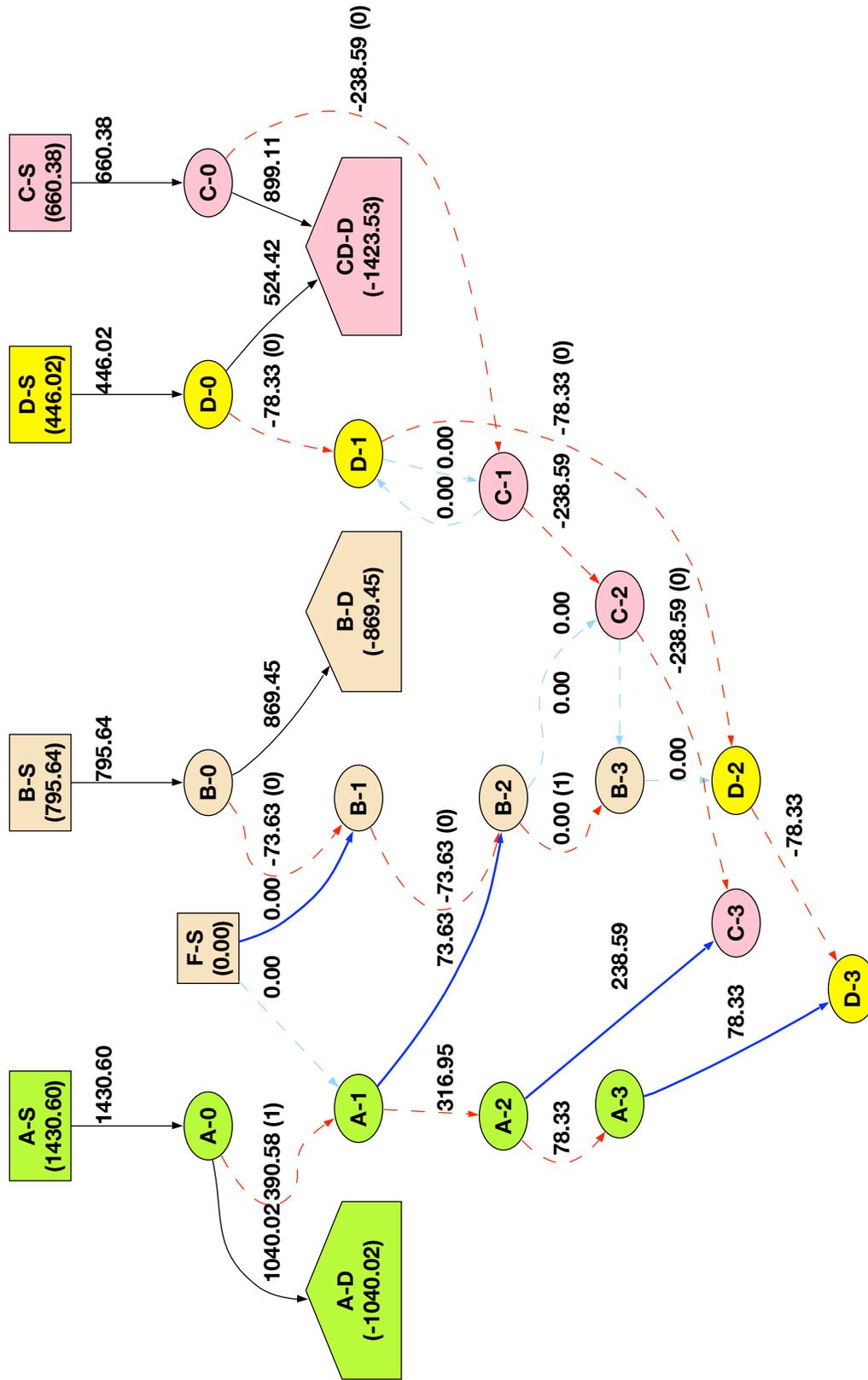
Figure 5-2: Dry gas maximization *base case* solution: State of production-sharing contracts

very end of the solution procedure ($>90\%$ of the solution time). A plausible explanation of this behavior may be that a very small subset of the hyperrectangle defined by bounds is feasible for certain values of the parameters and therefore unless a partition containing the actual solution gets small enough, the solution estimates fail to shrink. Therefore, global optimization algorithms are indispensable to solve this problem.

3. There are multiple globally optimal solutions to the problem with the same objective value due to the network structure of the problem where it is possible to deliver the same amount at the demand nodes with several different pressure, flowrate and production rate distributions in the network. This presents a problem in terms of choosing an operational state for the system. Some operational states may be more favorable than others (to a human operator) because of factors that may have not been represented in the model.

4. There is a dilemma between setting tight bounds and loose bounds. A problem with tighter bounds is easier to solve and the solution is more likely to be implementable on the network (because it is expected to be close to the current operating point). However, tight bounds can artificially restrict the feasible set of the problem, remove operational flexibility and therefore prevent discovery of unconventional and novel operational strategies. Finally, loose bounds indeed create problems for the convergence of the solution procedure.

5. The time required for solution can be very different even with a slight change in the parameters and the bounds.

The following are specific characteristics of the solution:

1. *Effect of Quality Constraints*: Between case (1) and (2), a major redistribution of production rates takes place to satisfy quality. Quality constraints force the production rates from fields with high CO_2 and H_2S concentrations, i.e., B11, M3 and F6 to decrease. Also GHV constraints and composition constraints on C_4 and C_5 force production from HL and M3 to drop since they contain less

C_1 and high levels of C_4 and C_5. On the other hand, production from fields such as F23, F23SW and BN increase, since they have high levels of C_1, which compensate for the drop from the other fields, although not fully since the overall production rate drops by 300 MMscfd. Although the difference between the objective values of case (1) and (2) is only 300 MMscfd, the solutions are actually very different and bear no resemblance to each other. A substantial drop occurs in LNG 3 delivery because the production from fields physically connected to LNG 3 drops.

2. *Effect of PSC Rules*: The incorporation of PSC rules results in further cuts in production from M1 and JN, which is only partially compensated by a rise in SC and M4 production rates and hence a net decrease in the delivery is observed. This decrease is also reflected by contracts B and C being in deficit. A redistribution of delivery between the LNG plants takes place due to PSC rules. In particular, there is a major decrease in contract B production and therefore LNG 2 supply. This leads to excess contract A supply being freed up to supply LNG 3 via contract C and D and leads to an increase in LNG 3 delivery.

3. Supply from contract F is close to zero due to quality constraints being in force since the field producing under F is a high CO_2 field.

5.5   Comparison of Different Objectives

The full problem can also solved with the other objectives described in Section 5.1. A summary of the results appear in Table 5.2. All these runs were started without any initial guess to test if convergence is possible in absence of a sensible initial guess. However, it is possible to use an initial guess from the dry gas maximization case for the other solutions since this solution is feasible for all other runs. It is also possible to add bounds to the objective variable as described in the previous section that can significantly accelerate the convergence. The NGL and priority field objectives fail

Table 5.2: Case study: Various optimization objectives

Objective	DryGas	NGL	Sour Gas	Best Possible	Time
	MMscfd	bpd	MMscfd		CPUs
Dry Gas Maximization	3,252 (3,333[a])	128,627	1,048	3,614	41,424[b]
NGL Maximization	3,204	141,584	1,086	149,757	24,238
Priority Maximization	3,276	136,489	1,144	1,248	16,042

[a] 1% gap, 19,237 CPUs

[b] The counter-intuitive observation that the solution time with 10% gap is greater than the base case (1% gap) stems from the fact that the base case solution procedure was initialized with an initial guess and included a lower bound on the objective.

to converge even after 24 hours with a 1% relative termination criterion. Hence, all the runs are solved with 10% relative termination criterion although as presented in previous section, dry gas maximization can be solved with a 1% relative termination criterion.

M4 and E11 production rates decrease in the NGL maximization case, while F6, SE, SC and M1 production rates increase since fields with high condensate-gas ratios are favored in this solution. BN and HL production rates decrease as well because they do not contribute to NGL production. The delivery rate to LNG 2 decreases substantially while delivery to LNG 1 increases. This is because contract C is in deficit due to a decrease in M4 production, forcing contract A to transfer gas to C and therefore less gas is available in contract A for transfer to contract B, decreasing LNG 2 rates.

The sour gas fields for the priority maximization run have been chosen to be the fields that produce gas with high CO_2 and H_2S content, i.e., B11, F6, E11, M1 and M4. Not all sour field production rates increase in the priority solution due to the quality specifications. Indeed, E11, M4 and B11 register small decreases from the base case, however the F6 and M1 production rate increases overcompensate this decrease resulting in a net increase in the objective. Delivery to LNG 3 increases mostly due to inter-contract transfers from contract A since contract A has excess gas available due to the decrease in LNG 1 demand rate as well as from an increase in F6 supply. It should be noted that sour gas maximization yields a higher dry gas delivery than

Table 5.3: Hierarchical multi-objective case study

Objective	DryGas MMscfd	NGL bpd	Sour Gas MMscfd	Best Possible	Time seconds
Base Case	3,252 (3,333[a])	128,627	1,048	3,614	41,424
NGL Maximization	3,252	131,960 (141,370[b])	1,048	146,648	86
Priority Maximization	3,276	138,502	1,144	1,211	4,717

[a] 1% gap, 19,237 CPUs

[b] 1% gap with a restart using objective cut as outlined in Section 5.4 and initial guess solution with 10% gap, 1,694 CPUs

the base case because of the 10% relative termination criterion employed.

5.6 Hierarchical Multi-Objective Case Study

As mentioned earlier, there are multiple globally optimal solutions to the problem. Moreover, there are multiple objectives for the operation of the system. This feature of the problem can be leveraged to obtain a solution that maximizes several objectives in a hierarchical way [121]. There is a clear hierarchy of objectives as follows:

1. The first priority is to supply the required amount of dry gas (represented by the maximum demand rates at the LNG plants), since this is mandated by production-sharing contracts. Therefore maximization of the dry gas production rate is the top priority for the upstream operator and only then a secondary production target can be considered.

2. The second priority is to maximize the NGL production rate because it is beneficial for the upstream operator from a revenue perspective. Moreover, NGL production is not governed by production-sharing contracts.

3. Finally. the last priority is to maximize the production rate from some subset of fields. This may be important to follow the long-term plans, however, it is not as important as the two other objective in the short-term. In this work,

this set is assumed to the same set of sour fields as described in the previous section.

Getting a Pareto-optimal solution for this problem is unnecessary and not very useful. For example, there is no explicit trade-off between gas and NGL production, and increasing gas production rate does not have any straightforward implications for NGL production rate due to the fact that a specific gas production rate may correspond to multiple NGL rates. Furthermore, as long as the gas production rate is below the combined maximum demand rate of all LNG plants, the top priority is maximization of gas production rate and it is meaningless for the upstream operator to optimize NGL production rate. Finally, determining a Pareto-optimal solution for this problem is prohibitively expensive due to nonconvexity.

This hierarchical multi-objective optimization is done as follows:

1. First, the MINLP is solved with the first objective, i.e., maximizing the dry gas production rate (minimizing z_g). Let the optimal solution value be z_g^{opt} (this is same as the base case).

2. Next the upper bound to z_g is set to z_g^{opt}

$$z_g^U = z_g^{opt}.$$

   Also upper bounds are added to the other objectives (z_L and z_{pr}) to get a better value than this solution (z_L^{p1} and z_{pr}^{p1} respectively). This also helps to accelerate convergence.

$$z_L^U = z_L^{p1},$$
$$z_{pr}^U = z_{pr}^{p1}.$$

   The MINLP is initialized with the previous solution point and is solved with the second objective, i.e., to maximize the NGL production rate (minimize z_L).

3. The upper bounds of z_g and z_L are set at z_g^{opt} and at z_L^{opt}, respectively, i.e., the

values of these variables in the solution of the second instance:

$$z_g^U = z_g^{opt},$$

$$z_L^U = z_L^{opt}.$$

Similarly the upper bound of z_{pr} is fixed at the value of this variable z_{pr}^{p2} from this solution

$$z_{pr}^U = z_{pr}^{p2}.$$

The MINLP is initialized with the solution point from second problem and is solved to minimize z_{pr}.

It should be noted that the solution obtained at each step is very different in terms of the pressure-flowrate distribution in the network.

All runs are solved with a 10% relative termination criterion because although the dry gas maximization step can be solved with 1% relative termination criterion, the NGL maximization fails to converge in 24 hours with a 1% relative termination criterion. A summary of the results is presented in Table 5.3.

The total NGL production rate after the second run is roughly 2.5% higher than the base case NGL production while maintaining the same production rate for dry gas. However the production rate distribution in this solution is almost the same as with the dry gas maximization. This is because the NGL production in the dry gas maximization solution is very close to being within 10% of the bounds and a minor change satisfies the convergence criteria. Since the gap is big enough, a second case with the NGL maximization objective was run with a 1% relative termination criterion and an initial guess and objective variable bounds from this solution. This offers an improvement of 10% in the solution value. This improvement is substantial in financial terms to the upstream operator in comparison to the base case solution. The priority field solution value shows an increase of 10% over the base case, increasing their share from 32% to 35% of total production.

This approach is a very powerful tool for planning operations in the system with multiple operational objectives and criteria, and in the presence of multiple optimal solutions. Innovative operational strategies may be constructed by carefully choosing the set of objectives. In a sense, this approach is more viable to narrow down the operational choices than employing symmetry-breaking constraints because the form and nature of such constraints from an operational perspective is unclear. Customized solution approaches that exploit the problem structure and avoid redundant computation for the multi-objective case study may be required to solve the problem efficiently.

Chapter 6

Comparison With the Existing Approach

The motivation for this analysis is to compare the proposed MINLP approach to the existing planning methodology to quantify the benefits of using the proposed approach. The approach proposed in the previous chapters, i.e., formulating the upstream planning problem as a mixed-integer nonlinear program and solving it with a global optimization algorithm, has been referred to as *the proposed approach* (also *the proposed model* or *the proposed framework*) as well as *the MINLP approach* (also *the MINLP model* or *the MINLP framework*). The approach that is currently in use for upstream planning is termed as *the existing approach*. The presentation of the comparison between the two approaches here is mostly qualitative in nature due to the business-sensitive nature of the information involved.

6.1   The Existing Approach: Overview

The existing approach employs a commercial software suite that is an integrated production modeling environment designed for oil and gas production and is intended to model the entire upstream system comprising reservoirs, wells, trunkline network and surface facilities. It comprises several subsystems:

1. A reservoir prediction tool that models the reservoir dynamics and generates

reservoir performance data.

2. A well performance component that models flow in wells.

3. A modeling framework for the surface network and facilities.

All components, i.e., well performance models and data, reservoir models and surface models can be integrated into a single model for simulation and optimization. It has a local nonlinear programming (NLP) optimization engine built into it that can locally optimize the entire model. All these feature are leveraged to create a full model of the system. The advantage of this approach is that there is a single software environment for simulation, optimization and calibration purposes.

The suite is primarily designed for oil production and the ability to simulate and optimize upstream gas production systems has only been added as an afterthought. The result is that there is no ability for full composition tracking throughout the network, a serious drawback because of the presence of CO_2 rich fields and gas quality specifications. There is also no direct support for the modeling of PSC and operational rules that are crucial to the operation of upstream systems. For example, the software cannot handle logical conditions on the production infrastructure, e.g., shutting down selected trunklines and facilities under certain conditions. These factors have led to development of ad hoc and complicated procedures inside the model to handle these issues. Such approaches are bound to create model maintainability issues in the long run. Finally, the optimization algorithms being employed are local NLP algorithms that are well-known to be unreliable for nonconvex problems as was shown in Section 2.4.1 (page 68). They are expected to perform even worse on a blackbox model with embedded procedures and iterative calculations that is likely to be nonsmooth or discontinuous.

6.2 Scope

Two entirely different approaches cannot be compared simply by comparing the results of the two models. A rigorous approach for reconciliation of the two models and a

metric for comparison must be defined. Towards this end, the following were defined as the objectives for this study:

1. Reconcile the proposed approach to the existing approach to the extent possible.

2. Ascertain the capability of the proposed approach to reproduce the solution of the existing approach for production infrastructure.

3. Check the solution of the existing model for feasibility in the proposed framework comprising gas quality specifications and PSC and operational rules.

4. Compare performance of the reconciled models in terms of dry gas production rate.

5. Ascertain the benefits of the capability of the proposed approach to deal with various operational objectives and compare solutions thus obtained with the solution obtained from the existing approach

6. Ascertain the potential value added by the hierarchical multi-objective optimization approach in comparison with the solution of the existing approach.

6.3 Model Reconciliation

The results from both models cannot be compared directly as they have not been calibrated to a common data-set. The proposed approach must be reconciled against the existing one. This section qualitatively describes the methodology used to perform this reconciliation. The formulation used for the proposed approach is the same as the alternative formulation presented in Section 3.6 (page 96).

6.3.1 Well-performance Model

There are two main components of the well-performance model that need to be reconciled: in-flow performance and vertical-lift performance.

In-Flow Performance

In-flow performance (IFP) in the proposed approach is given by the following pressure-flowrate relationship (of the same form as Equation (xxviii) page 105 in alternative formulation):

$$\alpha Q + \beta Q^2 = (\pi_r + \Delta \pi_r)^2 - P_b^2, \tag{IFPA}$$

where Q is well production rate and P_b is the bottom-hole pressure. Reservoir pressure π_r and pressure shift $\Delta \pi_r$ for all wells are directly available from the existing model and were used as given. The inflow performance parameters (Darcy coefficient α and non-Darcy coefficient β) for most wells are available in the existing model. However, for several wells, a "psuedo-pressure" formulation is used and therefore the usual pressure drop relationships cannot be used. A response table was constructed manually by using the software to evaluate bottom-hole pressure for different values of flowrates and a linear pressure drop relationship (same as Equation (xxix) in alternative formulation) as follows:

$$\kappa Q = \pi_r - P_b. \tag{IFPB}$$

Vertical-Lift Performance

Vertical-lift performance (VLP) in the proposed model is modeled as follows (same as Equation (xxx) page 105 in the alternative formulation):

$$\vartheta Q^2 = P_b^2 - \lambda P_t^2, \tag{VLP}$$

where P_b is bottom-hole pressure and P_t is well-head pressure. A calibration of VLP between the existing model and the proposed model is considerably more involved than the in-flow performance model. The suite uses look-up tables for VLP predictions. These tables list tubing-head pressure, CGR, WGR, well production rate, flowing bottom-hole pressure and other information. A look-up table cannot be used in conjunction with the proposed approach as global optimization algorithms require

explicit functional representation of relationships. Moreover, the above VLP model cannot accommodate flowrate dependencies on CGR, WGR and temperature. The data in look-up tables is listed for 2-5 different values for CGR and almost the same number values for WGR. The first step is to choose CGR and WGR values that are the closest to the CGR and WGR of the well under consideration, irrespective of the temperature. In most cases, the pressure range for bottom-hole pressure was also narrowed down since the table contained flow data points for very low bottom-hole pressure ($<$ 10-20 bar) or very high pressure ($>$ 500 bar) that were not likely to be observed during normal operation. Therefore, one ends up with a table of flowrates, bottom-hole pressures and tubing-head pressures for the requested CGR and WGR range. A linear regression (using GNU R [122] statistical package) was then used to fit ϑ and λ in the above equation. The above procedure was partially automated using a Perl script, but still required quite a bit of manual intervention in the filtering step. The fits are generally quite good for most wells.

Additional Parameters

Condensate Gas Ratio (CGR) and Water Gas Ratio (WGR) are available for all wells from the existing model and were used as given. Some wells also have a maximum rate constraint in the existing model which was enforced via well production rate bounds in the MINLP model.

6.3.2 Gas Composition

Composition is available individually for fluids from each well in the existing model, though it is not used to track composition through the network. The MINLP model is formulated to accept only a single composition per field. While composition for wells belonging to the same field is the same for most cases, it does differ among wells for a few fields. In such cases, compositions of the majority of wells was chosen as the composition of the field. In cases where a composition was not available in the existing model, i.e., associated gas fields and third party fields that are modeled as

fixed sources, composition data from the earlier SGPS model were used.

It may be possible to easily extend the MINLP model to accept differing composition for wells if it is beneficial to do so. The extra constraints and variables introduced by doing so will be linear, so presumably the resulting MINLP, though larger than the current one, may still be tractable.

The existing model has a complex surface network that has numerous units. It is neither desirable nor probably feasible to duplicate the network in the MINLP model exactly as it is in the existing model. Hence, a simpler trunkline network was derived from the detailed network that represents the system with the sufficient fidelity required for planning.

The standard gas pressure-flowrate relationship has been employed for the pressure-flowrate relationship drop in most pipelines

$$P_i^2 - P_j^2 = \kappa_{(i,j)} Q_{(i,j)}^2,$$

where P_i and P_j are upstream and downstream pressures, respectively. The issues involved in estimating the pressure drop coefficient $\kappa_{(i,j)}$ are similar to the problems outlined earlier in regressing VLP parameters. A look-up table is available from the software suite that lists a combination of several sets of CGR, WGR, upstream and downstream pressures, downstream temperature, volumetric flowrate, maximum line pressure, velocity and so on. Obviously the simplified standard gas flow relationship above cannot represent flowrate dependency on all these factors.

A range of CGR and WGR that should be feasible for a particular line was deduced from the CGR and WGR of the fields feeding a particular line. The look-up table was filtered corresponding to this feasible range of CGR and WGR (ignoring other factors) generating a data-set containing just the upstream and downstream pressure, and volumetric flowrate for a particular line. The pressure drop constant $\kappa_{(i,j)}$ can then be estimated using linear regression (using GNU R) with the standard gas flow equation. Again, a Perl script was deployed to automate this procedure partially, but filtering the data-set required substantial manual work. For most trunklines, the fit

164

is reasonable.

However, for some trunklines, the above single parameter equation does not provide good results. This may be due to the unique configuration of these lines. A two parameter equation (Equation (ii), page 98) similar in form to the VLP relationship is used for these lines that provides a relatively better fit. There is also a set of trunklines in the existing model that do not seem to model any pressure drop, e.g., lines from associated and third-party fields. These trunklines in the MINLP model have been modeled simply by enforcing a pressure inequality between the upstream and downstream pressure, and allowing any flowrate up to an upper bound. Also, most of these lines carry fixed flowrates as they are fed by associated and third party fields that have fixed source flowrate, and so they are mostly inconsequential to the planning problem.

Lines that can be closed or open during normal operation are modeled using Equations (iii) and (iv) as detailed in Section 3.6.2 (page 97). Complex platform configurations have been modeled using the approach outlined in Section 3.6.3 (page 100). This circumvents the ad hoc and complicated modeling paradigm used in the existing model to represent such configurations. Composition tracking combined with this formulation can automatically determine how the flow around such hubs must be routed so that quality constraints are met at LNG plants.

6.4   Issues in Reconciliation and Comparison

It is important to note that a complete reconciliation and a rigorous comparison of both models was not possible within the scope of this analysis. The following are some reasons for problems with reconciliation and comparison in this particular study:

1. The existing model relies on ad hoc and complicated approaches to divert flows and switch pressure inequalities at junctions and splitters. In principle, it is possible to functionally recreate these in the proposed model using extra constraints and integer variables. However, due to unnecessary complexity of the existing model, there is no rigorous way to guarantee that the functionality of

165

every relevant script and source has been duplicated in the MINLP model.

2. No attempt has been made to duplicate the paradigm used in the existing model to represent PSC rules since the MINLP model relies on the logical modeling of the PSC and operational rules and automatically satisfies them in the solution.

3. Compression parameters cannot be reconciled between the two models. The existing model uses a polytropic model that relates RPM of the compressor to flowrate, compression ratio and power. However, the explicit relationship is not available from the documentation. So the compression subcomponents of the models were not reconciled and the compression model and parameters from the earlier SGPS model were used in the proposed approach.

4. It is not clear how pressures at fixed source fields is being calculated or set. In absence of this information, pressures at these nodes were left free.

5. The suite used for the existing model makes the export of model parameters and data tables difficult. IFP parameters for wells had to manually copied one at a time. Formats of VLP and trunkline tables exported to text files had to be reverse engineered by comparing values in the software interface since copying them from the interface is not even possible. The results files were not easy to parse automatically. Due to the large amount of manual work involved there is always a possibility of errors, even though extreme care has been taken to avoid any.

6. Some of the relationships in the existing model were not clear. To get around this and proceed further, either they were reverse-engineered or certain assumptions about the form were made. An example of reverse-engineering is the inflow performance equation, where it was not clear what role parameter $\Delta\pi_r$ (in Equation (IFPA)) plays and was not available in the documentation. A plot of bottom-hole pressure had to be examined and compared to predictions of the various trial-and-error forms of the relationship before arriving at the correct one. Another example is that at least one well in the system lists two local

reservoir pressures. It was unclear in this case what should be the form of in-flow performance relationship. One set of parameters was chosen along with the normal in-flow relationship. In some cases, it is therefore conceivable that the relationships inferred might be incorrect.

7. There was also insufficient information about the existing model with regard to its deployment in production use. This includes issues such as what kind of inputs can be provided to the model, what inputs or parameters can or cannot be changed, what kind of different case studies are run and how the results are extracted and used.

8. The existing model is a much more detailed model than the proposed model. It is therefore difficult to ascertain if there is a constraint that is present in the existing model but has not been duplicated in the MINLP model. The only way to test this is to attempt to replicate the proposed MINLP solution in the existing model which was outside the scope for this study. Again a best attempt has been made to duplicate all constraints, but it is not possible to guarantee that everything has been replicated.

9. The existing model has lots of left over disabled elements (probably from his-torical model development and from development in progress). It is therefore difficult to reproduce the active part of the network. For example, in several cases, there are multiple paths between two points in the network of which only one is used. It is quite easy to not realize this and pick a wrong configuration, though extreme care was taken to avoid this.

To summarize, an entire reconciliation between the two models was not feasible since the existing model is designed as a system-wide simulation model and not ex-plicitly formulated as an optimization model. A complete reconciliation with all units may result in a model with a non-tractable MINLP with no added benefits in terms of improved predictions or forecasting. The existing model is an overly complicated solution for the purposes of operational planning.

6.5 Model Comparison: Reconciliation Tests

As a first step in comparison, it is instructive to test the solution of the existing model in the proposed model. This is done to accomplish the following two objectives:

1. To gauge whether the model reconciliation is sufficient with respect to the infrastructure component of both models and whether it is reasonable to make a meaningful comparison between the results from the two approaches.

2. In the case that the first test succeeds, it is important to ascertain if the existing model solution respects the PSC rules and gas quality specifications by running this solution through the MINLP model.

6.5.1 Solution of the Existing Model: Feasibility in the Proposed Infrastructure Model

In these runs, only the infrastructure component of the MINLP model is used to check if the production infrastructure components in both models agree with each other. The obvious approach is to satisfy all the degrees of freedom in the MINLP model using the solution of the existing model to test feasibility. However, this is almost certain to fail and end up with an infeasible model because the two models can never be calibrated to such an extent that the solution of one will be feasible in the other due to the differences in governing relationships being used in both models. So the next best possible alternative is to fix some natural set of variables (which is a strict subset of the total number of free variables) from the structure of the problem. There are several ways in which this can be accomplished. Three possible ones attempted here are as follows:

1. Fix the well production rates using the solution of the existing model: this fixes most variables on the well side but leaves routing and delivery profiles in the surface network free.

2. Bound the delivery rates and pressures using the existing model solution: this

fixes the delivery profile and part of the surface network, but leaves well allocation free.

3. Fix the well bottom-hole pressures from the existing model solution: this is similar to the above case of fixing well flowrates, but may numerically perform differently.

Well Production Profile Test

In the well production profile test, all well flowrates except a particular well in the MINLP model were fixed to the solution obtained from the existing model. The MINLP model slightly under-predicts the wellhead pressure in the aforementioned well at the existing model production rate compared to the pressure predicted by the existing model which results in a low delivery pressure at LNG plants and hence, the rate at this well is left free. In all the MINLP solutions discussed later in the well production rate test, production rate in this well is only slightly less than the rate at the existing model solution. The demands were bounded from above by Maximum Demand Rate (MDR) obtained from the existing model. The rates were left free from below to test if the delivery profile of the existing model solution can be reproduced. The model was solved for the dry gas maximization objective without quality and PSC constraints.

The delivery profile at LNG plants can be reproduced to within 10% in the proposed model without being forced by the bounds at demand nodes. This is quite impressive as it shows that in spite of calibration issues with wells and network, overall the surface network in the proposed framework is sufficiently calibrated to reproduce the delivery profile. However, there are also discrepancies between the solutions as follows:

1. There is a discrepancy between NGL production rate between the two models that is unexpected since well and fixed-source flowrates are fixed in the MINLP model to the same value as in the existing model and hence corresponding liquids productions should be the same under a constant CGR assumption. An

169

analysis reveals that it is due to discrepancies in NGL production rates for a couple of particular wells and a field.

2. The proposed model over-predicts the pressure drop in the system. Anything above a certain threshold is deemed infeasible. As is discussed later, this issue is unique to the "reproduction of well production profile" test in the proposed model and in general, the proposed model can generate higher delivery pressures. This is a reconciliation issue in the production infrastructure model, however, it is quite difficult to trace the origin of under-prediction.

To solve the issue related with liquids production for the two wells discussed above, CGR was recalculated from the actual production value and used in the MINLP model instead of the value from the existing model. The other discrepancy with a field cannot be addressed satisfactorily as it involves accepting a discrepancy in either natural gas production or in liquids production and the reconciliation of the models is based on gas productions.

With the exception of slugcatcher pressures, this solution is surprisingly close to the solution obtained from the existing model. The delivery rates at individual LNG plant have less than 3% discrepancy with the ones in existing model solution. The MINLP infrastructure model so calibrated has been used as the base case for further studies and all further references to the MINLP model refer to this corrected model.

Delivery Profile Test

A demand profile test can also be carried on the proposed model to further test the reconciliation. This test is expected to indicate whether the MINLP model can match the delivery profile, especially pressures at the slugcatchers in the solution of the existing model. The well profile test previously shows that the pressures at slugcatchers are under-predicted.

The well flowrate in the MINLP model are set free subject to maximum productions rates bounds if available and natural flowrate bounds otherwise. The demand delivery rates and pressures obtained from the solution of the existing model are

170

forced as bounds in the MINLP model and the model is solved for the dry gas maximization case. In this case, the MINLP model produces 9% higher total delivery rate than the existing model since the constraints force the model to deliver at least as much as the solution of the existing model at a pressure at least as high as the existing delivery pressure but otherwise allow it go above it both for delivery rate (subject to MDR) and pressure. The pressures at slugcatchers agree more closely for this case ($<$1% discrepancy)

Well Bottom-hole Pressure Profile Test

It is also theoretically possible to reproduce the solution of the existing model by fixing bottom-hole pressure in the MINLP model. However, when attempted this fails because bottom-hole pressures reported in the existing model for several wells (especially the ones producing at low flowrates) are greater than the (local) reservoir pressure which makes it infeasible in the existing model.

6.5.2 Existing Model Solution: Quality Specifications and PSC Rules Feasibility

The methodology in this test is similar to the well production profile test. The well production rates are fixed to the solution of the existing model in the MINLP model. The quality and PSC rules are switched on successively to test if the solution of the existing model violates any of them. The delivery bounds are not enforced in this case. Therefore, it is the earlier well production profile test with the PSC and quality constraints added. Following are the conclusions of this test:

1. The existing model solution is infeasible in the sulfur and C_{5+} quality constraint, i.e., composition exceed the threshold mg of sulfur per unit volume of gas.

2. The existing model solution meets all the other quality constraints, specifically, GHV, H_2S, CO_2, N_2, C_1, C_2, C_3 and C_4. However, it is unclear if this is a coincidence or is by design.

3. The solution of the existing model is feasible with respect to all PSC and operational rules except one operational rule.

The run of the full MINLP model with all well rates fixed and with the constraints corresponding to the two violated quality specifications and the violated operational rules dropped, is compared with the existing model solution. The agreement is excellent for delivery volumes, less than 1% discrepancy for total rate and 2.5% for individual LNG plant rates, but not so close for slugcatcher pressures (almost 10% discrepancy for one LNG plant and <5% for the other two). The disagreement in pressure was already expected from the earlier well production profile test.

All further references to the MINLP model refer to the MINLP model that has the corrected CGR values and the constraints corresponding to the quality and the operational rule dropped for a fair comparison to the existing model solution. It should be pointed out that although the operational rule is dropped out of the MINLP model, all further solutions do respect it.

6.6   Model Comparison: Performance Gains

Three objectives were chosen in the MINLP model to compare performance with the existing model solution. These are as follows:

1. maximize dry gas production,

2. maximize NGL production,

3. maximize production from sour fields. These fields are chosen in this case study to be high CO_2 fields.

The existing model is not flexible enough to handle a variety of objectives as the MINLP model. There are two comparison sets that were run to ascertain the performance of the MINLP model. The first test on various objectives is to run them independently of each other and see what performance gains are possible for each objective. The second test is to run various objectives through the MINLP hierarchical

Table 6.1: Comparison with the existing approach: Independent optimization objectives[a]

Objective	Dry Gas	NGL	Sour Gas
	%	%	%
Dry Gas Maximization	8.8	6.3	144.1
NGL Maximization	4.1	15.9	16.7
Priority Maximization	9.2	8.9	424.1

[a] The table shows percentage improvement over the existing solution. The actual rates cannot be disclosed due to their business-sensitive nature. The numbers presented are $(p - e_r)/e_r \times 100$, where p is the particular rate in a solution of the proposed approach and e_r is the same rate in the reference solution of the existing approach.

multi-objective approach. All tests are run on the full MINLP model with corrected CGR values and excluding the problematic quality specifications and operational rule as discussed earlier. The MINLP model has 1,026 continuous variables and 24 binary variables with 1,390 constraints. The model is solved with a global branch-and-reduce algorithm as implemented in GAMS 22.5/BARON 7.8.1 [123].

6.6.1 Case Study: Independent Optimization Objectives

All these runs were run independently of each other. A summary of the comparison is presented in Table 6.1, which compares the solution of each proposed model run (with a different objective) to a single reference solution of the existing model. It is important to note that all solutions are totally different in terms of field and well allocation. The termination gap was 10% for the NGL and sour gas maximization.

Following can also be noted about these solutions:

1. The MINLP model produces more gas than the existing model solution in dry gas (around 9%), NGL (around 4%) and sour gas (around 9%) maximization cases.

2. NGL production in the MINLP model is also higher in all cases, roughly 6%,

16% and 9% over the existing solution for dry gas, NGL and sour gas objectives respectively, even after discounting the discrepancy.

3. Sour gas production can be increased roughly five times with the same dry gas production in sour gas maximization, which is quite important for distributed depletion of all fields.

6.6.2 Case Study: Hierarchical Multi-Objective Scenarios

The hierarchical multi-objective scenario study intends to highlight the additional benefits that can be obtained with the MINLP model. The approach for hierarchical multi-objective scenario is as outlined in Section 5.6, i.e., successive solutions of MINLPs with objectives in order of their priorities. This procedure generates an entirely different rate-pressure profile for each solution that still has the same level of dry gas production rate but a better secondary objective value (NGL production rate in this case). This procedure can be repeated several times to successively narrow the operational choices through a wise choice of objectives. A summary of the comparison between the solution values of the proposed approach and the existing approach for the hierarchical multi-objective case study appears in Table 6.2. The table compares the solution of each proposed model run with a single reference solution of the existing model.

The following are the important features of this solution:

1. A hierarchical optimization involving the NGL maximization as the first secondary objective yields roughly a 25% increase in the NGL production over the existing model and roughly a 17% increase over the MINLP model with dry gas optimization objective. It is important to point out here is that the dry gas production rate can be maintained at the same level.

2. Similarly, a second run with sour gas maximization as a tertiary objective yields a sour gas production rate that is roughly 4 times the sour gas production rate in the existing model solution and 50% more than the MINLP model's dry gas maximization case.

Table 6.2: Comparison with the existing approach: Hierarchical multi-objective study[a]

Objective	Dry Gas	NGL	Sour Gas
	%	%	%
Dry Gas Maximization	8.8	6.3	144.1
NGL Maximization	9.2 (0.3)[b]	24.6 (17.2)	218.5 (30.5)
Priority Maximization	9.2 (0.3)	24.6 (17.2)	278.7 (55.1)

[a] The table shows percentage improvement over the existing solution. The actual rates cannot be disclosed due to their business-sensitive nature. The numbers presented are $(p - e_r)/e_r \times 100$, where p is the particular rate in a solution of the proposed approach and e_r is the same rate in the reference solution of the existing approach.

[b] Number in parenthesis are percentage improvements for the rates, defined similarly as above, but with respect to the rates in the proposed model solution for the dry gas maximization objective (instead of the existing approach solution).

It is easy to extend this approach further by choosing more objectives and solving problems at further levels, e.g., maximization of production from a certain field, choosing a PSC or operational rule to violate in case the problem in infeasible, maximizing pressures at certain nodes and so on.

6.7 Summary of Comparison

It is evident from the results that the proposed approach seems to be very promising and warrants a further consideration. However, there are issues with the model calibration and reconciliation. It may be interesting to reproduce the MINLP model results in the simulation model which was not attempted in this study due to insufficient expertise with the existing model and the software suite

However, the level of detail in the existing model, while making it a good system-wide simulation model, precludes its use as a reliable operational planning and optimization tool. It may be better to split these functions into two models that have different levels of fidelity. A model similar in nature to the proposed model that includes production infrastructure only to a detail sufficient for an optimization model

but, on the other hand, includes all customer requirements and production-sharing, commercial and operational rules, can be used for operational planning. A detailed model of the production infrastructure such as the current existing model can then use this solution to generate operational parameters, e.g., actuator control information to run the system and in the process, also validate that the optimization solution can actually be reproduced on the system. Of course, both models must be calibrated to each other and to the real system periodically. For the proposed model, it is possible to partly automate the calibration to the real system. This will result in a more maintainable, robust and a simpler tool than the current practices.

Chapter 7

Global Optimization of Algorithms: Applications to Upstream Gas Networks

Mathematical programs with embedded computer evaluated procedures defining constraints and objective functions arise in several areas where a sequential calculation is employed to calculate a quantity of interest. Such programs more often than not are expected to be nonconvex and therefore require a global optimization approach for their solution. This chapter discusses the application of such an approach to the upstream gas networks described earlier.

Global optimization requires formulation of lower and upper bounding procedures. The lower bounding approach here is based on the approach described in Mitsos et al. [124] that is based on formulating subgradient propagation rules for the McCormick relaxation theory [125] for factorable functions, and employing automatic (algorithmic) differentiation (AD) theory to propagate convex/concave relaxations and the corresponding subgradients. The resulting convex nonlinear programs are nonsmooth and are solved with a simple bundle algorithm which is used as a linearization heuristic for generating LP relaxations. The upper bounding problem is solved as a normal NLP using AD to generate gradients with respect to input variables, as described later. Finally, the entire framework is integrated with a Branch-and-Bound algorithm with

reduction heuristics for variable bounds. An application of such a solution approach and its advantages are demonstrated for upstream gas networks.

7.1 Motivation

Consider a general nonconvex nonlinear program with continuous variables as follows:

$$\min_{\mathbf{w}} \quad f(\mathbf{w})$$
$$\mathbf{g}(\mathbf{w}) \leq \mathbf{0}$$
$$\mathbf{h}(\mathbf{w}) = \mathbf{0}$$
$$\mathbf{w} \in \mathbb{W} \subset \mathbb{R}^n$$

Denote a partition of the decision variables $\mathbf{w}$ as $(\mathbf{x}, \mathbf{y})$ such that $\mathbf{x} \in \mathbb{X} \subset \mathbb{R}^{n-m}$ and $\mathbf{y} \in \mathbb{Y} \subset \mathbb{R}^m$. Assume that the problem has the following special structure:

$$\min_{\mathbf{x}, \mathbf{y}} \quad f(\mathbf{x}, \mathbf{y})$$
$$\mathbf{g}(\mathbf{x}, \mathbf{y}) \leq \mathbf{0}$$
$$h_i(\mathbf{x}, \mathbf{y}) = 0, \quad i = 1 \ldots m$$
$$\mathbf{x} \in \mathbb{X} \subset \mathbb{R}^{n-m}, \quad \mathbf{y} \in \mathbb{Y} \subset \mathbb{R}^m$$

Furthermore, assume that for all $\hat{\mathbf{x}} \in \mathbb{X}$, the system of equations

$$h_i(\hat{\mathbf{x}}, \mathbf{y}) = 0, \quad i = 1 \ldots m,$$

can be solved for $\mathbf{y}$. Moreover, this can be done with a non-iterative algorithm. This requires a special structure or feature in the system of equations above, e.g., $\mathbf{h}$ may be linear in $\mathbf{y}$ in which case it can be solved in fixed number of operations related to m, $\mathbf{h}$ may be a sequential calculation sequence such as a simple network or flowsheet calculation. Provided such a structure exists, the problem can be expressed as the

following *reduced problem*:

$$\min_{\mathbf{x}} \quad f(\mathbf{x}, \mathbf{y}(\mathbf{x}))$$

$$\mathbf{g}(\mathbf{x}, \mathbf{y}(\mathbf{x})) \leq \mathbf{0} \tag{AP}$$

$$\mathbf{x} \in \mathbb{X} \subset \mathbb{R}^{n-m}$$

It may also be possible that the full set of equality constraints do not have such a structure, but a large subset does. A flowsheet with a recycle is an example of such a system. In this case, the above problem will also have some equality constraints remaining after reduction (e.g., the *tear equations* in the flowsheeting example).

In general, a deterministic global optimization algorithm has worst-case exponential run-time performance (in the number of variables). While the original formulation has an exponent of n, the reduced problem will have an exponent of $n-m$. For a case when n and m are large but $n - m$ is relatively small, this can result in a significant saving of computational effort. This approach can therefore be termed a *reduced-space global optimization approach*.

Such problems are not too uncommon. Any system with few inputs and outputs, but a large number of internal state variables and corresponding governing equations that relate them is a candidate for such a reduction. Examples include chemical processes and unit operations, nonlinear networks, biological systems and so on.

In the current context, the resulting system of equations needs to have a structure that has following specific properties as per the assumptions outlined before:

1. It can be solved using a finite, non-iterative procedure.

2. It can be solved for every $\mathbf{x} \in \mathbb{X}$.

Algorithm is therefore used in a narrow context here: given an $\mathbf{x} \in \mathbb{X}$, a *finite calculation sequence* that calculates functions f and $\mathbf{g}$ in Program (AP). Such *calculation sequences* can be arbitrarily complex and involve any function provided each step is factorable and relaxations and their subgradients as well as derivatives to univariate intrinsic functions involved can be calculated. For example, Gauss elimination is an

179

example of an algorithm in this context because the operations performed are known a priori, but a Newton iteration is not an algorithm because it requires an unknown (a priori) number of steps to converge to a given tolerance.

7.2 Convex/Concave Relaxations of Computer Procedures

McCormick's work [125] is a well-known method for constructing convex and concave relaxations of *factorable functions*. A factorable function is a function that can be defined from the finite, recursive composition of a number of "simple" mathematical operations, i.e., binary and unary operations as well as univariate functional compositions. Convex and concave relaxations can be defined for each simple operation involved in such a definition. Propagation rules for the relaxations can be formulated for each binary or unary operation in simple terms, i.e., given the convex and concave relaxations of the individual terms, one can define how to construct the relaxation of the compound term. Finally, McCormick composition theorem [125] enables the construction of the relaxations of the composition of two functions given the relaxations of each.

Mitsos et al. [124] have formulated subgradient propagation rules that can be used to propagate subgradients along with the convex and concave relaxation of functions at each operation and composition. They have also proposed to combine this idea with automatic differentiation (AD) [126] concepts using operator overloading or source code transformation. An implementation of these principles using operator overloading, libMC, [127] is also presented in Mitsos et al. [124]. As a result, a computer procedure for evaluating a factorable function by replicating the steps in the composition of the function can also propagate convex and concave relaxations and the respective subgradients. This means that a lower bounding procedure can be constructed for a procedure evaluating such function, a crucial component for global optimization of mathematical programs with computer evaluated functions.

7.3 Application of Reduced-Space Methods to Upstream Gas Networks

Upstream gas production network models satisfy all the requirements (that were described in Section 7.1) for application of a reduced-space approach. There are large numbers of internal variables in upstream gas network models; pressures, flowrates, state of facilities and compositions. The inputs to the network model are the volumetric production-rates at wells and select variables in the trunkline network. The majority of the flowrates, pressure and composition variables form the internal variable of the network model and are given by pressure-flowrate relationships, molar balances and facility operating equations. Finally, there are only a few outputs from the model that include delivery states at the demand nodes and key states in the network.

Based on this, network variables can be classified into the following three categories: input variables, internal variables and output variables. *Input variables* are the variables that can be manipulated by the optimizer directly. Constraints and bounds can therefore be enforced on these variables to prevent the optimization procedure from stepping outside a valid range so as to make sure that the network can be solved for the internal and output variables. *Internal variables* (*intermediate variables*) are the variables that represent the internal state of the network and are not directly seen by the optimizer. *Output variables* represent the values of functions that are being calculated. For example, the delivery rate at a demand node is an output variable. Conditions on an intermediate variable can enforced only by adding an explicit constraint on it, so that it becomes an output variable.

The network structure permits a sequential calculation of the internal variables and output variables, given the input variables. One can start at the wells and progressively move towards the demand nodes calculating the values of the internal variables and the output variables. All equality constraints are incorporated into the network calculation procedure and therefore the resulting mathematical programming problem with embedded network calculation procedure has no equality constraints.

The following section describes the calculation sequence and the problem formulation. The assumptions for the model are identical to the infrastructure model formulation described in Chapter 3. The calculation sequence derived in the next section is based on the alternative infrastructure formulation presented in Section 3.6. The variable, set and parameter naming conventions are in line with Table 2.3 and, most variable and parameter symbols are consistent with and retain the same or a similar meaning as in the description of alternative formulation in Section 3.6.

7.4 Derivation of a Network Calculation Sequence

As a first step, the pressure variables in this model are transformed to the square of the actual pressures. Denote the *transformed pressure* as $\widehat{P}_{(.)}$ that is square of the actual pressure $P_{(.)}$:

$$\widehat{P}_{(.)} = P^2_{(.)}$$

Since most pressure variables in the network are represented by this transformed pressure, $\widehat{P}_{(.)}$ is referred to without qualification as just pressure later, as long as the context is clear from the notation.

This model is a pure infrastructure NLP model that does not include features of the alternative formulation that requires binary variables. As in the alternative infrastructure model, the following are the components of a calculation sequence:

1. well performance model,

2. field calculations,

3. compression calculations,

4. network model.

The actual derived calculation steps are numbered using Roman numerals. Non-numbered equations just represent intermediate manipulations or alternative calculation steps.

7.4.1 Well Calculation Sequence

The primary input variables to the network is the well volumetric production rates $Q_{w,i}$. $Q_{w,i}$ is constrained by the natural production bounds as explained in Section 5.2.1 (page 145). The natural bounds make sure that the key pressures in the well (bottom-hole and well-head pressure), do not fall below atmospheric pressure at the upper bound of the flow. Given the volumetric rate, the molar production rate from a well can be calculated using the following relationship:

$$F_{s,i,k} = \chi_{i,k}\,\phi\,Q_{w,i}, \quad \forall (i,k) \in \mathcal{W} \times \mathcal{S}, \tag{i}$$

where $\chi_{i,k}$ is the mole fraction of component k in gas from well i and ϕ is the volumetric to molar conversion factor as defined in Section 3.3.1 (page 89). The bottom-hole pressure can be calculated from the inflow performance relationship from the production rate for the wells:

$$\widehat{P}_{b,i} = \pi_{r,i}^2 - \alpha_w Q_{w,i} - \beta_w Q_{w,i}^2, \quad \forall i \in \mathcal{W}. \tag{ii}$$

Once the bottom-hole pressure is known, the well-head pressure can be calculated from the vertical lift performance equation as follows:

$$\widehat{P}_{t,i} = \frac{1}{\lambda_i}(\widehat{P}_{b,i} - \vartheta_w Q_{w,i}^2), \quad \forall i \in \mathcal{W}. \tag{iii}$$

The NGL production rate can be calculated from the volumetric production rate:

$$Q_{Lw,i} = \sigma_w Q_{w,i}, \quad \forall i \in \mathcal{W}. \tag{iv}$$

7.4.2 Field Balances

There are two different approaches for calculating pressure at the common header to which all wells belonging to a particular field produce. Let $\mathcal{F}_w$ denote the set of fields for which well performance is modeled. The most obvious is to perform a direct

calculation of pressure at the well headers as follows:

$$\widehat{P}_i = \min_{w \in \mathcal{W}_i} \widehat{P}_{t,w}, \quad \forall i \in \mathcal{F}_w. \tag{WF1}$$

On a close inspection, however, this formulation is not correct because it forces the field pressure to be minimum of all well-head pressures and therefore, violates the implicit choke assumption that allows for a finite nonzero pressure drop, i.e., it excludes the following operating possibility:

$$P_i = P_{t,w} - \Delta P_w, \quad \Delta P_w > 0, \quad w \in \mathcal{W}_i, \quad \forall i \in \mathcal{F}_w$$

where $P_{t,w}$ is the actual well-head pressure and P_i is the (actual) field pressure. Violation of the above operational possibility means that high-pressure fields can choke low-pressure fields since high-pressure field headers cannot be brought to the pressure of low-pressure fields using the relationship (WF1). Another option is to use the above formulation and accept ΔP_w as an input variable, however, not only does this result in an increase in the number of input variables, but it also introduces an additional calculation step involving $\sqrt{\widehat{P}}$, likely to make the problem ill-conditioned.

Another option is to accept field pressure $\widehat{P}_i$ as an input variables and enforce an inequality constraint of the following form:

$$\widehat{P}_i - \widehat{P}_{t,w} \leq 0, \quad w \in \mathcal{W}_i, \quad \forall i \in \mathcal{F}_w. \tag{v}$$

However, this formulation has the drawback to add additional constraints to the model and making it harder to locate a feasible point. This is the formulation used in the case study presented later.

A yet another option is to introduce an intermediate variable $\widehat{P}_{mh,i}$ for every field that is the minimum of all well-head pressures corresponding to the field. A single constraint for a field can then be used to enforce the field pressure inequality instead

of a constraint for each well, thereby reducing the number of constraints.

$$\widehat{P}_{mh,i} = \min_{w \in \mathcal{W}_i} \widehat{P}_{t,w}, \quad \forall i \in \mathcal{F}_w, \tag{WF3}$$

$$\widehat{P}_i - \widehat{P}_{mh,i} \leq 0, \quad \forall i \in \mathcal{F}_w.$$

Volumetric and species-wise molar production rate from fields can be calculated as follows:

$$F_{s,i,k} = \sum_{w \in \mathcal{W}_i} F_{s,w,k}, \quad \forall (i,k) \in \mathcal{F}_w \times \mathcal{S}, \tag{vi}$$

$$Q_{s,i} = \sum_{w \in \mathcal{W}_i} Q_{w,w}, \quad \forall i \in \mathcal{F}_w. \tag{vii}$$

The same is true for the total NGL production:

$$Q_{Ls,i} = \sum_{w \in \mathcal{W}_i} Q_{Lw,w}, \quad \forall i \in \mathcal{F}_w. \tag{viii}$$

7.4.3   Compression

The compression equation is the same as presented in Equation (vii) (page 99). There are two approaches to formulating the calculation sequence for compressors based on the choice of the input variable. One can choose either the power consumption of the compressor or choose the output pressure.

Consider the choice of compression power $W_{(i,j)}$ as input to calculate the output pressure $\widehat{P}_j$:

$$\widehat{P}_j = \widehat{P}_i \left[1 + \frac{W_{(i,j)}}{\omega_{(i,j)} Q_{a,(i,j)}} \right]^{2/\nu}, \quad \forall (i,j) \in \mathcal{A}_c.$$

This formulation is indeed more natural since in a real-world case, it is the power that is controlled which in turn dictates the output pressures. However, this formulation has a serious drawback of becoming ill-conditioned as $Q_{a,(i,j)}$ gets close to zero. There is no good physical guide to set bounds on $Q_{a,(i,j)}$. It is an internal network variable that is not visible to the optimizer so it cannot be manipulated directly. It is possible to set lower bounds on one or more well production rates to control the range of

185

$Q_{a,(i,j)}$. Again, there is no clear physical argument for doing so and therefore, it is not obvious how these lower bounds on well production rates should be set. It may be possible to reformulate this relationship by adding a small term to the denominator based on some physical argument.

The other option is to instead choose the output pressure of the compressor as an input variable and enforce the calculated power to be limited by the rated power of the compressor using a constraint. Output pressure $\widehat{P}_j$ is constrained by optimizer to respect a maximum limit. This results in the following constraint that models the compression:

$$\omega_{(i,j)} \, Q_{s,(i,j)} \left[\left(\frac{\widehat{P}_j}{\widehat{P}_i} \right)^{\nu/2} - 1 \right] - \Psi^U_{(i,j)} \leq 0, \quad \forall (i,j) \in \mathcal{A}_c. \qquad \text{(ix)}$$

Both $\widehat{P}_j$ and $\widehat{P}_i$ are inputs. $\Psi^U_{(i,j)}$ is the rated power of the compressor. The relaxation can be further strengthened by enforcing a constraint relating them

$$\widehat{P}_i - \widehat{P}_j \leq 0, \quad \forall (i,j) \in \mathcal{A}_c. \qquad \text{(x)}$$

The apparent drawback of this formulation is that in some sense, it seems to decouple the subsystems upstream and downstream of a compressor from a pressure perspective and thereby prevents the flow of pressure information from the downstream to the upstream system. On the other hand, it can also be argued that a real compressor actually does decouple the upstream and downstream parts of a real system.

7.4.4 Network Balances

One has to distinguish between nodes that are mixers and splitters for network balances as in the discussion on the infrastructure model earlier in Chapter 3. Nodes with multiple incoming arcs and only a single outgoing trunkline (arc) are defined as mixers in this context (set $\mathcal{N}_m$). Therefore, specieswise molar flowrates can be

calculated in a straightforward way by summing the incoming flowrates:

$$F_{a,(i,v),k} = F_{s,i,k} + \sum_{v:(v,i)\in\mathcal{A}} F_{a,(v,i),k}, \quad \forall(i,k) \in \mathcal{N}_{m,s} \times \mathcal{S}, \tag{xi}$$

$$F_{a,(i,v),k} = \sum_{v:(v,i)\in\mathcal{A}} F_{a,(v,i),k}, \quad \forall(i,k) \in \mathcal{N}_{m,J} \times \mathcal{S}. \tag{xii}$$

Here $\mathcal{N}_{m,s}$ is the set of mixers that have a source term while $\mathcal{N}_{m,J}$ is the set of mixers that are junctions.

A splitter by definition has strictly more than one outgoing arcs. There is no need to distinguish between splitters that are junctions and sources separately as was done in Chapter 3. One needs to consider only splitters that are junctions because a production node that is a splitter can be simply decomposed into a node that is a production node (handled as above) that is connected to a splitter junction. This is possible because volumetric and molar flowrate variables are internal variables that are not seen by the optimizer and any increase in their number is not expected to have any significant impact on the solution performance. At a splitter, one can define a set of outgoing arcs $\mathcal{A}_x$ in exactly the same fashion as described in Section 3.3.2 (all outgoing arcs except one). A split fraction is defined over set $\mathcal{A}_x$ that is an input variable for the optimizer. It is now possible to calculate the outgoing flow at the splitters (set $\mathcal{N}_x$) as follows:

$$F_{a,(i,j),k} = s_{(i,j)} \sum_{v:(v,i)\in\mathcal{A}} F_{a,(v,i),k}, \quad \forall((i,j),k) \in \{\mathcal{A}_x : i \in \mathcal{N}_x\} \times \mathcal{S}, \tag{xiii}$$

$$F_{a,(i,u),k} = \left(1 - \sum_{v:(i,v)\in\mathcal{A}_x} s_{(i,v)}\right) \sum_{v:(v,i)\in\mathcal{A}} F_{a,(v,i),k}, \quad \forall(i,u) \notin \mathcal{A}_x, \quad (i,k) \in \mathcal{N}_x \times \mathcal{S}.$$
$$\tag{xiv}$$

Split fractions must be forced to sum up to one explicitly

$$\sum_{v:(i,v)\in\mathcal{A}_x} s_{(i,v)} - 1 \le 0, \quad \forall i \in \mathcal{N}_x. \tag{xv}$$

An alternative option to enforcing constraint (xv) is to decompose a splitter having three or more outgoing arcs into multiple splitters each having only two outgoing arcs. This does away with the need for the above constraint as bounds are sufficient to enforce these.

The volumetric flowrate in an arc can be calculated as before by using the conversion factor ϕ:

$$Q_{a,(i,j)} = \frac{1}{\phi} \sum_{k \in \mathcal{S}} F_{a,(i,j),k}, \quad \forall (i,j) \in \mathcal{A}. \tag{xvi}$$

Outlet pressure for lines whose outlet is not a mixer can be calculated directly employing the standard gas flow relationship:

$$\widehat{P}_j = \widehat{P}_i - \kappa_{(i,j)} Q^2_{a,(i,j)}, \quad j \notin \mathcal{N}_m, \quad \forall (i,j) \in \mathcal{A}_p. \tag{xvii}$$

However, a similar pressure calculation for lines with mixer outlets (i.e., more than one arc is flowing into the node) runs into a similar problem as with field headers. If the pressure was calculated at this node using pressure-flowrate relationship as above, it may invalidate molar balances or pressure-flowrate relationships in other arcs. The only way to reconcile pressure calculations between multiple arcs in such instances is to have it enforced as constraints. The outlet pressure for all arcs incoming at a mixer can be calculated as:

$$\widehat{P}_o = \widehat{P}_i - \kappa_{(i,j)} Q^2_{a,(i,j)}, \quad j \in \mathcal{N}_m, \quad \forall (i,j) \in \mathcal{A}_p.$$

However, one does not need to define $\widehat{P}_o$ explicitly, instead the constraint can be enforced directly as:

$$\widehat{P}_j - \widehat{P}_i + \kappa_{(i,j)} Q^2_{a,(i,j)} \leq 0, \quad j \in \mathcal{N}_m, \quad \forall (i,j) \in \mathcal{A}_p. \tag{xviii}$$

In this case $\widehat{P}_j$ is an input variable being manipulated by the optimizer. Physically this implies that there is a valve to choke down the flow coming from a high pressure line so that it does not choke the flow from a low pressure line at a mixer. If such a situation

is constraining the system, optimization is expected to drive $\widehat{P}_o = \widehat{P}_j$. However, in a case when there are no valves on the system, this modeling approach can possibly result in a nonzero $\widehat{P}_o - \widehat{P}_j$ for some nodes which may be unphysical. In practice, by manipulating wellhead chokes, it is always possible to achieve $\widehat{P}_o = \widehat{P}_j$ which is in fact what happens in a conventional optimization approach as presented earlier. To incorporate this feature in the current context requires part of the calculation sequence to track back upstream from mixers to wells in a reverse direction to the flow on some arcs, which is complicated and has not been attempted here. A similar formulation to the one in Equations (WF3) that is based on calculating the minimum of incoming pressures can also be used here. An intermediate variable $\widehat{P}_{o,(i,j)}$ corresponding to the pressure at the outlet of an arc (i,j) can be calculated as follows:

$$\widehat{P}_{o,(i,j)} = \widehat{P}_i - \kappa_{(i,j)} Q^2_{a,(i,j)}, \quad j \in \mathcal{N}_m, \quad \forall (i,j) \in \mathcal{A}_p.$$

The minimum of all such intermediate outlet pressures $\widehat{P}_{o,(i,j)}$ can be calculated as another intermediate variable $\widehat{P}_{nm,j}$

$$\widehat{P}_{nm,j} = \min_{i:(i,j)\in\mathcal{A}_p} \widehat{P}_{o,(i,j)}, \quad j \in \mathcal{N}_m.$$

Pressure $\widehat{P}_j$ at node j is an input variable with the following constraint on it

$$\widehat{P}_j - \widehat{P}_{nm,j} \leq 0 \quad j \in \mathcal{N}_m.$$

This formulation is physically equivalent to the earlier formulation, i.e., it implies presence of a choke valve.

7.4.5 Demands

The delivery at the demand nodes (set $\mathcal{N}_D$) can be calculated using similar network balance relationships earlier:

$$F_{s,i,k} = \sum_{v:(v,i)\in\mathcal{A}} F_{a,(v,i),k}, \quad \forall(i,k) \in \mathcal{N}_D \times \mathcal{S}, \tag{xix}$$

$$Q_{s,i} = \frac{1}{\phi} \sum_{k\in\mathcal{S}} F_{s,i,k}, \quad \forall(i,k) \in \mathcal{N}_D. \tag{xx}$$

The delivery constraints at the demand nodes can be enforced once the relevant output variables are calculated. For example, once the pressures at the demand nodes are known, the pressure range as per the delivery specification can be enforced:

$$(\pi_i^L)^2 \leq \widehat{P}_i \leq (\pi_i^U)^2, \quad \forall i \in \mathcal{N}_D. \tag{xxi}$$

Similarly there are constraints for delivery rates can be enforced once $Q_{s,i}$ is known at demand nodes

$$\Lambda_i^L \leq Q_{s,i} \leq \Lambda_i^U, \quad \forall i \in \mathcal{N}_D. \tag{xxii}$$

Finally, the composition specification $\chi_{i,k}^s$ for a species k can be formulated as follows:

$$F_{s,i,k} - \chi_{i,k}^s \sum_{j\in\mathcal{S}} F_{s,i,j} \leq 0, \quad \forall i \in \mathcal{N}_D. \tag{xxiii}$$

An overview of the calculation procedure for a small example network is presented in Figure 7-1.

7.4.6 Overall Formulation

The objective function is simply the total production rate from the system:

$$\min \quad \sum_{i\in\mathcal{N}_D} -Q_{s,i} \tag{xxiv}$$

$$Q_{w,A} \qquad Q_{w,B} \qquad Q_{w,S} \qquad Q_{w,T}$$

$$\widehat{P}_{t,B} = \frac{1}{\lambda_B}\left(\widehat{P}_{b,B} - \vartheta_B Q_{w,B}^2\right)$$

$$\widehat{P}_{t,i}$$

$$\widehat{P}_{b,B} = \pi_{r,B}^2 - \alpha_B Q_{w,B} - \beta_B Q_{w,B}^2$$

$$\pi_{r,B}$$

Input $\quad Q_{w,B}$

$$F_{s,i,k} = \chi_{i,k}\,\phi\,Q_{w,i}$$

$$\widehat{P}_i - \widehat{P}_{t,w} \le 0 \qquad\qquad \widehat{P}_i - \widehat{P}_{t,w} \le 0$$

M4 $\quad \widehat{P}_{M4} \qquad\qquad \widehat{P}_{SE} \quad$ **SE**

$$\widehat{P}_j - \widehat{P}_i \le 0$$

$$\omega_{(i,j)}\,Q_{s,(i,j)}\left[\left(\frac{\widehat{P}_1}{\widehat{P}_{in}}\right)^{\nu/2} - 1\right] - \Psi_{(i,j)}^U \le 0 \qquad \widehat{P}_{in}$$

$$\widehat{P}_{in} - \widehat{P}_1 \le 0 \qquad \widehat{P}_1$$

$$\widehat{P}_2 = \widehat{P}_1 - \kappa_{(1,2)}Q_{a,(1,2)}^2$$

$$s_{(2,D1)}$$

$$F_{a,(2,D1),k} = s_{(2,D1)}F_{a,(1,2),k}, \quad \forall k \in \mathcal{S} \qquad\qquad F_{a,(2,D2),k} = (1 - s_{(2,D2)})F_{a,(1,2),k}, \forall k \in \mathcal{S}$$

$$s_{(2,D1)} - 1 \le 0$$

$$\widehat{P}_{D1} = \widehat{P}_2 - \kappa_{(2,D1)}Q_{a,(2,D1)}^2 \qquad\qquad \widehat{P}_{D2} = \widehat{P}_2 - \kappa_{(2,D2)}Q_{a,(2,D2)}^2$$

$$\widehat{P}_{D1} \ge \pi_{spec,D1}^2 \qquad\qquad \widehat{P}_{D2} \ge \pi_{spec,D2}^2$$

D1 $\qquad f_{obj} = -Q_{s,D1} - Q_{s,D2} \qquad$ **D2**

$$Q_{s,D1} \ge \Lambda_{D1} \qquad\qquad\qquad\qquad Q_{s,D2} \le \Lambda_{D2}$$
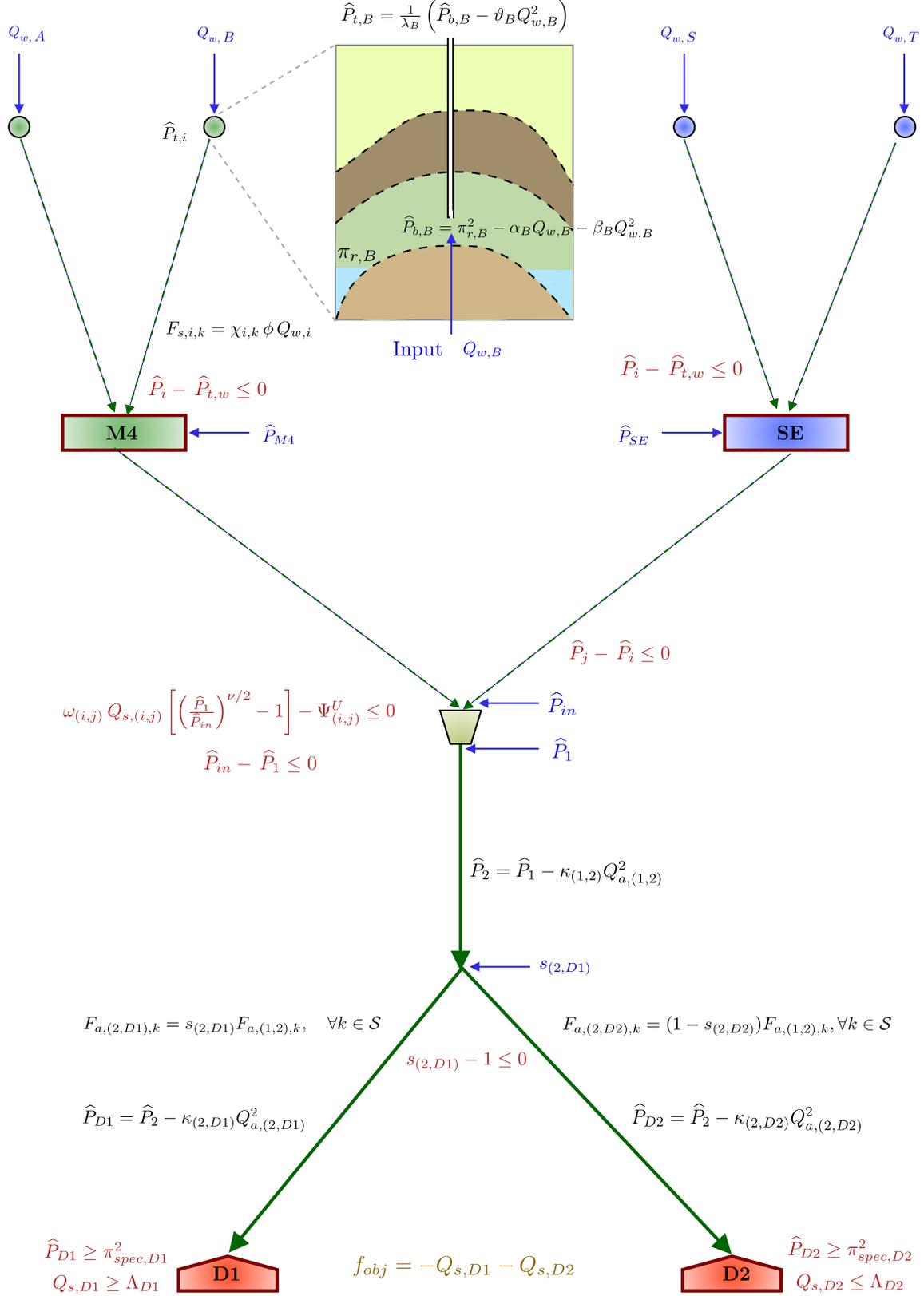
Figure 7-1: Network calculation algorithm: Schematic for an example network

The overall program has both objective function and constraints as algorithms

$$\min_{\mathbf{x}} \quad f(\mathbf{x}, \mathbf{y}(\mathbf{x}))$$

$$\mathbf{g}(\mathbf{x}, \mathbf{y}(\mathbf{x})) \leq \mathbf{0}$$

$$\mathbf{x} \in \mathbb{X} \subset \mathbb{R}^{n-m}.$$

Both f and $\mathbf{g}$ can be calculated in a single pass through the calculation procedure and a separate constraint or objective calculation is not required.

7.5 A Bundle Algorithm Implementation

Convex and/or concave relaxations obtained by the McCormick theory may not be differentiable everywhere in the host set. Therefore the lower bounding convex program is in general nonsmooth and cannot be solved with conventional NLP methods that rely on differentiability of all functions on (an open superset of) the host set.

7.5.1 Theoretical Background

A survey of convex analysis and nonsmooth programming methods is outside the scope of this work. Excellent theoretical treatments of convex analysis and nonsmooth optimization can be found in Hiriart-Urruty and Lemaréchal [128, 129]. Further details about nonsmooth algorithms can be found in Kiwiel [130] and Mäkelä and Neittaanmäki [131]. The material presented here and the corresponding notation follows closely and borrows heavily from these two works.

Consider the following convex program:

$$\min f(\mathbf{x}) \qquad\qquad\qquad (\text{CP})$$

$$\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$$

$$\mathbf{x} \in \mathbb{C} \subset \mathbb{R}^n.$$

Here $\mathbb{X} \subset \mathbb{R}^n$ is an open bounded convex set, $\mathbb{C} \subset \mathbb{X}$ is a compact convex set, and

$g_i : \mathbb{X} \to \mathbb{R}, \quad i = 1, \ldots, m$ and $f : \mathbb{X} \to \mathbb{R}$ are convex (though not necessarily differentiable) functions.

Definition 7.5.1 (Total Constraint Function). *Total constraint function is a scaler function $G : \mathbb{X} \to \mathbb{R}$ defined as follows:*

$$G(\mathbf{x}) = \max\{g_1(\mathbf{x}), g_2(\mathbf{x}), \ldots, g_m(\mathbf{x})\}.$$

Definition 7.5.2 (Improvement Function). *Improvement function $H(.\,;\,\mathbf{x}) : \mathbb{C} \to \mathbb{R}$, $\mathbf{x} \in \mathbb{C}$ corresponding to problem* (CP) *is defined as follows:*

$$H(\mathbf{y}\,;\,\mathbf{x}) = \max\{\, f(\mathbf{y}) - f(\mathbf{x}),\, G(\mathbf{y}) \,\}.$$

Assume that $\mathbf{x}$ is feasible, which in turn is equivalent to $H(\mathbf{x}; \mathbf{x}) = 0$. If $H(\mathbf{y}; \mathbf{x}) < H(\mathbf{x}; \mathbf{x})$ then $f(\mathbf{y}) - f(\mathbf{x}) < 0$ and $G(\mathbf{y}) < 0$ and therefore $\mathbf{y}$ is feasible and has a better objective function value than $\mathbf{x}$. This justifies the designation of H as an *improvement function* for Program (CP).

It is straightforward to show that if f and g_i are convex, the total constraint function G and the improvement function $H(.;\mathbf{x})$ corresponding to program (CP) at a point $\mathbf{x}$ are convex.

Consider the following program:

$$\min H(\mathbf{y}; \mathbf{x}) \tag{IF$_\mathbf{x}$}$$

$$\mathbf{y} \in \mathbb{C}$$

Theorem 7.5.3. *Assume that the Slater constraint qualification holds for Program* (CP). *Then, the following are equivalent*

1. $\mathbf{x}^*$ *is a solution of Program* (CP).

2. $\mathbf{x}^*$ *is optimal in* IF$_{\mathbf{x}^*}$ *and furthermore*
   $$\min_{\mathbf{y} \in \mathbb{C}} H(\mathbf{y}; \mathbf{x}^*) = H(\mathbf{x}^*; \mathbf{x}^*) = 0.$$

Proof. (1) $\implies$ (2): If $\mathbf{x}^*$ is a minimum of Program (CP), $g_i(\mathbf{x}^*) \leq 0$ and therefore

$$H(\mathbf{x}^*; \mathbf{x}^*) = \max\{f(\mathbf{x}^*) - f(\mathbf{x}^*), G(\mathbf{x}^*)\} = 0.$$

If $\mathbf{x}^*$ does not minimize Program IF$_{\mathbf{x}^*}$, then $\exists \hat{\mathbf{x}}$ such that $H(\hat{\mathbf{x}}; \mathbf{x}^*) < 0 = H(\mathbf{x}^*; \mathbf{x}^*)$ which in turn implies that $\hat{\mathbf{x}}$ is feasible in (CP) and $f(\hat{\mathbf{x}}) < f(\mathbf{x}^*)$ and therefore $\mathbf{x}^*$ does not minimize CP.

(2) $\Rightarrow$ (1): Assume that $\exists \hat{\mathbf{x}} \in \mathbb{C}$ such that $f(\hat{\mathbf{x}}) < f(\mathbf{x}^*)$ and $\mathbf{g}(\hat{\mathbf{x}}) \leq \mathbf{0}$ (i.e., $\mathbf{x}^*$ is not optimal for (CP)). Equivalently one can write $f(\hat{\mathbf{x}}) = f(\mathbf{x}^*) - \delta, \ \delta > 0$. Moreover, assume that Slater constraint qualification holds for $\bar{\mathbf{x}} \in \mathbb{C}$, i.e., $\mathbf{g}(\bar{\mathbf{x}}) < \mathbf{0}$.

For a sufficiently small $\lambda \in (0, 1)$, one has the following:

$$
\begin{align}
f((1 - \lambda)\hat{\mathbf{x}} + \lambda\bar{\mathbf{x}}) &\leq (1 - \lambda)f(\hat{\mathbf{x}}) + \lambda f(\bar{\mathbf{x}}) \tag{f1}\\
&= f(\hat{\mathbf{x}}) + \lambda\left(f(\bar{\mathbf{x}}) - f(\hat{\mathbf{x}})\right)\\
&= f(\mathbf{x}^*) - \delta + \lambda\left(f(\bar{\mathbf{x}}) - f(\hat{\mathbf{x}})\right)\\
&< f(\mathbf{x}^*).
\end{align}
$$

Similarly for a sufficiently small λ:

$$\mathbf{g}((1 - \lambda)\hat{\mathbf{x}} + \lambda\bar{\mathbf{x}}) \leq (1 - \lambda)\mathbf{g}(\hat{\mathbf{x}}) + \lambda\mathbf{g}(\bar{\mathbf{x}}) \leq \lambda\mathbf{g}(\bar{\mathbf{x}}) < \mathbf{0} \tag{g1}$$

The first inequality follows from convexity of $\mathbf{g}$, the second from the feasibility of $\hat{\mathbf{x}}$ (i.e., $(1 - \lambda)\mathbf{g}(\hat{\mathbf{x}}) \leq \mathbf{0})$) while the final strict inequality follows from the fact that $\bar{\mathbf{x}}$ is a Slater point.

From inequalities (f1) and (g1), one has for a sufficient small $\lambda \in (0, 1)$

$$
\begin{align}
H((1 - \lambda)\hat{\mathbf{x}} + \lambda\bar{\mathbf{x}}; \mathbf{x}^*) &= \max\{f((1 - \lambda)\hat{\mathbf{x}} + \lambda\bar{\mathbf{x}}) - f(\mathbf{x}^*), G((1 - \lambda)\hat{\mathbf{x}} + \lambda\bar{\mathbf{x}})\}\\
&< 0 = H(\mathbf{x}^*; \mathbf{x}^*),
\end{align}
$$

therefore $\mathbf{x}^*$ does not minimize (IF$_{\mathbf{x}^*}$) and (2) does not hold. $\qquad\square$

Remark 7.5.4. *This result can be further extended to derive necessary and sufficient optimality conditions for nonsmooth convex programs [131]. The starting point is to convert* (IF$_\mathbf{x}$) *into an unconstrained minimization by using the* indicator function *[128, 129] of set* $\mathbb{C}$ *and then using necessary and sufficient conditions for minimization of unconstrained nonsmooth convex functions in conjunction with properties of relevant functions and their subdifferentials.*

7.5.2 Algorithm Overview

The treatment presented here is exclusively for convex programs. From a convex programming point of view, the fundamental principle of these algorithms derives from the simple cutting plane approach for convex programming. The cutting plane methods involve constructing a polyhedral approximation of the feasible set and objective function. They, however, suffer from slow convergence behavior due to the fact that the polyhedral approximations poorly approximate the feasible set and objective function far away from the linearization points and, therefore, quite a bit of zigzagging is observed before the algorithm reaches anywhere near the optimum.

The basic idea of a bundle method is to create an approximation of the subdifferentials of the functions at every step and locate a descent direction based on this information. There are sophisticated strategies to refine the approximations at every step to keep the size of the direction finding problem manageable. Line search plays a very crucial role for these algorithms and sophisticated line search strategies are required for fast convergence. For convex programming, the idea of polyhedral approximation and subdifferential approximation are the same[1].

The approach is described for program (CP). Based on the definition of the improvement function and Theorem 7.5.3, the following is a very rough structure for a simple nonsmooth constrained optimization algorithm [130, 131]:

Step 1: Initialization: Set $k = 1$ and find a point $\mathbf{x}_1 \in \mathbb{C}$ such that $G(\mathbf{x}_1) \leq 0$.

---

[1]As is obvious, these arguments do not completely motivate nonsmooth nonconvex methods, although both convex and nonconvex methods employ very similar principles. A nonsmooth nonconvex method is attempting to approximate generalized gradients instead of subdifferentials.

Step 2: *Direction Finding*: Find a direction $\mathbf{d}_k$ that solves:

$$\min H(\mathbf{x}_k + \mathbf{d}; \mathbf{x}_k)$$

$$\mathbf{x}_k + \mathbf{d} \in \mathbb{C}.$$

Step 3: *Line search*: Find a step size $t_k > 0$ in the direction $\mathbf{d}_k$ such that

$H(\mathbf{x}_k + t_k\mathbf{d}_k; \mathbf{x}_k) < H(\mathbf{x}_k; \mathbf{x}_k)$. Set $\mathbf{x}_{k+1} = \mathbf{x}_k + t_k\mathbf{d}_k$.

Step 4: Set $k = k + 1$. Go to Step 2.

It is clear that the direction finding problem is in itself is a nonsmooth problem in the broad framework above and must be replaced by an approximation as is described in the following section. The above representation is of course a very coarse (and only one of the several possible) motivation for this class of algorithms.

7.5.3 The Direction Finding Problem

The approach described here closely follows the development of cutting plane and bundle methods in Kiwiel [130] and Mäkelä and Neittaanmäki [131]. The nonsmooth program is given as before by Program (CP). Assume that the host set $\mathbb{C}$ is defined as a n-dimensional interval:

$$\mathbb{C} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x}^L \leq \mathbf{x} \leq \mathbf{x}^U\}.$$

Using the definitions of set $\mathbb{C}$ and the total constraint function G, Program (CP) can be stated equivalently as:

$$\min_{\mathbf{x} \in \mathbb{R}^n} \quad f(\mathbf{x}) \qquad \text{(ECP)}$$

$$G(\mathbf{x}) \leq \mathbf{0}$$

$$\mathbf{x}^L \leq \mathbf{x} \leq \mathbf{x}^U.$$

Denote the iterate at iteration k of the procedure as $\mathbf{x}_k$. One can define an

improvement function for the program at the current iterate as follows:

$$H(\mathbf{x}; \mathbf{x}_k) = \max \left\{ f(\mathbf{x}) - f(\mathbf{x}_k), g_1(\mathbf{x}), \ldots, g_m(\mathbf{x}) \right\}.$$

It is assumed that at every iterate $\mathbf{x}_k$, one can evaluate functions f and G, and at least one element of their subdifferentials, i.e., a subgradient, at $\mathbf{x}_k$.

At every iterate, there is a collection of linearizations to the functions that is termed as a *bundle*. It forms the best polyhedral approximation (or alternatively best approximation of the subdifferential) of the functions at the point. One can also distinguish between the *objective bundle* (linearizations of f) and the *constraint bundle* (linearizations of G). Denote the index sets at the kth iteration of the bundles for the objective function f and the total constraint function G as J_f^k and J_G^k, respectively. The points of linearization are denoted as $\mathbf{y}_j \in \mathbb{C}$ where j runs over bundle indices. The subgradients at these points are denoted as follows:

$$\boldsymbol{\xi}_j^f \in \partial f(\mathbf{y}_j), \quad j \in J_f^k,$$
$$\boldsymbol{\xi}_j^G \in \partial G(\mathbf{y}_j), \quad j \in J_G^k,$$

where $\partial f(\mathbf{y}_j)$ and $\partial G(\mathbf{y}_j)$ are the set of all subgradients at $\mathbf{y}_j$ (i.e., the subdifferentials at $\mathbf{y}_j$) of f and G respectively. Linearizations to f and G at points $\mathbf{y}_j \in \mathbb{C}$ are defined as follows:

$$\bar{f}_j(\mathbf{x}) = f(\mathbf{y}_j) + \boldsymbol{\xi}_j^f(\mathbf{x} - \mathbf{y}_j), \quad j \in J_f^k, \tag{L1}$$
$$\bar{G}_j(\mathbf{x}) = G(\mathbf{y}_j) + \boldsymbol{\xi}_j^G(\mathbf{x} - \mathbf{y}_j), \quad j \in J_G^k.$$

Denote the value of linearizations for f and G at the current iterate $\mathbf{x}_k$ as f_j^k and G_j^k respectively:

$$f_j^k = \bar{f}_j(\mathbf{x}_k) = f(\mathbf{y}_j) + \boldsymbol{\xi}_j^f(\mathbf{x}_k - \mathbf{y}_j), \quad j \in J_f^k, \tag{L2}$$
$$G_j^k = \bar{G}_j(\mathbf{x}_k) = G(\mathbf{y}_j) + \boldsymbol{\xi}_j^G(\mathbf{x}_k - \mathbf{y}_j), \quad j \in J_G^k.$$

Subtracting the two relations (L1) and (L2) does away with the need to store linearization points:

$$\bar{f}_j(\mathbf{x}) - f_j^k = \boldsymbol{\xi}_j^f(\mathbf{x} - \mathbf{x}_k), \quad j \in J_f^k,$$
$$\bar{G}_j(\mathbf{x}) - G_j^k = \boldsymbol{\xi}_j^G(\mathbf{x} - \mathbf{x}_k), \quad j \in J_G^k.$$

For a direction finding problem, one is interested in the step $\mathbf{d}$ at the current iterate $\mathbf{x}_k$, i.e., $\mathbf{x} = \mathbf{x}_k + \mathbf{d}$. Substituting this into the relationships:

$$\bar{f}_j(\mathbf{x}_k + \mathbf{d}) = f_j^k + \boldsymbol{\xi}_j^f \mathbf{d}, \quad j \in J_f^k,$$
$$\bar{G}_j(\mathbf{x}_k + \mathbf{d}) = G_j^k + \boldsymbol{\xi}_j^G \mathbf{d}, \quad j \in J_G^k.$$

Finally, substituting $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{d}$, where $\mathbf{d}$ is the actual step after the line search and recalling the definition of f_j^k and G_j^k, the following update formula is obtained for the bundle when stepping from iterate k to $k + 1$:

$$f_j^{k+1} = f_j^k + \boldsymbol{\xi}_j^f(\mathbf{x}_{k+1} - \mathbf{x}_k), \quad j \in J_f^k, \tag{LU}$$
$$G_j^{k+1} = G_j^k + \boldsymbol{\xi}_j^G(\mathbf{x}_{k+1} - \mathbf{x}_k), \quad j \in J_G^k$$

Definition 7.5.5. *Define the polyhedral representation of f and G at iteration k as being given as:*

$$\widehat{f}^k(\mathbf{x}) = \max\{\bar{f}_j(\mathbf{x}) : j \in J_f^k\},$$
$$\widehat{G}^k(\mathbf{x}) = \max\{\bar{G}_j(\mathbf{x}) : j \in J_G^k\}.$$

The polyhedral approximation to the improvement function at the current iterate $H(.;\mathbf{x}_k)$ is defined by

$$\widehat{H}^k(\mathbf{x}) = \max\{\widehat{f}^k(\mathbf{x}) - f(\mathbf{x}_k), \widehat{G}^k(\mathbf{x})\}.$$

Lemma 7.5.6. *The polyhedral approximation function $\widehat{H}^k$ is convex. If Program*

198

(ECP) *is convex, then*

$$\widehat{H}^k(\mathbf{x}) \leq H(\mathbf{x}; \mathbf{x}_k).$$

Proof. Convexity follows from the fact that polyhedral approximation of f and G are convex (linear) and the polyhedral approximation of H is the maximum of a finite collection of linear approximations and hence is convex. The lower bounding nature is obtained by the fact that it is an outer approximation of the (convex) feasible set and objective function. $\square$

The standard direction finding problem for a proximal bundle method is given by:

$$\min_{\mathbf{d} \in \mathbb{R}^n} \quad \widehat{H}^k(\mathbf{x}_k + \mathbf{d}) + \frac{u_k}{2}\|\mathbf{d}\|_2^2 \qquad \text{(PBDFP)}$$

$$\mathbf{x}_k + \mathbf{d} \in \mathbb{C}.$$

The $\|\mathbf{d}\|^2$ term originates from a combination of the trust region concept and the cutting plane strategies. In the most basic trust region approach in differentiable optimization, constraints similar to the form $\|\mathbf{d}\|^2 \leq \delta_k^2$ serve to limit an algorithm from taking steps outside the "region of trust" defined by the constraint, where the function approximations (usually quadratic for "Newton-like" methods) may no longer be sufficiently accurate. In the direction finding problem (PBDFP) this constraint is incorporated into objective function using a penalty term since quadratic programs (QP) are easier to solve than quadratically-constrainted programs (QCP). In the end, $\frac{u_k}{2}\|\mathbf{d}\|^2$ serves exactly the same purpose here to limit the stepsize to a certain extent. The weight parameter u_k can be modified at every iteration based on the subgradient information and problem geometry based on complicated strategies.

In this work, the intention was to keep the direction finding problem as a linear program. Two obvious options are to use a one-norm $\widehat{H}^k(\mathbf{x}_k + \mathbf{d}) + u_k\|\mathbf{d}\|_1$ or an infinity-norm $\widehat{H}^k(\mathbf{x}_k + \mathbf{d}) + u_k\|\mathbf{d}\|_\infty$ so as to avoid large steps and prevent the next iterate from ending up too far away. However, cursory numerical experimentation failed to show any benefit from using such terms. Again, this is insufficient to discard the described approach because sophisticated strategies are required to update the

weight parameter at every iteration that were not implemented or tried.

Therefore, the direction finding problem here is the polyhedral approximation to the improvement function and therefore is equivalent to a linear program:

$$\min_{\mathbf{d}\in\mathbb{R}^n} \quad \widehat{H}^k(\mathbf{x}_k + \mathbf{d}) \tag{DFH}$$

$$\mathbf{x}_k + \mathbf{d} \in \mathbb{C}.$$

Lemma 7.5.7. *Program* (DFH) *is equivalent to the following program:*

$$\min_{v,\mathbf{d}\in\mathbb{R}^n} \quad v \tag{DF}$$

$$v - \boldsymbol{\xi}_j^f \mathbf{d} \geq f_j^k - f(\mathbf{x}_k), \quad \forall j \in J_f^k$$

$$v - \boldsymbol{\xi}_j^G \mathbf{d} \geq G_j^k, \qquad \forall j \in J_G^k$$

$$\mathbf{x}^L - \mathbf{x}_k \leq \mathbf{d} \leq \mathbf{x}^U - \mathbf{x}_k.$$

Proof. (DFH) can be represented as:

$$\min_{\mathbf{x}_k+\mathbf{d}\in\mathbb{C}} \max \left\{ \widehat{f}^k(\mathbf{x}_k + \mathbf{d}) - f(\mathbf{x}_k), \widehat{G}^k(\mathbf{x}_k + \mathbf{d}) \right\}$$

which can be rewritten using the definitions of $\widehat{f}^k$ and $\widehat{G}^k$ as:

$$\min_{\mathbf{x}_k+\mathbf{d}\in\mathbb{C}} \max_{i\in J_f^k, j\in J_G^k} \left\{ \bar{f}_i(\mathbf{x}_k + \mathbf{d}) - f(\mathbf{x}_k), \bar{G}_j(\mathbf{x}_k + \mathbf{d}) \right\}$$

and again as:

$$\min_{\mathbf{x}_k+\mathbf{d}\in\mathbb{C}} \max_{i\in J_f^k, j\in J_G^k} \left\{ f_i^k + \boldsymbol{\xi}_i^f \mathbf{d} - f(\mathbf{x}_k), G_j^k + \boldsymbol{\xi}_j^G \mathbf{d} \right\}$$

Finally considering the definition of set $\mathbb{C}$ as an interval, it is straightforward to formulate this program as (DF). $\qquad\square$

7.5.4 Subgradient Selection and Aggregation

Once a solution of the direction finding problem has been found, new linearizations can be added to the constraint and objective bundles to incorporate new information at the updated iterate and construct an updated direction finding problem. If nothing else is done, the problem size (as in the number of constraints) will continue to grow with each iteration. A strategy is therefore required to keep the size of the bundle under control by dropping selected linearizations that may no longer be sufficiently accurate and somehow making sure that doing this does not adversely impact the approximations or result in loss of information. The *Subgradient Aggregation and Selection* strategy [130] serves to accomplish exactly this objective. The core idea is that constraints can be aggregated to generate new constraints. As a result some of the past linearizations can be dropped without any loss of information about the problem. This section first describes aggregation and then selection.

Consider the solution of the direction finding problem at the kth iteration. Let ν_j^k denote the Lagrange multipliers corresponding to $j \in J_f^k$ (the objective bundle) and μ_j^k denote the Lagrange multipliers corresponding to $j \in J_G^k$ (the constraint bundle) at the kth solution of the direction finding problem (DF). Denote the sums of Lagrange multipliers as $\nu_f^k = \sum_{j \in J_f^k} \nu_j^k$ and $\mu_G^k = \sum_{j \in J_G^k} \mu_j^k$. Define the scaled Lagrange multipliers as follows:

$$\bar{\nu}_j^k = \begin{cases} \nu_j^k/\nu_f^k & \nu_f^k > 0 \\ 1/|J_f^k| & \nu_f^k = 0 \end{cases} \qquad \forall j \in J_f^k, \tag{AG1}$$

$$\bar{\mu}_j^k = \begin{cases} \mu_j^k/\mu_G^k & \mu_G^k > 0 \\ 1/|J_G^k| & \mu_G^k = 0 \end{cases} \qquad \forall j \in J_G^k. \tag{AG2}$$

Define:

$$\boldsymbol{\sigma}_f^k = \sum_{j \in J_f^k} \bar{\nu}_j^k \boldsymbol{\xi}_j^f, \qquad f_\sigma^k = \sum_{j \in J_f^k} \bar{\nu}_j^k f_j^k, \tag{AG3}$$

$$\boldsymbol{\sigma}_G^k = \sum_{j \in J_G^k} \bar{\mu}_j^k \boldsymbol{\xi}_j^G, \qquad G_\sigma^k = \sum_{j \in J_G^k} \bar{\mu}_j^k G_j^k.$$

Definition 7.5.8 (Aggregated Linearizations). *The aggregated linearizations are defined as follows:*

$$f_\sigma(\mathbf{d}) = f_\sigma^k + \boldsymbol{\sigma}_f^k \mathbf{d}$$
$$G_\sigma(\mathbf{d}) = G_\sigma^k + \boldsymbol{\sigma}_G^k \mathbf{d}.$$

Define strict subsets $\widehat{J}_f^k \subset J_f^k$ and $\widehat{J}_G^k \subset J_G^k$ so that the new bundle defined by $\widehat{J}_f^k$ and $\widehat{J}_G^k$ contains only a subset of the constraints. The reduced problem is defined by:

$$\min_{v,\mathbf{d}\in\mathbb{R}^n} \quad v \tag{DFA}$$
$$v - \boldsymbol{\xi}_j^f \mathbf{d} \geq f_j^k - f(\mathbf{x}_k), \quad \forall j \in \widehat{J}_f^k$$
$$v - \boldsymbol{\xi}_j^G \mathbf{d} \geq G_j^k, \quad \forall j \in \widehat{J}_G^k$$
$$v - \boldsymbol{\sigma}_f^k \mathbf{d} \geq f_\sigma^k - f(\mathbf{x}_k)$$
$$v - \boldsymbol{\sigma}_G^k \mathbf{d} \geq G_\sigma^k$$
$$\mathbf{x}^L - \mathbf{x}_k \leq \mathbf{d} \leq \mathbf{x}^U - \mathbf{x}_k.$$

Theorem 7.5.9. *Every optimal solution of Program* (DF) *is optimal in Program* (DFA).

Proof. For the purposes of this proof, represent the bounds as $\mathbf{Bd} \geq \mathbf{c}$ to simplify the notation.

Consider a primal-dual solution pair $(v_k^*, \mathbf{d}_k^*, \widetilde{\boldsymbol{\nu}}_k, \widetilde{\boldsymbol{\mu}}_k, \widetilde{\boldsymbol{\lambda}}_k)$ at the kth solution of Program (DF). The necessary and sufficient conditions for optimality for this pair are given by:

Primal Feasibility:

$$v_k^* - \boldsymbol{\xi}_j^f \mathbf{d}_k^* \geq f_j^k - f(\mathbf{x}_k), \quad \forall j \in J_f^k,$$
$$v_k^* - \boldsymbol{\xi}_j^G \mathbf{d}_k^* \geq G_j^k, \quad \forall j \in J_G^k,$$
$$\mathbf{Bd}^* \geq \mathbf{c}.$$

Complementary Slackness:

$$\widetilde{\nu}_j^k(v_k^* - \boldsymbol{\xi}_j^f \mathbf{d}_k^* - f_j^k + f(\mathbf{x}_k)) = 0, \quad \forall j \in J_f^k,$$

$$\widetilde{\mu}_j^k(v_k^* - \boldsymbol{\xi}_j^G \mathbf{d}_k^* - G_j^k) = 0, \qquad \forall j \in J_G^k,$$

$$\widetilde{\boldsymbol{\lambda}}_k(\mathbf{B}\mathbf{d}^* - \mathbf{c}) = \mathbf{0}.$$

Dual Feasibility:

$$\sum_{j \in J_f^k} \widetilde{\nu}_j^k + \sum_{j \in J_G^k} \widetilde{\mu}_j^k = 1,$$

$$-\sum_{j \in J_f^k} \widetilde{\nu}_j^k \boldsymbol{\xi}_j^f - \sum_{j \in J_G^k} \widetilde{\mu}_j^k \boldsymbol{\xi}_j^G + \sum_j \widetilde{\lambda}_j^k \mathbf{B}_j = \mathbf{0},$$

$$\widetilde{\boldsymbol{\nu}}_k \geq \mathbf{0}, \ \widetilde{\boldsymbol{\mu}}_k \geq \mathbf{0}, \ \widetilde{\boldsymbol{\lambda}}_k \geq \mathbf{0}.$$

Denote the sums of multipliers as $\widetilde{\nu}_f^k = \sum\limits_{j \in J_f^k} \widetilde{\nu}_j^k$ and $\widetilde{\mu}_G^k = \sum\limits_{j \in J_G^k} \widetilde{\mu}_j^k$. Consider program (DFA) and set the multipliers as follows. Assign the multipliers corresponding to the aggregated constraints f_σ and G_σ as $\widetilde{\nu}_f^k$ and $\widetilde{\mu}_G^k$ respectively. For the reduced objective bundle (indexed by $\widehat{J}_f^k$) and for the reduced constraint bundle (indexed by $\widehat{J}_G^k$), define the multipliers in the following way:

$$\widehat{\nu}_j^k = 0, \quad \forall j \in \widehat{J}_f^k, \qquad \widehat{\mu}_j^k = 0, \quad \forall j \in \widehat{J}_G^k.$$

Consider $(v_k^*, \mathbf{d}_k^*, \widehat{\boldsymbol{\nu}}_k, \widehat{\boldsymbol{\mu}}_k, \widetilde{\boldsymbol{\lambda}}_k, \widetilde{\nu}_f^k, \widetilde{\mu}_G^k)$ as a candidate for the primal-dual solution pair of program (DFA). It is obvious that primal feasibility holds for $\forall j \in \widehat{J}_f^k \subset J_f^k$ and $\forall j \in \widehat{J}_G^k \subset J_G^k$. Multiply constraints over $j \in J_f^k$ and $j \in J_G^k$ with $\bar{\nu}_j^k$ and $\bar{\mu}_j^k$, respectively ($\bar{\nu}_j^k$ and $\bar{\mu}_j^k$ are defined as in Equations (AG1) and (AG2), page 201), and

sum over $j \in J_f^k$ and $j \in J_G^k$, respectively:

$$v_k^* \sum_{j \in J_f^k} \bar{\nu}_j^k - (\sum_{j \in J_f^k} \bar{\nu}_j^k \boldsymbol{\xi}_j^f) \mathbf{d}_k^* \geq \sum_{j \in J_f^k} (\bar{\nu}_j^k f_j^k) - \bar{\nu}_j^k f(\mathbf{x}_k),$$

$$v_k^* \sum_{j \in J_G^k} \bar{\mu}_j^k - (\sum_{j \in J_G^k} \bar{\mu}_j^k \boldsymbol{\xi}_j^G) \mathbf{d}_k^* \geq (\sum_{j \in J_G^k} \bar{\mu}_j^k G_j^k).$$

Considering that $\sum_{j \in J_f^k} \bar{\nu}_j^k = 1$, $\sum_{j \in J_G^k} \bar{\mu}_j^k = 1$ and definitions of $\boldsymbol{\sigma}_f^k$, $\boldsymbol{\sigma}_G^k$, f_σ^k and G_σ^k (page 201), it is obvious that primal feasibility of $(v_k^*, \mathbf{d}_k^*)$ holds in the aggregated constraints. Complementary slackness holds for $\forall j \in \widehat{J}_f^k$ and $\forall j \in \widehat{J}_G^k$ by the definition of the multipliers $\widehat{\boldsymbol{\nu}}_k$ and $\widehat{\boldsymbol{\mu}}_k$. If $\widetilde{\nu}_f^k = 0$ or $\widetilde{\mu}_G^k = 0$, complementary slackness for the corresponding aggregated constraint already holds. So one can assume that $\widetilde{\nu}_f^k > 0$ and $\widetilde{\mu}_G^k > 0$ for demonstrating the complementary slackness condition for aggregated constraints. Summing and manipulating the complementary slackness conditions for $\forall j \in J_f^k$ and $\forall j \in J_G^k$:

$$\widetilde{\nu}_f^k \left(v_k^* \sum_{j \in J_f^k} \frac{\widetilde{\nu}_j^k}{\widetilde{\nu}_f^k} - \mathbf{d}_k^* (\sum_{j \in J_f^k} \frac{\widetilde{\nu}_j^k}{\widetilde{\nu}_f^k} \boldsymbol{\xi}_j^f) - (\sum_{j \in J_f^k} \frac{\widetilde{\nu}_j^k}{\widetilde{\nu}_f^k} f_j^k) + f(\mathbf{x}_k) \sum_{j \in J_f^k} \frac{\widetilde{\nu}_j^k}{\widetilde{\nu}_f^k} \right) = 0,$$

$$\widetilde{\mu}_G^k \left(v_k^* \sum_{j \in J_G^k} \frac{\widetilde{\mu}_j^k}{\widetilde{\mu}_G^k} - \mathbf{d}_k^* (\sum_{j \in J_G^k} \frac{\widetilde{\mu}_j^k}{\widetilde{\mu}_G^k} \boldsymbol{\xi}_j^G) - (\sum_{j \in J_G^k} \frac{\widetilde{\mu}_j^k}{\widetilde{\mu}_G^k} G_j^k) \right) = 0.$$

From definitions (AG1), (AG2) and (AG3), it is clear that complementary slackness holds for the aggregated constraints. Finally consider the dual feasibility conditions.

$$\sum_{j \in \widehat{J}_f^k} \widehat{\nu}_j^k + \sum_{j \in \widehat{J}_G^k} \widehat{\mu}_j^k + \widetilde{\nu}_f^k + \widetilde{\mu}_G^k = \sum_{j \in J_f^k} \widetilde{\nu}_j^k + \sum_{j \in J_G^k} \widetilde{\mu}_j^k = 1$$

Assume $\widetilde{\nu}_f^k > 0$ and $\widetilde{\mu}_G^k > 0$ holds:

$$\mathbf{0} = -\sum_{j \in J_k^f} \widetilde{\nu}_j^k \boldsymbol{\xi}_j^f - \sum_{j \in J_k^G} \widetilde{\mu}_j^k \boldsymbol{\xi}_j^G + \sum_j \widetilde{\lambda}_j^k \mathbf{B}_j =$$

$$-\widetilde{\nu}_f^k \sum_{j \in J_k^f} \frac{\widetilde{\nu}_j^k}{\widetilde{\nu}_f^k} \boldsymbol{\xi}_j^f - \widetilde{\mu}_G^k \sum_{j \in J_k^G} \frac{\widetilde{\mu}_j^k}{\widetilde{\mu}_G^k} \boldsymbol{\xi}_j^G + \sum_j \widetilde{\lambda}_j^k \mathbf{B}_j =$$

$$-\widetilde{\nu}_f^k \boldsymbol{\sigma}_f^k - \widetilde{\mu}_G^k \boldsymbol{\sigma}_G^k - \sum_{j \in \widehat{J}_k^f} \widehat{\nu}_j^k \boldsymbol{\xi}_j^f - \sum_{j \in \widehat{J}_k^G} \widehat{\mu}_j^k \boldsymbol{\xi}_j^G + \sum_j \widetilde{\lambda}_j^k \mathbf{B}_j.$$

If $\widetilde{\nu}_f^k = 0$, it implies $\widetilde{\nu}_j^k = 0, \forall j \in \widehat{J}_f^k$. By dropping terms corresponding to $\widetilde{\nu}_j^k$ and $\widetilde{\nu}_f^k$ the above argument holds. The same is true if $\widetilde{\mu}_G^k = 0$. Also note that both $\widetilde{\nu}_f^k = 0$ and $\widetilde{\mu}_G^k = 0$ cannot hold simultaneously due to the dual feasibility conditions for (DF). This implies $(v_k^*, \mathbf{d}_k^*)$ is optimal in Program (DFA). $\qquad \square$

Remark 7.5.10. *The set of optimal solutions of* (DF) *and* (DFA) *must be the same for being able to replace program* (DF) *with* (DFA)*. The previous result only shows that the optimal solution set of* (DF) *is a subset of the solution set of* (DFA)*. Therefore, it does not provide a sufficient theoretical argument for aggregation.*

The core reason for this issue seems to be the fact that the direction finding problem is a linear program (LP). A quadratic program (QP) has a unique solution, which means that a similar result to the previous one for a QP would immediately imply that the reduced direction finding problem (with aggregated constraint) is equivalent to the original problem. On the other hand, an LP can have multiple optimal solutions and therefore only a subset relationship can be shown. This still does not preclude aggregation being valid for an LP direction finding problem since the converse of the previous result may be true under certain assumptions. However, no attempt has been made to analyze this further. Finally, although, aggregation has been implemented in the formal statement of the algorithm (Section 7.5.6, page 212), it can be turned off.

Subgradient Selection

The cardinality of sets J_f^k and J_G^k is bounded by a maximum size to keep the bundle size manageable and therefore to manage the computational effort required to solve the direction finding problem. To choose the constraints to drop from the sets J_f^k and J_G^k (and therefore construct $\widehat{J}_f^k$ and $\widehat{J}_G^k$) the value of the corresponding Lagrange multipliers is used. A constraint with a zero multiplier can be dropped as it is not active at the current optimum and therefore is likely to be a poor approximation of the feasible set and objective function at this point.

The size of the bundle is an important tunable parameter of the algorithm. There is an obvious trade-off between the work per iteration (or equivalently per direction finding problem solution) and the number of iterations (direction finding problems solved). A large bundle size results in a relatively more accurate representation of the local problem geometry and therefore hopefully a better descent direction. This means that a relatively smaller number of solutions of direction finding problems may be required, however the work per iteration may be larger. On the other hand, a smaller direction finding problem represents a coarser representation of problem geometry and therefore a relatively larger number of direction finding problem may need to be solved before declaring convergence. However, the work per iteration in this case will be lower than the former one.

7.5.5 Line Search

Once a potential descent direction is available from the line search, a stepsize t_k must be calculated using a line search in the direction $\mathbf{d}_k$ to update the current iterate, i.e. $\mathbf{x}_{k+1} = \mathbf{x}_k + t_k \mathbf{d}_k$. The exposition and logic of the line search presented here is due to Hiriart-Urruty and Lemaréchal [128].

Define a scalar function $h : \mathbb{R}^+ \to \mathbb{R}$ as $h(t) = f(\mathbf{x}_k + t\mathbf{d}_k)$. One is interested in a $t > 0$ such that $h(t) < h(0)$. Given a step size t^* and information about the nature of the objective function, a *test* can be designed that generates one of three possible answers:

1. (T) t^* is appropriate, terminate the line search,

2. (R) t^* is not appropriate and no suitable $t > t^*$ is possible,

3. (L) t^* is not appropriate and no suitable $t < t^*$ is possible.

Note that cases 2 and 3 provide a bound on the stepsize from one side. The generic framework of the line search is then as follows:

1. Set $t_L = 0$, $t_R = +\infty$ and select $t \in (t_L, t_R)$.

2. Perform the *test* on t. If case (T) stop.

3. In case (L), set $t_L = t$, In case (R), set $t_R = t$.

4. Select $t \in (t_L, t_R)$ and go to step 2.

Finding the exact minimum of the scalar function h can be a nontrivial and computationally intensive task. It is realized that a line search is an intermediate problem and it is not worthwhile to spend too much computational effort on solving it. Roughly, a line search should not generate too "large" a step which can potentially lead to zigzagging because the next iterate may be too far away from the current one and therefore the current polyhedral approximations may be poor and unreliable. On the other hand too "small" a step can unnecessarily delay the convergence of the algorithm. Defining "large" and "small" is the essence of case (T). The following is an example of a test used for cases (R) and (L) from Hiriart-Urruty and Lemaréchal [128]:

1. (R) $t > 0$ is not too "large" when:

$$h(t) \leq h(0) + m\,t\,h'(0), \tag{R}$$

where m is a line search parameter in $(0, 1)$ usually less than 0.5. This guarantees that $h(t) < h(0)$ and there is sufficient descent in the step. A stepsize that fails to satisfy the above is declared too large and therefore is unacceptable.
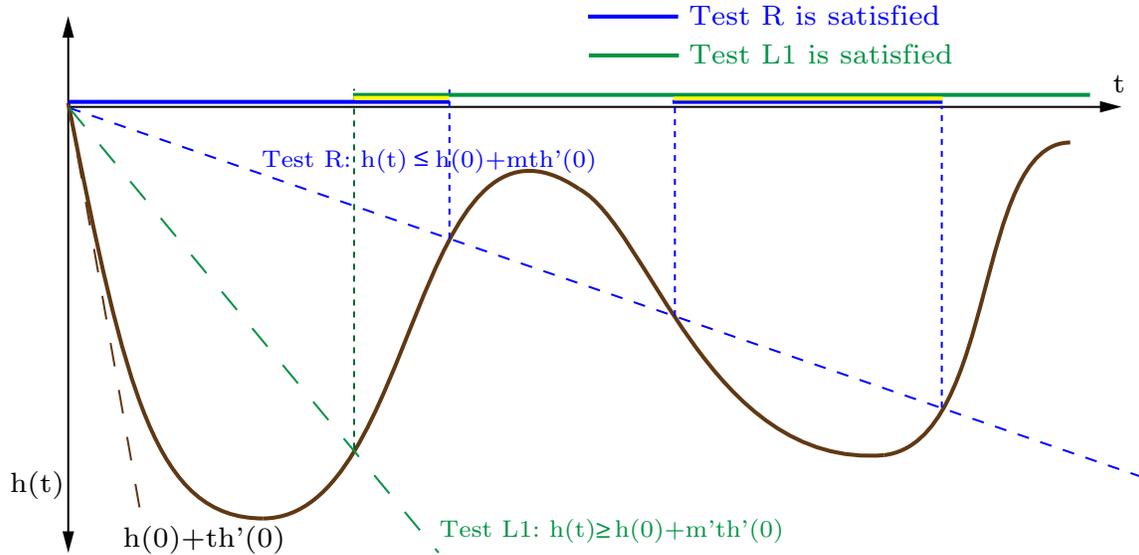
Figure 7-2: Line search tests (R) and (L1) (based on a figure from Hiriart-Urruty and Lemaréchal [128])

Finally, it is not necessary to calculate the derivative value $h'(0)$ exactly and an approximation may be used.

2. (L) A similar condition is possible for rejecting a small t. Two examples of conditions used are as follows:

$$h(t) \geq h(0) + m'th'(0), \tag{L1}$$

or

$$h(t) \geq m'h'(0), \tag{L2}$$

where m' is a line search parameter in $(m, 1)$ (m is the parameter that was discussed before for test (R)).

A stepsize is acceptable, i.e., case (T), when it satisfies (R) and (L1) or (L2). A stepsize is too large if it fails to satisfy (R). A stepsize is regarded as too "small" if it fails to satisfy (L1) or (L2). Figure 7-2 illustrates the geometric description of tests (R) and (L1).

The following is a simpler version of the result presented in Kiwiel [130] and Mäkelä

and Neittaanmäki [131] which motivates the line search test used in this work.

Theorem 7.5.11. *Assume that the current iterate* $\mathbf{x}_k$ *is feasible in Program* (ECP). *Let* $\widehat{f}(\mathbf{x})$ *represent the polyhedral representation of* f *as defined earlier. Let* $(v_k, \mathbf{d}_k)$ *be the* k^{th} *solution of the direction finding problem formulated at the current iterate* $\mathbf{x}_k$. *Then the following holds:*

1. $\widehat{f}(\mathbf{x}_k + \mathbf{d}_k) - f(\mathbf{x}_k) \leq v_k \leq 0,$

2. $\widehat{f}(\mathbf{x}_k + t\mathbf{d}_k) \leq f(\mathbf{x}_k) + tv_k, \quad \forall t \in [0, 1].$

Proof. This proof considers the program (DF) instead of (DFA) to keep the notation simple. However, all arguments hold for program (DFA) as it is a matter of simply forming a positive linear combination of inequalities to generate aggregated constraints.

1. Recalling the definition of $\widehat{f}$:

$$\widehat{f}^k(\mathbf{x}_k + \mathbf{d}_k) - f(\mathbf{x}_k) = \max_{j \in J_f^k} \{\bar{f}_j(\mathbf{x}_k + \mathbf{d}_k) - f(\mathbf{x}_k)\} = \max_{j \in J_f^k} \{f_j^k + \boldsymbol{\xi}_j^f \mathbf{d}_k - f(\mathbf{x}_k)\}.$$

From feasibility of v_k in (DF):

$$v_k \geq f_j^k + \boldsymbol{\xi}_j^f \mathbf{d}_k - f(\mathbf{x}_k), \quad \forall j \in J_f^k \quad \Longleftrightarrow \quad v_k \geq \max_{j \in J_f^k} \{f_j^k + \boldsymbol{\xi}_j^f \mathbf{d}_k - f(\mathbf{x}_k)\},$$

and therefore

$$v_k \geq \widehat{f}^k(\mathbf{x}_k + \mathbf{d}_k) - f(\mathbf{x}_k),$$

which shows one part of the result. Substituting $(v, \mathbf{d}) = (0, \mathbf{0})$ in the constraints of (DF)

$$f(\mathbf{x}_k) \geq f_j^k, \quad \forall j \in J_f^k,$$
$$0 \geq G_j^k, \quad \forall j \in J_G^k,$$
$$\mathbf{x}^L - \mathbf{x}_k \leq \mathbf{0} \leq \mathbf{x}^U - \mathbf{x}_k.$$

209

Given $\mathbf{x}_k$ is feasible, the last set of (bound) inequalities holds. Also, $G(\mathbf{x}_k) \leq 0$ because $\mathbf{x}_k$ is feasible. The first and second inequalities follow from the fact that linearizations of f and G at $\mathbf{x}_k$ should underestimate $f(\mathbf{x}_k)$ and $G(\mathbf{x}_k)$, i.e. $f_j^k \leq f(\mathbf{x}_k)$ and $G_j^k \leq G(\mathbf{x}_k)$. This implies that $(0, \mathbf{0})$ is feasible in (DF). Since v_k is optimal in (DF), it implies that $v_k \leq 0$ and this completes the proof of the first set of inequalities.

2. Consider a $t \in [0, 1]$:

$$\widehat{f}(\mathbf{x}_k + t\mathbf{d}_k) = \widehat{f}((1 - t)\mathbf{x}_k + t(\mathbf{d}_k + \mathbf{x}_k))$$
$$\leq (1 - t)\widehat{f}(\mathbf{x}_k) + t\widehat{f}(\mathbf{d}_k + \mathbf{x}_k)$$
$$\leq (1 - t)\widehat{f}(\mathbf{x}_k) + t(\widehat{f}(\mathbf{x}_k) + v_k)$$
$$\leq \widehat{f}(\mathbf{x}_k) + tv_k$$
$$\leq f(\mathbf{x}_k) + tv_k.$$

This completes the proof of inequality 2. $\qquad\qquad\qquad\qquad$ □

The second inequality in Theorem 7.5.11 above inspires a rough interpretation of v_k as an approximate directional derivative of f at $\mathbf{x}_k$ [130]. This in combination with the earlier description of test (R) for ensuring sufficient descent yields the following line search test: find the largest t_k such that

$$f(\mathbf{x}_k + t_k\mathbf{d}_k) \leq f(\mathbf{x}_k) + m\, t_k\, v_k, \quad t_k \in [0, 1].$$

Parameter m usually is chosen to be less than 0.5 and serves to provide an additional control on the step. It is also possible to define a line search test for case (L) described before. However, in this work a minimum stepsize has not been defined. The notation for line search used here and later (in the description of the implementation) is from Kiwiel [130]. Following this notation, t_k is labeled t_L. Therefore, the above test is restated as: find the largest t_L such that

$$f(\mathbf{x}_k + t_L\mathbf{d}_k) \leq f(\mathbf{x}_k) + m\, t_L\, v_k, \quad t_L \in [0, 1]. \tag{LC1}$$

One also needs to make sure that $\mathbf{x}_k + t_L\mathbf{d}_k$ is feasible given that $\mathbf{x}_k$ is feasible. So a second test required is:

$$G(\mathbf{x}_k + t_L\,\mathbf{d}_k) \leq 0. \tag{LC2}$$

In theory, $\mathbf{d}_k$ obtained from the direction finding problem is guaranteed to respect bounds due to the fact that these bounds are included in the direction finding problem. A full step $\mathbf{x}_k + \mathbf{d}_k$ should lead to a next iterate still within bounds and therefore, line search need not test against bounds. However, in practice, a full step can potentially violate a bound by the magnitude of the bound satisfaction tolerance in the LP solution procedure. When the functions involved in the convex program are McCormick relaxations (that are valid only within the interval on which they are constructed), violation by even a small magnitude can result in incorrect or undefined function evaluations. Therefore a test on bounds should also be included in the line search to avoid a full step violating the bounds.

$$\mathbf{x}^L \leq \mathbf{x}_k + t_L^k\mathbf{d}_k \leq \mathbf{x}^U. \tag{LC3}$$

Set a stepsize threshold $\bar{t} > 0$. There are three possibilities that can be considered:

1. A $t_L > \bar{t}$ satisfying both (LC1), (LC2) and (LC3) is found. In this case a *long serious step* is possible. Usually this means that there is significant decrease in the objective function value. In a *long serious step*, the next iterate $\mathbf{x}_{k+1} = \mathbf{x}_k + t_L\mathbf{d}_k$ and the next linearization point $\mathbf{y}_{k+1} = \mathbf{x}_{k+1}$ are the same point.

2. A $0 < t_L < \bar{t}$ satisfying the above conditions is found. This is termed a *short serious step*. In this case, a stepsize t_R is also found that violates one or more of the conditions above. For a *short serious step*, set the next iterate to $\mathbf{x}_{k+1} = \mathbf{x}_k + t_L\mathbf{d}_k$ and the next linearization point to $\mathbf{y}_{k+1} = \mathbf{x}_k + t_R\mathbf{d}_k$.

3. Finally, if $t_L = 0$, i.e., no $t_L > 0$ satisfies (LC1), (LC2) and (LC3), a *null step* is declared. Again in this case, a $t_R > 0$ is known that violates one or more conditions. In this case, set the next iterate to $\mathbf{x}_{k+1} = \mathbf{x}_k$ and the next linearization point to $\mathbf{y}_{k+1} = \mathbf{x}_k + t_R\mathbf{d}_k$. In a *null step*, the current iterate stays

the same and a new point is added to the bundle.

In both short serious step and null steps, there is a possible "kink" in one or more of the functions (equivalently a discontinuity in the gradients) and the line search must make sure that the new linearization point is on the other side of the "kink". Addition of the new linearization will lead to a substantial modification of the search direction [131].

In closing, it is important to note that line search plays a much more vital role in nonsmooth algorithms than in differentiable optimization. This is primarily because there is no guarantee that a direction obtained from the direction finding problem is indeed a descent direction for the problem. The strategies implemented here are still simplistic and more sophisticated strategies must be employed for a faster and more robust convergence.

7.5.6   Formal Statement

Step 1:   *Initialization*: Assume that a feasible point $\mathbf{x}_1 \in \mathbb{C}$ is available, i.e., assume that

$$G(\mathbf{x}_1) \leq 0.$$

(a) Set the maximum bundle sizes $|J_f^k|_{max}$ and $|J_G^k|_{max}$. Initialize the problem as follows:

$$\text{Set } k = 1, \qquad \mathbf{y}_1 = \mathbf{x}_1,$$
$$f_1^1 = f(\mathbf{y}_1), \qquad G_1^1 = G(\mathbf{y}_1),$$
$$\boldsymbol{\sigma}_f^1 = \boldsymbol{\xi}_1^f \in \partial f(\mathbf{y}_1), \qquad \boldsymbol{\sigma}_G^1 = \boldsymbol{\xi}_1^G \in \partial G(\mathbf{y}_1),$$
$$J_f^1 = \{1\}, \quad J_G^1 = \{1\}.$$

(b) Set the line search parameters: $m \in (0,1)$ a measure of decrease, usually less than 0.5, $\bar{t} \in (0,1]$ a threshold for a *long serious step*, $t_{null} > 0$ a minimum stepsize at which a step is declared a null step, and $t_{max} \in [\bar{t}, 1]$ a maximum stepsize permitted for both serious and

null steps.

Step 2: *Direction Finding Problem*: Solve the following linear program:

$$\min_{v,\mathbf{d}} \quad v$$

$$v - \boldsymbol{\xi}_j^f \mathbf{d} \geq f_j^k - f(\mathbf{x}_k), \quad \forall j \in J_f^k$$

$$v - \boldsymbol{\xi}_j^G \mathbf{d} \geq G_j^k, \quad \forall j \in J_G^k$$

$$v - \boldsymbol{\sigma}_f^k \mathbf{d} \geq f_\sigma^k - f(\mathbf{x}_k)$$

$$v - \boldsymbol{\sigma}_G^k \mathbf{d} \geq G_\sigma^k$$

$$\mathbf{d}_k^L \leq \mathbf{d} \leq \mathbf{d}_k^U$$

where the bounds on the direction are set to:

$$\mathbf{d}_k^L = \mathbf{x}^L - \mathbf{x}_k, \qquad \mathbf{d}_k^U = \mathbf{x}^U - \mathbf{x}_k.$$

Denote the solution as $(v_k, \mathbf{d}_k)$. If infeasible, terminate immediately.

Step 3: *Line search*:

(a) Find largest number $t_L^k \in \{t_{max}, \frac{t_{max}}{2}, \frac{t_{max}}{4}, \frac{t_{max}}{8}, \ldots\}$ that satisfies

    i. $f(\mathbf{x}_k + t_L^k \mathbf{d}_k) \leq f(\mathbf{x}_k) + m t_L^k v_k,$

    ii. $G(\mathbf{x}_k + t_L^k \mathbf{d}_k) \leq 0,$

    iii. $\mathbf{x}^L \leq \mathbf{x}_k + t_L^k \mathbf{d}_k \leq \mathbf{x}^U,$

    iv. $t_L^k \geq \bar{t}.$

(b) If such a $t_L^k > 0$ exists, take a *long serious step*. Set $\mathbf{x}_{k+1} = \mathbf{x}_k + t_L^k \mathbf{d}_k$, $\mathbf{y}_{k+1} = \mathbf{x}_{k+1}$. Set $t_R^k = t_L^k$.

(c) If (i) and (ii) hold, but $t_{\text{null}} \leq t_L^k < \bar{t}$ accept a *short serious step* $\mathbf{x}_{k+1} = \mathbf{x}_k + t_L^k \mathbf{d}_k$, $\mathbf{y}_{k+1} = \mathbf{x}_k + t_R^k \mathbf{d}_k$ where t_R^k is known as below.

(d) If $t_L^k < t_{\text{null}}$ implying (i) and/or (ii) are violated for all $t_L^k > t_{\text{null}}$, accept a *null step*. Set $t_L^k = 0$, $\mathbf{x}_{k+1} = \mathbf{x}_k$, $\mathbf{y}_{k+1} = \mathbf{x}_k + t_R^k \mathbf{d}_k$ where t_R is known as below.

213

In both of above cases, $t_R^k \in [\bar{t}, t_{max}]$ violates at least one of (i) and (ii) and has been found in the search for t_L^k above.

Step 4: *Update RHS of linearizations*: Update the RHS of linearizations as follows:

$$f_j^{k+1} = f_j^k + t_L^k \boldsymbol{\xi}_j^f \mathbf{d}_k, \quad \forall j \in J_f^k$$

$$\hat{f}_\sigma^{k+1} = f_\sigma^k + t_L^k \boldsymbol{\sigma}_f^k \mathbf{d}_k$$

$$G_j^{k+1} = G_j^k + t_L^k \boldsymbol{\xi}_j^G \mathbf{d}_k, \quad \forall j \in J_G^k$$

$$\hat{G}_\sigma^{k+1} = G_\sigma^k + t_L^k \boldsymbol{\sigma}_G^k \mathbf{d}_k$$

$$\mathbf{d}_{k+1}^L = \mathbf{x}^L - \mathbf{x}_{k+1}$$

$$\mathbf{d}_{k+1}^U = \mathbf{x}^U - \mathbf{x}_{k+1}$$

Update $f_j^k - f(\mathbf{x}_k)$ term on the RHS of the objective function bundle (i.e., $\forall j \in J_f^k$) to $f_j^{k+1} - f(\mathbf{x}_{k+1})$. Similarly, replace $f_\sigma^k - f(\mathbf{x}_k)$ by $f_\sigma^{k+1} - f(\mathbf{x}_{k+1})$ on the RHS of the aggregated constraint corresponding to the objective function bundle.

Step 5: Re-solve the direction finding problem with the updated RHS with a (dual) simplex warm start. Define the multipliers at this step as $\nu_k, k \in J_f^k$, $\mu_k, k \in J_G^k$, ν_σ^k and μ_σ^k respectively.

Step 6: *Calculation of Aggregate Linearizations:* Calculate normalized multipliers as follows:

$$\bar{\nu}_j^k = \begin{cases} \nu_j^k / \nu_f^k & \nu_f^k > 0 \\ 1/|J_f^k| & \nu_f^k = 0 \end{cases} \quad \forall j \in J_f^k,$$

$$\bar{\mu}_j^k = \begin{cases} \mu_j^k / \mu_G^k & \mu_G^k > 0 \\ 1/|J_G^k| & \mu_G^k = 0 \end{cases} \quad \forall j \in J_G^k,$$

where $\nu_f^k = \sum_{j \in J_f^k} \nu_j^k$ and $\mu_G^k = \sum_{j \in J_G^k} \mu_j^k$.

Calculate aggregate linearizations as follows:

$$\boldsymbol{\sigma}_f^{k+1} = \sum_{j \in J_f^k} \bar{\nu}_j^k \boldsymbol{\xi}_j^f, \qquad\qquad f_\sigma^{k+1} = \sum_{j \in J_f^k} \bar{\nu}_j^k f_j^{k+1},$$

$$\boldsymbol{\sigma}_G^{k+1} = \sum_{j \in J_G^k} \bar{\mu}_j^k \boldsymbol{\xi}_j^G, \qquad\qquad G_\sigma^{k+1} = \sum_{j \in J_G^k} \bar{\mu}_j^k G_j^{k+1}.$$

Step 7: *Subgradient Selection:*

(a) If $|J_f^k| \leq |J_f^k|_{max}$, skip this step. Find the first index $p \in J_f^k$ such that $\nu_p^k = 0$. Reset $J_f^k = J_f^k / \{p\}$.

(b) If $|J_G^k| \leq |J_G^k|_{max}$, skip this step. Find the first index $p \in J_G^k$ such that $\mu_p^k = 0$. Reset $J_G^k = J_G^k / \{p\}$.

Step 8: *Cut Addition:* Calculate $\boldsymbol{\xi}_{k+1}^f \in \partial f(\mathbf{y}_{k+1})$, $\boldsymbol{\xi}_{k+1}^G \in \partial G(\mathbf{y}_{k+1})$

$$f_{k+1}^{k+1} = f(\mathbf{y}_{k+1}) + \boldsymbol{\xi}_j^f(\mathbf{x}_{k+1} - \mathbf{y}_{k+1}) = f(\mathbf{y}_{k+1}) + (t_L^k - t_R^k)\boldsymbol{\xi}_j^f \mathbf{d}_k$$

$$G_{k+1}^{k+1} = G(\mathbf{y}_{k+1}) + \boldsymbol{\xi}_j^G(\mathbf{x}_{k+1} - \mathbf{y}_{k+1}) = G(\mathbf{y}_{k+1}) + (t_L^k - t_R^k)\boldsymbol{\xi}_j^G \mathbf{d}_k$$

Set $J_f^{k+1} = \{k + 1\} \cup J_f^k, \quad J_G^{k+1} = \{k + 1\} \cup J_G^k.$

Step 9: *Termination Test:* Terminate if either of the following conditions is true

(a) $\|\mathbf{d}_k\| < \epsilon_d$, $v_k \leq 0$ and $k > N_{\min}$,

(b) $k > N_{\max}$.

Step 10: *Continue:* Set the LP solution procedure to primal warm start (since new constraints have been added to the problem). Set $k = k + 1$. Go to *Step 2*.

Notes

1. In the literature implementations of similar algorithms, only a single direction finding problem is solved instead of the two here. The main reason for this is to exploit the LP simplex warm start by first updating the RHS of the constraints (dual simplex warm start) and then adding new constraints (primal simplex warm start). An additional reason is to simplify the implementation segments

associated with data storage and LP solver queries.

2. As was already described, most literature methods solve a QP obtained from Program (PBDFP) (page 199). The standard practice is to solve the dual of this QP. In the QP formulation, the descent direction is a linear combination of objective and constraint bundle subgradients (and also linear constraints if present) at the current iteration[1]. This provides a direct parallel with simple gradient based methods like steepest descent.

3. The aggregation of linearization can be turned off and in numerical experiments does not seem to impact the performance, possibly because of the theoretical issues outlined in Remark 7.5.10 (page 205) or because the bundle is sufficiently large already.

4. The minimum iteration limit is to make sure that the bundle is full to avoid false convergence in the first few steps.

5. If the functions involved in the original convex program are McCormick relaxations, function evaluation are only valid within the interval in which the relaxations are constructed. Therefore, each step of the algorithm should rule out the possibility of even small bound violations for the iterates and linearization points. Minor bound violations (on the order of 10^{-8}) have been observed to severely impact the performance of algorithm. The possibility of a bound violation when making a full step is the reason why an additional parameter t_{max} (to limit the maximum possible step) and redundant bounds checks are employed in the line search procedure.

6. The algorithm behavior is extremely sensitive to scaling. In particular, an unscaled problem can result in slow convergence.

---

[1]This follows from the KKT conditions for QP. However, the optimality conditions for LP do not yield this result. A possible route to proving this result for LP may be the representation theorem for polyhedrons.

7.5.7 Finding a Feasible Point and Detection of Infeasibility

As can be noted from the formal statement of the algorithm, a feasible point is required to start the algorithm. Moreover, in a global optimization framework, a certain number of subproblems will eventually be infeasible and this needs to be robustly detected. Locating a feasible point and detecting infeasibility remain the biggest challenges for the presented framework. The following two options have been investigated:

1. Convert program (ECP) into a penalty representation and apply the bundle method to this program instead of original program

$$\min f(\mathbf{x}) + c \max\{0, G(\mathbf{x})\} \qquad \text{(PCP)}$$
$$\mathbf{x} \in \mathbb{C} \subset \mathbb{R}^n$$

   where $c > 0$ is a penalty parameter. This approach altogether does away with the requirement of a bundle algorithm for nonlinear constrained optimization. However, it has been found to be not very reliable, possibly because of the simplistic formulation used here. Using an unconstrained bundle method in this way has been observed to result in the iterates tending to oscillate between feasibility and infeasibility. A barrier function can be implemented in conjunction with the above formulation to avoid the iterate leaving the feasible region (that was not attempted).

   The problem above belongs to the general class of penalty methods. There are various choices of penalty functions and methods for updating penalty parameter c_k in sequences of penalty problem solutions. It is possible to integrate penalty parameter updates directly into the direction finding problem in the bundle algorithm, i.e., formulate the direction finding problem corresponding to program (PCP) and at the direction finding problem k employ a different penalty parameter c_k depending on the information available so far. This approach was cursorily attempted unsuccessfully, but was not investigated seri-

ously. Nevertheless, it does seem promising and attractive and may be useful to explore theoretically and numerically.

One of the serious drawbacks associated with these methods is controlling the magnitude of the penalty parameter. The problem can become severely ill-conditioned if the magnitude of the penalty parameter is too large (because the subgradient to the objective is $\boldsymbol{\xi}_j^f + c\boldsymbol{\xi}_j^G$ outside the feasible region). The direction finding problems have been observed to fail in such cases. On the other hand, a small penalty parameter may be simply ineffective to drive the problem into the feasible region. Therefore, a tight control is required on the penalty parameter magnitude that is difficult to achieve in practice. Finally, as indicated earlier, this can be combined with a barrier method. Once feasible, the penalty term can be replaced by a barrier function.

2. Another approach is to solve a standalone feasibility problem and to switch to a normal bundle approach once a feasible point is detected. The form of the feasibility problem is as follows:

$$\min G(\mathbf{x}) \tag{FP}$$
$$\mathbf{x} \in \mathbb{C} \subset \mathbb{R}^n$$

If the optimal value of this program is positive then the problem is infeasible. If a feasible point is detected, the main solver can be started.

The latter option is employed in this work. The feasibility detection algorithm is almost the same as the main algorithm presented with a different termination criteria.

Statement

Step 1: Initialization:

    (a) Get the maximum bundle size $|J_G^k|_{max}$ (same as the original algorithm).

Initialize the problem as follows:

$$\text{Set } k = 1, \qquad \mathbf{x}_1 \in [\mathbf{x}^L, \mathbf{x}^U], \qquad \mathbf{y}_1 = \mathbf{x}_1,$$

$$G_1^1 = G(\mathbf{y}_1), \qquad \boldsymbol{\sigma}_G^1 = \boldsymbol{\xi}_1^G \in \partial G(\mathbf{y}_1), \qquad J_G^1 = \{1\}.$$

(b) Set line search parameters: $m \in (0,1)$ a measure of decrease, usually less than 0.5, $\bar{t} \in (0,1]$ a threshold for a *long serious step*, $t_{null} > 0$ minimum stepsize at which a step is declared a null step, and $t_{max} \in [\bar{t}, 1]$ a maximum stepsize permitted for both serious and null steps.

Step 2: *Direction Finding Problem*: Solve the following linear program:

$$\min_{v, \mathbf{d}} \quad v$$

$$v - \boldsymbol{\xi}_j^G \mathbf{d} \geq G_j^k, \quad \forall j \in J_G^k$$

$$v - \boldsymbol{\sigma}_G^k \mathbf{d} \geq G_\sigma^k$$

$$\mathbf{d}_k^L \leq \mathbf{d} \leq \mathbf{d}_k^U.$$

Denote the solution as $(v_k, \mathbf{d}_k)$. If infeasible, terminate immediately.

Step 3: *Line search*: Find largest number $t_L^k \in \{t_{max}, \frac{t_{max}}{2}, \frac{t_{max}}{4}, \frac{t_{max}}{8}, \ldots\}$ that satisfies

(a) $G(\mathbf{x}_k + t_L^k \mathbf{d}_k) \leq G(\mathbf{x}_k) + m t_L^k v_k,$

(b) $\mathbf{x}^L \leq \mathbf{x}_k + t_L^k \mathbf{d}_k \leq \mathbf{x}^U,$

(c) $t_L^k \geq \bar{t}.$

Decide on a (short or long) serious and null step as in the previous algorithm.

Step 4: *Update RHS of linearizations*: Update RHS of linearizations as follows:

$$G_j^{k+1} = G_j^k + t_L^k \boldsymbol{\xi}_j^G \mathbf{d}_k, \quad \forall j \in J_G^k$$

$$\hat{G}_\sigma^{k+1} = G_\sigma^k + t_L^k \boldsymbol{\sigma}_G^k \mathbf{d}_k$$

$$\mathbf{d}_{k+1}^L = \mathbf{x}^L - \mathbf{x}_{k+1}$$

$$\mathbf{d}_{k+1}^U = \mathbf{x}^U - \mathbf{x}_{k+1}$$

Step 5: Re-solve the direction finding problem with the updated RHS with a (dual) simplex warm start. Define the multipliers at this step as $\mu_k, k \in J_G^k$ and μ_σ^k respectively.

Step 6: *Calculation of Aggregate Linearizations:* Calculate normalized multipliers as follows:

$$\bar{\mu}_j^k = \begin{cases} \mu_j^k / \mu_G^k & \mu_G^k > 0 \\ 1/|J_G^k| & \mu_G^k = 0 \end{cases} \quad \forall j \in J_G^k.$$

where $\mu_G^k = \sum_{j \in J_G^k} \mu_j^k$. Calculate aggregate linearizations as follows:

$$\boldsymbol{\sigma}_G^{k+1} = \sum_{j \in J_G^k} \bar{\mu}_j^k \boldsymbol{\xi}_j^G, \qquad G_\sigma^{k+1} = \sum_{j \in J_G^k} \bar{\mu}_j^k G_j^{k+1}$$

Step 7: *Subgradient Selection:* If $|J_G^k| \leq |J_G^k|_{max}$, skip this step. Find the first index $p \in J_G^k$ such that $\mu_p^k = 0$. Reset $J_G^k = J_G^k/\{p\}$.

Step 8: *Cut Addition:* Calculate $\boldsymbol{\xi}_{k+1}^G \in \partial G(\mathbf{y}_{k+1})$

$$G_{k+1}^{k+1} = G(\mathbf{y}_{k+1}) + \boldsymbol{\xi}_j^G(\mathbf{x}_{k+1} - \mathbf{y}_{k+1}) = G(\mathbf{y}_{k+1}) + (t_L^k - t_R^k)\boldsymbol{\xi}_j^G \mathbf{d}_k$$

Set $J_G^{k+1} = \{k+1\} \cup J_G^k$.

Step 9: *Termination Test:*

(a) If $G(\mathbf{x}_{k+1}) \leq 0$: Problem is feasible. Preserve J_G^k. Call the main algorithm and pass $\mathbf{x}_{k+1}$ and J_G^k.

(b) If $v_k > 0$ and $k > N_{min}$, problem is infeasible because the direction

finding problem underestimates the value of program (FP). Declare infeasibility.

Step 10: *Continue:* Set LP solution procedure to primal warm start (since new constraints have been added to the problem). Set $k = k + 1$. Go to *Step 2*.

Notes

1. A cycling behavior has been occasionally observed in the line search procedure: the direction finding solution and line search results in a null step, a new linearization is added, however, the next iteration generates exactly the same direction again and algorithm is unable to move. One of the possible solution for this problem that works in some cases is to reduce t_R as follows if more than one null step with same direction occurs, so that a different linearization point is generated. This strategy has also been incorporated in the main solver line search.

$$\{t_{max}, \bar{t} + \frac{t_{max} - \bar{t}}{2}, \bar{t} + \frac{t_{max} - \bar{t}}{4}, \ldots\}$$

This procedure does have an important potential benefit. If the problem is feasible, not only does it locate a feasible point but also populates the constraint bundle J_G^k, so that the main algorithm can be "warm-started".

7.5.8   Implementation

The above algorithms have been implemented in C++. CPLEX 11.1 is used as the LP solver. The algorithm implementation uses CPLEX Concert Technology for easy manipulation of constraints. CPLEX Concert Technology data structures are employed to store the constraints. Warm starts are as per CPLEX defaults and not explicitly specified as in the statement of the algorithm. It is trivial to extend the implementation to be able to solve QP direction finding problems. In fact it is only a matter of updating the objective. However, sophisticated strategies are required for the weight-updating procedure (for updating u_k in Program (PBDFP), page 199) and hence, this was not attempted in this work.

For large-scale problems which may involve solution of tens of thousands subproblems, it may eventually become important to re-implement (a part of the/the entire) algorithm in plain C/FORTRAN style instead of the C++ based Concert interface to allow for further code optimization. However, for the time being there are substantial improvements to be made in the algorithm itself and hence, this option is worth considering only if code-profiling indicates substantial gains are possible by doing so and there are no further easier avenues for improvement in the procedure.

7.5.9 Future Work

1. The line search strategies need to be improved. As has already been pointed out, a direction obtained from the direction finding problem need not be a descent direction. Moreover, there is potential for algorithm to get stuck at kinks in the functions and gradients. The nonsmooth optimization literature contains several sophisticated and complicated line search strategies to overcome some of these hurdles. Such line search strategies can improve the robustness of algorithms.

2. The direction finding problem can be formulated as a quadratic program instead of an LP using trust region concepts. This provides an additional algorithm parameter that can be tuned using problem geometry to limit step size. Moreover, this can also resolve some of the theoretical and convergence difficulties associated with using LP for the direction finding problem.

3. There is substantial work left for robust detection of infeasibility of the original nonsmooth convex program as well as location of a feasible point for the program if the user provided initial guess is infeasible in the nonlinear constraints. Infeasibility of the original program does not always result in infeasibility of direction finding problem. Instead, in several instances it can result in a complete breakdown of the procedure and no way to update the current iterate, e.g., $v_k > 0$ and $\mathbf{d} = \mathbf{0}$, possibly because the core assumption to the improvement function formulation no longer hold. A feasibility phase has been implemented

as discussed before to solve $\min G(\mathbf{x})$. The apparent advantage of this procedure is that if it does succeed in finding a feasible point, the constraint bundle is already populated so that the performance of the main algorithm may be better. Unfortunately, numerical experiments indicate that this approach is not very robust, although it does work in most cases. It suffers from several problem, e.g., cycling as described earlier and inability to improve objective value. Improved heuristics in the feasibility phase may hold the key for robustness. For example, a simple heuristic to add linearization points to the bundle at random corners of the hyper-rectangle defined by bounds was implemented to break out of the zero direction norm loops in the feasibility phase and thereby ascertain infeasibility for such problems, which seems to work at least in some cases. A separate customized line search procedure for the feasibility phase may also help to certain extent to avoid breakdown.

It may be possible to test $v_k > 0$ to detect infeasibility under certain assumptions depending on the nature of the functions involved. This can improve the robustness of the algorithm. However, it is important to check if this assumption holds for the specific class of problem. Usually the sequence $\{v_1, v_2, v_3 \ldots\}$ starts below zero (recall that $H(\mathbf{y}; \mathbf{x}_k) < H(\mathbf{x}_k; \mathbf{x}_k) = 0$ if $\mathbf{y} - \mathbf{x}_k$ is a descent direction) and increases monotonically with each iteration and eventually should approach zero as algorithm converges (because $H(\mathbf{x}^*; \mathbf{x}^*) = 0$ and no descent direction is available). If the sequence crosses zero or starts above zero for a well-behaved problem, it is usually a strong indication that the problem is infeasible. However, this is not guaranteed by definition of the improvement function. It may be possible to deploy a heuristic based on this.

4. The convergence of the algorithm is sensitive to the scaling of the problem as indicated earlier. For example, a slow rate of convergence (and often an increase in time required for convergence by several times) is observed for problems that have variables varying over just two orders of magnitude. For most engineering problem, a variation over a couple of orders of magnitude is not something

unusual. Furthermore, this issue needs to be addressed within the context of using this class of algorithms as a lower bounding procedure in a global optimization framework because even a well-conditioned problem can become ill-conditioned after several branching operations depending on the branching heuristics used. It is therefore preferable to implement an automatic scaling and conditioning of the problem within the algorithm itself.

5. There is substantial literature[1] on variants of cutting plane methods. Several ideas used to accelerate convergence and enhance robustness of these methods can be extended and used to improve the algorithm presented earlier.

6. Subgradient aggregation does not seem to offer any significant benefits in practice and in fact has been turned off for case studies described later. There may be theoretical reasons for aggregation being ineffective as outlined in Remark 7.5.10 (page 205), which need to be resolved before putting further effort into an aggregation strategy. Alternatively, lack of any benefits from aggregation may be attributed to a larger than required bundle being employed (although a smaller bundle does seem to impact the convergence adversely). A fine tuning of the algorithm may lead to a smaller bundle and therefore may see benefits of aggregation as well as cheaper direction finding solutions. However, there seems to be another problem with subgradient aggregation, which seems to indicate that a more sophisticated handling than simple summing up the constraints is required. The simple approach can make the direction finding problem ill-conditioned resulting in occasional CPLEX failures. This seems to indicate a more intelligent strategy for aggregation is needed: some kind of processing or scaling of multipliers, selective aggregation instead of including all constraints, based on a threshold of multiplier magnitudes and so on.

---

[1]For example, see Hiriart-Urruty and Lemaréchal [129], Chapter XV. Acceleration of Cutting Plane Algorithm: Primal Forms of Bundle Methods, page 275.

7.6 Lower Bounding Problem

The lower bounding calculation sequence is implemented in terms of libMC objects to calculate convex relaxations and the corresponding subgradients. The resulting convex but nonsmooth problem is underestimated by constructing its polyhedral approximation to obtain a lower bounding linear programming (LP) relaxation to the original nonconvex program. The bundle algorithm in this context is used primarily as a method to construct a polyhedral approximation of the feasible set and objective function and, therefore, to generate a linear programming (LP) relaxation and not directly as a method to solve the nonsmooth convex program. Simplex-based solution methods for LP provide a guarantee to detect optimality or infeasibility for the lower bounding problem (unboundedness is not possible in LP relaxations because the host set $\mathbb{C}$ is bounded) and therefore more robust and reliable than the nonlinear convex relaxations.

The maximum iteration count for the solver N_{max} is set to a small number of iterations, usually, a number larger than (but comparable to) the maximum bundle size, and roughly an order of magnitude lower than required for a positive convergence test for the bundle method. The bundle obtained at the end of the bundle method termination can be transformed to the following LP relaxation (polyhedral approximation) of the nonsmooth convex program in direction finding form[1]:

$$\min_{v,\mathbf{d}} \quad v \tag{LPR}$$

$$v - \boldsymbol{\xi}_j^f \mathbf{d} \geq f_j^k, \quad \forall j \in J_f^k$$

$$-\boldsymbol{\xi}_j^G \mathbf{d} \geq G_j^k, \quad \forall j \in J_G^k$$

$$\mathbf{x}^L - \mathbf{x}_k \leq \mathbf{d} \leq \mathbf{x}^U - \mathbf{x}_k.$$

Lemma 7.6.1. *The solution value v^* of Program* (LPR) *bounds from below the op-*

---

[1]Aggregated constraints are not shown here to simplify presentation. However, the relaxation argument in Lemma 7.6.1 applies to aggregated constraints (transformed to the same form as the corresponding constraints in Program (LPR)) as well since they are simply positive linear combinations of corresponding constraints in Program (LPR).

timal solution value of Program (ECP) *(page 196). Denote an optimal solution of Program* (LPR) *as* $\mathbf{d}^*$. *Consider an LP relaxation of Program* (ECP) *that is constructed by linearizing objective function at* $\{\mathbf{y}_j : j \in J_f^k\}$ *and linearizing total constraint function at* $\{\mathbf{y}_j : j \in J_G^k\}$. *Then an optimal solution of this relaxation can be constructed by taking a full step,* $\mathbf{x}_k + \mathbf{d}^*$, *where* $\mathbf{x}_k$ *was the last iterate (in the algorithm presented in 7.5.6), which was used to update RHS of bundle linearizations in the direction finding problem.*

Proof. Program (ECP) (page 196) can be stated equivalently as:

$$\min_{v, \mathbf{x} \in \mathbb{R}^n} v$$

$$v \geq f(\mathbf{x})$$

$$0 \geq G(\mathbf{x})$$

$$\mathbf{x}^L \leq \mathbf{x} \leq \mathbf{x}^U.$$

Construct a polyhedral approximation to the above program by linearizing f at points corresponding to objective function bundle index J_f^k and for G at points corresponding to constraint bundle index J_G^k at k^{th} iteration.

$$\min_{v, \mathbf{x} \in \mathbb{R}^n} v$$

$$v \geq f(\mathbf{y}_j) + \boldsymbol{\xi}_j^f(\mathbf{x} - \mathbf{y}_j), \quad \forall j \in J_f^k$$

$$0 \geq G(\mathbf{y}_j) + \boldsymbol{\xi}_j^G(\mathbf{x} - \mathbf{y}_j), \quad \forall j \in J_G^k$$

$$\mathbf{x}^L \leq \mathbf{x} \leq \mathbf{x}^U.$$

The above is an LP relaxation to convex program (ECP) and therefore its optimal value v^* underestimates the solution value of (ECP). Denote a solution of the above program as $\mathbf{x}^*$. Using the definitions of f_j^k and G_J^k from Relationship L2 (page 208),

it can be equivalently stated as:

$$\min_{v,\mathbf{x}\in\mathbb{R}^n} v$$

$$v \geq f_j^k + \boldsymbol{\xi}_j^f(\mathbf{x} - \mathbf{x}_k), \quad \forall j \in J_f^k$$

$$0 \geq G_j^k + \boldsymbol{\xi}_j^G(\mathbf{x} - \mathbf{x}_k), \quad \forall j \in J_G^k$$

$$\mathbf{x}^L \leq \mathbf{x} \leq \mathbf{x}^U.$$

Define $\mathbf{d} = \mathbf{x} - \mathbf{x}_k$. Substituting and rearranging:

$$\min_{v,\mathbf{d}\in\mathbb{R}^n} v$$

$$v - \boldsymbol{\xi}_j^f\mathbf{d} \geq f_j^k, \quad \forall j \in J_f^k$$

$$-\boldsymbol{\xi}_j^G\mathbf{d} \geq G_j^k, \quad \forall j \in J_G^k$$

$$\mathbf{x}^L - \mathbf{x}_k \leq \mathbf{d} \leq \mathbf{x}^U - \mathbf{x}_k.$$

This is Program (LPR). Therefore, the constructed LP relaxation can be transformed to Program (LPR). Direction $\hat{\mathbf{d}} = \mathbf{x}^* - \mathbf{x}_k$ is clearly feasible in the above program with an objective value v^*. Assume however, that $\hat{\mathbf{d}}$ is not optimal in (LPR), i.e., $\exists \bar{\mathbf{d}} \in \mathbb{R}^n$ that is feasible in (LPR) and has a solution value $\bar{v} < v^*$. If so, one can construct a new solution $\bar{\mathbf{x}} = \mathbf{x}_k + \bar{\mathbf{d}}$ to the LP relaxation with a solution value $\bar{v} < v^*$. This violates the assumption that $\mathbf{x}^*$ is an optimal solution to LP relaxation. Therefore, optimal solution value of (LPR) is v^* and any solution of (LPR), e.g., $\mathbf{d}^*$ corresponds to an optimal solution $\mathbf{x}_k + \mathbf{d}^*$ of the LP relaxation. This completes the proof. □

The LP relaxation is also used to detect infeasibility. This relaxation should become infeasible (either immediately or after few branching steps) if the original lower bounding convex problem is infeasible. The bundle algorithm implementation can detect infeasibility in certain circumstances, however it may also get stuck at an iterate unable to find a reasonable stepsize or descent direction as described earlier. In such cases, it is better to simply extract the bundle, solve an LP relaxation and detect infeasibility based on it.
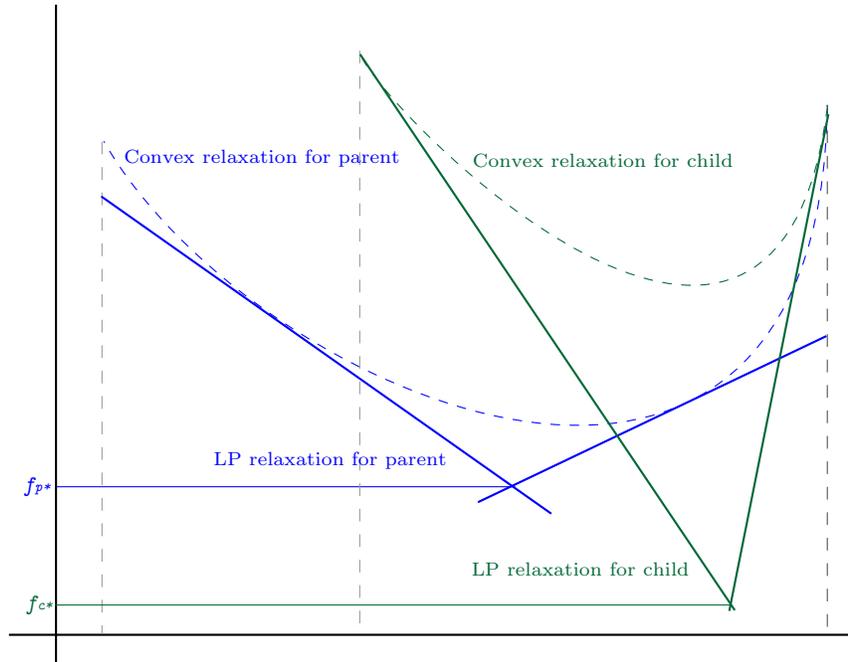
Figure 7-3: Need to propagate linearizations for LP relaxations: LP Lower bound for a child can be worse than the parent even when convex relaxation on the child is better than the parent

One of the problems associated with using LP relaxations in this manner is that there is no guarantee that the lower bound will actually improve after a single branching operation (although, after a large number of branching operations, i.e., in the limit, the linearizations for children will indeed improve). For example, a situation shown in Figure 7-3 is possible when a child LP relaxation solution is worse than the parent even though the convex relaxation for the child node is far tighter than its parent. This behavior is indeed observed in practice. The best lower bound in the branch-and-bound tree will improve till a certain point in the algorithm iteration and then drop back again. This can happen several times at intermittent iterations with successive drops becoming smaller as partitions become smaller. As a result the algorithm does converge but only very slowly.

7.6.1 Propagation of Linearizations in Branch-and-Bound Tree

The solution to this problem is to propagate the linearizations from parent to the child node. The constraints above for the LP relaxations are stored in the node data structure and are propagated to its children when branching. Let index sets J_f^p and J_G^p denote the bundle inherited from the parent node. At the entry point into the lower bounding procedure in a child, The LP relaxation generated by the index sets J_f^p and J_G^p is formed with the new child bounds. The LP relaxation is solved and a full step is taken to update the initial point (provided by the branch-and-bound procedure), so that it is feasible in the child bounds. An explicit check is performed to make sure that the point so generated is indeed within child bounds and in case, a component is outside its upper or lower bound by the LP solver bound satisfaction tolerance or less, it is reset to the corresponding upper or lower bound. The point may be still infeasible in the original convex program at the child because the objective and constraints functions of the convex program are convex relaxations are expected to change in the child node. This point is passed to the main solver routine. If the LP relaxation is infeasible, the node is declared infeasible.

There are two options for handling these inherited linearizations in the direction finding problems at the child node. They can be added as normal linear constraints. The second option is to transform them into their bundle form for the direction finding problem as follows:

$$v - \boldsymbol{\xi}_j^{f_p} \mathbf{d} \geq f_{pj}^{k} - f(\mathbf{x}_k), \quad \forall j \in J_f^p,$$
$$v - \boldsymbol{\xi}_j^{G_p} \mathbf{d} \geq G_{pj}^{k}, \quad \forall j \in J_G^p.$$

The second option was chosen based on the relative simplicity of its implementation. The argument for doing so is as follows. Consider the kth direction finding step

obtained from minimization of the improvement function at the current node:

$$\min_{\mathbf{y} \in \mathbb{C}} \quad H(\mathbf{y}; \mathbf{x}_k) = \begin{cases} \min\limits_{\mathbf{y} \in \mathbb{C}, v} v \\ v \geq f(\mathbf{y}) - f(\mathbf{x}_k) \\ v \geq G(\mathbf{y}). \end{cases}$$

Let the f_p and G_p denote the relaxations at the parent node. Then any $(v, \mathbf{y})$ feasible in the above convex program (with the feasible set corresponding to current node) is not cut off from the direction finding problem if one adds additional constraints of the form

$$v \geq f_p(\mathbf{y}) - f(\mathbf{x}_k),$$
$$v \geq G_p(\mathbf{y}),$$

provided the relaxations on parent and child satisfy the following property[1]

$$f(\mathbf{y}) \geq f_p(\mathbf{y}), \qquad G(\mathbf{y}) \geq G_p(\mathbf{y}), \quad \forall \mathbf{y} \in \mathbb{C} \subset \mathbb{C}_p,$$

where $\mathbb{C}$ is the host set at the current node and $\mathbb{C}_p$ is the superset at the parent node. Therefore an equivalent representation of the convex program minimizing the improvement function at $\mathbf{x}_k$ is as follows:

$$\min_{\mathbf{y} \in \mathbb{C}, v} v$$
$$v \geq f(\mathbf{y}) - f(\mathbf{x}_k)$$
$$v \geq G(\mathbf{y})$$
$$v \geq f_p(\mathbf{y}) - f(\mathbf{x}_k)$$
$$v \geq G_p(\mathbf{y}).$$

---

[1]This has been shown to be true for McCormick relaxations by Joseph K. Scott (currently at Process Systems Engineering Laboratory, MIT), however, the result is unpublished at the time of writing of this thesis.

The polyhedral representation of the above can be represented using the notation in Section 7.5.2:

$$\min_{\mathbf{y} \in \mathbb{C}, v} v$$

$$v \geq f(\mathbf{y}) - f(\mathbf{x}_k) \geq \widehat{f}^k(\mathbf{y}) - f(\mathbf{x}_k),$$

$$v \geq G(\mathbf{y}) \geq \widehat{G}^k(\mathbf{y}),$$

$$v \geq f_p(\mathbf{y}) - f(\mathbf{x}_k) \geq \widehat{f}_p^k(\mathbf{y}) - f(\mathbf{x}_k),$$

$$v \geq G_p(\mathbf{y}) \geq \widehat{G}_p^k(\mathbf{y}),$$

and finally obtain the following problem for the direction finding problem:

$$\min_{v, \mathbf{d}} v$$

$$v - \boldsymbol{\xi}_j^f \mathbf{d} \geq f_j^k - f(\mathbf{x}_k), \quad \forall j \in J_f^k$$

$$v - \boldsymbol{\xi}_j^G \mathbf{d} \geq G_j^k, \quad \forall j \in J_G^k$$

$$v - \boldsymbol{\xi}_j^{f_p} \mathbf{d} \geq f_{pj}^{\ k} - f(\mathbf{x}_k), \quad \forall j \in J_f^p$$

$$v - \boldsymbol{\xi}_j^{G_p} \mathbf{d} \geq G_{pj}^{\ k}, \quad \forall j \in J_G^p$$

$$\mathbf{d}_k^L \leq \mathbf{d} \leq \mathbf{d}_k^U.$$

Inherited linearizations are propagated as is through the direction finding problem and not aggregated or dropped even if they are inactive. Also note that the RHS of the inherited constraints requires an update as per *Step 4* (page 214) in the formal statement of the algorithm (Section 7.5.6). Sets J_f^k and J_G^k are passed to the children of the current node on branching while the inherited linearizations are dropped.

7.6.2   Linearization Propagation and Feasibility Phase

The inherited constraint linearizations are passed on to the feasibility phase if the solution of the initial LP constructed from inherited parent linearization (on entry to the lower bounding procedure) yields a solution point that is infeasible in nonlinear constraints. The constraint linearizations are passed back to the main solver if the

problem is found feasible. If the feasibility phase exceeds the maximum number of iterations without locating a feasible point, the constraint bundle is extracted and combined with the linearizations inherited from the parent node to generate an LP relaxation at this node. This LP relaxation is solved to generate a lower bounding value for the node. If this LP relaxation is infeasible, the node is declared infeasible.

7.7   Upper Bounding Problem

The upper bounding problem is implemented as exactly the same computer calculation sequence as the lower bounding problem, however, now ADOL-C [132] objects are used for implementing the calculation sequence (instead of libMC objects). ADOL-C [132] is used to generate gradients of the objective function and constraints with respect to input variables using operator overloading. Simple "tapeless" forward mode in ADOL-C is used to generate the derivative for simplicity of implementation. SNOPT [119] is used for local optimization of the resulting reduced NLP. It must be noted that any information about the sparsity of the problem is lost in this operation and the Jacobian of the system has to be assumed dense. Therefore, this approach may not directly scale to very large-scale problems (i.e., problem with several hundred or several thousand input variables). However, ADOL-C does permit the calculation of sparsity in tape mode but this has not been explored in this work.

It is also possible to avoid solving upper bounding problem at every node. For example, it is possible to attempt an upper bounding solution periodically every few nodes (instead of each node which has not been proven infeasible or fathomed by value dominance). Alternatively, for nodes, where lower bounding solution is feasible in the upper bounding feasible set, a simple objective function evaluation can provide an upper bound for the node.

7.8 Branch-and-Bound Algorithm

The Branch-and-Bound is implemented in C++. The node data structures are able to propagate the objective function and constraint bundles from parent to child node. Simple range reduction heuristics are implemented for bound constraints [133] using the duality multiplier corresponding to bound constraints to tighten the bounds. This scheme cuts off part of the feasible regions that is value-dominated, i.e., where the lower bound value is greater than the current best upper bound available in the tree. Let λ_i^L and λ_i^U denote the duality multipliers corresponding to lower and upper bounds respectively corresponding to x_i. Let L be the solution of the lower bounding problem at this node and let U_b be the current best upper bound available in the tree. Then the bounds can be tightened in the following way for variable x_i

$$x_i^U = \min\left\{\bar{x}_i^U, \quad \left(\bar{x}_i^L - \frac{L - U_b}{\lambda_i^L}\right)\right\},$$
$$x_i^L = \max\left\{\bar{x}_i^L, \quad \left(\bar{x}_i^U + \frac{L - U_b}{\lambda_i^U}\right)\right\},$$

where $\bar{x}_i^U$ and $\bar{x}_i^L$ are the original upper and lower bounds corresponding to the current node. Note that the actual bounds reduction is carried out on bound on $\mathbf{d}$ since the LP relaxations is in terms of $\mathbf{d}$ from which the above relationships can be deduced.

7.9 Implementation

An overview of the implementation is shown in Figure 7-4. The network and NLP data structures are implemented as C++ templates. Exactly the same calculation sequence and the corresponding data structures are implemented in terms of different underlying objects, i.e., libMC and ADOL-C for lower and upper bounding implementations respectively. C++ templates avoid the duplication of the same code between the upper and lower bounding problems for different objects or more precisely, they allows for automatic generation of the same code by the C++ compiler for objects from libMC and ADOL-C for lower and upper bounding problems. The actual cal-

233
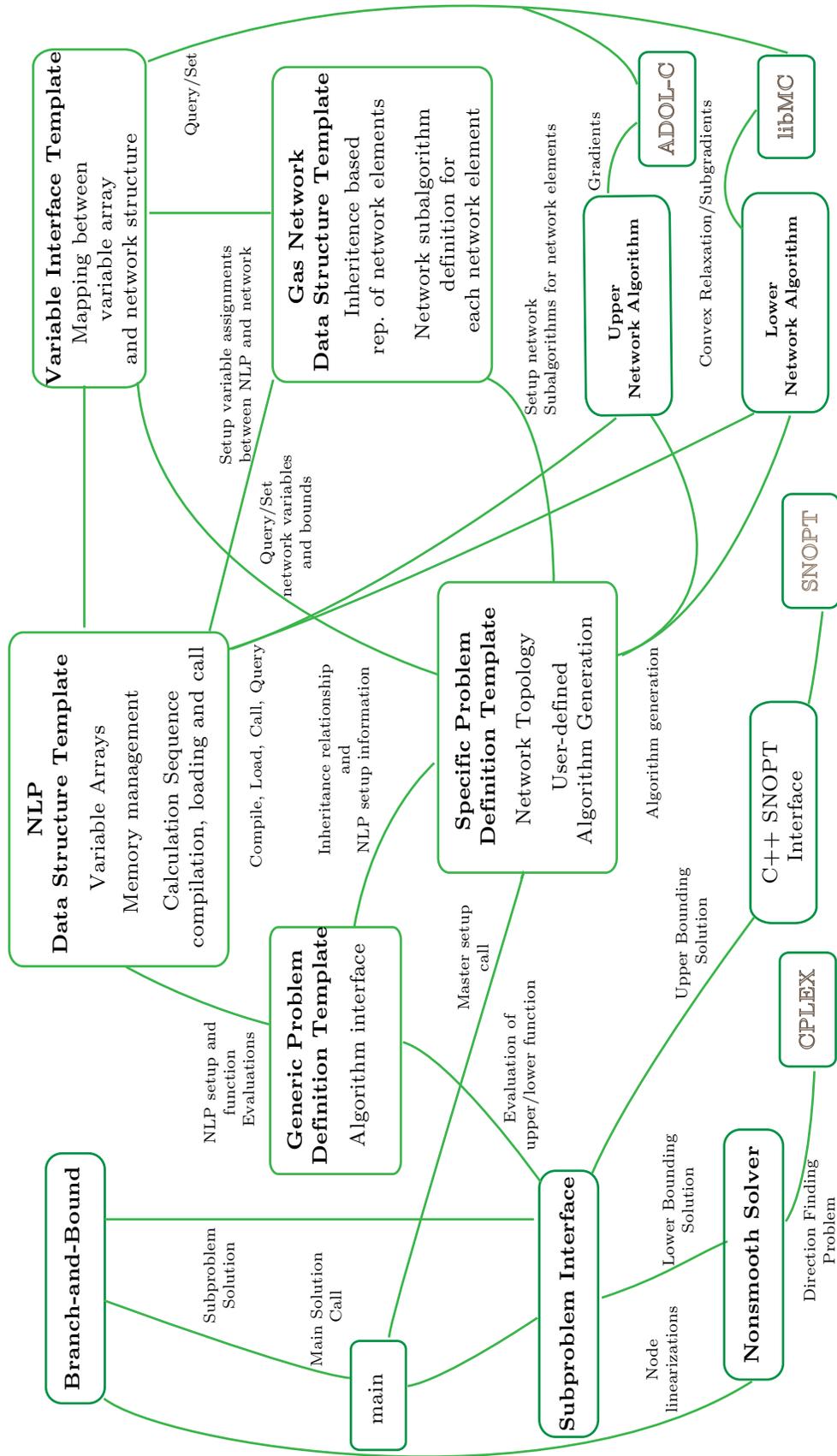
Figure 7-4: Implementation overview: Global optimization of algorithms applied to upstream gas networks

culation sequences are implemented with simple libMC and ADOL-C object arrays to speed up computation. The book keeping to map the plain array variables that are allocated in NLP data structures to the actual network variables is done by the object-oriented representation of the network. The calculation sequence is generated, compiled and loaded on the fly using calculation fragments for each network element defined in the network data structures.

7.10   Preliminary Case Studies

Two preliminary case studies are presented to demonstrate the approach presented so far. The networks corresponding to case study A and B are presented in Figure 7-5 and 7-6 respectively. Best lower bound is used as a node selection heuristic and largest absolute diameter is used for branching variable selection heuristics in the branch-and-bound algorithm. The maximum number of iterations, objective bundle size and constraint bundle size for the bundle solver are all set to $n + 5$ where n is number of variables. $m = 0.5$, $\bar{t} = 0.1$ and $t_{min} = 10^{-12}$ are used as line search parameters. The results are presented in Table 7.1. The CPU times shown in the table are on an Intel Core Duo 2.16 GHz processor running Linux kernel 2.6. The source code was compiled using GCC 4.2 with optimizations.

The number of nominal variables is the number of variables that will be required to model the problem in a conventional NLP framework. This number is estimated as follows. The gas composition is assumed to vary between wells belonging to the same field (as opposed to the infrastructure model presented earlier). Therefore, 12 variables are required for modeling of a well (volumetric production rate, 8 specieswise molar rates, 2 pressures, NGL production rates). Each arc (including the compressor) requires 9 variables (one volumetric flowrate and 8 molar flowrates). Each field and demand requires 10 variables (volumetric rate, 8 molar rates and a pressure). Each compressor inlet and outlet, and junctions require a pressure. Finally, each demand requires 10 variables, volumetric and molar delivery rates and a pressure. The actual number of variables is the actual number of input variables in the reduced-NLP.
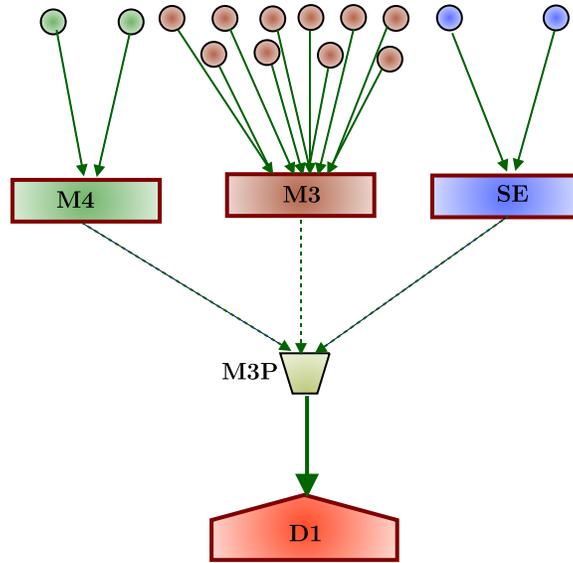
Figure 7-5: Network corresponding to case study A
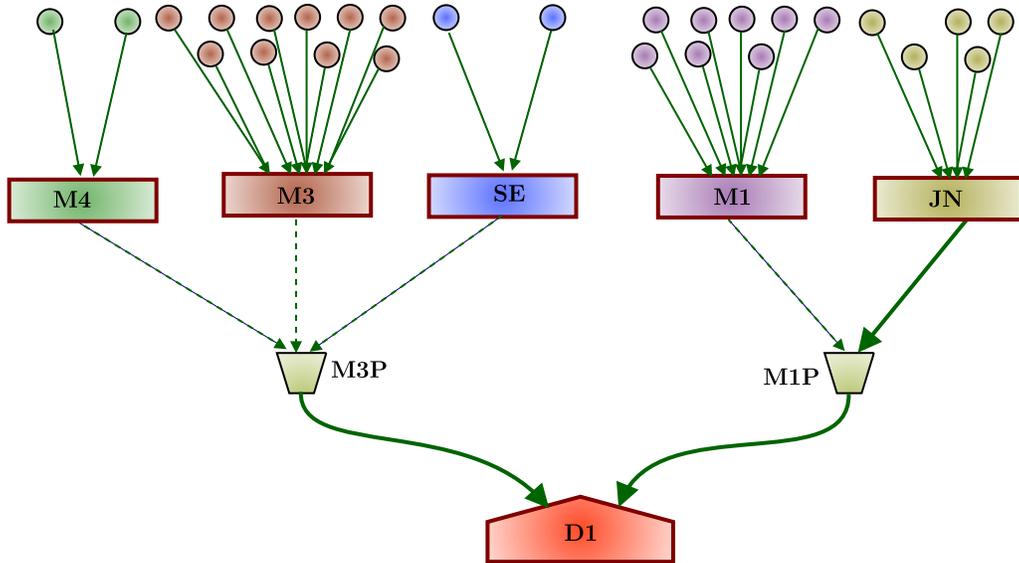
Figure 7-6: Network corresponding to case ctudy B

Table 7.1: Relaxation of algorithms applied to upstream gas networks: Preliminary case studies

Case study	Number of variables		CPUs	Delivery
	Nominal	Actual		hm^3
A (14 Wells, 3 fields, 1 demand)	228	19	12[a]	30.42
B (27 Wells, 5 fields, 1 demand)	433	37	3,280[b]	73.83

[a] 2% termination gap
[b] 3% termination gap

Case study A has only delivery pressure constraints at the demand node. Case study B has H_2S and C_2 quality constraints as well as delivery pressure constraints at the demand nodes. Further details on the parameters and the constraints for the case studies can be found in Appendix E (page 297).

Chapter 8

Conclusions and Future Work

8.1 Conclusions

Natural gas supplied a fifth of global energy demand in 2005 and it is forecasted to continue to supply a similar fraction in the coming decades. Natural gas demand growth is expected to be especially strong in power generation. It is the cleanest of all fossil fuels and is therefore also expected to play the role of a transition fuel in the near future as alternative and greener sources of energy come online and various technical and business issues associated with these sources are resolved.

Natural gas supply chains have unique characteristics (when compared to oil) due to the low volumetric energy density of natural gas. The production, transportation and storage infrastructure is capital-intensive to build, incurs high operational costs and is specific to natural gas. Natural gas demand is volatile and prone to daily, weekly and seasonal fluctuations because a large share of consumption is contributed by electricity generation and commercial/residential heating. It is difficult to alleviate supply shocks due to the difficulties in storage and transportation. Transportation and consumption infrastructure are sensitive to gas quality. The entire natural gas supply chain, right from upstream systems to local distribution networks, is operated on the basis of contractual agreements. Contracts play a central role in gas markets. The growth of liquefied natural gas (LNG) trade and spot markets are welcome developments that are working toward faster development of resources and more flexibility

in markets.

Supply chain planning in natural gas systems is critically important because the entire supply chain is much more interdependent and coupled. Disruptions can propagate through the chain and can lead to supply shocks in markets. Supply chain modeling frameworks can potentially play a crucial role in evaluating scenarios and ensuring that the effect of disruptions and breakdowns in one part of the system can be limited in extent and localized. A supply chain planning problem is an inherently short-term (i.e., over several weeks to several months) planning problem.

A crucial component of any supply chain planning framework is the model of the upstream production system. Upstream systems are usually centrally operated and governed by production-sharing contracts between multiple stakeholders and gas sales agreements with the consumer facilities. Planning frameworks in upstream systems can help to maximize production infrastructure utilization, aid asset management, minimize costs, increase returns, honor governing rules and ensure reliable supplies.

A model for operational planning in the upstream natural gas supply chain has been developed and was presented in Chapters 3 and 4. The model features and requirements are inspired by the Sarawak Gas Production System (SGPS) in East Malaysia. The SGPS is used as a real-world case study to demonstrate the application and benefits of the proposed modeling methodology. This is the first attempt (to the best knowledge of previous works) to formulate a comprehensive modeling framework for an upstream production system that includes not only the production infrastructure model but also a methodology to incorporate the governing rules for upstream systems, e.g., production-sharing contracts, customer specifications and operational rules into the modeling framework. The model has two components: *the infrastructure model* and *the contractual model*. The infrastructure model is the model of the physical system, i.e., wells, trunkline networks and facilities. The contractual model is model of the governing rules, e.g., customer specifications and production-sharing contracts. The model formulation and objectives are from the perspective of the upstream operator.

The infrastructure model incorporates the capability to track multiple qualities of

240

gas through the trunkline network and, therefore, to be able to route and blend gas in the network in such a way that the customer quality specifications are satisfied at LNG plants. This is an important concern with the rising number of sour fields being developed throughout the world. Nonlinear pressure-flowrate relationships in wells and trunkline networks are included for a realistic pressure-flowrate profile because the infrastructure network is controlled by regulating pressure at certain nodes. Modeling of complex platform configurations with reversible lines and lines that can be opened and closed in normal operation provides realistic routing for blending.

The contractual model represents a framework for modeling customer specifications, complicated production-sharing contracts (PSC) and operational rules that are central to the operation of the system. PSC modeling is a two-fold challenge: accounting for volumes and converting the logical rules from the system operational manuals to their mathematical representation. A PSC network representation is proposed to account for volumes and the interactions between different PSC. A formal approach is proposed to express PSC rules as binary constraints by first defining atomic propositions to represent excess, priority and transfers states of the PSC, converting PSC rules to logical expressions in terms of these states and finally converting these logical expressions to binary constraints. The PSC states are linked to the flow on the PSC network by governing constraints that force the flow to be in line with the states. Additional logical constraints are required to model the inference of the rules. Operational rules can also be modeled within the same framework.

Although the model has been inspired by SGPS features, it is general enough to handle most upstream gas production systems with large consumer facilities. The modeling framework results in a relatively large nonconvex mixed-integer nonlinear program (MINLP) with tens of binary variables and several hundred to a few thousand continuous variables and several hundred nonlinear equality constraints. Numerical experiments indicate that local solution methods can even fail to detect feasibility in (continuous versions of) such models reliably. This necessitates the use of global optimization algorithms to solve the problem. The contractual modeling framework introduces substantial complexity both from a modeling standpoint and

241

computational time. A reproducible case study was presented to demonstrate the entire framework. Three operational objectives: maximization of dry gas, natural gas liquids (NGL) and sour gas rates have been considered.

The resulting MINLP can be solved with current state-of-the-art approaches [117], provided close attention is paid to modeling details. The MINLP has multiple solutions with the same optimal solution value. The upstream planning problem has multiple operational objectives which have priorities as dictated by the contractual and operational rules. A hierarchical multi-objective approach is proposed to exploit the fact that there are multiple solutions to the mathematical program and hence, a lower priority objective can be optimized over the solution set of the higher priority objective to obtain a win-win scenario. As an example, a 10% increase in the NGL production rate has been demonstrated for the case study with the same dry gas production, yielding a substantial increase in revenues for the upstream operator.

The existing approach involves a laborious trial-and-error procedure to satisfy quality specifications, PSC and operational rules. Also, the modeling approaches used in the existing framework possibly violate assumptions inherent in the local solution algorithms and there is little guarantee for the quality of the solutions obtained with such approaches. Hence, compared with the existing approach, the proposed approach is a theoretically and practically better alternative. A preliminary comparison with the existing approach indicates that substantial gains may be possible by using the proposed approach. A dual model approach of having a simulation model (the existing approach) and a separate planning tool (the proposed approach) calibrated to each other is recommended as the most promising way forward for the overall implementation.

The application of reduced-space global optimization to the upstream gas networks has been demonstrated. This can significantly lower the number of variables in the branch-and-bound algorithm. The lower bounding problem was implemented using libMC and solved by implementing a bundle solver as an iterative linearization tool. The upper bounding problem was implemented using ADOL-C and SNOPT. A branch-and-bound algorithm with reduction heuristics and linearization propagation

242

was used for the global optimization. It has been demonstrated that the number of variables in the branch-and-bound can be reduced by a factor of 10 in upstream network problems and these methods are competitive with current state-of-the-art approaches.

8.2 Future Work

8.2.1 Conventional Solution Methods

Efficient solution methods for the short-term upstream natural gas planning problem should be one of the most important areas for future work. Current state-of-the-art methods can only just barely solve these problems. A minor variation in a bound or a change of objective is enough to cause a model instance from being solvable within few hours to not converging in several days. This is especially true for problems that are just barely feasible, i.e., when the system is operating near its maximum potential. Given that determining availability under such conditions is one of the most important applications for such models, this is a major drawback.

One of the most promising approaches in this regard seems to be a successive solution strategy that involves starting with a strict subset of the constraints so as to make the problem relatively "easy to solve" and then successively adding more constraints, bounding the objective using the previous solution value and restarting the solution procedure from the previous solution point. As shown in Tables 5.1 (page 148) and 5.2 (page 153) for the dry gas rate maximization objective, a direct convergence to within 10% relative gap takes 41,424 CPU seconds (CPUs), while a convergence to 1% gap can be achieved in approximately half this time, 19,679 CPUs, using three successive restarts. The indication that this approach has potential is also supported by the results for the hierarchical multi-objective optimization presented in Table 5.3 (page 154) where the first stage objective takes 41,424 CPUs while the other two combined take only 4,803 CPUs. It is possible to further expand, formalize and refine this strategy. The success of this strategy is probably related to the structure of

the problem. The most plausible reason may be that the tight bounding of objective achieved in such a strategy combined with the structure of the solution helps to prune a large part of the branch-and-bound tree. If the problem structure permits an easier or even tighter calculation of these bounds, a successive solution approach can be further speeded-up. It is also possible to integrate this idea within the branch-and-bound tree directly instead of multiple branch-and-bound solutions as was done in Section 5.4. A succession of passes can be made over a single tree each time adding a subset of constraints and pruning the tree based on value dominance and infeasibility. The initial passes are expected to be quite fast, as in Table 5.1, since the first two solution steps take very little time.

A lot of experimentation is needed to come up with heuristics suited to this class of problems within a branch-and-bound framework by exploiting problem structure. Not only that, the subproblems solution can also be speeded-up using some of these customized heuristics. Following are some of the possible options to exploit. A subset of the constraints in the problem are the classical network constraints for which fast LP methods are available. The direct relationship between volumetric rates, molar rates and pressures permits an easy calculation of all these quantities provided a few of them are known. In fact this is what the algorithmic relaxation approach presented in Chapter 7 exploits. Even in a conventional branch-and-bound approach, it should be possible to deduce information about the subproblem on a node much higher up in the tree based on these features. Another subset of constraints in the infrastructure model originate from the pooling problem component (mixers and splitters in the network). Pooling problem formulations and solution methods have been extensively addressed in the literature and some of these may apply to gas networks. A substantial amount of literature exists on the simulation and optimization of natural gas transmission systems, some of which may be applicable to upstream networks. Getting a feasible contractual binary realization is possible by designing simple heuristics to adjust infrastructure model bounds and factor the binary variable out from the flows on the contract network. Additionally, information from the contractual model can be used to immediately rule out some infrastructure solutions higher up in the tree and deduce

244

additional constraints. As an example, quality specifications can permit calculation of an additional cut on sour field production rate using system-wide molar balances and maximum delivery rates that may lead to stronger relaxations higher in the tree. However, some redundant constraints can result in linear dependence in the problem which can adversely affect some local solution algorithms. A preprocessing step should be able to resolve linear dependence at least for linear constraints. A preprocessing step can also reduce the number of variables in the problem by performing a structural analysis of constraints and symbolic substitutions. However, a good understanding of the algorithms and their relaxation procedures is required in doing so. The contractual rules themselves can be processed further using Boolean algebra to deduce additional integer constraints.

It is also possible to integrate the hierarchical multi-objective approach into branch-and-bound approach. A first pass with the highest priority objective can made. In the process of finding the optimal solution, a part of the tree is pruned. The second pass only needs to consider the left over (both explored and unexplored) part of the tree. In each successive pass, the part of the tree that remains is reduced having been pruned by earlier objectives. The implementation of this strategy is expected to be similar to the implementation of the successive solution strategy in branch-and-bound framework which was discussed earlier.

Substantial overhead is involved in making multiple solver calls with modeling languages such as GAMS or AMPL and therefore, any of the strategies suggested above are unlikely to be efficiently implementable in a higher level modeling language. They require in-house implementation of a solution procedure that can be modified and customized at will. From an implementation perspective, all strategies requiring multiple branch-and-bound passes require sophisticated memory management due to the fact that first few passes may require thousands of nodes which may be expensive to store if all the required information for subproblem warm-starts is included. Another advantage of an in-house implementation is that there is a wider choice of subproblem solvers available. For example, BARON cannot use current state-of-the-art interior point local solvers.

8.2.2 Global Optimization of Algorithms

Relaxation of algorithms seems to be a promising approach for solving such problems. However, there is lot of work done to be able to apply these methods to the upstream production planning problem robustly. The bundle method implementation needs to be improved a lot and some key areas for improvement were already identified in Section 7.5.9 (page 222). For solving bigger problem, source code transformation and an more efficient calculation of subgradients (reverse mode) may be required. Note, for example, that although only one subgradient is needed for updating the constraint bundle, the entire subgradient matrix is being propagated in the current approach. The upper bounding approach can also improved to exploit sparsity.

A better calculation sequence needs to be derived from a structural analysis of the system of equations. A unidirectional calculation sequence (in the direction of flow) does not seem to be the best sequence possible. A better sequence may be derived by traversing the network in both directions. The modeling approach can also be easily extended to the calculation of contractual volumes by incorporating the contractual network in the calculation sequence.

Again, there is much experimentation to be done to come up with improved branch-and-bound heuristics for reduced-space approaches in general and the application to natural gas networks in particular. Probing and other advanced bounds reduction techniques [117] can also be implemented to further accelerate convergence.

Combining the reduced-space approach with mixed-integer programming to handle discrete decisions is another challenge. It is possible to take binary variables as input (fixed on a node) and modify the calculation sequences (for that node) as dictated by the binary vector. However, it is not clear what must be done with the binary variables which have not been branched upon (possibly retain a conventional formulation till the variable are branched upon). A network calculation procedure can also aid in generating a feasible binary realization because some of the outcomes of a discrete decision can be taken as input and a calculation (some kind of reverse calculation) can determine other variables.

8.2.3 Modeling Aspects

Determining variable bounds continues to be another big challenge. Their importance was already discussed in Section 5.2. Two broad approaches, namely, estimation from physical arguments and from historical operating data were discussed (in the context of well production bounds) in Section 5.2. Given the fact that the computational time is extremely sensitive to bounds, varying over orders of magnitude with bounds variation, and, the sensitivity of the solution point to these variations, this is a crucial area for further improvements.

There continues to be a substantial amount of work left to refine the modeling framework. Better pressure-rate relationships that are accurate enough for planning purposes while simple enough to be handled within global optimization frameworks are an important focus. For example, sophisticated reservoir and multiphase flow simulators can be used to generate response data that can be used to obtain better well performance models. It may also be required to incorporate a simple multiphase flow model for lines with high condensate flow to account for the effect of liquid transport on gas flow. More detailed modeling of well and riser platform configurations would help to predict actuator inputs accurately in the planning solution that can be fed to a lower level control system. It can therefore allow for preferential routing of gas in the network for effective blending and result in a routing that is more realistic and closer to implementation.

A model of the consumer facility (facilities) can be included in the upstream planning problem if the mode of operation of the system allows for a coordinated control of the upstream system and the consumer facility. For example, a representation of the LNG plant in the model can help the LNG plants to respond to upstream fluctuations. Although the last improvement is not directly applicable to the SGPS, as the LNG plants are owned by a third party, it is true for several other gas production systems where the upstream operator also owns the liquefaction facility. A combined operation of the LNG plant and the upstream system can help the LNG plant and shipping facilities to adapt to the changing state of the upstream system

and therefore move to a new steady state. For example, if there is a breakdown in a field, the delivery amount, pressure, condensate amount and composition of feed gas will change and LNG plant operating conditions need to undergo a corresponding change. Equally possible is the flow of information from the LNG plant back to the upstream system. For example, if the CO_2 processing capacity of the plant is reduced due to a breakdown, the upstream system can temporarily respond by cutting production from CO_2 rich fields so that the LNG plant can maintain LNG production rates without sacrificing quality constraints and without any operational difficulties. If the consumer facility has an LPG plant, a simplified model of the LPG plant in the model would help to incorporate the objective of maximizing LPG production (this also is of interest to the upstream operator in the SGPS).

An economic representation of the system could be built on top of the model presented here where the contractual modeling framework is extended to include complex commercial and economic rules. Several upstream systems have complicated contractual clauses that imposes penalties on the operator if the contractual volume or quality requirements are not met. These can be incorporated in the upstream planning models to prevent these penalties and to minimize the penalties in the case when an operational difficulty makes it impossible to meet all the requirements.

More innovative objectives can be considered in this framework. The main product from an upstream system coupled with a liquefaction (or GTL) system is LNG (or liquid fuels). If the mode of operation of the system permits, it is more sensible to target plant operational/economic objectives than the upstream system objective. This can potentially result in higher returns from the entire system. An objective based on contractual violations can indicate which contractual violation can lead to increased production rates. Several other operational objectives such as pressures at certain nodes, certain qualities and production from certain fields can also be analyzed. Finally, economic objectives can be included if the model includes an economic representation of the system.

8.2.4 Variable Transformations

A network is composed of repeating elements that are more or less similar to each other with respect to governing equations. Therefore, the constraints representing these elements and the corresponding "motifs" (i.e., form of terms or cluster of terms in decision variables) repeat multiple times in the MINLP formulation. It is possible to use transformed variable definitions to simplify the terms that occur multiple times (e.g., convert a nonlinear term to a linear term), thereby tightening relaxations. Any repeating functional form of variables is a candidate for such transformation.

An example is as follows. In the infrastructure model, the pressure variable appears as a square term more often than as linear term. Therefore, the transformation of pressure (replace pressure variable by its square) presented in the derivation of the calculation sequence in Section 7.4 can also be applied within conventional MINLP approach. Such a transformation reduces the nonlinear terms in the model and therefore may make the relaxations tighter. However, the approach presented in Section 7.4 runs into a problem when there are linear equality constraints on pressure. In one candidate formulation, a square root of the transformed pressure variable appears (making a previously linear constraint as nonlinear) which can potentially become ill-conditioned at low pressures (though, all pressures are bounded below by atmospheric pressure). In superficial experimentation, this formulation only marginally differs in computational time from the original formulation. However, a more serious investigation is warranted to ascertain performance. If there are only a few linear constraints involving pressure, a second candidate formulation is possible. A linear constraint can be formulated in terms of an auxiliary variable and the square of this auxiliary variable can be set equal to the transformed pressure variable. The benefits of such an approach are questionable and uncertain since it increases the number of variables in the problem.

It may be possible to come up with creative transformations that simplify the MINLP considerably.

8.2.5   Sensitivity Analysis

A sensitivity analysis needs to be carried out to ascertain the variation of optimal solution value and solution point with respect to parameter values and variable bounds. Intuitively, the solution is expected to be less sensitive to minor variations in parameter values in pressure-flowrate relationships because there is enough slack in the network to adjust pressure-flowrate profiles as long as bounds permit it. However, there is a strong dependence of the solution point on variable bounds for the problem. Not only that, the computational effort required for the problem can vary over orders of magnitude depending on bounds.

Hence, sensitivity of solution with respect to bounds is one of the important areas to investigate (and in some sense is also the "easiest" because bounds form RHS of the constraints, the theoretical foundation for which are available from convex nonlinear programming). Such an analysis will help by identifying the most sensitive bounds and thereby direct the effort to estimate them more accurately.

The following are some representative works in the area. Differentiability of the solution for parametric convex programs appears in Dempe [134]. However, this is a challenging theoretical problem for nonconvex NLPs and MINLPs. One of the earliest discussions of shadow prices for nonconvex NLP appears in Gauvin [135]. A discussion of properties of NLP solutions appears in Gauvin and Janin [136] and Shapiro [137]. Another recent treatment of local sensitivity analysis of multi-valued solution maps appears in Levy and Mordukhovich [138]. A detail discussion of sensitivity analysis for optimization problem appears in Bonnans and Shapiro [139].

8.2.6   Implementation Issues

Although implementation of this work in an industrial environment is not strictly a research problem, there are some outstanding issues that fall between research and application that need to be addressed. Heuristics for implementations that generate an "answer" under every possible input with a graceful fallback to local solvers or feasibility-phase procedures is an important usability concern that cannot be ad-

equately addressed without a deep knowledge of solution algorithms and problem structure. A systematic mechanism to trace infeasibility of the model to specific delivery specifications and contractual rules is needed for a good implementation because determining the source of infeasibility is not always obvious because the effect of constraints can propagate through the network to appear far away from the concerned constraint. An approach to do automatic conversion of the complicated logical rules to the most favorable integer programming formulation and generation of a maximum number of redundant constraints (to strengthen relaxations) requires a good understanding of mixed-integer modeling and algorithms. A proper mechanism for calibrating the model automatically using historical operating data also needs to be explored.

Bibliography

[1] *World Energy Outlook 2007 - China and India Insights*. International Energy Agency, 2007.

[2] International energy outlook 2008 (highlights). Technical report, Energy Information Administration, U.S. Department of Energy, 2008.

[3] *International Energy Outlook 2007*. Energy Information Administration, U. S. Department of Energy, 2007.

[4] *BP Statistical Review of World Energy*. BP, June 2008.

[5] *Natural Gas Market Review 2007 : Security in a globalizing market to 2015*. International Energy Agency, 2007.

[6] *Natural Gas Market Review 2006 : Towards a global gas market*. International Energy Agency, 2006.

[7] International natural gas and liquefied natural gas (LNG) imports and exports: Energy Information Administration, U.S. Department of Energy. http://www.eia.doe.gov/emeu/international/gastrade.html.

[8] *Security of Gas Supply in Open Markets: LNG and power at a turning point*. International Energy Agency, 2004.

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

[10] C.A. Haverly. Behavior of recursion model - More studies. *ACM SIGMAP Bulletin*, 26:22–28, 1979.

[11] C.A. Haverly. Studies of the behavior of recursion for the pooling problem. *ACM SIGMAP Bulletin*, 25:19–28, 1978.

[12] D. L. Katz and R. L. Lee. *Natural Gas Engineering: Production and Storage.* McGraw-Hill Publishing Company, 1990.

[13] Chi U. Ikoku. *Natural Gas Engineering: A Systems Approach.* PennWell Books, Tulsa Oklahoma, 1980.

[14] William C. Lyons and Gary J. Plisga, editors. *Standard Handbook of Petroleum and Natural Gas Engineering.* Elsevier, second edition, 2005.

[15] Mehmet Rasin Tek. *Natural Gas Underground Storage: Inventory and Deliverability.* PennWell Publishing Co., 1996.

[16] E. L. Dougherty. Application of optimization methods to oilfield problems: proved, probable, possible. In *47th Annual Fall Meeting of SPE*, San Antonio, Texas, Oct 1972. Society of Petroleum Engineers Inc., SPE 3978.

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

[18] J. Van Dam. Planning of optimum production from a natural gas field. *Journal of the Institute of Petroleum*, 54(531):55–67, March 1968.

[19] E. L. Dougherty, D. Dare, P. Hutchison, and E. Lombardino. Optimizing SANTOS's gas production and processing operations in central Australia using the decomposition method. *Interfaces*, 17(1):65–93, 1987.

[20] Kunal Dutta-Roy, Santanu Barua, and Adel Heiba. Computer-aided gas field planning and optimization. In *SPE Production Operations Symposium*, Oklahoma City, Oklahoma, Mar 1997. SPE, SPE 37447.

[21] Allan Mortimer. North Sea plants get process, scheduling tools. *Oil and Gas Journal*, 96(46):59–60, Nov 1998.

[22] A. B. Bitsindou and M. G. Kelkar. Gas well production optimization using dynamic nodal analysis. In *SPE Mid-Continent Operations Symposium*, Oklahoma City, Oklahoma, Mar 1999. Society Petroleum Engineers Inc., SPE 52170.

[23] A. Tomasgard, F. Rømo, M. Fodstad, and K. Midthun. Optimization models for the natural gas value chain. Technical report, Norwegian University of Science and Technology, SINTEF, Trondheim, Norway, 2005.

[24] Model documentation: Natural gas transmission and distribution module of the National Energy Modeling System. Technical Report DOE/EIA-M062 , Energy Information Administration, U. S. Department of Energy, Oct 2007.

[25] Documentation of oil and gas supply module (OGSM). Technical Report DOE/EIA-M063 , Energy Information Administration, U. S. Department of Energy, Feb 2007.

[26] Kjetil Trovik Midthun. *Optimization models for liberalized natural gas markets*. PhD thesis, Norwegian University of Science and Technology, October 2007.

[27] T. L. Mason, C. Emelle, J. Van Berkel, A. M. Bagirov, F. Kampas, and J. D. Pintér. Integrated production system optimization using global optimization techniques. *Journal of Industrial and Management Optimization*, 3(2):257–277, May 2007.

[28] A.J. Osiadacz. *Simulation and Analysis of Gas Networks*. E. & F.N. Spon, London, 1987.

[29] Jaroslav Kraálik, Petr Stiegler, Zdeněk Vostrý, and Jiří Závorka. *Dynamic Modeling of Large-scale Networks with Application to Gas Distribution*. Elsevier, New York, 1988.

[30] R.I. Il'kaev, V.E. Seleznev, V.V. Aleshin, and G.S. Klishin. *Numerical Simulation of Gas Pipeline Networks: Theory, Computational Implementation and Industrial Applications.* URSS, Moscow, 2005.

[31] A. Peretti and P. Toth. Optimization of a pipe-line for the natural gas transportation. *European Journal of Operational Research*, 11(3):247–254, 1982.

[32] D. Marqués and M. Morari. On-line optimization of gas pipeline networks. *Automatica*, 24(4):455–469, 1988.

[33] A. J. Osiadacz and D. J. Bell. A simplified algorithm for optimization of large-scale gas networks. *Optimal Control Applications and Methods*, 7(1):95–104, 1986.

[34] B. P. Furey. A sequential quadratic programming-based algorithm for optimization of gas networks. *Automatica*, 29(6):1439–1450, 1993.

[35] D. De Wolf and Y. Smeers. Optimal dimensioning of pipe networks with application to gas transmission networks. *Operations Research*, 44(4):596–608, 1996.

[36] D. De Wolf and Y. Smeers. The gas transmission problem solved by an extension of the simplex algorithm. *Management Science*, 46(11):1454–1465, Nov 2000.

[37] H. J. Dahl, F. Rømo, and A. Tomasgard. Optimisation model for rationing-efficient allocation of capacity in a natural gas transportation network. In *Conference proceedings: International Association for Energy Economics*, Prague, Jun 2003.

[38] Panote Nimmanonda, Varanon Uraikul, Christine W. Chan, and Paitoon Tontiwachwuthikul. Computer-aided simulation model for natural gas pipeline network system operations. *Industrial and Engineering Chemistry Research*, 43: 990–1002, 2004.

[39] R. Z. Ríos-Mercado, S. Wu, L. R. Scott, and E. A. Boyd. A reduction technique for natural gas transmission network optimization problems. *Annals of Operations Research*, 117(1-4):217–234, 2002.

[40] A. Martin, M. Möller, and S. Moritz. Mixed integer models for the stationary case of gas network optimization. *Mathematical Programming*, 105(2-3):563–582, Feb 2006.

[41] R. Z. Ríos-Mercado, S. Kim, and E. A. Boyd. Efficient operation of natural gas transmission systems: A network-based heuristic for cyclic structures. *Computers and Operations Research*, 33(8):2323–2351, 2006.

[42] Diana Cobos-Zaleta and Roger Z. Ríos-Mercado. A MINLP model for a minimizing fuel consumption on natural gas pipeline networks. In *Memorias del XI Congreso Latino Iberoamericano de Investigación de Operaciones (CLAIO)*, Concepción, Chile, Oct 2002.

[43] Roger Z. Ríos-Mercado. Natural gas pipeline optimization. In Panos M. Pardalos and Mauricio G. C. Resende, editors, *Handbook of Applied Optimization*, pages 813–826. Oxford University Press, 2002.

[44] Suming Wu, Roger Z. Ríos-Mercado, E. Andrew Boyd, and L. Ridgway Scott. Model relaxations for the fuel cost minimization of steady-state gas pipeline networks. Technical report, Department of Computer Science, University of Chicago, Chicago, Jan 1999.

[45] Suming Wu, R. Z. Ríos-Mercado, E. A. Boyd, and L. R. Scott. Model relaxations for the fuel cost minimization of steady-state gas pipeline networks. *Mathematical and Computer Modelling*, 31:197–220, 2000.

[46] Roger Z. Ríos-Mercado, Suming Wu, L. Ridgway Scott, and E. Andrew Boyd. Preprocessing on natural gas transmission networks. Technical report, Universidad Autónoma de, Nuevo León, Nov 2000.

[47] Arsegianto, Edy Soewono, and Mochamad Apri. Non-linear optimization model for gas transmission system: A case of Grissik - Duri pipeline. In *SPE Asia Pacific Oil and Gas Conference and Exhibition*, Jakarta, Indonesia, Apr 2003. Society of Petroleum Engineers Inc., SPE 80506.

[48] Helmuth Cremer, Farid Gasmi, and Jean-Jacques Laffont. Access to pipelines in competitive gas markets. *Journal of Regulatory Economics*, 24(1):5–33, Jul 2003.

[49] Rachel A. Davidson, Arthur J. Lembo Jr., June Ma, Linda K. Nozick, and Thomas D. O'Rourke. Optimization of investments in natural gas distribution networks. *Journal Of Energy Engineering*, 132(2):52–60, 2006.

[50] Ioannis S. Papadakis and Paul R. Kleindorfer. Optimizing infrastructure network maintenance when benefits are interdependent. *OR Spectrum*, 27:63–84, 2005.

[51] M. Abbaspour, P. Krishnaswami, and K.S. Chapman. Transient optimization in natural gas compressor stations for linepack operations. *Journal of Energy Resources Technology*, 129:314–324, Dec 2007.

[52] Alireza Kabirian and Mohammad Reza Hemmati. A strategic planning model for natural gas transmission networks. *Energy Policy*, 35:5656–5670, 2007.

[53] Yue Wu, Kin Keung Lai, and Yongjin Liu. Deterministic global optimization approach to steady-state distribution gas pipeline networks. *Optimization and Engineering*, 8:259–275, 2007.

[54] R. A. Wattenbarger. Maximizing seasonal withdrawals from gas storage reservoirs. *Journal of Petroleum Technology*, 22(8):994–998, Aug 1970.

[55] R. P. O'Neill, M. Williard, B. Wilkins, and R. Pike. A mathematical programming model for allocation of natural gas. *Operations Research*, 27(5):857–873, 1979.

[56] R. R. Levary and B. V. Dean. A natural gas flow model under uncertainty in demand. *Operations Research*, 28(6):1360–1374, 1980.

[57] R. E. Brooks. Using generalized networks to forecast natural gas distribution and allocation during periods of shortage. *Mathematical Programming Study*, 15:23–42, May 1981.

[58] A. E. Bopp, V. R. Kannan, S. W. Palocsay, and S. P. Stevens. An optimization model for planning natural gas purchases, transportation, storage and devliverability. *Omega The International Journal of Management Science*, 24(5): 511–522, 1996.

[59] J. Guldmann. Supply, storage and service reliability decisions by gas distribution utilities: a chance constrained approach. *Management Science*, 29(8): 884–906, Aug 1983.

[60] Jean-Michel Guldmann. A marginal-cost pricing model for gas distribution utilities. *Operations Research*, 34(6):851–863, 1986.

[61] W. Avery, G. G. Brown, J. A. Rosenkranz, and R. K. Wood. Optimization of purchase, storage and transmission contracts for natural gas utilities. *Operations Research*, 40(3):446–462, 1992.

[62] Louis Chin and Thomas E. Vollmann. Decision support models for natural gas dispatch. *Transportation Journal*, 32(2):38–45, 1992.

[63] Jean-Michel Guldmann and Fahui Wang. Optimizing the natural gas supply mix of local distribution utilities. *European Journal of Operational Research*, 112:598–612, 1999.

[64] A. Haurie, Y. Smeers, and G. Zaccour. Towards a contract portfolio management model for a gas producing firm. *INFOR*, 30(3):257–273, Aug 1992.

[65] James C. Butler and James S. Dyer. Optimizing natural gas flows with linear programming and scenarios. *Decision Sciences*, 30(2):563–580, 1999.

[66] L. Contesse, J. C. Ferrer, and S. Maturana. A mixed-integer programming model for gas purchase and transportation. *Annals of Operations Research*, 139 (1):39–63, Oct 2005.

[67] S. A. Gabriel, S. Kiet, and J. Zhuang. A mixed complementarity-based equilibrium model of natural gas markets. *Operations Research*, 53(5):799–818, 2006.

[68] Jifang Zhuang and Steven A. Gabriel. A complementarity model for solving stochastic natural gas market equilibria. *Energy Economics*, 30, 2008.

[69] Ruud Egging, Steven A. Gabriel, Franziska Holz, and Jifang Zhuang. A complementarity model for the European natural gas market. *Energy Policy*, 36: 2385–2414, 2008.

[70] Franziska Holz, Christian von Hirschhausen, and Claudia Kemfert. A strategic model of European gas supply (GASMOD). *Energy Economics*, 30:766–788, 2008.

[71] Hanjie Chen and Ross Baldick. Optimizing short-term natural gas supply portfolio for electric utility companies. *IEEE Transactions on Power Systems*, 22 (1):232–239, Feb 2007.

[72] S. Gil and J. Deferrari. Generalized model of prediction of natural gas consumption. *Journal of Energy Resources Technology*, 126:90–98, 2004.

[73] Eugenio Fco, Sánchez-Úbeda, and Ana Berzosa. Modeling and forecasting industrial end-use natural gas consumption. *Energy Economics*, 29:710–742, 2007.

[74] J. D. Huppler. Scheduling gas field production for maximum profit. *Society of Petroleum Engineers Journal*, 14(3):279–294, Jun 1974.

[75] O. Flanigan. Constrained derivatives in natural gas pipeline system optimization. *Journal of Petroleum Technology*, 24(5):549–556, May 1972.

[76] P. M. O'Dell, N. W. Steubing, and J. W. Gray. Optimization of gas field operations. *Journal of Petroleum Technology*, 25(4):419–425, Apr 1973.

[77] J. E. Murray III and T. F. Edgar. Optimal scheduling of production and compression in gas-fields. *Journal of Petroleum Technology*, 30:109–116, Jan 1978.

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

[79] E. M. L. Beale. A mathematical programming model for the long-term development of an offshore field. *Discrete Applied Mathematics*, 5(1):1–9, 1983.

[80] E. M. L. Beale. Optimization methods in oil and gas exploration. *IMA Journal of Applied Mathematics*, 36(1):1–10, 1986.

[81] D. Haugland, Å. Hallefjord, and H. Asheim. Models for petroleum field exploitation. *European Journal of Operational Research*, 37(1):58–72, 1988.

[82] B. Nygreen, M. Christiansen, K. Haugen, T. Bjørkvoll, and Ø. Kristiansen. Modeling Norwegian petroleum production and transportation. *Annals of Operations Research*, 82:251–267, 1998.

[83] S. A. van den Heever, I. E. Grossmann, S. Vasantharajan, and K. Edwards. A Lagrangean decomposition heuristic for the design and planning of offshore hydrocarbon field infrastructures with complex economic objectives. *Industrial and Engineering Chemistry Research*, 40(13):2857–2875, 2001.

[84] W. Gao, P. Hartley, and R. C. Sickles. Optimal dynamic production policy: The case of a large oil field in Saudi Arabia. Technical report, James A. Baker III Institute for Public Policy, Rice University, Feb 2004.

[85] K. R. Ayda-zade and A. G. Bagirov. On the problem of spacing of oil wells and control of their production rates. *Automation and Remote Control*, 67(1): 44–53, 2006.

[86] K. K. Haugen. A stochastic dynamic programming model for scheduling of offshore petroleum fields with resource uncertainty. *European Journal of Operational Research*, 88(1):88–100, 1996.

[87] V. Goel, I. E. Grossmann, A. S. El-Bakry, and E. L. Mulkay. A novel branch and bound algorithm for optimal development of gas fields under uncertainty in reserves. *Computers and Chemical Engineering*, 30(6-7):1076–1092, 2006.

[88] T. W. Jonsbråten. Oil field optimization under price uncertainty. *Journal of the Operational Research Society*, 49(8):1998, Aug 1998.

[89] P. Yusgiantoro and F. S. T. Hsiao. Production-sharing contracts and decision-making in oil production - The case of Indonesia. *Energy Economics*, 15(4): 245–256, Oct 1993.

[90] Robert C. Dawson and J. David Fuller. A mixed integer nonlinear program for oilfiled production planning. *INFOR*, 37(2):121–140, May 1999.

[91] V. D. Kosmidis, J. D. Perkins, and E. N. Pistikopoulos. Optimization of well oil rate allocations in petroleum fields. *Industrial and Engineering Chemistry Research*, 43(14):3513–3527, 2004.

[92] Lyes Khezzar and Abdennnour Seibi. Production system modeling, calibration and optimization in practice. *Petroleum Science and Technology*, 2000.

[93] A. Ortíz-Gómez, V. Rico-Ramirez, and S. Hernández-Castro. Mixed-integer multiperiod model for the planning of oilfield production. *Computers and Chemical Engineering*, 26(4-5):703–714, 2002.

[94] N. V. Queipo, L. E. Zerpa, J. V. Goicochea, A. J. Verde, S. A. Pintos, and A. Zambrano. A model for the integrated optimization of oil production systems. *Engineering with Computers*, pages 130–141, 2003.

[95] V. Barragán-Hernández, R. Vázquez-Román, L. Rosales-Marines, and F. García-Sánchez. A strategy for simulation and optimization of gas and oil production. *Computers and Chemical Engineering*, 30(2):215–227, 2005.

[96] Janie M. Chermak, James Crafton, Suzanne M. Norquist, and Robert H. Patrick. A hybrid economic-engineering model for natural gas production. *Energy Economics*, 21:67–94, 1999.

[97] Elizabeth Olmst Teisberg and Thomas J Teisberg. The value of commodity purchase contracts with limited price risks. *Energy Journal*, 12(3):109–135, 1991.

[98] James L. Paddock, Daniel R. Siegel, and James L. Smith. Option valuation of claims on real assets: The case of offshore petroleum leases. *The Quarterly Journal of Economics*, 103:479–508, August 1988.

[99] James E. Smith and Kevin F. McCardle. Options in the real world: Lessons learned in evaluating oil and gas investments. *Operations Research*, 47(1):1–15, 1999.

[100] F. H. Murphy, R. C. Sanders, S. H. Shaw, and R. L. Thrasher. Modeling natural gas regulatory proposals using the project independence evaluation system. *Operations Research*, 29(5):876–902, 1981.

[101] Zhuliang Chen and Peter A. Forsyth. A semi-Lagrangian approach for natural gas storage valuation and optimal operation. *SIAM Journal on Scientific Computing*, 30(1):339–368, 2007.

[102] Helmuth Cremer and Jean-Jacques Laffont. Competition in gas markets. *European Economic Review*, 46:928–935, 2002.

[103] Ana Quelhas, Esteban Gil, James D. McCalley, and Sarah M. Ryan. A multiperiod generalized network flow model of the U.S. integrated energy system: Part I - Model description. *IEEE Transactions on Power Systems*, 22(2):829–836, May 2007.

[104] Ana Quelhas and James D. McCalley. A multiperiod generalized network flow model of the U.S. integrated energy system: Part II - Simulation results. *IEEE Transactions on Power Systems*, 22(2):837–844, May 2007.

[105] Ertunga C. Özelkan, Alfred D'Ambrosio, and S. Gary Teng. Optimizing liquefied natural gas terminal design for effective supply-chain operations. *International Journal of Production Economics*, 111:529–542, 2008.

[106] Argyris G. Kagiannas, Dimitris Th. Askounis, and John Psarras. Power generation planning: a survey from monopoly to competition. *International Journal of Electrical Power and Energy Systems*, 26(6):413–421, Jul 2004.

[107] Benjamin F. Hobbs. Optimization methods for electric utility resource planning. *European Journal of Operational Research*, 83:1–20, 1995.

[108] D. P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont MA, 1999.

[109] M. S. Bazaraa, H. D. Sherali, and C.M. Shetty. *Nonlinear Programming: Theory and Algorithms*. John Wiley and Sons Inc., second edition, 1993.

[110] I. E. Grossmann. Review of nonlinear mixed-integer and disjunctive programming techniques. *Optimization and Engineering*, 3:227–252, 2002.

[111] F. Glover. Improved linear integer programming formulations of nonlinear integer problems. *Management Science*, 22(4):455–460, Dec 1975.

[112] ISO6976:1995. Natural gas - calculation of calorific values, density, relative density and Wobbe index from composition. Technical report, International Organization for Standardization, 1995.

[113] S. Post. Reasoning with incomplete and uncertain knowledge as an integer linear program. In *Proceeding Avignon 1987: Expert Systems and Their Applications*, volume 2, pages 1361–1377, Avignon, France, May 1987.

[114] R. Raman and I. E. Grossmann. Relation between MILP modelling and logical inference for chemical process synthesis. *Computers and Chemical Engineering*, 15(2):73–84, 1991.

[115] General Algebraic and Modeling System. http://www.gams.com.

[116] M. Tawarmalani and N. V. Sahinidis. *Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming: Theory, Algorithms, Software and Applications*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.

[117] M. Tawarmalani and N. V. Sahinidis. Global optimization of mixed-integer nonlinear programs: A theoretical and computational study. *Mathematical Programming*, 99(3):563–591, Apr 2004.

[118] N. V. Sahinidis and M. Tawarmalani. *BARON 7.5: Global Optimization of Mixed-Integer Nonlinear Programs,* User's Manual, 2005. Available at http://www.gams.com/dd/docs/solvers/baron.pdf.

[119] P. E. Gill, W. Murray, and M. A. Saunders. SNOPT: An SQP algorithm for large-scale constrained optimization. *SIAM Review*, 47(1):99–131, 2005.

[120] CPLEX. http://ilog.com/products/cplex.

[121] B. S. Ahmad and P. I. Barton. Process-wide integration of solvent mixtures. *Computers and Chemical Engineering*, 23(10):1365–1380, 1999.

[122] The R project for statistical computing. http://www.r-project.org/.

[123] N. V. Sahinidis and M. Tawarmalani. *BARON 7.8: Global Optimization of Mixed-Integer Nonlinear Programs,* User's Manual, 2007. Available at http://www.gams.com/dd/docs/solvers/baron.pdf.

[124] Alexander Mitsos, Benoît Chachuat, and Paul I. Barton. McCormick-based relaxation of algorithms. *SIAM Journal on Optimization*, In Press, 2008.

[125] G. P. McCormick. Computability of global solutions to factorable nonconvex programs: Part I. Convex underestimating problems. *Mathematical Programming*, 10:147–175, 1976.

[126] Andreas Griewank. *Evaluating Derivatives: Principles and techniques of algorithmic differentiation*. Frontiers in Applied Mathematics. SIAM, Philadelphia, 2000.

[127] Benoît Chachuat. libMC: A numeric library for McCormick relaxation of factorable functions. http://yoric.mit.edu/libMC, 2008.

[128] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Convex Analysis and Minimization Algorithms I - Fundamentals.* Springer-Verlag, 1996.

[129] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Convex Analysis and Minimization Algorithms II - Advanced Theory and Bundle Methods.* Springer-Verlag, 1996.

[130] Krzysztof C. Kiwiel. *Methods of Descent for Nondifferentiable Optimization.* Number 1133 in Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1985.

[131] Marko M. Mäkelä and Pekka Neittaanmäki. *Nonsmooth Optimization.* World Scientific, Singapore, 1992.

[132] Andrea Walther, Andreas Kowarz, and Andreas Griewank. *ADOL-C: A Package for the Automatic Differentiation of Algorithms Written in C/C++ Version 1.1*, July 2005.

[133] H. S. Ryoo and N. V. Sahinidis. Global optimization of nonconvex NLPs and MINLPs with applications in process design. *Computers and Chemical Engineering*, 19(5):551–566, 1995.

[134] S. Dempe. Directional differentiability of optimal solutions under Slater's conditions. *Mathematical Programming*, 59:49–69, 1993.

[135] Jacques Gauvin. Shadow prices in nonconvex mathematical programming. *Mathematical Programming*, 19:300–312, 1980.

[136] Jacques Gauvin and Robert Janin. Directional behaviour of optimal solutions in nonlinear mathematical programming. *Mathematics of Operations Research*, 13(4):629–649, November 1988.

[137] Alexander Shapiro. Sensitivity analysis of nonlinear programs and differentiability properties of metric projections. *SIAM Journal on Control and Optimization*, 26(3):628–645, 1988.

[138] Adam B. Levy and Boris S. Mordukhovich. Coderivatives in parametric optimization. *Mathematical Programming Series A*, 99:311–327, 2004.

[139] J. Frédéric Bonnans and Alexander Shapiro. *Perturbation Analysis of Optimization Problems.* Springer-Verlag, New York, 2000.

Appendix A

Set Definitions

This section presents the set definitions used in the case study. Figure 3-1 (page 85, Chapter 3) presents a schematic of the infrastructure network as defined by the arc and node set definitions presented here. Figure 4-2 (page 118, Chapter 4) does the same for the PSC network.

<p align="center">Table A.1: Node set definitions (Infrastructure network)</p>

Set	$\subset$	Elements
$\mathcal{N}$	-	F23, F23P, F23SW, F6, E11, SC, E11P, RA, RB, RC, M1, M3, M3P, M4, T, SE, JN, B11, HL, SC1, SC2, SC3, LNG1, LNG2, LNG3, BN, D35, BY
$\mathcal{N}_p$	$\mathcal{N}$	F23P, F23, F23SW, F6, E11P, SC, E11, RA, RB, RC, M1, M3P, T, JN, B11, HL, SC1, SC2, SC3, LNG1, LNG2, LNG3, BN, D35, BY
$\mathcal{N}_s$	$\mathcal{N}$	F23, F23P, F23SW, F6, E11, SC, E11P, M1, M3, M3P, M4, SE, JN, B11, HL, LNG1, LNG2, LNG3, BN, D35, BY
$\mathcal{N}_q$	$\mathcal{N}_s$	F23P, F6, E11P, M1, M3P, JN, B11, HL, LNG1, LNG2, LNG3, BN, D35, BY
$\mathcal{N}_D$	$\mathcal{N}_q$	LNG1, LNG2, LNG3
$\mathcal{N}_J$	$\mathcal{N}_p$	T, RA, RB, RC, SC1, SC2, SC3

<p align="right">Continued on next page</p>

Table A.1: Node set definitions (Infrastructure network)

Set	$\subset$	Elements
$\mathcal{N}_{sc}$	$\mathcal{N}_J$	SC1, SC2, SC3
$\mathcal{F}$	$\mathcal{N}_s$	F23, F23SW, F6, E11, SC, M1, M3, M4, SE, JN, B11, HL, BN, D35, BY
$\mathcal{F}_w$	$\mathcal{F}$	F23, F23SW, F6, E11, SC, M1, M3, M4, SE, JN, B11
$\mathcal{F}_{w,nc}$	$\mathcal{F}_w$	F23, F23SW, SC, E11, JN
$\mathcal{F}_{pr}$	$\mathcal{F}$	B11, F6, E11, M1, M4
$\mathcal{N}_{wp}$	$\mathcal{N}_s$	F23P, F6, E11P, M1, M3P, JN, B11, HL, BN, D35, BY
$\mathcal{N}_{wp,m}$	$\mathcal{N}_{wp}$	F23P, E11P, M3P
$\mathcal{N}_{wp,c}$	$\mathcal{N}_{wp}$	F6, M1, M3P, B11
$\mathcal{N}_x$	$\mathcal{N}$	RA, RB, RC, SC2, M1
$\mathcal{N}_{x,J}$	$\mathcal{N}_J, \mathcal{N}_x$	RA, RB, RC, SC2
$\mathcal{N}_{x,q}$	$\mathcal{N}_q, \mathcal{N}_x$	M1
$\mathcal{N}_{D,HSP}$	$\mathcal{N}_D$	LNG1, LNG2
$\mathcal{N}_{D,HSM}$	$\mathcal{N}_D$	LNG3

Table A.2: Platforms serving multiple fields

$i \in \mathcal{N}_{wp,m}$	$\mathcal{F}_i$
F23P	F23SW, F23
E11P	E11, SC
M3P	M3, M4, SE

Table A.3: Arc set definitions (Infrastructure network)

Set	⊂	Elements
$\mathcal{A}$	-	(F6,RA), (F23SW,F23P), (F23,F23P), (F23P,RA), (SC,E11P), (E11,E11P), (E11P,RA), (BN,RA), (D35,BY), (BY,LNG1), (RA,RB), (RB,RC), (M4,M3P), (SE,M3P), (M3,M3P), (M3P,T), (M1,T), (T,RB), (M1,RC), (JN,M1), (HL,B11), (B11,RB), TL1[1], TL2[1], TL3[2], TL4[2], TL5[3], TL6[3], (SC2,LNG3), (SC1,LNG1), (SC2,LNG2), (SC3,LNG3)
$\mathcal{A}_q$	$\mathcal{A}$	(F6,RA), (F23P,RA), (E11P,RA), (BN,RA), (D35,BY), (BY,LNG1), (RA,RB), (RB,RC), (M3P,T), (M1,T), (T,RB), (M1,RC), (JN,M1), (HL,B11), (B11,RB), TL1, TL2, TL3, TL4, TL5, TL6, (SC2,LNG3), (SC1,LNG1), (SC2,LNG2), (SC3,LNG3)
$\mathcal{A}_i$	$\mathcal{A}$	(F6,RA), (F23SW,F23P), (F23,F23P), (F23P,RA), (SC,E11P), (E11,E11P), (E11P,RA), (BN,RA), (D35,BY), (BY,LNG1), (M3P,T), (M1,T), (T,RB), (M1,RC), (JN,M1), (HL,B11), (B11,RB), TL1, TL3, TL6, (SC2,LNG3), (SC1,LNG1), (SC2,LNG2), (SC3,LNG3)
$\mathcal{A}_p$	$\mathcal{A}_q$	(F6,RA), (F23P,RA), (BN,RA), (D35,BY), (BY,LNG1), (M3P,T), (M1,T), (T,RB), (M1,RC), (JN,M1), (HL,B11), (B11,RB), TL1, TL2, TL3, TL4, TL5, TL6
$\mathcal{A}_y$	$\mathcal{A}_q$	(RA,RB), (RB,RC), (SC2,LNG3)
$\mathcal{A}_{sc}$	$\mathcal{A}_q$	(SC1,LNG1), (SC2,LNG2), (SC3,LNG3)
$\mathcal{A}_x$	$\mathcal{A}_q$	(M1,T), TL1, TL2, TL3, TL4, TL5, (SC2,LNG3)

---

[1] (RA, SC1)
[2] (RB, SC2)
[3] (RC, SC3)

271

Table A.4: Node set definitions (PSC network)

Set	$\subset$	Elements
$\mathcal{L}$	-	A_s, A_0, A_1, A_2, A_3, B_s, F_s, B_0, B_1, B_2, B_3, C_s, C_0, C_1, C_2, C_3, D_s, D_0, D_1, D_2, D_3, A_d, B_d, CD_d
$\mathcal{L}_s$	$\mathcal{L}$	A_s, F_s, B_s, C_s, D_s, A_d, B_d, CD_d
$\mathcal{L}_d$	$\mathcal{L}_s$	A_d, B_d, CD_d
$\mathcal{L}_l$	$\mathcal{L}$	A_0, A_1, A_2, A_3, B_0, B_1, B_2, B_3, C_0, C_1, C_2, C_3, D_0, D_1, D_2, D_3

Table A.5: Arc set definitions (PSC network)

Set	$\subset$	Elements
$\mathcal{E}$	-	(A_s,A_0), (A_0,A_d), (A_0,A_1), (A_1,A_2), (A_2,A_3), (B_s,B_0), (B_0,B_d), (B_0,B_1), (B_1,B_2), (B_2,B_3), (C_s,C_0), (C_0,CD_d), (C_0,C_1), (C_1,C_2), (C_2,C_3), (D_s,D_0), (D_0,CD_d), (D_0,D_1), (D_1,D_2), (D_2,D_3), (A_1,B_2), (A_2,C_3), (A_3,D_3), (B_2,C_2), (B_3,D_2), (C_2,B_3), (C_1,D_1), (D_1,C_1), (F_s, A_1), (F_s,B_1)
$\mathcal{E}_l$	$\mathcal{E}$	(A_0,A_1), (A_1,A_2), (A_2,A_3), (B_0,B_1), (B_1,B_2), (B_2,B_3), (C_0,C_1), (C_1,C_2), (C_2,C_3), (D_0,D_1), (D_1,D_2), (D_2,D_3)
$\mathcal{E}_{l,a}$	$\mathcal{E}_l$	(A_0,A_1), (B_0,B_1), (B_1,B_2), (C_0,C_1), (D_0,D_1)
$\mathcal{E}_{l,S}$	$\mathcal{E}_l$	(B_2,B_3), (D_1,D_2), (C_2,C_3)
$\mathcal{E}_t$	$\mathcal{E}$	(A_1,B_2), (A_2,C_3), (A_3,D_3), (B_2,C_2), (B_3,D_2), (C_2,B_3), (C_1,D_1), (D_1,C_1), (F_s,A_1), (F_s,B_1)
$\mathcal{E}_s$	$\mathcal{E}$	(A_s,A_0), (A_0,A_d), (B_s,B_0), (B_0,B_d) (C_s,C_0), (C_0,CD_d), (D_s,D_0), (D_0,CD_d)

Appendix B

Model Parameters

As described in Chapter 5 on the case study, the values of the model parameters in this section are not related to the actual SGPS model parameters to preserve business sensitive information.

Table B.1: Well-performance model parameters

$\mathcal{W}$	$\pi_{r,w}$	α_w	β_w	λ_w	ϑ_w	σ_w	ς_w
	bar	bar^2.d/hm^3	bar^2.d^2/hm^6		bar^2.d^2/hm^6	m^3/hm^3	m^3/hm^3
B11A	75.39	2.163×10^{-2}	5.616×10^{-4}	3.534	$7.285 \times 10^{+2}$	49.50	16.61
B11B	78.46	2.287×10^{-2}	5.605×10^{-4}	3.204	$6.787 \times 10^{+2}$	47.04	17.18
B11C	78.62	2.266×10^{-2}	5.206×10^{-4}	3.628	$6.668 \times 10^{+2}$	49.79	18.22
B11D	71.45	2.045×10^{-2}	5.190×10^{-4}	3.568	$7.987 \times 10^{+2}$	45.65	17.27
E11A	54.74	6.160×10^{-1}	9.266×10^{-5}	2.555	$3.410 \times 10^{+3}$	13.64	13.52
E11B	56.68	6.630×10^{-1}	9.918×10^{-5}	2.78	$3.558 \times 10^{+3}$	14.94	14.36
E11C	60.05	6.074×10^{-1}	9.082×10^{-5}	2.706	$3.700 \times 10^{+3}$	13.93	14.52
E11D	52.69	6.701×10^{-1}	9.837×10^{-5}	2.595	$3.340 \times 10^{+3}$	14.51	14.40
E11E	56.15	6.241×10^{-1}	8.899×10^{-5}	2.433	$3.320 \times 10^{+3}$	12.54	12.70
E11F	56.50	5.960×10^{-1}	9.096×10^{-5}	2.525	$3.255 \times 10^{+3}$	14.40	12.21
E11G	49.89	5.831×10^{-1}	9.180×10^{-5}	2.603	$3.661 \times 10^{+3}$	13.20	13.41
E11H	59.71	6.503×10^{-1}	9.728×10^{-5}	2.648	$3.567 \times 10^{+3}$	13.32	12.80
E11I	56.77	6.153×10^{-1}	8.749×10^{-5}	2.558	$3.224 \times 10^{+3}$	14.45	14.37
E11J	54.47	6.661×10^{-1}	8.810×10^{-5}	2.362	$3.745 \times 10^{+3}$	14.11	12.51
F23A	247.67	$1.591 \times 10^{+0}$	2.356×10^{-6}	1.576	$4.962 \times 10^{+2}$	104.55	2.93
F23B	231.76	$1.658 \times 10^{+0}$	2.170×10^{-6}	1.455	$4.653 \times 10^{+2}$	108.73	2.80
F23C	266.73	$1.603 \times 10^{+0}$	2.525×10^{-6}	1.488	$5.112 \times 10^{+2}$	109.70	2.74
F23D	227.98	$1.720 \times 10^{+0}$	2.162×10^{-6}	1.444	$4.854 \times 10^{+2}$	114.70	3.08

Table B.1: Well-performance model parameters

$\mathcal{W}$	$\pi_{r,w}$	α_w	β_w	λ_w	ϑ_w	σ_w	ς_w
	bar	bar^2.d/hm^3	bar^2.d^2/hm^6		bar^2.d^2/hm^6	m^3/hm^3	m^3/hm^3
F23E	244.14	$1.521\times10^{+0}$	2.160×10^{-6}	1.568	$4.607\times10^{+2}$	114.78	2.87
F23F	245.95	$1.522\times10^{+0}$	2.545×10^{-6}	1.712	$5.236\times10^{+2}$	102.39	3.02
F23G	265.48	$1.540\times10^{+0}$	2.582×10^{-6}	1.493	$4.953\times10^{+2}$	98.16	2.77
F23H	226.21	$1.735\times10^{+0}$	2.189×10^{-6}	1.555	$4.971\times10^{+2}$	99.82	3.00
F23I	271.81	$1.699\times10^{+0}$	2.136×10^{-6}	1.432	$4.485\times10^{+2}$	104.40	3.16
F23J	252.13	$1.703\times10^{+0}$	2.503×10^{-6}	1.428	$4.888\times10^{+2}$	107.87	2.70
F23K	235.67	$1.734\times10^{+0}$	2.559×10^{-6}	1.502	$4.981\times10^{+2}$	97.33	3.05
F23L	266.15	$1.543\times10^{+0}$	2.278×10^{-6}	1.701	$4.536\times10^{+2}$	103.03	2.92
F23M	252.16	$1.515\times10^{+0}$	2.234×10^{-6}	1.441	$4.915\times10^{+2}$	97.98	2.83
F23N	264.42	$1.633\times10^{+0}$	2.147×10^{-6}	1.723	$4.833\times10^{+2}$	114.02	2.84
F23SW	239.26	$1.645\times10^{+0}$	2.358×10^{-6}	1.688	$4.895\times10^{+2}$	105.92	3.06
F6A	44.81	3.673×10^{-2}	8.976×10^{-4}	1.613	$6.049\times10^{+2}$	145.16	12.42
F6B	47.11	3.397×10^{-2}	8.247×10^{-4}	1.527	$5.993\times10^{+2}$	137.00	11.50
F6C	41.81	3.781×10^{-2}	9.758×10^{-4}	1.553	$5.708\times10^{+2}$	136.05	13.44
F6D	44.95	3.344×10^{-2}	9.623×10^{-4}	1.654	$6.066\times10^{+2}$	138.11	13.29
F6E	40.47	3.578×10^{-2}	9.846×10^{-4}	1.558	$5.835\times10^{+2}$	138.56	12.61
F6F	40.83	4.016×10^{-2}	8.238×10^{-4}	1.691	$6.467\times10^{+2}$	141.61	13.41
F6G	43.24	3.770×10^{-2}	8.531×10^{-4}	1.622	$5.926\times10^{+2}$	158.50	12.56
F6H	47.84	3.612×10^{-2}	8.448×10^{-4}	1.496	$5.553\times10^{+2}$	144.54	12.88
F6I	46.02	3.444×10^{-2}	9.426×10^{-4}	1.486	$6.637\times10^{+2}$	145.87	13.00
F6J	47.91	3.644×10^{-2}	8.938×10^{-4}	1.627	$6.399\times10^{+2}$	147.88	12.32
F6K	44.49	3.742×10^{-2}	9.648×10^{-4}	1.578	$6.195\times10^{+2}$	142.81	13.55
F6L	44.19	3.390×10^{-2}	9.100×10^{-4}	1.703	$5.889\times10^{+2}$	135.04	13.11
F6M	45.43	3.568×10^{-2}	8.411×10^{-4}	1.676	$5.666\times10^{+2}$	140.55	12.14
JNA	142.21	1.258×10^{-1}	1.504×10^{-7}	0.724	$7.807\times10^{+2}$	284.16	2.25
JNB	150.50	1.356×10^{-1}	1.468×10^{-7}	0.696	$7.518\times10^{+2}$	274.25	2.13
JNC	145.96	1.378×10^{-1}	1.520×10^{-7}	0.684	$7.853\times10^{+2}$	299.10	2.38
JND	129.31	1.303×10^{-1}	1.375×10^{-7}	0.729	$7.548\times10^{+2}$	278.09	2.16
JNE	152.84	1.292×10^{-1}	1.422×10^{-7}	0.773	$7.329\times10^{+2}$	280.22	2.30
M1A	113.51	$8.472\times10^{+0}$	2.545×10^{-3}	1.868	$1.580\times10^{+2}$	361.62	2.97
M1B	113.81	$8.484\times10^{+0}$	2.651×10^{-3}	1.974	$1.476\times10^{+2}$	358.52	2.79
M1C	102.24	$8.608\times10^{+0}$	2.691×10^{-3}	2.013	$1.729\times10^{+2}$	370.98	2.74
M1D	121.17	$8.948\times10^{+0}$	2.659×10^{-3}	1.761	$1.727\times10^{+2}$	355.33	2.89
M1E	109.27	$8.026\times10^{+0}$	2.358×10^{-3}	1.942	$1.483\times10^{+2}$	370.83	3.09
M1F	107.60	$7.784\times10^{+0}$	2.509×10^{-3}	1.78	$1.598\times10^{+2}$	379.66	3.20
M1G	111.87	$8.510\times10^{+0}$	2.498×10^{-3}	1.695	$1.617\times10^{+2}$	391.33	2.72

Table B.1: Well-performance model parameters

$\mathcal{W}$	$\pi_{r,w}$	α_w	β_w	λ_w	ϑ_w	σ_w	ς_w
	bar	$bar^2.d/hm^3$	$bar^2.d^2/hm^6$		$bar^2.d^2/hm^6$	m^3/hm^3	m^3/hm^3
M1H	105.22	$8.074 \times 10^{+0}$	2.325×10^{-3}	1.926	$1.566 \times 10^{+2}$	345.96	2.99
M3A	69.33	1.888×10^{-1}	3.445×10^{-4}	1.482	$1.090 \times 10^{+3}$	57.75	8.67
M3B	82.08	1.662×10^{-1}	3.691×10^{-4}	1.621	$1.215 \times 10^{+3}$	61.18	9.08
M3C	78.36	1.816×10^{-1}	3.481×10^{-4}	1.539	$1.048 \times 10^{+3}$	58.32	8.21
M3D	73.13	1.657×10^{-1}	3.695×10^{-4}	1.53	$1.076 \times 10^{+3}$	62.13	9.18
M3E	79.77	1.627×10^{-1}	3.159×10^{-4}	1.642	$1.209 \times 10^{+3}$	61.00	8.39
M3F	80.46	1.818×10^{-1}	3.815×10^{-4}	1.446	$1.160 \times 10^{+3}$	61.50	8.40
M3G	82.30	1.825×10^{-1}	3.513×10^{-4}	1.663	$1.056 \times 10^{+3}$	67.66	9.56
M3H	76.71	1.749×10^{-1}	3.555×10^{-4}	1.617	$1.193 \times 10^{+3}$	60.95	9.11
M3I	79.49	1.824×10^{-1}	3.375×10^{-4}	1.621	$1.033 \times 10^{+3}$	58.05	9.71
M3J	72.10	1.577×10^{-1}	3.502×10^{-4}	1.673	$1.148 \times 10^{+3}$	63.25	8.57
M4A	75.69	1.724×10^{-1}	3.474×10^{-4}	1.573	$1.143 \times 10^{+3}$	62.21	8.98
M4B	81.38	1.753×10^{-1}	3.316×10^{-4}	1.484	$1.219 \times 10^{+3}$	64.58	8.28
SCA	142.30	9.924×10^{-1}	6.815×10^{-3}	3.638	$2.285 \times 10^{+3}$	162.82	12.54
SCB	146.68	$1.054 \times 10^{+0}$	7.021×10^{-3}	3.577	$2.345 \times 10^{+3}$	161.38	11.39
SEA	153.40	$1.050 \times 10^{+0}$	4.298×10^{-3}	3.515	$3.853 \times 10^{+2}$	401.44	10.00
SEB	141.22	$1.154 \times 10^{+0}$	4.566×10^{-3}	3.813	$3.845 \times 10^{+2}$	403.21	9.77

Table B.2: Gas compositions (mole percent)

Field	CO_2	N_2	H_2S	C_1	C_2	C_3	C_4	C_5+
E11	9.2341	2.8907	0.0015	73.2353	8.8970	3.0058	0.6258	2.1098
F23	1.6427	1.0500	0.0004	89.1064	2.8060	3.6792	0.9170	0.7983
F6	3.4121	1.2627	0.0038	79.9876	7.6318	4.8226	1.0002	1.8793
M1	5.0408	0.4280	0.0033	81.0281	5.4633	3.5071	2.7518	1.7776
M3	0.9488	0.4465	0.0036	76.2553	7.3721	6.9870	1.1547	6.8320
M4	2.3048	0.2579	0.0048	82.2489	7.2965	3.6886	3.1960	1.0025
B11	8.8511	1.2626	0.0520	80.5107	5.3962	0.5317	0.6486	2.7471
SC	0.2687	0.6999	0.0000	96.1775	1.7790	0.4955	0.1158	0.4635
F23SW	0.6840	0.4068	0.0010	91.4870	4.4357	2.1390	0.4961	0.3504
JN	2.6347	0.1439	0.0003	88.7193	2.6646	3.7095	1.2948	0.8329
SE	2.4263	0.1808	0.0006	87.6063	3.7230	1.4481	2.3271	2.2879
HL	1.5894	0.9714	0.0000	70.0081	5.6934	5.4388	6.5830	9.7160
BY	0.8782	0.2160	0.0000	91.3967	4.5313	1.1332	0.3970	1.4476
D35	0.7177	0.5513	0.0000	83.8315	7.3682	4.6499	0.7337	2.1477
BN	1.4483	0.5193	0.0000	81.0205	4.4047	6.6060	4.2640	1.7372

Table B.3: Trunkline parameters

$(i,j) \in \mathcal{A}_p$	$\kappa_{(i,j)}$
	$\mathrm{bar^2.d^2/hm^6}$
(RA,SC1) TL1	2.46
(RA,SC1) TL2	5.06
(RB,SC2) TL3	6.10
(RB,SC2) TL4	5.43
(RC,SC3) TL5	7.65
(RC,SC3) TL6	5.11
(F6,RA)	5.33
(F23P,RA)	4.40
(B11,RB)	12.78
(HL,B11)	35.58
(D35,BY)	3062.62
(BN,RA)	97.51
(BY,SC1D)	254.77
(M3P,T)	0.39
(M1,T)	1.17
(T,RB)	2.59
(JN,M1)	3.17
(M1,RC)	21.84

Table B.4: Demand rate bounds

$i \in \mathcal{N}_D$	$\Lambda_{d,i}^L$	$\Lambda_{d,i}^U$
	MMscfd	MMscfd
SC1D	700	1100
SC2D	600	1300
SC3D	800	1600

Table B.5: Maximum reservoir pressure

$i \in \mathcal{F}_w$	$\pi_{r,i}^M$
	bar
E11	187
B11	83
F23	273
F23SW	273
JN	157
M4	84
M3	84
SC	165
F6	50
M1	125
SE	169

Table B.6: Compression power bounds

$i \in \mathcal{N}_{wp,c}$	$\Psi_{L,i}$	$\Psi_{U,i}$
	MW	MW
F6	0.01	22.0
B11	0.01	27.0
M3P	0.01	27.0
M1	0.01	20.0

Table B.7: Slugcatcher pressure bounds

$i \in \mathcal{N}_{sc}$	$\pi_{d,i}^L$	$\pi_{d,i}^U$
	bar	bar
SC1	60	70
SC2	60	70
SC3	60	70

Table B.8: Miscellaneous parameter values

Parameter	Value	Unit	Remarks
π_{sc}	1.013	bar	-
θ_{sc}	288.15	K	-
$\Delta\pi_{(i,j)}$	5.0	bar	$\forall (i,j) \in \mathcal{A}_{sc}$
η_i	0.75	-	$\forall i \in \mathcal{N}_{wp,c}$
$\theta_{m,i}$	315	K	$\forall i \in \mathcal{N}_{wp,c}$
ζ	1.5	-	
ϕ	42.2845	Mmole/hm^3	-
ϱ_g	0.0283168	hm^3/MMscfd	-
ϱ_L	0.158987	m^3/barrel	-

Table B.9: Heating values and Molecular weights

$k \in \mathcal{S}$	γ_k [a]	μ_k
	MJ/kg	kg/mole
CO_2	-	44.010×10^{-3}
N_2	-	28.020×10^{-3}
H_2S	-	34.082×10^{-3}
C_1	55.574	16.043×10^{-3}
C_2	51.95	30.070×10^{-3}
C_3	50.37	44.097×10^{-3}
C_4	49.47	58.123×10^{-3}
C_{5+}	48.72	86.177×10^{-3}

---

[a]from [112], only for $k \in \mathcal{S}_h$

Appendix C

Bounds

As pointed out in Section 5.2 (page 144) in Chapter 5, the variable bounds in the SGPS model have been obtained from design capacities and historical operating data that cannot be disclosed. These values therefore have been changed from the values used in the SGPS model. The bounds can be divided into two categories:

1. Derived Bounds: these bounds are derived from other variables or parameters.

2. Variable specific bounds: these bounds are set individually.

C.1   Derived Bounds

C.1.1   Bounds: Infrastructure Model

The derived bounds are as follows:

1. The bounds on the rates at the LNG plants are as given in Section 4.1 (page 108).
$$-\Lambda_{d,i}^{U} \leq Q_{s,i} \leq -\Lambda_{d,i}^{L}, \quad \forall i \in \mathcal{N}_D.$$

2. The bounds on the pressures at the slugcatchers are also given as per Section 4.1 (page 108).
$$\pi_{d,i}^{L} \leq P_i \leq \pi_{d,i}^{U}, \quad \forall i \in \mathcal{N}_{sc}.$$

3. The molar rate bounds are set using the standard volumetric rate bounds. For nodes with production term

$$\chi_{i,k}\,\phi\,Q_{s,i}^{L} \leq F_{s,i,k} \leq \chi_{i,k}\,\phi\,Q_{s,i}^{U}, \quad \forall(i,k) \in \mathcal{F} \times \mathcal{S},$$

$$\phi\,Q_{s,i}^{L} \leq F_{s,i,k} \leq 0, \qquad\qquad \forall(i,k) \in \mathcal{N}_D \times \mathcal{S},$$

$$0 \leq F_{s,i,k} \leq \phi\,Q_{s,i}^{U}, \quad \forall(i,k) \in \mathcal{N}_{wp,m}\times \in \mathcal{S}.$$

The relationship between standard volumetric flowrates and molar flowrates in arcs in given by

$$0 \;\leq\; F_{a,(i,j),k} \;\leq\; \phi\,Q_{a,(i,j)}^{U}, \quad \forall((i,j),k) \in \mathcal{A}_q \times \mathcal{S}.$$

4. For the variables at wells the following bounds are used

$$0 \leq Q_{Lw,w} \leq \sigma_w Q_{w,w}^{U}, \quad \forall w \in \mathcal{W},$$

where $Q_{w,w}^{U}$ is calculated as in the Section 5.2.1 (page 145). The pressure bounds at wells are given by

$$\pi_{atm} \leq P_{b,w} \leq \pi_{r,i}^{M}, \quad \forall w \in \mathcal{W}_i, \quad i \in \mathcal{F}_w,$$

$$\pi_{atm} \leq P_{t,w} \leq \pi_{r,i}^{M}, \quad \forall w \in \mathcal{W}_i, \quad i \in \mathcal{F}_w.$$

5. Fields for which well performances is modeled:

$$0 \leq Q_{s,i} \leq \sum_{w \in \mathcal{W}_i} Q_{w,w}^{U}, \quad \forall i \in \mathcal{F}_w,$$

$$0 \leq Q_{Ls,i} \leq \sum_{w \in \mathcal{W}_i} Q_{Lw,w}^{U}, \quad \forall i \in \mathcal{F}_w,$$

6. For the well platforms serving various fields,

$$0 \le Q_{s,i} \le \sum_{j \in \mathcal{F}_i} Q_{s,j}^U, \quad \forall i \in \mathcal{N}_{wp,m}.$$

7. The split fractions should lie between 0 and 1

$$0 \le s_{(i,j)} \le 1, \quad \forall (i,j) \in \mathcal{A}_x.$$

8. The bounds on the compression variables are given as per Section 3.5 (page 95).

$$\Psi_{L,i} \le W_i \le \Psi_{U,i}, \quad \forall i \in \mathcal{N}_{wp,c}.$$

9. The bounds on the dummy variables for reformulation of constraints in set $\mathcal{A}_y$

$$0 \le w_{u,(i,j)} \le P_i^U, \quad \forall (i,j) \in \mathcal{A}_y,$$
$$0 \le w_{d,(i,j)} \le P_j^U, \quad \forall (i,j) \in \mathcal{A}_y.$$

C.1.2 Bounds: Production-sharing Contracts (PSC) Model

The following is a discussion of the bounds for the PSC model:

1. The supply rates for PSC are bounded above by the production rate of the corresponding fields

$$0 \le q_{c,i} \le \sum_{j \in \mathcal{F}_i} Q_{s,j}^U, \quad \forall i \in \mathcal{C}^S.$$

2. The source and sink production rates in the PSC network are set using the supply and demand rate bounds

$$0 \le q_{s,p_s} \le \sum_{j \in \mathcal{C}_i^s} q_{c,j}^U, \quad \forall p \in \mathcal{C},$$
$$Q_{s,i}^L \le q_{s,u_i} \le Q_{s,i}^U, \quad \forall i \in \mathcal{N}_D.$$

where $u_i \in \mathcal{L}_s$ is the demand node in the PSC network corresponding to the demand node $i \in \mathcal{N}_D$.

3. The source and sink production rate bounds are used to set arc flowrate bounds in the demand and supply arcs

$$q^L_{s,p_s} \leq q_{a,(p_s,p_0)} \leq q^U_{s,p_s}, \quad \forall p \in \mathcal{C},$$

$$-q^U_{s,p_d} \leq q_{a,(p_0,p_d)} \leq -q^L_{s,p_d}, \quad \forall p \in \{\text{A,B}\},$$

$$0 \leq q_{a,(p_0,\text{CD}_d)} \leq -q^L_{s,\text{CD}_d}, \quad \forall p \in \{\text{C,D}\}.$$

4. The bounds on the flowrates in arcs representing levels are set using the supply and demand rates of that production-sharing contract:

$$e^U_{(p_i,p_{i+1})} = q^U_{a,(p_s,p_0)}, \quad \forall(p_i,p_{i+1}) \in \mathcal{E}_l\backslash\{(\text{B}_1,\text{B}_2),(\text{B}_2,\text{B}_3)\},$$

$$e^L_{(p_i,p_{i+1})} = -q^U_{a,(p_0,p_d)}, \quad \forall(p_i,p_{i+1}) \in \mathcal{E}_l.$$

However the upper bound is not applicable for (B_1,B_2) and (B_2,B_3) as there is a transfer arc terminating at node B_1

$$e^U_{(p_i,p_{i+1})} = q^U_{a,(\text{B}_s,\text{B}_0)} + t^U_{(\text{F}_s,\text{B}_1)}, \quad \forall(p_i,p_{i+1}) \in \{(\text{B}_1,\text{B}_2),(\text{B}_2,\text{B}_3)\}.$$

5. The transfer rate bounds are set using the upstream level arc flowrate bounds from the receiving production-sharing contract:

$$t^L_{(p_i,q_j)} = 0, \quad \forall(p_i,q_j) \in \mathcal{E}_t,$$

$$t^U_{(p_i,q_j)} = e^U_{(p_{i-1},p_i)}, \quad \forall(p_i,q_j) \in \mathcal{E}_t, \; i \neq s,$$

$$t^U_{(p_i,q_j)} = q^U_{s,p_s}, \quad \forall(p_i,q_j) \in \mathcal{E}_t, \; i = s.$$

282

C.2 Variable Specific Bounds

Table C.1: Trunkline flow bounds

Trunkline	$Q_{a,(i,j)}^L$	$Q_{a,(i,j)}^U$
	MMscfd	MMscfd
(F6,RA)	0	1000
(F23P,RA)	0	950
(RA,RB)	0	950
(RB,RC)	0	720
(T,RB)	0	1700
(M1,RC)	0	900
(HL,B11)	0	500
(B11,RB)	0	500
TL1	0	$-Q_{s,\mathrm{SC1D}}^L$
TL2	0	$-Q_{s,\mathrm{SC1D}}^L$
TL3	0	850
TL4	0	850
TL5	0	900
TL6	0	900
(SC2,SC3D)	0	800
(BN,RA)	$Q_{s,\mathrm{BN}}^L$	$Q_{s,\mathrm{BN}}^U$
(D35,BY)	$Q_{s,\mathrm{D35}}^L$	$Q_{s,\mathrm{D35}}^U$
(BY,SC1D)	$Q_{s,\mathrm{BY}}^L$	$Q_{s,\mathrm{BY}}^U + Q_{s,\mathrm{D35}}^U$
(M3P,T)	0	$Q_{s,\mathrm{M3}}^U + Q_{s,\mathrm{M4}}^U + Q_{s,\mathrm{SE}}^U$
(M1,T)	$\varrho_g 500$	$Q_{a,(\mathrm{T,RB})}^U$
(JN,M1)	0	$Q_{s,\mathrm{JN}}^U$
(E11P,RA)	$Q_{s,\mathrm{E11P}}^L$	$Q_{s,\mathrm{E11P}}^U$
(SC1,SC1D)	$Q_{a,(\mathrm{TL1})}^L$	$Q_{a,(\mathrm{TL1})}^U + Q_{a,(\mathrm{TL2})}^U$
(SC2,SC2D)	$Q_{a,(\mathrm{TL3})}^L$	$Q_{a,(\mathrm{TL3})}^U + Q_{a,(\mathrm{TL4})}^U$
(SC3,SC3D)	$Q_{a,(\mathrm{TL5})}^L$	$Q_{a,(\mathrm{TL5})}^U + Q_{a,(\mathrm{TL6})}^U$

Table C.2: Pressure bounds

Node	P_i^L	P_i^U
	bar	bar
SC1D	50	70
SC2D	50	70
SC3D	50	70
F23	π_{atm}	$\pi_{r,\mathrm{F23}}^M$
F23SW	π_{atm}	$\pi_{r,\mathrm{F23SW}}^M$
E11	π_{atm}	$\pi_{r,\mathrm{E11}}^M$
SC	π_{atm}	$\pi_{r,\mathrm{SC}}^M$
M1	π_{atm}	150
JN	π_{atm}	$\pi_{r,\mathrm{JN}}^M$
B11	π_{atm}	150
F23P	π_{atm}	$\pi_{r,\mathrm{F23}}^M$
E11P	π_{atm}	$\pi_{r,\mathrm{E11}}^M$
F6	π_{atm}	150
T	π_{atm}	150
M3P	π_{atm}	150
HL	π_{atm}	150
BN	π_{atm}	120
D35	π_{atm}	100
BY	π_{atm}	95
RA	π_{atm}	110
RB	π_{atm}	110
RC	π_{atm}	110

Table C.3: Compression inlet pressure bounds

Platform	$P_{c,i}^L$	$P_{c,i}^U$
	bar	bar
B11	20	$\pi_{r,\mathrm{B11}}^M$
F6	π_{atm}	$\pi_{r,\mathrm{F6}}^M$
M3P	30	$\pi_{r,\mathrm{SE}}^M$
M1	π_{atm}	$\pi_{r,\mathrm{M1}}^M$

Table C.4: Field rate bounds (no well performance)

Fields	$Q_{s,i}^L$	$Q_{s,i}^U$
	MMscfd	MMscfd
HL	0	600
BN	50	155
D35	40	130
BY	40	125

Appendix D

Base Case: Optimal Solution

The base case infrastructure solution is graphically shown in Figure 5-1 (page 149) and the base case PSC solution is shown in Figure 5-2 (page 150).

Total Dry Gas Production : 94.38 hm^3/d.
Total Priority Production : 30.37 hm^3/d.
Total NGL Production   : 21440 m^3/d.

D.1   Trunkline Network

(RA,RB) is open.
(RB,RC) is open.
(SC2,LNG3) is open.

M1 production is greater than 500 MMscfd.
All of JN production is being diverted into (M1,RC).

LNG1

Total delivery rate : -29.45 hm^3/d

Delivery pressure   : 62.6 bar

$F_{s,\text{LNG1},CO_2}$: -1.803×10^{+1} Mmol/d

$F_{s,\text{LNG1},H_2S}$: -9.869×10^{-3} Mmol/d

$F_{s,\text{LNG1},N_2}$: -9.792×10^{+0} Mmol/d

$F_{s,\text{LNG1},C_1}$: -1.097×10^{+3} Mmol/d

$F_{s,\text{LNG1},C_2}$: -5.081×10^{+1} Mmol/d

$F_{s,\text{LNG1},C_3}$: -4.356×10^{+1} Mmol/d

$F_{s,\text{LNG1},C_4}$: -1.380×10^{+1} Mmol/d

$F_{s,\text{LNG1},C_5+}$: -1.227×10^{+1} Mmol/d

LNG2

Total delivery rate : -24.62 hm^3/d

Delivery pressure   : 55 bar

$F_{s,\text{LNG2},CO_2}$: -3.073×10^{+1} Mmol/d

$F_{s,\text{LNG2},H_2S}$: -1.689×10^{-2} Mmol/d

$F_{s,\text{LNG2},N_2}$: -4.215×10^{+0} Mmol/d

$F_{s,\text{LNG2},C_1}$: -8.890×10^{+2} Mmol/d

$F_{s,\text{LNG2},C_2}$: -4.514×10^{+1} Mmol/d

$F_{s,\text{LNG2},C_3}$: -3.226×10^{+1} Mmol/d

$F_{s,\text{LNG2},C_4}$: -2.240×10^{+1} Mmol/d

$F_{s,\text{LNG2},C_5+}$: -1.717×10^{+1} Mmol/d

LNG3

Total delivery rate : -40.31 hm^3/d

Delivery pressure   : 55 bar

$F_{s,\text{LNG3},CO_2}$: -5.729×10^{+1} Mmol/d

$F_{s,\text{LNG3},H_2S}$: -3.076×10^{-2} Mmol/d

$F_{s,\text{LNG3},N_2}$: -6.374×10^{+0} Mmol/d

$F_{s,\text{LNG3},C_1}$: -1.447×10^{+3} Mmol/d

$F_{s,\text{LNG3},C_2}$: -7.415×10^{+1} Mmol/d

$F_{s,\text{LNG3},C_3}$: -5.576×10^{+1} Mmol/d

$F_{s,\text{LNG3},C_4}$: -3.689×10^{+1} Mmol/d

$F_{s,\text{LNG3},C_5+}$: -2.668×10^{+1} Mmol/d

B11

Dry gas production rate :
   $3.586×10^{-3}$ hm^3/d

NGL production rate :
   $1.688×10^{-1}$ m^3/d

Pressure : 71.99 bar

Compression power : 0.01 MW

Compression inlet pressure : 20 bar

$F_{s,\text{B11},CO_2}$: 1.342×10^{-2} Mmol/d

$F_{s,\text{B11},H_2S}$: 7.885×10^{-5} Mmol/d

$F_{s,\text{B11},N_2}$: 1.915×10^{-3} Mmol/d

$F_{s,\text{B11},C_1}$: 1.221×10^{-1} Mmol/d

$F_{s,\text{B11},C_2}$: 8.183×10^{-3} Mmol/d

$F_{s,\text{B11},C_3}$: 8.063×10^{-4} Mmol/d

$F_{s,\text{B11},C_4}$: 9.835×10^{-4} Mmol/d

$F_{s,\text{B11},C_5+}$: 4.166×10^{-3} Mmol/d

BN

Dry gas production rate :
   $4.389×\ 10^{+0}$ hm^3/d

Pressure: 83.96 bar

F_{s,BN,CO_2}: 2.688×10^{+0} Mmol/d

F_{s,BN,H_2S} : 0.000×10^{+0} Mmol/d

F_{s,BN,N_2} : 9.638×10^{-1} Mmol/d

F_{s,BN,C_1} : 1.504×10^{+2} Mmol/d

F_{s,BN,C_2} : 8.175×10^{+0} Mmol/d

F_{s,BN,C_3} : 1.226×10^{+1} Mmol/d

F_{s,BN,C_4} : 7.914×10^{+0} Mmol/d

F_{s,BN,C_5+}: 3.224×10^{+0} Mmol/d

BY

Dry gas production rate :

$1.774 \times 10^{+0}$ hm^3/d

Pressure : 77.92 bar

$F_{s,\text{BY},\text{CO}_2}$: 6.587×10^{-1} Mmol/d

$F_{s,\text{BY},\text{H}_2\text{S}}$: $0.000 \times 10^{+0}$ Mmol/d

$F_{s,\text{BY},\text{N}_2}$: 1.620×10^{-1} Mmol/d

$F_{s,\text{BY},\text{C}_1}$: $6.856 \times 10^{+1}$ Mmol/d

$F_{s,\text{BY},\text{C}_2}$: $3.399 \times 10^{+0}$ Mmol/d

$F_{s,\text{BY},\text{C}_3}$: 8.500×10^{-1} Mmol/d

$F_{s,\text{BY},\text{C}_4}$: 2.978×10^{-1} Mmol/d

$F_{s,\text{BY},\text{C}_5+}$: $1.086 \times 10^{+0}$ Mmol/d

D35

Dry gas production rate :

$1.133 \times 10^{+0}$ hm^3/d

Pressure : 100 bar

$F_{s,\text{D35},\text{CO}_2}$: 3.437×10^{-1} Mmol/d

$F_{s,\text{D35},\text{H}_2\text{S}}$: $0.000 \times 10^{+0}$ Mmol/d

$F_{s,\text{D35},\text{N}_2}$: 2.640×10^{-1} Mmol/d

$F_{s,\text{D35},\text{C}_1}$: $4.015 \times 10^{+1}$ Mmol/d

$F_{s,\text{D35},\text{C}_2}$: $3.529 \times 10^{+0}$ Mmol/d

$F_{s,\text{D35},\text{C}_3}$: $2.227 \times 10^{+0}$ Mmol/d

$F_{s,\text{D35},\text{C}_4}$: 3.514×10^{-1} Mmol/d

$F_{s,\text{D35},\text{C}_5+}$: $1.029 \times 10^{+0}$ Mmol/d

E11

Dry gas production rate :

$0.000 \times 10^{+0}$ hm^3/d

NGL production rate :

$0 \times 10^{+0}$ m^3/d

Pressure : 92.9 bar

$F_{s,\text{E11},\text{CO}_2}$: $0.000 \times 10^{+0}$ Mmol/d

$F_{s,\text{E11},\text{H}_2\text{S}}$: $0.000 \times 10^{+0}$ Mmol/d

$F_{s,\text{E11},\text{N}_2}$: $0.000 \times 10^{+0}$ Mmol/d

$F_{s,\text{E11},\text{C}_1}$: $0.000 \times 10^{+0}$ Mmol/d

$F_{s,\text{E11},\text{C}_2}$: $0.000 \times 10^{+0}$ Mmol/d

$F_{s,\text{E11},\text{C}_3}$: $0.000 \times 10^{+0}$ Mmol/d

$F_{s,\text{E11},\text{C}_4}$: $0.000 \times 10^{+0}$ Mmol/d

$F_{s,\text{E11},\text{C}_5+}$: $0.000 \times 10^{+0}$ Mmol/d

E11P

Dry gas production rate :

$1.927 \times 10^{+0}$ hm^3/d

Pressure : 71.91 bar

$F_{s,\text{E11P},\text{CO}_2}$: 2.189×10^{-1} Mmol/d

$F_{s,\text{E11P},\text{H}_2\text{S}}$: $0.000 \times 10^{+0}$ Mmol/d

$F_{s,\text{E11P},\text{N}_2}$: 5.703×10^{-1} Mmol/d

$F_{s,\text{E11P},\text{C}_1}$: $7.837 \times 10^{+1}$ Mmol/d

$F_{s,\text{E11P},\text{C}_2}$: $1.450 \times 10^{+0}$ Mmol/d

$F_{s,\text{E11P},\text{C}_3}$: 4.037×10^{-1} Mmol/d

$F_{s,\text{E11P},\text{C}_4}$: 9.436×10^{-2} Mmol/d

$F_{s,\text{E11P},\text{C}_5+}$: 3.777×10^{-1} Mmol/d

F23

Dry gas production rate :

$1.751 \times 10^{+1}$ hm^3/d

NGL production rate :

1.879×10^{3} m^3/d

Pressure : 91.41 bar

$F_{s,\text{F23},\text{CO}_2}$: $1.217 \times 10^{+1}$ Mmol/d

$F_{s,\text{F23},\text{H}_2\text{S}}$: 2.962×10^{-3} Mmol/d

$F_{s,\text{F23},\text{N}_2}$: $7.776 \times 10^{+0}$ Mmol/d

$F_{s,\text{F23},\text{C}_1}$: $6.599 \times 10^{+2}$ Mmol/d

$F_{s,\text{F23},\text{C}_2}$: $2.078 \times 10^{+1}$ Mmol/d

$F_{s,\text{F23},\text{C}_3}$: $2.725 \times 10^{+1}$ Mmol/d

$F_{s,\text{F23},\text{C}_4}$: $6.791 \times 10^{+0}$ Mmol/d

$F_{s,\text{F23},\text{C}_{5+}}$: $5.912 \times 10^{+0}$ Mmol/d

F23P

Dry gas production rate :

$2.690 \times 10^{+1}$ hm^3/d

Pressure : 91.41 bar

$F_{s,\text{F23P},\text{CO}_2}$: $1.488 \times 10^{+1}$ Mmol/d

$F_{s,\text{F23P},\text{H}_2\text{S}}$: 6.931×10^{-3} Mmol/d

$F_{s,\text{F23P},\text{N}_2}$: $9.391 \times 10^{+0}$ Mmol/d

$F_{s,\text{F23P},\text{C}_1}$: $1.023 \times 10^{+3}$ Mmol/d

$F_{s,\text{F23P},\text{C}_2}$: $3.839 \times 10^{+1}$ Mmol/d

$F_{s,\text{F23P},\text{C}_3}$: $3.574 \times 10^{+1}$ Mmol/d

$F_{s,\text{F23P},\text{C}_4}$: $8.760 \times 10^{+0}$ Mmol/d

$F_{s,\text{F23P},\text{C}_{5+}}$: $7.303 \times 10^{+0}$ Mmol/d

F23SW

Dry gas production rate :

$9.386 \times 10^{+0}$ hm^3/d

NGL production rate :

9.942×10^2 m^3/d

Pressure : 91.41 bar

$F_{s,\text{F23SW},\text{CO}_2}$: $2.715 \times 10^{+0}$ Mmol/d

$F_{s,\text{F23SW},\text{H}_2\text{S}}$: 3.969×10^{-3} Mmol/d

$F_{s,\text{F23SW},\text{N}_2}$: $1.615 \times 10^{+0}$ Mmol/d

$F_{s,\text{F23SW},\text{C}_1}$: $3.631 \times 10^{+2}$ Mmol/d

$F_{s,\text{F23SW},\text{C}_2}$: $1.761 \times 10^{+1}$ Mmol/d

$F_{s,\text{F23SW},\text{C}_3}$: $8.490 \times 10^{+0}$ Mmol/d

$F_{s,\text{F23SW},\text{C}_4}$: $1.969 \times 10^{+0}$ Mmol/d

$F_{s,\text{F23SW},\text{C}_{5+}}$: $1.391 \times 10^{+0}$ Mmol/d

F6

Dry gas production rate :

$4.387 \times 10^{+0}$ hm^3/d

NGL production rate :

6.155×10^2 m^3/d

Pressure : 72.62 bar

Compression power : 13.91 MW

Compression inlet pressure :

17.55 bar

$F_{s,\text{F6},\text{CO}_2}$: $6.330 \times 10^{+0}$ Mmol/d

$F_{s,\text{F6},\text{H}_2\text{S}}$: 7.050×10^{-3} Mmol/d

$F_{s,\text{F6},\text{N}_2}$: $2.343 \times 10^{+0}$ Mmol/d

$F_{s,\text{F6},\text{C}_1}$: $1.484 \times 10^{+2}$ Mmol/d

$F_{s,\text{F6},\text{C}_2}$: $1.416 \times 10^{+1}$ Mmol/d

$F_{s,\text{F6},\text{C}_3}$: $8.947 \times 10^{+0}$ Mmol/d

$F_{s,\text{F6},\text{C}_4}$: $1.856 \times 10^{+0}$ Mmol/d

$F_{s,\text{F6},\text{C}_{5+}}$: $3.486 \times 10^{+0}$ Mmol/d

HL

Dry gas production rate :

9.654×10^{-1} hm^3/d

Pressure : 72.22 bar

$F_{s,\text{HL},\text{CO}_2}$: 6.488×10^{-1} Mmol/d

$F_{s,\text{HL},\text{H}_2\text{S}}$: $0.000 \times 10^{+0}$ Mmol/d

$F_{s,\text{HL},\text{N}_2}$: 3.965×10^{-1} Mmol/d

$F_{s,\text{HL},\text{C}_1}$: $2.858 \times 10^{+1}$ Mmol/d

$F_{s,\text{HL},\text{C}_2}$: $2.324 \times 10^{+0}$ Mmol/d

$F_{s,\text{HL},\text{C}_3}$: $2.220 \times 10^{+0}$ Mmol/d

$F_{s,\text{HL},\text{C}_4}$: $2.687 \times 10^{+0}$ Mmol/d

$F_{s,\text{HL},\text{C}_{5+}}$: $3.966 \times 10^{+0}$ Mmol/d

JN

Dry gas production rate :

$1.428 \times 10^{+1}$ hm^3/d

NGL production rate :

4.044×10^3 m^3/d

Pressure : 101.4 bar

F_{s,JN,CO_2}: $1.591\times10^{+1}$ Mmol/d

F_{s,JN,H_2S}: 1.812×10^{-3} Mmol/d

F_{s,JN,N_2} : 8.691×10^{-1} Mmol/d

F_{s,JN,C_1} : $5.358\times10^{+2}$ Mmol/d

F_{s,JN,C_2} : $1.609\times10^{+1}$ Mmol/d

F_{s,JN,C_3} : $2.240\times10^{+1}$ Mmol/d

F_{s,JN,C_4} : $7.820\times10^{+0}$ Mmol/d

F_{s,JN,C_5+}: $5.030\times10^{+0}$ Mmol/d

M1

Dry gas production rate :
$$2.253\times10^{+1} \text{ hm}^3/\text{d}$$

NGL production rate :
$$8.296\times10^3 \text{ m}^3/\text{d}$$

Pressure : 98.11 bar

Compression power : 20 MW

Compression inlet pressure :
$$61.33 \text{ bar}$$

$F_{s,\text{M1},CO_2}$: $4.802\times10^{+1}$ Mmol/d

$F_{s,\text{M1},H_2S}$: 3.144×10^{-2} Mmol/d

$F_{s,\text{M1},N_2}$: $4.077\times10^{+0}$ Mmol/d

$F_{s,\text{M1},C_1}$: $7.719\times10^{+2}$ Mmol/d

$F_{s,\text{M1},C_2}$: $5.205\times10^{+1}$ Mmol/d

$F_{s,\text{M1},C_3}$: $3.341\times10^{+1}$ Mmol/d

$F_{s,\text{M1},C_4}$: $2.621\times10^{+1}$ Mmol/d

$F_{s,\text{M1},C_5+}$: $1.693\times10^{+1}$ Mmol/d

SE

Dry gas production rate :
$$1.263\times10^{+1} \text{ hm}^3/\text{d}$$

NGL production rate :
$$5.081\times10^3 \text{ m}^3/\text{d}$$

F_{s,SE,CO_2}: $1.296\times10^{+1}$ Mmol/d

F_{s,SE,H_2S}: 3.205×10^{-3} Mmol/d

F_{s,SE,N_2} : 9.657×10^{-1} Mmol/d

F_{s,SE,C_1} : $4.679\times10^{+2}$ Mmol/d

F_{s,SE,C_2} : $1.988\times10^{+1}$ Mmol/d

F_{s,SE,C_3} : $7.734\times10^{+0}$ Mmol/d

F_{s,SE,C_4} : $1.243\times10^{+1}$ Mmol/d

F_{s,SE,C_5+}: $1.222\times10^{+1}$ Mmol/d

M3

Dry gas production rate :
$$3.855\times10^{-6} \text{ hm}^3/\text{d}$$

NGL production rate :
$$2.608\times10^{-4} \text{ m}^3/\text{d}$$

$F_{s,\text{M3},CO_2}$: 1.546×10^{-6} Mmol/d

$F_{s,\text{M3},H_2S}$: $0.000\times10^{+0}$ Mmol/d

$F_{s,\text{M3},N_2}$: $0.000\times10^{+0}$ Mmol/d

$F_{s,\text{M3},C_1}$: 1.243×10^{-4} Mmol/d

$F_{s,\text{M3},C_2}$: 1.202×10^{-5} Mmol/d

$F_{s,\text{M3},C_3}$: 1.139×10^{-5} Mmol/d

$F_{s,\text{M3},C_4}$: 1.882×10^{-6} Mmol/d

$F_{s,\text{M3},C_5+}$: 1.114×10^{-5} Mmol/d

M3P

Dry gas production rate :
$$1.608\times10^{+1} \text{ hm}^3/\text{d}$$

Pressure : 95.57 bar

Compression power : 27 MW

Compression inlet pressure :
$$41.49 \text{ bar}$$

$F_{s,\text{M3P},CO_2}$: $1.632\times10^{+1}$ Mmol/d

$F_{s,\text{M3P},H_2S}$: 1.021×10^{-2} Mmol/d

$F_{s,\text{M3P},N_2}$: $1.342\times10^{+0}$ Mmol/d

$F_{s,\text{M3P},C_1}$: $5.880\times10^{+2}$ Mmol/d

$F_{s,\text{M3P},C_2}$: $3.054\times10^{+1}$ Mmol/d

$F_{s,\text{M3P},C_3}$: $1.312\times10^{+1}$ Mmol/d

$F_{s,\text{M3P},C_4}$: $1.709\times10^{+1}$ Mmol/d

$F_{s,\text{M3P},\text{C}_5+}$: $1.368{\times}10^{+1}$ Mmol/d

M4

Dry gas production rate :
$$3.452{\times}10^{+0} \text{ hm}^3/\text{d}$$
NGL production rate :
$$2.191{\times}10^2 \text{ m}^3/\text{d}$$
$F_{s,\text{M4},\text{CO}_2}$: $3.364{\times}10^{+0}$ Mmol/d

$F_{s,\text{M4},\text{H}_2\text{S}}$: $7.007{\times}10^{-3}$ Mmol/d

$F_{s,\text{M4},\text{N}_2}$: $3.765{\times}10^{-1}$ Mmol/d

$F_{s,\text{M4},\text{C}_1}$: $1.201{\times}10^{+2}$ Mmol/d

$F_{s,\text{M4},\text{C}_2}$: $1.065{\times}10^{+1}$ Mmol/d

$F_{s,\text{M4},\text{C}_3}$: $5.384{\times}10^{+0}$ Mmol/d

$F_{s,\text{M4},\text{C}_4}$: $4.665{\times}10^{+0}$ Mmol/d

$F_{s,\text{M4},\text{C}_5+}$: $1.463{\times}10^{+0}$ Mmol/d

SC

Dry gas production rate :
$$1.927{\times}10^{+0} \text{ hm}^3/\text{d}$$
NGL production rate :
$$3.121{\times}10^2 \text{ m}^3/\text{d}$$
Pressure : 71.91 bar

$F_{s,\text{SC},\text{CO}_2}$: $2.189{\times}10^{-1}$ Mmol/d

$F_{s,\text{SC},\text{H}_2\text{S}}$: $0.000{\times}10^{+0}$ Mmol/d

$F_{s,\text{SC},\text{N}_2}$: $5.703{\times}10^{-1}$ Mmol/d

$F_{s,\text{SC},\text{C}_1}$: $7.837{\times}10^{+1}$ Mmol/d

$F_{s,\text{SC},\text{C}_2}$: $1.450{\times}10^{+0}$ Mmol/d

$F_{s,\text{SC},\text{C}_3}$: $4.037{\times}10^{-1}$ Mmol/d

$F_{s,\text{SC},\text{C}_4}$: $9.436{\times}10^{-2}$ Mmol/d

$F_{s,\text{SC},\text{C}_5+}$: $3.777{\times}10^{-1}$ Mmol/d

Table D.1: Trunkline volumetric flowrates, molar flowrates and pressures

$\mathcal{A}_q$	$Q_{a,(i,j)}$ hm^3/d	P_i bar	P_j bar	$s_{(i,j)}$	$F_{a,(i,j),CO_2}$ Mmol/day	$F_{a,(i,j),H_2S}$ Mmol/day
(B11,RB)	0.97	71.99	71.91	-	6.622×10^{-1}	7.885×10^{-5}
(BN,RA)	4.39	83.96	71.91	-	$2.688 \times 10^{+0}$	$0.000 \times 10^{+0}$
(BY,LNG1)	2.91	77.92	62.60	-	$1.002 \times 10^{+0}$	$0.000 \times 10^{+0}$
(D35,BY)	1.13	100.00	77.92	-	3.437×10^{-1}	$0.000 \times 10^{+0}$
(E11P,RA)	1.93	71.91	71.91	-	2.189×10^{-1}	$0.000 \times 10^{+0}$
(F23P,RA)	26.90	91.41	71.91	-	$1.488 \times 10^{+1}$	6.931×10^{-3}
(F6,RA)	4.39	72.62	71.91	-	$6.330 \times 10^{+0}$	7.050×10^{-3}
(HL,B11)	0.97	72.22	71.99	-	6.488×10^{-1}	$0.000 \times 10^{+0}$
(JN,M1)	14.28	101.40	98.11	-	$1.591 \times 10^{+1}$	1.812×10^{-3}
(M1,RC)	14.28	98.11	71.91	-	$2.481 \times 10^{+1}$	1.290×10^{-2}
(M1,T)	22.53	98.11	95.04	0.612	$3.913 \times 10^{+1}$	2.035×10^{-2}
(M3P,T)	16.08	95.57	95.04	-	$1.632 \times 10^{+1}$	1.021×10^{-2}
(RA,RB)	11.06	71.91	71.91	-	$7.093 \times 10^{+0}$	4.112×10^{-3}
(RB,RC)	17.58	71.91	71.91	-	$2.194 \times 10^{+1}$	1.207×10^{-2}
(SC1,LNG1)	26.55	67.60	62.60	-	$1.702 \times 10^{+1}$	9.869×10^{-3}
(SC2,LNG2)	24.62	60.00	55.00	-	$3.073 \times 10^{+1}$	1.689×10^{-2}
(SC2,LNG3)	8.44	60.00	55.00	0.2553	$1.054 \times 10^{+1}$	5.792×10^{-3}
(SC3,LNG3)	31.87	60.00	55.00	-	$4.675 \times 10^{+1}$	2.497×10^{-2}
TL1	15.64	71.91	67.60	0.4159	$1.003 \times 10^{+1}$	5.815×10^{-3}
TL2	10.91	71.91	67.60	0.29	$6.994 \times 10^{+0}$	4.054×10^{-3}
TL3	16.05	71.91	60.00	0.3169	$2.003 \times 10^{+1}$	1.101×10^{-2}
TL4	17.01	71.91	60.00	0.3359	$2.123 \times 10^{+1}$	1.167×10^{-2}
TL5	14.33	71.91	60.00	0.4497	$2.103 \times 10^{+1}$	1.123×10^{-2}
TL6	17.53	71.91	60.00	-	$2.573 \times 10^{+1}$	1.374×10^{-2}
(T,RB)	38.61	95.04	71.91	-	$5.545 \times 10^{+1}$	3.056×10^{-2}

Table D.2: Trunkline molar flowrates (continued)

$\mathcal{A}_q$	$F_{a,(i,j),N_2}$ Mmol/day	$F_{a,(i,j),C_1}$ Mmol/day	$F_{a,(i,j),C_2}$ Mmol/day	$F_{a,(i,j),C_3}$ Mmol/day	$F_{a,(i,j),C_4}$ Mmol/day	$F_{a,(i,j),C_5+}$ Mmol/day
(B11,RB)	3.984×10^{-1}	$2.870 \times 10^{+1}$	$2.332 \times 10^{+0}$	$2.221 \times 10^{+0}$	$2.688 \times 10^{+0}$	$3.970 \times 10^{+0}$
(BN,RA)	9.638×10^{-1}	$1.504 \times 10^{+2}$	$8.175 \times 10^{+0}$	$1.226 \times 10^{+1}$	$7.914 \times 10^{+0}$	$3.224 \times 10^{+0}$
(BY,LNG1)	4.261×10^{-1}	$1.087 \times 10^{+2}$	$6.928 \times 10^{+0}$	$3.077 \times 10^{+0}$	6.492×10^{-1}	$2.114 \times 10^{+0}$
(D35,BY)	2.640×10^{-1}	$4.015 \times 10^{+1}$	$3.529 \times 10^{+0}$	$2.227 \times 10^{+0}$	3.514×10^{-1}	$1.029 \times 10^{+0}$
(E11P,RA)	5.703×10^{-1}	$7.837 \times 10^{+1}$	$1.450 \times 10^{+0}$	4.037×10^{-1}	9.436×10^{-2}	3.777×10^{-1}
(F23P,RA)	$9.391 \times 10^{+0}$	$1.023 \times 10^{+3}$	$3.839 \times 10^{+1}$	$3.574 \times 10^{+1}$	$8.760 \times 10^{+0}$	$7.303 \times 10^{+0}$
(F6,RA)	$2.343 \times 10^{+0}$	$1.484 \times 10^{+2}$	$1.416 \times 10^{+1}$	$8.947 \times 10^{+0}$	$1.856 \times 10^{+0}$	$3.486 \times 10^{+0}$
(HL,B11)	3.965×10^{-1}	$2.858 \times 10^{+1}$	$2.324 \times 10^{+0}$	$2.220 \times 10^{+0}$	$2.687 \times 10^{+0}$	$3.966 \times 10^{+0}$
(JN,M1)	8.691×10^{-1}	$5.358 \times 10^{+2}$	$1.609 \times 10^{+1}$	$2.240 \times 10^{+1}$	$7.820 \times 10^{+0}$	$5.030 \times 10^{+0}$
(M1,RC)	$1.919 \times 10^{+0}$	$5.074 \times 10^{+2}$	$2.644 \times 10^{+1}$	$2.166 \times 10^{+1}$	$1.321 \times 10^{+1}$	$8.522 \times 10^{+0}$
(M1,T)	$3.027 \times 10^{+0}$	$8.003 \times 10^{+2}$	$4.170 \times 10^{+1}$	$3.416 \times 10^{+1}$	$2.083 \times 10^{+1}$	$1.344 \times 10^{+1}$
(M3P,T)	$1.342 \times 10^{+0}$	$5.880 \times 10^{+2}$	$3.054 \times 10^{+1}$	$1.312 \times 10^{+1}$	$1.709 \times 10^{+1}$	$1.368 \times 10^{+1}$
(RA,RB)	$3.902 \times 10^{+0}$	$4.118 \times 10^{+2}$	$1.828 \times 10^{+1}$	$1.687 \times 10^{+1}$	$5.477 \times 10^{+0}$	$4.232 \times 10^{+0}$
(RB,RC)	$3.010 \times 10^{+0}$	$6.349 \times 10^{+2}$	$3.224 \times 10^{+1}$	$2.304 \times 10^{+1}$	$1.600 \times 10^{+1}$	$1.227 \times 10^{+1}$
(SC1,LNG1)	$9.366 \times 10^{+0}$	$9.884 \times 10^{+2}$	$4.389 \times 10^{+1}$	$4.048 \times 10^{+1}$	$1.315 \times 10^{+1}$	$1.016 \times 10^{+1}$
(SC2,LNG2)	$4.215 \times 10^{+0}$	$8.890 \times 10^{+2}$	$4.514 \times 10^{+1}$	$3.226 \times 10^{+1}$	$2.240 \times 10^{+1}$	$1.717 \times 10^{+1}$
(SC2,LNG3)	$1.445 \times 10^{+0}$	$3.048 \times 10^{+2}$	$1.548 \times 10^{+1}$	$1.106 \times 10^{+1}$	$7.682 \times 10^{+0}$	$5.889 \times 10^{+0}$
(SC3,LNG3)	$4.929 \times 10^{+0}$	$1.142 \times 10^{+3}$	$5.868 \times 10^{+1}$	$4.470 \times 10^{+1}$	$2.921 \times 10^{+1}$	$2.079 \times 10^{+1}$
TL1	$5.518 \times 10^{+0}$	$5.823 \times 10^{+2}$	$2.586 \times 10^{+1}$	$2.385 \times 10^{+1}$	$7.746 \times 10^{+0}$	$5.985 \times 10^{+0}$
TL2	$3.847 \times 10^{+0}$	$4.060 \times 10^{+2}$	$1.803 \times 10^{+1}$	$1.663 \times 10^{+1}$	$5.401 \times 10^{+0}$	$4.173 \times 10^{+0}$
TL3	$2.748 \times 10^{+0}$	$5.796 \times 10^{+2}$	$2.943 \times 10^{+1}$	$2.103 \times 10^{+1}$	$1.461 \times 10^{+1}$	$1.120 \times 10^{+1}$
TL4	$2.912 \times 10^{+0}$	$6.143 \times 10^{+2}$	$3.119 \times 10^{+1}$	$2.229 \times 10^{+1}$	$1.548 \times 10^{+1}$	$1.187 \times 10^{+1}$
TL5	$2.217 \times 10^{+0}$	$5.137 \times 10^{+2}$	$2.639 \times 10^{+1}$	$2.010 \times 10^{+1}$	$1.314 \times 10^{+1}$	$9.349 \times 10^{+0}$
TL6	$2.712 \times 10^{+0}$	$6.286 \times 10^{+2}$	$3.229 \times 10^{+1}$	$2.460 \times 10^{+1}$	$1.607 \times 10^{+1}$	$1.144 \times 10^{+1}$
(T,RB)	$4.369 \times 10^{+0}$	$1.388 \times 10^{+3}$	$7.224 \times 10^{+1}$	$4.728 \times 10^{+1}$	$3.792 \times 10^{+1}$	$2.713 \times 10^{+1}$

D.2 Wells

Table D.3: Well results

$\mathcal{W}$	$Q_{w,w}$	$P_{b,w}$	$P_{t,w}$	$Q_{Lw,w}$
	hm^3/d	bar	bar	m^3/d
B11A	4.566×10^{-5}	75.39	40.10	2.260×10^{-3}
B11B	3.540×10^{-3}	78.46	43.83	1.665×10^{-1}
B11C	$0.000\times 10^{+0}$	78.62	41.28	$0.000\times 10^{+0}$
B11D	$0.000\times 10^{+0}$	71.45	37.83	$0.000\times 10^{+0}$
E11A	$0.000\times 10^{+0}$	154.70	96.81	$0.000\times 10^{+0}$
E11B	$0.000\times 10^{+0}$	156.70	93.97	$0.000\times 10^{+0}$
E11C	$0.000\times 10^{+0}$	160.00	97.30	$0.000\times 10^{+0}$
E11D	$0.000\times 10^{+0}$	152.70	94.79	$0.000\times 10^{+0}$
E11E	$0.000\times 10^{+0}$	156.20	100.10	$0.000\times 10^{+0}$
E11F	$0.000\times 10^{+0}$	156.50	98.49	$0.000\times 10^{+0}$
E11G	$0.000\times 10^{+0}$	149.90	92.90	$0.000\times 10^{+0}$
E11H	$0.000\times 10^{+0}$	159.70	98.15	$0.000\times 10^{+0}$
E11I	$0.000\times 10^{+0}$	156.80	98.02	$0.000\times 10^{+0}$
E11J	$0.000\times 10^{+0}$	154.50	100.50	$0.000\times 10^{+0}$
F23A	7.147×10^{-1}	247.70	196.90	$7.472\times 10^{+1}$
F23B	4.598×10^{-1}	231.80	192.00	$4.999\times 10^{+1}$
F23C	$1.072\times 10^{+1}$	266.70	91.41	$1.175\times 10^{+3}$
F23D	4.598×10^{-1}	228.00	189.50	$5.274\times 10^{+1}$
F23E	$0.000\times 10^{+0}$	244.10	195.00	$0.000\times 10^{+0}$
F23F	2.230×10^{-1}	245.90	187.90	$2.283\times 10^{+1}$
F23G	$1.000\times 10^{+0}$	265.50	216.50	$9.820\times 10^{+1}$
F23H	4.392×10^{-1}	226.20	181.20	$4.385\times 10^{+1}$
F23I	4.581×10^{-1}	271.80	227.00	$4.783\times 10^{+1}$
F23J	4.551×10^{-1}	252.10	210.80	$4.909\times 10^{+1}$
F23K	$1.017\times 10^{+0}$	235.70	191.40	$9.896\times 10^{+1}$
F23L	6.683×10^{-1}	266.10	203.80	$6.885\times 10^{+1}$
F23M	4.395×10^{-1}	252.20	209.90	$4.307\times 10^{+1}$
F23N	4.648×10^{-1}	264.40	201.30	$5.299\times 10^{+1}$
F23SW	$9.386\times 10^{+0}$	239.20	91.41	$9.942\times 10^{+2}$
F6A	$0.000\times 10^{+0}$	44.81	35.28	$0.000\times 10^{+0}$
F6B	$0.000\times 10^{+0}$	47.11	38.12	$0.000\times 10^{+0}$
F6C	$1.491\times 10^{+0}$	41.81	17.55	$2.029\times 10^{+2}$
F6D	$0.000\times 10^{+0}$	44.95	34.95	$0.000\times 10^{+0}$
F6E	$0.000\times 10^{+0}$	40.47	32.42	$0.000\times 10^{+0}$

Table D.3: Well results

$\mathcal{W}$	$Q_{w,w}$	$P_{b,w}$	$P_{t,w}$	$Q_{Lw,w}$
	$\mathrm{hm^3/d}$	bar	bar	$\mathrm{m^3/d}$
F6F	$0.000\times 10^{+0}$	40.83	31.40	$0.000\times 10^{+0}$
F6G	$0.000\times 10^{+0}$	43.24	33.95	$0.000\times 10^{+0}$
F6H	$0.000\times 10^{+0}$	47.84	39.11	$0.000\times 10^{+0}$
F6I	4.052×10^{-1}	46.02	36.77	$5.911\times 10^{+1}$
F6J	3.908×10^{-1}	47.91	36.75	$5.779\times 10^{+1}$
F6K	$1.553\times 10^{+0}$	44.49	17.55	$2.217\times 10^{+2}$
F6L	5.474×10^{-1}	44.19	32.30	$7.393\times 10^{+1}$
F6M	$0.000\times 10^{+0}$	45.43	35.09	$0.000\times 10^{+0}$
JNA	$2.674\times 10^{+0}$	142.20	142.20	$7.598\times 10^{+2}$
JNB	$3.026\times 10^{+0}$	150.50	150.50	$8.300\times 10^{+2}$
JNC	$2.928\times 10^{+0}$	146.00	146.00	$8.757\times 10^{+2}$
JND	$2.843\times 10^{+0}$	129.30	120.70	$7.906\times 10^{+2}$
JNE	$2.812\times 10^{+0}$	152.80	150.70	$7.879\times 10^{+2}$
M1A	$6.062\times 10^{+0}$	113.30	61.33	$2.192\times 10^{+3}$
M1B	$5.916\times 10^{+0}$	113.60	62.60	$2.121\times 10^{+3}$
M1C	$1.114\times 10^{+0}$	102.20	71.28	$4.134\times 10^{+2}$
M1D	$2.865\times 10^{+0}$	121.10	86.71	$1.018\times 10^{+3}$
M1E	$0.000\times 10^{+0}$	109.30	78.41	$0.000\times 10^{+0}$
M1F	$0.000\times 10^{+0}$	107.60	80.65	$0.000\times 10^{+0}$
M1G	$6.135\times 10^{+0}$	111.60	61.33	$2.401\times 10^{+3}$
M1H	4.360×10^{-1}	105.20	75.70	$1.508\times 10^{+2}$
SEA	$6.734\times 10^{+0}$	153.40	41.49	$2.703\times 10^{+3}$
SEB	$5.897\times 10^{+0}$	141.20	41.49	$2.378\times 10^{+3}$
M3A	$0.000\times 10^{+0}$	69.33	56.95	$0.000\times 10^{+0}$
M3B	$0.000\times 10^{+0}$	82.08	64.47	$0.000\times 10^{+0}$
M3C	$0.000\times 10^{+0}$	78.36	63.16	$0.000\times 10^{+0}$
M3D	$0.000\times 10^{+0}$	73.13	59.12	$0.000\times 10^{+0}$
M3E	$0.000\times 10^{+0}$	79.77	62.25	$0.000\times 10^{+0}$
M3F	$0.000\times 10^{+0}$	80.46	66.91	$0.000\times 10^{+0}$
M3G	3.855×10^{-6}	82.30	63.82	2.608×10^{-4}
M3H	$0.000\times 10^{+0}$	76.71	60.32	$0.000\times 10^{+0}$
M3I	$0.000\times 10^{+0}$	79.49	62.43	$0.000\times 10^{+0}$
M3J	$0.000\times 10^{+0}$	72.10	55.74	$0.000\times 10^{+0}$
M4A	$1.626\times 10^{+0}$	75.69	41.49	$1.011\times 10^{+2}$
M4B	$1.827\times 10^{+0}$	81.38	41.49	$1.180\times 10^{+2}$
SCA	7.928×10^{-1}	142.30	71.91	$1.291\times 10^{+2}$

Continued on next page

Table D.3: Well results

$\mathcal{W}$	$Q_{w,w}$	$P_{b,w}$	$P_{t,w}$	$Q_{Lw,w}$
	hm^3/d	bar	bar	m^3/d
SCB	$1.134\times 10^{+0}$	146.70	71.91	$1.830\times 10^{+2}$

D.3 PSC Status

Table D.4: Production-sharing contracts: Supply rate status

Contracts	Production
	hm^3/d
A (Total of AM, X and Y)	40.51
AM	23.83
X	9.386
Y	7.296
B	22.53
C	18.7
D	12.63
F	0

Table D.5: Excess flowrates in the PSC network

$\mathcal{E}_l$	$e_{(i,j)}$	Status $(y^e_{(i,j)})$	Priority $(y^s_{(i,j)})$
	hm^3/d		
(A_0, A_1)	+11.060	Excess (1)	-
(A_1, A_2)	+8.975	-	-
(A_2, A_3)	+2.218	-	-
(B_0, B_1)	−2.085	Deficit (0)	-
(B_1, B_2)	−2.085	Deficit (0)	-
(B_2, B_3)	+0.000	-	Satisfied (1)
(C_0, C_1)	−6.756	Deficit (0)	-
(C_1, C_2)	−6.756	-	-
(C_2, C_3)	−6.756	-	Deficit (0)
(D_0, D_1)	−2.218	Deficit (0)	-
(D_1, D_2)	−2.218	-	Deficit (0)
(D_2, D_3)	−2.218	-	-

Table D.6: Production-sharing contracts: Transfer status

$\mathcal{E}_t$	$t_{(i,j)}$	Status $(y^t_{(i,j)})$
	hm^3/d	
(A, B)	2.085	Active (1)
(A, C)	6.756	Active (1)
(A, D)	2.218	Active (1)
(B, C)	0.000	Inactive (0)
(B, D)	0.000	Inactive (0)
(C, D)	0.000	Inactive (0)
(C, B)	0.000	Inactive (0)
(D, C)	0.000	Inactive (0)
(F, A)	0.000	Inactive (0)
(F, B)	0.000	Active (1)

Appendix E

Global Optimization of Algorithms: Preliminary Case Studies

The objective function in both case studies is to maximize the delivery rate at demand D1 (expressed as a minimization problem). The scaling of variables used is identical to the one described in Section 5.3 (page 146).

E.1   Case Study A

The network corresponding to case study A is presented in Figure 7-5 (page 236).

E.1.1   Parameters

Compressor M3P:

    Rated power : 27.0 MW

    Maximum pressure at outlet : 200 bar

Trunkline pressure-drop coefficient for (M3P,D1) : 2.46 bar^2.d^2/hm^6

The following delivery pressure constraint (presented here in unscaled form) is enforced

$$30^2 \leq \widehat{P}_{\text{D1}} \leq 80^2.$$

297

Table E.1: Case study A: Wells

Fields	No. of Wells	Wells
M3	10	M3A, M3B, M3C, M3D, M3E, M3F, M3G, M3H, M4I, M3J
M4	2	M4A, M4B
SE	2	SEA, SEB

The wells belonging to each field are presented in Table E.1. The parameters for the well-performance model are the same as in Table B.1 (page 275). Gas from each well is assumed to have the same composition as the field to which it belongs. The composition for fields are as in Table B.2 (page 275). Upper bounds on well production rates are calculated as in Section 5.2.1 (Page 145) and the lower bounds are set to a small number (10^{-9}) to avoid zero flowrates through the compressor.

E.2 Case Study B

The network corresponding to case study B is presented in Figure 7-6 (page 236).

E.3 Parameters

Compressor M3P:

Rated power : 27.0 MW

Maximum pressure at outlet : 200 bar

Compressor M1P:

Rated power : 20.0 MW

Maximum pressure at outlet : 200 bar

Trunkline pressure-drop coefficient for (M3P,D1) : 2.46 $\text{bar}^2.\text{d}^2/\text{hm}^6$

Trunkline pressure-drop coefficient for (JN,M1P) : 3.17 $\text{bar}^2.\text{d}^2/\text{hm}^6$

(JN, M1P) is connected to the outlet of compressor M1P and not the inlet.

Trunkline pressure-drop coefficient for (M1P,D1) : 5.06 $\text{bar}^2.\text{d}^2/\text{hm}^6$

The constraint enforced at demand D1 are as follows:

Table E.2: Case study B: Wells

Fields	No. of Wells	Wells
M1	8	M1A, M1B, M1C, M1D, M1E, M1F, M1G, M1H
M3	10	M3A, M3B, M3C, M3D, M3E, M3F, M3G, M3H, M4I, M3J
M4	2	M4A, M4B
JN	5	JNA, JNB, JNC, JND, JNE
SE	2	SEA, SEB

1. H_2S content must be less than 270 ppmV (i.e., $\chi^s_{D1,H_2S} = 270$)

$$F_{s,D1,H_2S} - 10^{-6}\chi^s_{D1,H_2S}\phi\, Q_{s,D1} \leq 0,$$

2. C_2 content should be greater than 3.4% (i.e., $\chi^s_{D1,C_2} = 0.034$).

$$10^{-2}\chi^s_{D1,C_2}\sum_{j\in\mathcal{S}\backslash CO_2} F_{s,D1,j} - F_{s,D1,C_2} \leq 0,$$

3. Delivery pressure must be within 10-80 bar range

$$10^2 \leq \widehat{P}_{D1} \leq 80^2.$$

The wells belonging to each field are presented in Table E.2. The parameters for the well-performance model are the same as in Table B.1 (page 275). Gas from each well is assumed to have the same composition as the field to which it belongs. The composition for fields are as in Table B.2 (page 275). Upper bounds on well production rates are calculated as in Section 5.2.1 (Page 145) and the lower bounds are set to a small number (10^{-9}) to avoid zero flow through compressors.

Nomenclature

The nomenclature presents the symbols used in the descriptions of the standard formulation of infrastructure model (as presented in Sections 3.2-3.5) and the contract modeling framework (Chapter 4). Although, symbols used in the descriptions of the alternative formulation of infrastructure model (presented in Section 3.6) and the applications of global optimization of algorithms to upstream gas networks (presented in Chapter 7), are more or less consistent with this nomenclature, there are minor differences in subscripts and superscripts used, and the definitions of relevant sets. Therefore, these symbols have not been listed here to avoid confusion.

Atomic Propositions

$C_{(.)}$ *Propositions denoting some operational condition*

$E_{(p_i,p_{i+1})}$ *True if contract p is in excess (defined using flowrate on arc (p_i, p_{i+1}))*

$S_{(q_j,q_{j+1})}$ *True if an inter-contract transfer terminating at q_j fulfilled the deficit in contract q (defined using flowrate on arc (q_j, q_{j+1}))*

$T_{p,q}$ *True if contract p supplies contract q*

Parameters

α_w *Darcy flow constant for a well $w \in \mathcal{W}$ ($bar^2.d/hm^3$)*

β_w *Non Darcy flow constant for a well $w \in \mathcal{W}$ ($bar^2.d^2/hm^6$)*

$\chi_{i,k}^s$ *Specification on the component k at demand node $i \in \mathcal{N}_D$ (Unit depends on the component under consideration)*

$\chi_{i,k}$ *Mole fraction of component k in field i*

$\Delta \pi_{(i,j)}$ *Pressure drop across arcs in set $\mathcal{A}_{sc}$ (bar)*

η_i	Compression efficiency for compressor i
Γ_i^s	Specification on the gross heating value of feed gas at the demand node $i \in \mathcal{N}_D$ (MJ/kg)
γ_k	Superior calorific value of component $k \in \mathcal{S}_h$ (MJ/kg)
$\kappa_{(i,j)}$	Pressure drop constant for arc (i,j), $(\forall (i,j) \in \mathcal{A}_p)$ $(bar^2.d^2/hm^6)$
$\Lambda_{d,i}^L$	Minimum demand rate at LNG plant $i \in \mathcal{N}_D$ (hm^3/day)
$\Lambda_{d,i}^U$	Maximum demand rate at LNG plant $i \in \mathcal{N}_D$ (hm^3/day)
λ_w	Coefficient of $P_{t,w}^2$ in VLP equation for a well $w \in \mathcal{W}$ (dimensionless)
μ_k	Molecular weight of a species k $(kg/mole)$
ν	Exponential factor for compressor (same for all compressors)
ω_i	Premultiplier for compression power equation for compressor i
ϕ	Number of moles per unit volume of dry gas at standard conditions $(42.2845\ Mmole/hm^3)$
$\pi_{d,i}^L$	Minimum delivery pressure at slugcatchers (bar)
$\pi_{r,i}^M$	Maximum reservoir pressure among all wells corresponding to a field $i \in \mathcal{F}_w$ (bar)
$\pi_{d,i}^U$	Maximum delivery pressure at slugcatchers (bar)
π_{atm}	Atmospheric pressure (bar)
$\pi_{r,w}$	Reservoir pressure at well $w \in \mathcal{W}$ (bar)
π_{sc}	Pressure at standard conditions (1 atmosphere, 1.013 bar)
Ψ_i^L	Lower bound for the compression power i (MW) (> 0)
Ψ_i^U	Upper bound for the compression power i (MW)
σ_w	Condensate gas ratio for well $w \in \mathcal{W}$ (m^3/hm^3)
τ_{sec}	Number of seconds in a day $(84600\ s/d)$
$\theta_{m,i}$	Mean operating temperature for compressor i (K)
θ_{sc}	Temperature at standard conditions $(15\ ^oC,\ 288.15\ K)$
ϱ_g	Conversion factor from MMscfd to hm^3/day $(0.0283168\ hm^3/MMscfd)$
ϱ_L	Conversion factor from barrel(bbl) to m^3/day $(0.158987\ m^3/bbl)$

ς_w *Water gas ratio for well $w \in \mathcal{W}$ (m^3/hm^3)*

ϑ_w *Coefficient of $Q^2_{w,w}$ in VLP equation for a well $w \in \mathcal{W}$ $(bar^2.d^2/hm^6)$*

ζ *Compression polytropic constant*

R *Universal gas constant (8.314 J/K.mole)*

Sets

$\mathcal{A}$ *Superset for all arcs in the graph representation $(\mathcal{N}, \mathcal{A})$ of the trunkline network*

$\mathcal{A}_p$ *Arcs for which Weymouth pressure-flowrate relationship is used*

$\mathcal{A}_q$ *Arcs for which a flowrate variable is defined*

$\mathcal{A}_x$ *Set of all arcs that are immediately downstream of splitters excluding exactly one arbitrary downstream arc per splitter*

$\mathcal{A}_y$ *Arc for which flow is controlled just by pressure inequality and binary variable*

$\mathcal{A}_i$ *Arcs for which pressure inequalities between inlet and outlet should be enforced*

$\mathcal{A}_{sc}$ *Arcs across which a fixed pressure drop is assumed*

$\mathcal{C}$ *Set of contracts in the system (represented on contract network)*

$\mathcal{C}^S$ *Superset of contracts in the system (not necessarily represented on contract network)*

$\mathcal{E}$ *Set of arcs in contract network representation*

$\mathcal{E}_l$ *Set of arcs in contract network representation that denote excess/deficit levels (equivalently flowrate between levels for a contract) $(\subset \mathcal{E})$*

$\mathcal{E}_s$ *Supply and demand arcs in the contract network representation (i.e., $\forall (i,j) \in \mathcal{E}$ s.t. $i \in \mathcal{L}_s$ or $j \in \mathcal{L}_s$)*

$\mathcal{E}_t$ *Set of arcs in the contract network representing* inter-contract transfer rules $(\subset \mathcal{E})$

$\mathcal{E}_{l,a}$ *Subset of excess/deficit arcs in the contract network representation over which binary variables to indicate contract excess are defined. $\subset \mathcal{E}_l$*

$\mathcal{E}_{l,S}$ *Set of arcs in contract network representation over which binary variables to indicate priorities are defined $(\subset \mathcal{E}_l)$*

$\mathcal{F}$ *Set of all fields $(\subset \mathcal{N}_s)$*

$\mathcal{F}_{pr}$ *Fields that have high priority of production $(\subset \mathcal{F})$*

$\mathcal{F}_{w,nc}$ *Fields with well performance modeling that do not have compression ($\subset \mathcal{F}_w$)*

$\mathcal{F}_w$ *Fields for which well performance is modeled ($\subset \mathcal{F}$)*

$\mathcal{L}$ *Nodes in contract network representation*

$\mathcal{L}_D$ *Set of demand nodes in contract network representation ($\subset \mathcal{L}_s$)*

$\mathcal{L}_l$ *Nodes in contract network that represent levels of availability ($\subset \mathcal{L}$)*

$\mathcal{L}_s$ *Source/sink nodes in contract network representation ($\subset \mathcal{L}$)*

$\mathcal{N}$ *Superset for nodes in the graph representation ($\mathcal{N}, \mathcal{A}$) of the trunkline network*

$\mathcal{N}_D$ *Set of demand nodes (equivalently set of nodes in the network that are sinks) ($\subset \mathcal{N}_s$)*

$\mathcal{N}_J$ *Set of nodes that are junctions (for which production term does not exist)($\subset \mathcal{N}_p$)*

$\mathcal{N}_p$ *Set of nodes over which pressure variable P_i is defined ($\subset \mathcal{N}$)*

$\mathcal{N}_q$ *A subset of source/sink nodes whose elements are on the sub- network defined by $\mathcal{A}_q$ ($\subset \mathcal{N}_s$)*

$\mathcal{N}_s$ *Set of nodes that are sources or sinks (i.e., they have a production term) ($\subset \mathcal{N}$)*

$\mathcal{N}_{D,HSM}$ *Demand nodes for which H_2S specification is expressed in mg/m^3*

$\mathcal{N}_{D,HSP}$ *Demand nodes for which H_2S specification is expressed in $ppmV$*

$\mathcal{N}_{sc}$ *Nodes that are slugcatchers $\subset \mathcal{N}_J$*

$\mathcal{N}_{wp,c}$ *Set of well platforms that have compression ($\subset \mathcal{N}_{wp}$)*

$\mathcal{N}_{wp,m}$ *Set of well platforms that receive gas from multiple fields ($\subset \mathcal{N}_{wp}$)*

$\mathcal{N}_{wp,nc}$ *Set of well platforms that do not have compression ($\subset \mathcal{N}_{wp,w}$)*

$\mathcal{N}_{wp}$ *Set of well platforms ($\subset \mathcal{N}_s$)*

$\mathcal{N}_{x,J}$ *Set of nodes that are splitters and junctions ($\mathcal{N}_x \cap \mathcal{N}_J$)*

$\mathcal{N}_{x,q}$ *Set of nodes that are splitters and production nodes ($\mathcal{N}_x \cap \mathcal{N}_q$)*

$\mathcal{N}_x$ *Nodes that are splitters ($\subset \mathcal{N}$)*

$\mathcal{S}$ *Set of chemical species*

$\mathcal{S}_h$ *Species that are used for gross heating value calculation $\subset \mathcal{S}$*

$\mathcal{W}$ *Set of all wells in the system*

$\mathcal{W}_i$ *Wells that belong to a field i or a well platform i ($\subset \mathcal{W}$)*

Decision Variables

$e_{(p_i, p_{i+1})}$ *Excess/deficit flowrates in the contract network representation $(p_i, p_{i+1}) \in \mathcal{E}_l$ (hm^3/day)*

$F_{a,(i,j),k}$ *Molar flowrate of component k in arc $(i,j) \in \mathcal{A}_q$ (Mmole/day)*

$F_{s,i,k}$ *Molar production rate (negative for demand) of component k for node $i \in \mathcal{N}_s$ (Mmole/day)*

P_i *Pressure at node $i \in \mathcal{N}_p$ (bar)*

$P_{b,w}$ *Bottom-hole pressure at well w (bar)*

$P_{c,i}$ *Compression inlet pressure for well platform i (equivalently common header pressure) ($i \in \mathcal{N}_{wp,c}$) (bar)*

$P_{t,w}$ *Flowing tubing-head pressure at well w (bar)*

$Q_{a,(i,j)}$ *Volumetric flowrate at standard conditions in arc (i,j), $((i,j) \in \mathcal{A}_q)$ (hm^3/day)*

$q_{a,(u,v)}$ *Flowrate in the arcs connecting the supply/demand nodes $((u,v) \in \mathcal{E}_s)$ (hm^3/day)*

$q_{c,i}$ *Production rate for contract $i \in \mathcal{C}^S$ (hm^3/day)*

$Q_{Ls,i}$ *NGL production rate from field $i \in \mathcal{F}_w$ (m^3/day)*

$Q_{Lw,w}$ *NGL production rate from well $w \in \mathcal{W}$ (m^3/day)*

$Q_{s,i}$ *Production rate at source or sink node $i \in \mathcal{N}_s$ (hm^3/day)*

$q_{s,i}$ *Production rate for nodes in contract network representation (negative for sinks) $i \in \{p_s, p_d\}$, $p \in \mathcal{C}$ (hm^3/day)*

$Q_{w,w}$ *Dry gas production from well w (hm^3/day)*

$s_{(i,j)}$ *Split fraction $(i,j) \in \mathcal{A}_x$*

$t_{(p_i,q_j)}$ *inter-contract transfer rates from contract p to contract q (equivalently flowrate from p_i to q_j in the contract network $((p_i, q_j) \in \mathcal{E}_t)$) ($hm^3/day$)*

W_i *Daily average power consumption of the compressor $i \in \mathcal{N}_{wp,c}$ (MW)*

$w_{d,(i,j)}$ *Dummy pressure variable for reformulation of constraints for arcs in set $\mathcal{A}_y$ (Downstream pressure)*

$w_{u,(i,j)}$ *Dummy pressure variable for reformulation of constraints for arcs in set $\mathcal{A}_y$ (Upstream pressure)*

$y^c_{(.)}$ — *Binary variable representing the operational atomic proposition $C_{(.)}$*

$y^e_{(p_i,p_{i+1})}$ — *Binary variable representing excess in the contract p ($y^e_{(p_i,p_{i+1})} = 1 \Rightarrow E_{p,(p_i,p_{i+1})}$), $(p_i, p_{i+1}) \in \mathcal{E}_{l,a}$*

$y^l_{(i,j)}$ — *Binary variable to indicate if an arc $(i,j) \in \mathcal{A}_y$ is open ($y^l_{(i,j)} = 1$) or closed ($y^l_{(i,j)} = 0$)*

$y^s_{(q_i,q_{i+1})}$ — *Binary variable to indicate whether a transfer succeeded in fulfilling the deficit of the destination contract q ($(q_i, q_{i+1}) \in \mathcal{E}_{l,S}$)*

$y^t_{p,q}$ — *Binary variable representing the status of the contract transfer from contract p to contract q ($(p_i, q_j) \in \mathcal{E}_t$)*

z_g — *Production rate from fields that have priority (hm^3/day)*

z_g — *Total dry gas production rate from the system (hm^3/day)*

z_L — *Total NGL production rate from the system (m^3/day)*