Multistage Stochastic Optimization Under Exogenous and Endogenous Uncertainty

by

Farough Motamed Nasab

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Process Control

Department of Chemical and Materials Engineering
University of Alberta

Abstract

Stochastic optimization is a field in mathematics that deals with decision making under uncertain conditions. Traditional methods for solving stochastic optimization problems such as scenario-tree methods usually result in large-size problems that suffer from the curse of dimensionality and require significant amount of time and computational resources. This study aims at investigating alternative methods referred to as decision-rule methods in order to reduce the problem size while preserving the solution optimality.

Uncertain parameters are generally classified into two categories: exogenous and endogenous uncertainty. Endogenous uncertain parameters are revealed by decisions taken in the course of the problem while exogenous uncertain parameters are independent of the problem decisions. This study investigates both types of uncertainty and employs linear decision rule to model the uncertainty. Two main solution methods are used in this study: partitioning and lifting. In partitioning method, the uncertainty set is partitioned into rectangular segments. At each partition, the binary variable is fixed and the continuous variable is a linear combination of uncertain parameters. In lifting method, the uncertain parameter is lifted to a higher dimensional uncertainty set. The binary variable is formulated using 0-1 indicator functions that results in an adaptive binary variable. The continuous variable is a linear combination of lifted uncertain parameters that results in a piecewise linear solution. At each problem, after applying linear decision rule to the constraints, duality theorem for linear problems is used to convert the constraints to their robust deterministic counterpart.

Chapter 2 introduces a novel hybrid method that combines traditional scenario-tree method

with linear decision rule to reduce the run time and computational expense in large-scale multistage problems. On the other hand, when uncertain coefficients are multiplied at adaptive variables, using linear decision rules results in non-linear constraints with respect to uncertain parameters. In such constraints, duality method for linear constraints can not be employed to convert the constraint to its deterministic dual counterpart. The hybrid method proposed in this chapter introduces a new method to resolve the described problems efficiently.

Chapter 3 addresses multistage stochastic binary problems. Two methods are presented to solve binary problems: lifting method and partitioning method. In the lifting method, the uncertain parameter is lifted to a higher dimension uncertainty set based on pre-defined breakpoints that results in an adaptive binary solution. In the partitioning method, the original uncertain parameters are segmented into rectangular partitions using pre-defined breakpoints. A new technique for break point location optimization is proposed for both lifting and partitioning methods and their performance is compared using an inventory control problem.

Chapter 4 presents a novel piecewise linear decision rule for adaptive continuous variables that results in adaptive discontinuous solution. The presented framework in this chapter can incorporate both adaptive binary and adaptive continuous variables for multistage problems. Two methods are studies in this chapter: lifting and partitioning. The partitioning method provides a better solution quality for smaller problems while the lifting method provides significant computational efficiency for large-scale problems.

Chapter 5 introduces a new general problem formulation where both continuous and binary variables depend on both types of endogenous and exogenous uncertainty. The continuous variable is a linear combination of both types of uncertain parameters. The binary variable is a multiplication of two binary variables, one for endogenous uncertainty formulated using partitioning method and one for exogenous uncertainty formulated using lifting method.

Chapter 6 presents the linear decision rule method for both endogenous and exogenous uncertainty and applies this method to a long term problem for infrastructure construction and production planning of an oil field. The oil price is the exogenous uncertainty and the uncertain oil flow rate is the endogenous uncertainty that depends on oil well drilling and well exploration decisions.

Preface

This thesis was completed under supervision of Dr. Zukui Li. This is a paper-based thesis where each chapter is independent and self-explanatory. The reader can obtain all the required background for each chapter by reading the same chapter. The computer programming is mainly done in MATLAB and GAMS and the script is validated by close collaboration between me and Dr. Li. The submission details and contributions of each researcher is provided below.

Chapter 2 was accepted as Motamed Nasab, Farough, and Zukui Li. "Multistage adaptive optimization using hybrid scenario and decision rule formulation" *AIChE Journal* 65.12 (2019): e16764. Dr. Li proposed ideas about hybrid problem formulation, I simulated the case study and wrote the MATLAB/GAMS script and contributed in the development of the mathematical theory.

Chapter 3 was submitted as Motamed Nasab, Farough, and Zukui Li. "Multistage Adaptive Binary Optimization: Uncertainty set Lifting versus Partitioning through Breakpoint Optimization" *Journal of Global Optimization*, July 2019. The paper is still under review. Dr. Li suggested some ideas about partitioning method and break point location optimization for lifting method. I recommended some ideas about break point optimization for partitioning method and wrote the MATLAB/GAMS script.

Chapter 4 was submitted as Motamed Nasab, Farough, Z. Li. "Multistage Adaptive Stochastic Mixed Integer Optimization through Piecewise Decision Rule Approximation" *Journal of Computers and Chemical Engineering*, April 2020. The paper is undergoing the revision process. Dr. Li presented some ideas about combining binary and continuous variables using the lifting method. I simulated the case studies, wrote the GAMS/MATLAB script and resolved the mathematical difficulties at each stage.

A journal paper is under preparation based on chapter 5 and it's titled as: " Multistage Robust Mixed Integer Optimization Under Endogenous and Exogenous Uncertainty". Dr. Li suggested ideas about combining the endogenous and exogenous uncertainty into a single framework. I resolved the mathematical problems at each step, simulated the case studies and wrote the script.

Chapter 6 was accepted as Nasab, Farough Motamed, Hossein Shahandeh, and Zukui Li. "Strategic planning of steam assisted gravity drainage reservoir development under reservoir production and oil price uncertainty." *Industrial & Engineering Chemistry Research* 57.29 (2018): 9513-9526. Dr. Li recommended ideas about using linear decision rule to model endogenous and exogenous uncertainty, I modeled the SAGD case study and wrote the GAMS/MATLAB script. Some ideas about modeling the SAGD process were taken from the previous work done by my colleague Dr. Hossein Shahandeh.

Dedicated to my parents for their lifetime love, kindness and encouragement.

Acknowledgement

I am sincerely thankful to my supervisor Dr. Zukui Li for providing me with this great opportunity and for his continuous support and supervision. It was my pleasure to learn various aspects of optimization under his guidance.

I am wholeheartedly grateful to my parents and family for their love, support and inspiration.

I would like to sincerely thank all the staff in the University of Alberta who contributed in creating a pleasant environment.

The researchers in this study greatly acknowledge the financial contribution from NSERC during the course of this study.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Stochastic programming is an important field in optimization that aims at making the best decisions under uncertain conditions. It has applications in various fields of science and engineering such as operations management [4, 5], control engineering [6, 7], finance [8, 9] and process systems engineering [10, 11].

The traditional method in stochastic programming is referred to as scenario-tree method. In this method, the uncertain parameters are discretized and a deterministic problem formulation is obtained. In multistage stochastic optimization, the uncertain parameters are revealed over time and decisions are made at each stage based on the revealed uncertainty. Traditional scenario-tree methods result in a large problem size specially in multistage problems and require significant amount of run time and computational resources. This research study addresses this problem by proposing alternative methods that result in reduced problem size that require less computational resources and consequently shorter run time while the solution quality is mainly preserved.

Decision rule method has been employed in this study as an alternative or complementary method to traditional scenario-tree method. Decision rules are categorized as linear and nonlinear. In this study linear decision rules for continuous and binary variables have been employed. Linear decision rules, due to their simple structure, provide the scalability required for large scale problems. This simple structure results in a reduced problem size and shorter run time but suboptimal solution. In order to address the sub-optimality of the solution in linear decision rules, two different methods are mainly used in this study: lifting and partitioning. In the lifting method, the original uncertain parameter is lifted to a higher dimensional uncertainty set. Two different lifting methods are utilized for binary and continuous variables that result in adaptive binary and adaptive continuous variables. In the partitioning method, the original uncertainty set is parti-

tioned in rectangular segments. in each partition, the binary variables are fixed and the continuous variables are linear functions of the original uncertain parameters. The performance of lifting and partitioning methods is evaluated in several case studies.

1.2 Literature review

The traditional method to solve stochastic optimization problmes is called scenario-tree method. In this method, the uncertain parameters are discretized and presented as scenarios. Based on these scenarios, the stochastic problem is reformulated as a determinisitc problem. This approach results in a very large problem size particularly in multistage settings. In order to address this problem, researchers have proposed the decision rule method. In decision rule method, the problem variables are formulated as functions of uncrertain parameters. Decision rules are classified as linear and non-linear. There are two main types of uncertainty: exogenous and endogenous. Endogenous uncertain parameters are revealed based on the decision taken in the course of the problem while exogenous parameters are independent of the problem decisions. In multistage stochastic problems, decisions are made sequentially at each stage based on the revealed uncertainty. In the following, a comprehensive literature review about different methods in stochastic optimization is presented.

Applications of scenario-tree method are abundant in the literature. To name a few, Ahmed et al. [12] addressed a multi-period investment model for capacity expansion of a chemical plant under uncertain demand and cost. They used a heuristic scheme and a branch and bound algorithm to solve the problem to optimality. Birge and Rosa [13] studied the problem of greenhouse gas policy decision-making under economic uncertainty. Escudero et al. [14] investigated production and capacity planning problems to assist in raw material supply sourcing decision under demand uncertainty. Takriti et al. [15] developed a model and a solution method for power generation decision making when demand is uncertain. For a review on stochastic optimization, the reader can refer to [16, 17, 18] .

Since scenario-based multistage problems generally suffer from the curse of dimensionality (exponential growth of model size for multistage problems), researches have tried to address this problem using different methods. Apap and Grossmann, 2017 [19] proposed a sequential scenario decomposition heuristic and a Lagrangean decomposition method. Goel and Grossmann, 2006 [20] presented a Lagrangean branch and bound algorithm to reduce the model size. Gupta and Grossmann, 2014 [21] developed a new Lagrangean decomposition algorithm for solving large-scale stochastic problems with endogenous uncertainty that can reduce the computational expense by

reducing the number of non-anticipativity constraints. Colvin and Maravelias, 2010 [11] developed novel branch and cut algorithms where non-anticipativity constraints that are unlikely to be active are removed from the formulation and added only if they are violated within the search tree.

Recent advances in robust optimization has enabled the researchers to use decision rule methods in the context of stochastic optimization. In decision rule methods, the adaptive variables are functions of uncertain parameters. Among different decision rule methods, linear decision rule has received considerable attention since it provides the scalability required for large-scale multistage problems while solution optimality is sacrificed. In this formulation, the adaptive variables are linear functions of uncertain parameters. Ben-Tal et al. were among the first who applied linear decision rules in the context of adaptive robust optimization [22]. They reformulated the stochastic problem into a robust deterministic counterpart by applying duality method for linear problems. Some application instances for linear decision rules is reviewed here. Skaf and Boyd[7] designed a general affine controller in which the control input is an affine function of all previous measurements to minimize a convex objective. They illustrated the method with applications in supply chain management optimization and dynamic portfolio optimization. Calafiore [8] addressed a portfolio optimization problem using an affine parametrization of the recourse policy that provides a sub-optimal but explicit formulation that can be used to solve multistage problems with many constraints and periods. Atamtürk and Zhang [23] presented a two-stage robust optimization method for network flow with uncertain demand and provide applications for lot-sizing and location transportation problems and compared the results with single-stage robust optimization and two-stage scenario-based optimization. Decision rule based method has also received attention in the process systems engineering community. Lappas and Gounaris [24] developed a multistage adjustable robust optimization framework that accounts for inherent endogenous uncertainty in process scheduling by employing decision-dependent uncertainty sets. Ning and You [25] presented a data-driven framework for adaptive optimization using data and applied the proposed framework on two industrial applications of process scheduling and process network planning. Zhang et al. [26] developed a scheduling model for continuous industrial processes that provide interruptible load. The uncertainty in the timing of the load reduction request is modeled by an adjustable robust optimization approach that integrates recourse decisions using linear decision rules.

As described above, the simple structure of linear decision rules results in a sub-optimal solution. In order to address this problem, different researchers have proposed non-linear decision rules for adaptive real-valued variables. The purpose of non-linear decision rules is to improve the solution optimality while preserving the scalability required for multistage problems. Chen et al. [27] proposed new decision rule structures, segregated and deflected linear decision rules and demonstrated that these proposed methods can outperform sampling approaches when limited information about the underlying probability distributions is available. Chen and Zhang [28]

presented a splitting-based extended affinely adjustable formulation and showed that their method is tractable and scalable to multistage problems. Georghiou et al. [29] proposed piecewise linear continuous decision rule based on axial segmentation and lifting of the uncertain parameters. They also proposed a method for piecewise linear continuous decision rule based on general segmentation. The authors recommended ideas for nonlinear continuous decision rules such as quadratic, power, monomial, inverse monomial and multilinear liftings and compared different methods on a dynamic inventory control problem. Goh and Sim [30] studied distributionally robust optimization problem and applied the linear decision rule method to get tractable approximation. Bampou and Kuhn [31] presented a polynomial decision rule to solve multistage stochastic problems. They estimated the suboptimality of the decision rule by solving a dual version of the problem in polynomial decision rule. Bertsimas et al. [32] introduced a hierarchy of near-optimal polynomial policies and showed that these policies can be computed by solving a single semidefinite programming problem. They evaluated the framework using three classical applications in the context of inventory management and robust regulation of a suspension system. Recently, Avraamidou and Pistikopoulos [33] proposed a method based on generalized affine decision rules for linear mixed integer robust optimization problems using multi-parametric programming and showed that the method can find the exact global solution.

Although different methods have been proposed for real-valued non-linear decision rule structures, the available literature for decision rules addressing adaptive binary variables is limited. Bertsimas and Caramanis [34] proposed a linear decision rule and approximated the semi-infinite optimization problem using a sampling algorithm. Bertsimas and Georghiou [35] suggested a structure that can provide near-optimal solutions but restricted scalability. Hanasusanto et al. [36] presented a binary decision rule based on the previous work done by Bertsimans and Caramanis [37] that can only be applied to two-stage problems. Postek and den Hertog [38] presented a method for iterative splitting of uncertainty set that can be used for adaptive integer variables and demonstrated the advantage of their method on a capital budgeting and a lot sizing problem. Recently, Bertsimas and Georghiou [39] proposed a binary decision rule that lifts the original uncertain parameters using 0-1 indicator functions. The trade-off between the solution optimality and scalability can be adjusted based on the segmentation resolution over the uncertainty set. The authors suggested a scalable formulation that can be used for large-scale multistage problems. For a comprehensive review on adaptive optimization, the reader can refer to Bertsimas et al., [40], Gabrel et al. [41].

Most of the available studies in the literature address only exogenous or endogenous uncertainty. There are few studies that address both types of uncertainty in a single framework. In the context of stochastic programming, two notable studies deal with this type of problems. Goel and Grossmann, 2006 [20] presented a mixed integer programming formulation and applied Lagrangean

duality based branch and bound algorithm to reduce to number of non-anticipativity constraints. Apap and Grossmann, 2017 [19] proposed a composite scenario tree that captures both types of uncertainty and discussed two solution approaches to reduce the number of non-anticipativity constraints and solve the problem. A sequential scenario decomposition heuristic and a Lagrangean decomposition method. Dupačová (2006) [42] provided only a general description of problems under both types of uncertainty but did not present a solution framework or numerical results.

Although not addressed in this study, another method that has been used to solve multistage dynamic problems is approximate dynamic programming (ADP). To mention few examples, Jay Lee and Jong Lee [43] applied ADP method to solve Markov decision processes (MDPs) for control and scheduling of process problems. The same authors in a different study [44], proposed two strategies based on APD method for data-driven control of non-linear processes. For a general study of the ADP method and its applications in different contexts, the reader can refer to [45] and for a review on applicability of ADP in process control please refer to [46].

The decision rule approach employed in this study have inherent similarities with the ADP method in that both methods use linear combinations of basis functions to approximate nonlinear functions. The decision rule approach applies these approximations to future adaptive decisions while the ADP method applies the approximations to cost functions. While the ADP method has the flexibility to solve non-convex problems, the decision rule method used in this study, can only be applied to convex problems. Although not investigated in this study, one advantage of the decision rule method is that error bounds can be calculated for the obtained solution [47, 48, 49, 50], however studies that address similar bounds for ADP method are rare.

This thesis proposes novel solution methods for endogenous and exogenous uncertainty independently and developes new frameworks that integrate both types of uncertainty into a signle mathematical framewrok.

1.3 Thesis contribution

This study endeavors to address the available gap in the literature in the field of stochastic optimization. In brief, chaper 2 presents a new hybrid method that combines scenario-tree and linear decision rule. Chapter 3 proposes a new breakpoint location optimization technique for lifting and partitioning methods. Chapter 4 combines the lifting method for binary and continuous variables into a signle framework and proposes a new decision rule for continuous variables that results in a piece-wise discointuous solution with significant computational efficiency. Chapter 5 proposes a new framework that integrates the exogenous and endogenous uncertainty into a single framework using lifting and partitioning methods. Chapter 6 applies linear decision rule to a long term case

study about infrastrusutre construction and production planning of an oil field developed using SAGD method. The contributions of this work are expalined in the following.

1.3.1 Multistage Adaptive Optimization Using Hybrid Scenario and Decision Rule Rormulation

Chapter 2 introduces a new hybrid method that combines traditional scenario-tree method with linear decision rules. The proposed method addresses problems where random coefficients are multiplied at adaptive variables. In such cases, traditional scenario-tree method results in a very large problem size. Decision rule method results in a non-linear problem with respect to uncertain parameters and thus it would not be possible to apply duality method to convert the problem into its robust deterministic counterpart. The developed hybrid method resolves these issues and reduces the run time and computational expense in large-scale multistage problems.

1.3.2 Multistage Adaptive Binary Optimization: Uncertainty Set Lifting Versus Partitioning Through Breakpoint Optimization

Chapter 3 introduces a novel breakpoint location optimization method for adaptive binary variables. The uncertainty is modeled using two different methods: lifting and partitioning. The breakpoint optimization method can be applied to both lifting and partitioning methods. Compared to fixed breakpoint formulation, it provides better solution quality using fewer number of breakpoints for small size problems.

1.3.3 Multistage Adaptive Stochastic Mixed Integer Optimization Through Piecewise Decision Rule Approximation

Chapter 4 proposes a novel mathematical framework that integrates adaptive continuous and adaptive binary variables into a single framework based on the lifting method. The proposed decision rule for adaptive continuous variables is very flexible and can provide discontinuous piecewise linear solution. The proposed framework is compared to the partitioning method using an inventory control case study. The results demonstrates that for large-scale multi-stage problems, the proposed framework provides significant computational efficiency compared to partitioning method while the partitioning method can provide better solution quality for small size problems.

1.3.4 Multistage Stochastic Mixed Integer Optimization Under Endogenous and Exogenous Uncertainty

Chapter 5 proposes a new framework that incorporates both exogenous and endogenous uncertainty into a single framework for both binary and continuous variables. The exogenous uncertainty is modelled using lifting method and endogenous uncertainty is modeled using partitioning method. The continuous variable is a linear combination of both exogenous and endogenous parameters and the binary variable is a multiplication of two binary variables that depend on exogenous and endogenous uncertainty independently.

1.3.5 Strategic Planning of SAGD Reservoir Development Under Reservoir Production and Oil Price Uncertainty

Chapter 6 presents a case study about infrastructure construction and production planning of an oil filed extracted using the Steam Assisted Gravity Drainage (SAGD) method. In this case study, the linear decision rule is used to model the exogenous oil price uncertainty and the endogenous oil production uncertainty. Obtained results demonstrate that linear decision rule can successfully be applied to multistage stochastic problems.

Chapter 2

Multistage Adaptive Optimization Using Hybrid Scenario and Decision Rule Formulation

Abstract

Scenario-based stochastic programming and linear decision rule-based robust optimization are prevalent methods for solving multistage adaptive optimization problems. In practical applications such as capacity expansion planning of chemical processes, often multiple sources of uncertainty affect the problem which introduces challenges to traditional stochastic optimization methods. While a large number of uncertain parameters exist in the problem, using scenario-based method results in very large problem size and the solution becomes computationally expensive. In addition, when the constraints include multiplication of uncertain parameters and adaptive variables, the constraints are not linear with respect to uncertain parameters when the linear decision rule method is used. In order to address these challenges, we propose two different hybrid methods where scenario and decision rule methods are combined to solve the multistage adaptive optimization problem. The paper demonstrates the computational performance of the proposed hybrid methods using two chemical process planning examples.

2.1 Introduction

Multistage adaptive optimization (MSAP) is an important technique for addressing dynamic decision making under uncertainty. It has received lots of applications in process systems engineering, including planning, scheduling and supply chain management [51]. Compared to deterministic optimization approach where the uncertainty is ignored in the problem modeling, adaptive optimization accounts for uncertainty and therefore results in a more reliable solution.

Two main approaches are used in the literature to solve MSAP problems: scenario-based stochastic programming method and decision rule-based robust optimization method. In scenario formulation, the uncertainty is presented using scenario or scenario tree and the problem size depends on the number of uncertain parameters [52]. The obtained solution is feasible for all the predefined uncertainty scenarios. Different studies have used scenario-based method to solve stochastic problems. Jonsbråten [53] presented the problem of optimal development of an oil field under uncertain future oil prices. Goel and Grossmann [54] developed an optimal investment and operational planning algorithm for gas field development under uncertainty in gas reserves. Gupta and Grossmann [10] proposed a strategic model for offshore oilfield development problem where the objective is to maximize the NPV over a long-term horizon. Colvin and Maravelias [11] discussed methods for solving a multi-stage stochastic problem in pharmaceutical research for resource-constrained scheduling of clinical trials. In general, scenario tree-based approximation method suffers from the issue of dimensionality. The number of scenarios may quickly increase such that the problem is computationally intractable; on the other hand, using very few number of scenarios can result in suboptimal or impractical solutions.

In order to address this problem in multistage adaptive optimization, linear decision rule (LDR) technique was proposed by Ben-Tal et al. [22] as an alternative. In linear decision rule method, uncertainty-dependent variables (adjustable variables) are modeled as affine functions of uncertain parameters. Subsequently, the constraints are converted to their deterministic counterparts using duality theorem. This approach reduces the problem size significantly. The solution is feasible for any realization of uncertainty within the predefined uncertainty set.

Decision rule techniques have been successfully applied within the field of multistage adaptive optimization. For instance, Shapiro and Nemirovski [55] presented a discussion on the complexity of multi-stage stochastic problems using linear decision rule. Goh and Sim [30] developed new piecewise linear decision rules that allow a more flexible reformulation of the original problem. Kuhn et al. [56] presented methods to estimate the approximation error introduced by linear decision rules and claimed that their method remains applicable for stochastic problems with random recourse and can be extended to problems that include ambiguous probability distributions.

Vayanos et al. [57] proposed a technique to solve stochastic problems where the uncertainty revelation depends on the decisions. They introduced binary variables to indicate the information revealing status and presented a method for approximating binary decisions by piecewise constant functions and continuous decisions by piecewise linear functions of uncertainty. Ben-Tal et al.[58] employed the affinely adjustable robust counterpart method to study a supply chain problem under uncertain demand know as retailer-supplier flexible commitment problem. Lorca et al.[59] presented a multistage adaptive optimization model for unit commitment (UC) of power systems under uncertain electricity loads which is the most critical daily operational problem of power systems. The authors employed the affine policy method to enable them to deal with large scale problems and developed a solution method based on constraint generation. They demonstrated that the proposed method can significantly outperform the deterministic and two-stage robust UC models. Zhang et al.[26] developed a scheduling model for continuous industrial processes that provide interruptible load. The uncertainty in the timing of load reduction request is modeled by an adjustable robust optimization approach that integrates recourse decisions using linear decision rules.

Various decision rule-based and scenario-based methods have been proposed for multistage adaptive problems. Some recent research efforts in this context are summarized as in the following. Adaptive binary or adaptive integer decisions has been addressed by some authors. Bertsimas and Georghiou [35] discussed that the existing decision rules are restricted by their a priori design and not incorporating adaptive binary decisions in their modeling. In order to address these shortcomings, they derived a new decision rule structure that models continuous variables as piecewise linear and binary variables as piecewise constant. They also proposed a method for optimal design of the decision rule that utilizes finite number of pieces and solves the problem using mixed integer optimization. In another work, Bertsimas and Georghiou [39] proposed a linearly parametrized binary decision rule structure that can be reformulated to its deterministic counterpart and can be used along with real-valued continuous decision rules available in the literature. For problems where the problem size grows exponentially with respect to problem data, they proposed a systematic method that trades off scalability and optimality. They explained that the proposed method is highly scalable and can practically be used for large scale problems while the available binary decision rules in the literature suffer from limited scalability and are mainly confined to worst-case problems. Postek and Den Hertog [38] presented a method for constructing decision rules for continuous and integer variables in multi-period linear optimization problems. They provided theoretical evidence showing that splitting the uncertainty set can improve the worst-case objective and proposed several splitting heuristics based on this theory.

Other researchers addressed stochastic problems where uncertainty revelation is decision dependent (endogenous uncertainty). Lappas and Gounaris [24] developed a multistage adjustable robust

optimization framework that accounts for inherent endogenous uncertainty in process scheduling by employing decision-dependent uncertainty sets. In other work [60], the same authors presented algorithmic considerations for decision-dependent uncertainty sets to reach guaranteed optimality and demonstrated that this approach can eliminate conservatism with respect to optimal decisions.

Data-driven approaches are addressed by other authors. Ning and You [61] presented a data-driven approach for multistage adaptive robust optimization. The proposed framework incorporates distributional information to avoid over-conservatism. Probability distributions are extracted from uncertain data by employing kernelized weighted least square algorithm. The uncertainty set is data-driven such that the bounds of uncertain parameters are defined by quantile functions to integrate the uncertainty set into the optimization framework. In a related work [25], the same authors proposed a data-driven framework for adaptive optimization that utilizes big data available in process industries. The authors mention that the proposed algorithm accounts for correlation, asymmetry and multimode of uncertainty data, therefore it generates less conservative solutions and the solution is robust to parameter variations and anomalous measurements. They demonstrated the advantage of the proposed framework on two industrial applications of process scheduling and process network planning.

Reducing the problem size and improving the existing techniques have been the main focus of other authors. Christian and Cremaschi [62] proposed two variants of branch and bound algorithm that reduces the problem size and resource requirements for solving large scale multistage stochastic problems with endogenous uncertainty in order to enable their application for real-world size problems. Bertsimas and Dunning [63] proposed a new partition-and-bound method for multistage mixed integer problems. This method first analyses the optimal solution to a static (nonadaptive) model of the multistage mixed integer problem in order to obtain insights into which regions of the uncertainty set restrict the objective value. This information is used to construct partitions in the uncertainty set that results in a finite adaptable formulation and also to determine a lower bound on the fully adaptable solution. This procedure is repeated to improve the objective and to reach a desired gap. Bertsimas and Caramanis [37] proposed an adaptability model with discrete second stage variables for multistage problems. In this method, a hierarchy of increasing adaptability bridges the gap between static robust formulations and the fully adaptable formulation. For a review on recent advances in optimization under uncertainty, the reader can refer to these references [18, 17, 16].

In practical applications, often multiple sources of uncertainty affect the problem. Uncertain parameters can appear as the objective coefficient uncertainty, constraint left-hand-side coefficient or right-hand-side parameter uncertainty. In such cases, using only scenario formulation can result in a very large problem size due to the presence of multiple sources of uncertainty and the solution can become computationally prohibitive. On the other hand, in problems where uncertain coef-

ficients and variables are multiplied together at the constraints, the problem is non-convex with respect to the uncertainty parameters and strong duality theorem cannot be used. Therefore, LDR method can not be employed to solve the problem. As explained above, the presence of multiple sources of uncertainty restricts the traditional scenario and LDR approaches used to solve multi-stage stochastic programming problems. This type of problems that include multiplication of static uncertain parameters and adaptive decision variables are rarely addressed in the literature [64]. In order to address these challenges, hybrid strategies that combine scenario and LDR methods are proposed in this study.

In the first method, time-dependent uncertainties are modeled using scenario tree and the corresponding adjustable variables are scenario-dependent. The remaining static (time-independent) uncertainties are modeled using uncertainty set. The resulting semi-infinite constraints with respect to static uncertainty, are reformulated to their deterministic counterparts using duality theorem. In this method, both binary and continuous variables are scenario-dependent. This approach results in a scenario-dependent dynamic solution that provides a solution that is statically robust with respect to static uncertainty at each node of the tree.

In the second hybrid method, time-dependent uncertainty is modeled using uncertainty set and the corresponding variables are modeled using linear decision rules with respect to time-dependent uncertainties. Static (time-independent) uncertainty is modeled as scenarios in finite sets and the corresponding constraints are enforced for all these scenarios. Subsequently, the resulting semi-infinite constraints are reformulated to their deterministic robust counterparts with respect to time-dependent uncertainty. In this method, binary variables are scenario-dependent and continuous variables are affine functions of time-dependent uncertainty. This method results in a dynamic decision rule solution that is feasible for any realization of time-dependent uncertainty within the predefined uncertainty set and for all scenarios of static uncertainty.

The paper is organized as follows. We first present the general hybrid formulation for problems that include only adaptive continuous variables and demonstrate the application of the developed hybrid methods on a simple problem with only adaptive continuous variables. Then, we present the general hybrid formulation for problems that include both continuous and binary variables and apply the developed hybrid method on a computationally demanding case. The results are discussed and conclusions are made in the last section.

2.2 MSAP with adaptive continuous variables

In this section, the following multistage adaptive optimization problem (MSAP) with only continuous variables is studied:

$$\min \quad \rho\left(\sum_t c_t(\xi_{[t]})^\top x_t(\xi_{[t]})\right) \tag{2.1a}$$

$$s.t. \quad \sum_{\tau=1}^{t} A_{t,\tau}(\omega)^\top x_\tau(\xi_{[\tau]}) \leq b_t(\xi_{[t]}) \quad \forall t, \xi \in \Xi, \omega \in \Omega \tag{2.1b}$$

where $\rho(\cdot)$ is a general risk measure, ω denotes uncertainty that is not modeled as a function of time (which may happen due to practical restrictions on the information availability), ξ_t denotes the uncertain parameter of time period t, and $\xi_{[t]} = [1, \xi_1, ..., \xi_t]$ represents the observed uncertainty up to time t (where element 1 is used for intercept term in linear decision rule). Notice that the decision variables are denoted as adaptive function of the observed uncertainty $\xi_{[t]}$, which means that the decisions are adjustable with respect to ξ. Also, notice that the constraints are enforced for any possible realization of the uncertainty within predefined set. In this sense, we are searching for an optimal solution that is adaptively robust with respect to ξ and statically robust with respect to ω. We further made the following assumptions for the above formulation:

- Constraint coefficients on the left-hand-side are linearly dependent on static uncertainty (time-independent uncertainty): $A_{t,\tau}(\omega) = A_{t,\tau}\omega$
- Constraint coefficients on the right-hand-side depend only on time-dependent uncertainty: $b_t(\xi_{[t]}) = B_t\xi_{[t]}$, and ξ is independent of ω
- Objective coefficients depend on time-dependent uncertainty: $c_t(\xi_{[t]}) = C_t\xi_{[t]}$

Then the MSAP model can be written as:

$$\min \quad \rho\left(\sum_t \xi_{[t]}^\top C_t^\top x_t(\xi_{[t]})\right) \tag{2.2a}$$

$$s.t. \quad \sum_{\tau=1}^{t} \omega^\top A_{t,\tau}^\top x_\tau(\xi_{[\tau]}) \leq B_t\xi_{[t]} \quad \forall t, \xi \in \Xi, \omega \in \Omega \tag{2.2b}$$

In the following subsections, by applying different risk measures $\rho(\cdot)$, the above model is cast into two different contexts: multistage stochastic programming (MSSP) and multistage robust optimization (MSRO).

2.2.1 MSSP formulation

We present two hybrid scenario and decision rule formulations to address the multistage adaptive optimization problem. In the first method (Hybrid 1), time-dependent uncertainty is modeled using scenario tree and the corresponding adjustable variables are modeled as scenario-dependent (i.e., node-dependent). The remaining static (time-independent) uncertainties are modeled using uncertainty set. The resulting semi-infinite constraints with respect to static uncertainty are reformulated to their deterministic counterparts using duality theorem. This approach results in a node-dependent dynamic solution that is robust with respect to static uncertainty at each node of the tree. Using the proposed Hybrid 1 method, the problem formulation can be stated as in the following:

$$\min \quad \sum_{s\in S} p_s \sum_t c_{t,s} x_{t,s} \tag{2.3a}$$

$$s.t. \quad \sum_{\tau=1}^{t} \omega^\top A_{t,\tau}^\top x_{\tau,s} \le b_{t,s} \quad \forall t, s, \omega \in \Omega \tag{2.3b}$$

$$x_{t,s} = x_{t,s'} \quad \forall \{t, s, s'\} \in SP \tag{2.3c}$$

where p_s is the probability of scenario s, $x_{t,s}$ denotes the scenario-dependent variable, $c_{t,s}$ and $b_{t,s}$ denote the parameter values under scenario s. Set SP defines the set where non-anticipative constraint should be applied. That is, at stage t, the decision x_t is determined over scenarios s and s' that share the same path up to stage t.

In the second method (Hybrid 2), time-dependent uncertainty is modeled using uncertainty set and the corresponding variables are modeled using linear decision rules. Static (time-independent) uncertainty is modeled as samples in a finite set: $l \in \mathcal{L}$. The corresponding constraints are enforced for all these samples. Subsequently, the resulting semi-infinite constraints are reformulated to their deterministic counterparts with respect to time-dependent uncertainty. This method results in a dynamic decision rule solution that is feasible for any realization of time-dependent uncertainty within predefined uncertainty set and for all samples of static uncertainty. In Hybrid 2 model, the observed uncertainty is related to the vector of all uncertain parameters using observation matrix P_t where $\xi_{[t]} = P_t \xi$ and linear decision rule is applied to the continuous variable: $x_t(\xi_{[t]}) = X_t \xi_{[t]} = X_t P_t \xi$. It should be noted that in this work, all linear decision rules include an intercept term. With this, the corresponding model can be written as:

$$\min \quad \mathbb{E}\left(\sum_t \xi^\top P_t^\top C_t^\top X_t P_t \xi\right) \tag{2.4a}$$

$$s.t. \quad \sum_{\tau=1}^{t} \omega_l^\top A_{t,\tau}^\top X_\tau P_\tau \xi \le B_t P_t \xi \qquad \forall t, l, \xi \in \Xi \tag{2.4b}$$

where the uncertainty set Ξ can be any type (e.g., polyhedral, ellipsoidal, etc.) that makes the linear semi-infinite constraint traceable. The model can be further equivalently rewritten as:

$$\min \quad \mathrm{Tr}\left(\sum_t P_t \mathbb{E}[\xi\xi^\top] P_t^\top C_t^\top X_t\right) \tag{2.5a}$$

$$s.t. \quad \sum_{\tau=1}^{t} \omega_l^\top A_{t,\tau}^\top X_\tau P_\tau \xi \le B_t P_t \xi \qquad \forall t, l, \xi \in \Xi \tag{2.5b}$$

where $\mathbb{E}(\cdot)$ and $\mathrm{Tr}(\cdot)$ are the expectation and trace operator, respectively.

2.2.2 MSRO formulation

Next, we present the formulation under the worst-case performance risk measure. The corresponding problem can be written as:

$$\min \quad \max_{\xi\in\Xi}\left(\sum_t \xi_{[t]}^\top C_t^\top x_t(\xi_{[t]})\right) \tag{2.6a}$$

$$s.t. \quad \sum_{\tau=1}^{t} \omega^\top A_{t,\tau}^\top x_\tau(\xi_{[\tau]}) \le B_t \xi_{[t]} \qquad \forall t, \xi \in \Xi, \omega \in \Omega \tag{2.6b}$$

It is equivalently reformulated as:

$$\min \quad z \tag{2.7a}$$

$$s.t. \quad \sum_{\tau=1}^{t} \omega^\top A_{t,\tau}^\top x_\tau(\xi_{[\tau]}) \le B_t \xi_{[t]} \quad \forall t, \xi \in \Xi, \omega \in \Omega \tag{2.7b}$$

$$\sum_t \xi_{[t]}^\top C_t^\top x_t(\xi_{[t]}) \le z \qquad \forall \xi \in \Xi \tag{2.7c}$$

Following the same idea behind MSSP, the MSRO model in hybrid 1 method can be written as:

$$\min \quad z \tag{2.8a}$$

$$s.t. \quad \sum_{\tau=1}^{t} \omega^\top A_{t,\tau}^\top x_{\tau,s} \le b_{t,s} \quad \forall t, s, \omega \in \Omega \tag{2.8b}$$

$$\sum_t c_{t,s} x_{t,s} \le z \qquad \forall s \tag{2.8c}$$

$$x_{t,s} = x_{t,s'} \qquad \forall \{t, s, s'\} \in SP \tag{2.8d}$$

and the MSRO model in Hybrid 2 method is formulated as:

$$\min \quad z \tag{2.9a}$$

$$s.t. \qquad \sum_{\tau=1}^{t} \omega_l^\top A_{t,\tau}^\top X_\tau P_\tau \xi \leq B_t P_t \xi \quad \forall t, l, \xi \in \Xi \tag{2.9b}$$

$$\sum_{t} \xi^\top P_t^\top C_t^\top X_t P_t \xi \leq z \qquad \forall \xi \in \Xi \tag{2.9c}$$

In the above models, the semi-infinite constraints can be converted to their deterministic robust counterpart and finally the problem can be solved using a deterministic solver. Note that the second constraint (2.9c) of the above model (9) contains a quadratic function of the uncertainty and semi-definite programming (SDP) type of robust counterpart is available for ellipsoidal uncertainty set.

The above two hybrid methods are summarized in Table 4.2.

Table 2.1: Summary of the proposed hybrid formulation characteristics

Method	Uncertainty model	Continuous Variable
Hybrid 1	Static: Set Time-dependent: Scenario tree	Scenario dependent
Hybrid 2	Static: Samples Time-dependent: Set	Decision Rule

2.3 Illustrating example

An illustrative example that includes only continuous variables is employed to demonstrate the application of Hybrid 1 and 2 methods. In this problem, there are two technologies (T_1,T_2) that produce a single product A (Fig. 2.1). A yield factor γ_i is associated with each technology that indicates its contribution to production. The problem consists of two time periods. At the beginning of each time period t, an installation decision $x_{i,t}$ is made. If the demand of the corresponding time period d_t is not met, then some outsourcing decision y_t is made (Fig. 2.2). There is also a restriction on the total amount of production using both technologies at each time period. The objective is to minimize the overall cost of production and outsourcing while the demand at each time period is met. Outsourcing cost is higher than production cost, therefore it is preferred to use production rather than outsourcing to reduce the total cost.

The optimization model is given as follows. Eq. 2.10a is the objective function which consists of the installation and outsourcing costs, where the expectation operator is applied to uncertainty vector ξ. Eqs. 2.10b and 2.10c indicate the production restriction for the first and second time periods. Eqs. 2.10d and 2.10e apply the demand constraint for the first and second time periods,

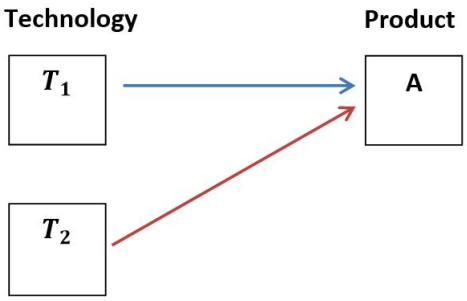


Figure 2.1: Technology and product relationship

respectively.

$$\min \quad \sum_i \alpha_i x_{i,1} + \mathbb{E}_\xi \left[\sum_i \alpha_i x_{i,2}(\xi) + \sum_{t=1}^{2} \beta y_t(\xi) \right] \tag{2.10a}$$

$$s.t. \quad \sum_{i=1}^{2} x_{i,1} \leq 100 \tag{2.10b}$$

$$\sum_{i=1}^{2} x_{i,2}(\xi) \leq 100 \qquad \forall \xi \in \Xi \tag{2.10c}$$

$$\sum_i \gamma_i(\omega) x_{i,1} + y_1(\xi) \geq d_1(\xi) \qquad \forall \xi \in \Xi, \omega \in \Omega \tag{2.10d}$$

$$\sum_i \gamma_i(\omega)(x_{i,1} + x_{i,2}(\xi)) + y_2(\xi) \geq d_2(\xi) \qquad \forall \xi \in \Xi, \omega \in \Omega \tag{2.10e}$$

$$x_{i,t}(\xi) \geq 0 \qquad \forall i, t, \forall \xi \in \Xi \tag{2.10f}$$

$$y_t(\xi) \geq 0 \qquad \forall t, \forall \xi \in \Xi \tag{2.10g}$$

Unit installation cost α_i for technologies T_1, T_2 is 1, 3, respectively. Nominal yield factor γ_i for technologies T_1, T_2 is 1 and 3.4, respectively. Outsourcing cost β is 7. The nominal demand at time periods 1 and 2 is 540 and 615, respectively. In the stochastic model, the yield $\gamma_i(\omega)$ and demand $d_t(\xi)$ parameters are considered to be uncertain. As a result, installation $x_{i,t}(\xi)$ and outsourcing $y_t(\xi)$ decisions depend on uncertainty. The parameter ξ represents the uncertainty vector which includes demand uncertainty.

The stochastic problem is solved using two different methods. In Hybrid 1, yield uncertainty is modeled by constructing a set and demand uncertainty is modeled by generating time-dependent scenarios. In Hybrid 2, an uncertainty set is built for demand uncertainty and time-independent scenarios are considered for yield uncertainty. Hybrid 1 and Hybrid 2 methods are explained in detail in sections 2.3.1 and 2.3.2, respectively.

2.3.1 Hybrid 1 method

In this illustrating example, two sources of uncertainty exist: demand and yield. Demand is time-dependent and yield is static. Demand uncertainty is modeled using scenario-tree representation (Fig. 2.3a) and the decision variables (installation $x_{i,k}$ and outsourcing y_k) are node-dependent. Yield is a static uncertainty and it is modeled using uncertainty set.

Demand uncertainty: As Fig. 2.3a illustrates, a scenario tree is constructed to represent demand uncertainty where each node corresponds to a demand scenario at a certain time period. Node $k = 1$ corresponds to $t = 0$, nodes $k = 2, 3, 4$ correspond to $t = 1$ and nodes $k = 5, 6, \cdots, 13$ correspond to $t = 2$. The probability of occurrence is $\frac{1}{3}$ for the three branches of each node. Table 2.2 provides the demand values for each node of the tree.

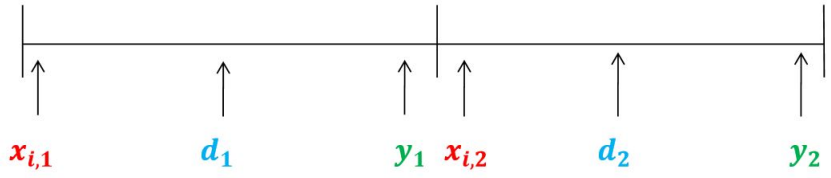


Figure 2.2: Decision making and uncertainty revelation sequence

Table 2.2: Demand for product A at node k, d_k

k	1	2	3	4	5	6	7	8	9	10	11	12	13
d_k	-	520	540	560	540	615	690	540	615	690	540	615	690

Yield uncertainty: The uncertain yield parameter is modeled as 10% perturbation on some nominal yield values $\gamma_i = \bar{\gamma}_i(1 + \omega_i)$. Eqs. 2.11 and 2.12 describe the yield uncertainty set and Fig. 2.3b illustrates the set.

$$|\omega_i| \leq 0.1 \qquad \forall i \tag{2.11}$$

$$\sum_i |\omega_i| \leq 0.15 \qquad \forall i \tag{2.12}$$

To simplify robust counterpart derivation, we use the vector notation of uncertain parameters $\omega = [\omega_1, \omega_2]^\top$. For this purpose, a truncate operator O_i is used such that $O_i\ \omega = \omega_i$. O_i is a row vector in which the i-th element is 1 and all the rest elements are zero (e.g. $O_2 = [0, 1]$ for $i = 2$); therefore yield can be written as:

$$\gamma_i = \bar{\gamma}_i(1 + O_i\omega) \qquad \forall i, \omega \in \Omega \tag{2.13}$$

Eqs. 2.11 and 2.12 can be written in the following inequality form which is a polyhedral set:

$$\Omega = \{\omega : W\omega \leq V\} \tag{2.14}$$

Based on the above uncertainty modeling, the stochastic model can be written as:

$$\min \quad \sum_i \alpha_i x_{i,1} + \frac{1}{3}\sum_{k=2}^{4}\left[\sum_i \alpha_i x_{i,k} + \beta y_k\right] + \frac{1}{9}\sum_{k=5}^{13}\beta y_k \tag{2.15a}$$

$$s.t. \quad \sum_{i=1}^{2} x_{i,1} \leq 100 \tag{2.15b}$$

$$\sum_{i=1}^{2} x_{i,k} \leq 100 \qquad \forall k \in \{2,3,4\} \tag{2.15c}$$

$$\sum_i \bar{\gamma}_i(1 + O_i\omega)x_{i,1} + y_k \geq d_k \qquad \forall k \in \{2,3,4\}, \omega \in \Omega \tag{2.15d}$$

$$\sum_i \bar{\gamma}_i(1 + O_i\omega)(x_{i,1} + x_{i,a(k)}) + y_k \geq d_k \qquad \forall k \in \{5,...,13\}, \omega \in \Omega \tag{2.15e}$$

$$x_{i,k} \geq 0 \qquad \forall i \in \{1,2\}, k \in \{1,2,3,4\} \tag{2.15f}$$

$$y_k \geq 0 \qquad \forall k \in \{2,...,13\} \tag{2.15g}$$

In the model, d_k, $x_{i,k}$ and y_k are the demand parameter, installation and outsourcing variables at each node of the tree respectively. In this formulation, node representation of time-dependent uncertainty is employed that inherently accounts for non-anticipativity; therefore non-anticipativity constraint is not required. The derivation procedure for robust counterparts of semi-infinite constraints 2.15d and 2.15e with respect to yield uncertainty is provided in the appendix. Eqs. 2.16a to 2.16g summarize the final Hybrid 1 formulation.

$$\min \quad \sum_i \alpha_i x_{i,1} + \frac{1}{3}\sum_{k=2}^{4}\left[\sum_i \alpha_i x_{i,k} + \beta\ y_k\right] + \frac{1}{9}\sum_{k=5}^{13}\beta y_k \tag{2.16a}$$

$$s.t. \quad \sum_{i=1}^{2} x_{i,1} \leq 100 \tag{2.16b}$$

$$\sum_{i=1}^{2} x_{i,k} \leq 100 \qquad \forall k \in \{2,3,4\} \tag{2.16c}$$

$$\sum_i \bar{\gamma}_i x_{i,1} - V^\top \theta_k + y_k \geq d_k \qquad \forall k \in \{2,3,4\} \tag{2.16d}$$

$$-W^\top \theta_k = (\sum_i \bar{\gamma}_i x_{i,1} O_i)^\top \qquad \forall k \in \{2,3,4\}$$

$$\theta_k \geq 0 \qquad \forall k \in \{2,3,4\}$$

$$\sum_i \bar{\gamma}_i \ (x_{i,1} + x_{i,a(k)}) - V^\top \phi_k + y_k \geq d_k \qquad \forall k \in \{5, ..., 13\} \tag{2.16e}$$

$$-W^\top \phi_k = (\sum_i \bar{\gamma}_i (x_{i,1} + x_{i,a(k)}) O_i)^\top \qquad \forall k \in \{5, ..., 13\}$$

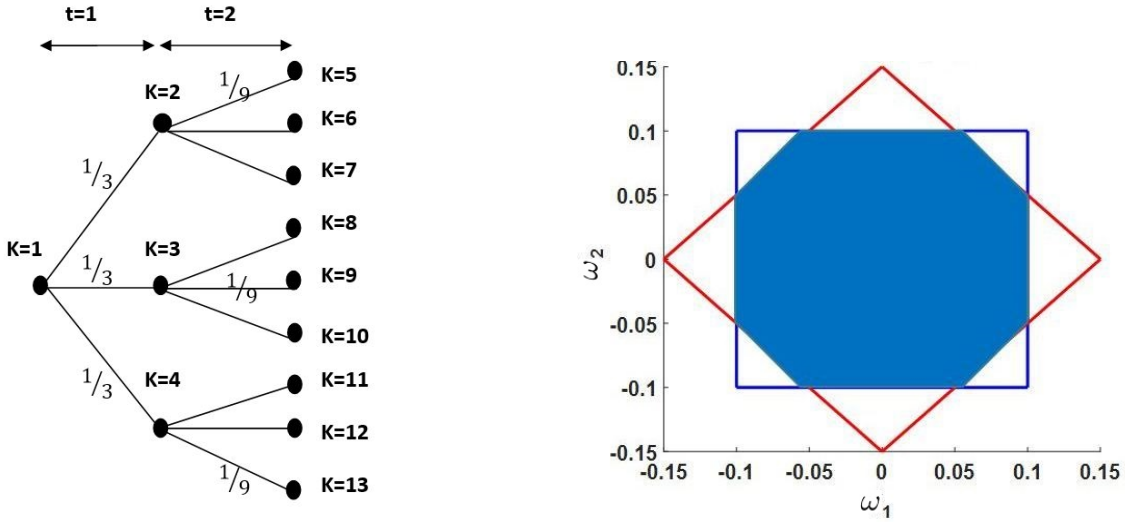
$$\phi_k \geq 0 \qquad \forall k \in \{5, ..., 13\}$$

$$x_{i,k} \geq 0 \qquad \forall i \in \{1, 2\}, k \in \{1, 2, 3, 4\} \tag{2.16f}$$

$$y_k \geq 0 \qquad \forall k \in \{2, ..., 13\} \tag{2.16g}$$

2.3.2 Hybrid 2 method

In hybrid 2 formulation, uncertainty is modeled in the following way: time-independent scenarios for yield parameter $\gamma_i(\omega)$; uncertainty set for demand parameters $d(\xi)$. Finally, linear decision rule for adjustable decisions $x(\xi), y(\xi)$.



(a) Demand uncertainty scenario tree

(b) Yield uncertainty set

Figure 2.3: Modeling of uncertainty in Hybrid 1 method

Demand Uncertainty: Demand is a time-dependent parameter. Demand at time periods 1 and 2 is modeled as a perturbation on some nominal demand values:

$$d_1(\xi) = \bar{d}_1(1 + \xi_1) \tag{2.17}$$

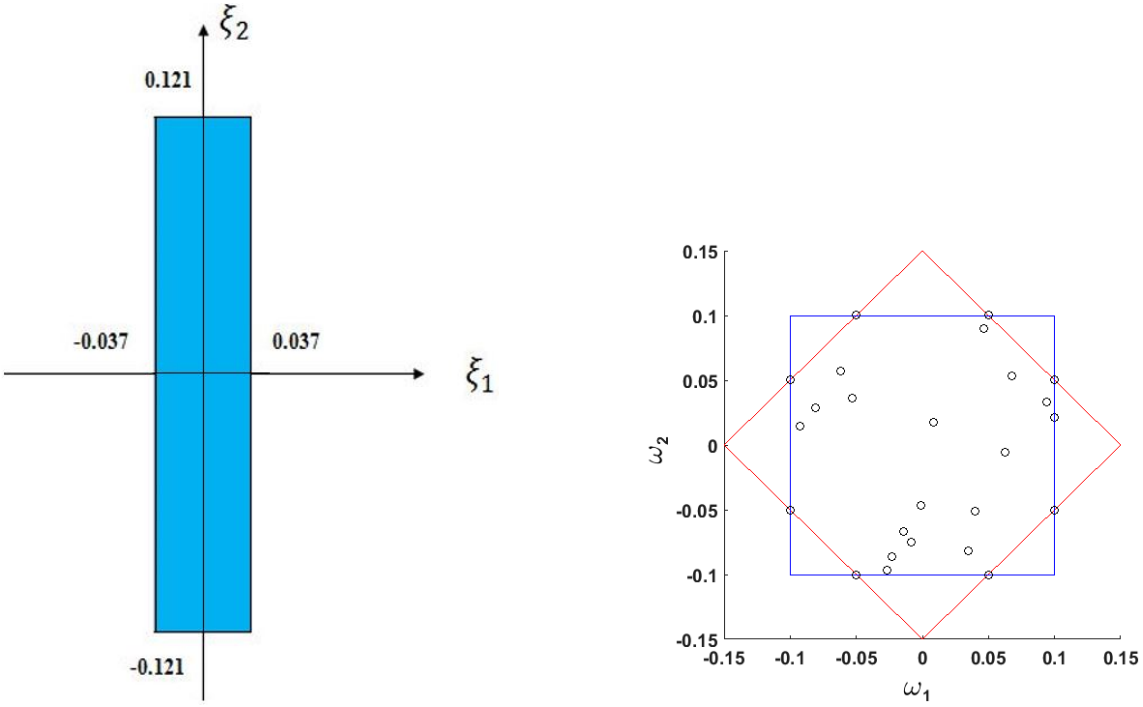
$$d_2(\xi) = \bar{d}_2(1 + \xi_2) \tag{2.18}$$

where $\bar{d}_1$, $\bar{d}_2$ are the nominal demand values ($\bar{d}_1 = 540$, $\bar{d}_2 = 615$), and ξ_1, ξ_2 are demand uncertainty parameters at time period 1 and 2, respectively.

For demand uncertainty, a box uncertainty set is defined as (Fig. 2.4a):

$$\Xi = \{\xi : |\xi_1| \leq 0.037, \ |\xi_2| \leq 0.121\} \tag{2.19}$$

The boundary values for ξ_1 and ξ_2 are designed so that demand uncertainty model of Hybrid 1 and 2 methods match. In Hybrid 2 method, the uncertainty vector is defined as $\xi = [1,\ \xi_1,\ \xi_2]^\top$, and the uncertainty set can be formulated as a polyhedral set using the compact representation $W\xi \leq V$.



(a) Uncertainty set for demand

(b) Yield uncertainty samples

Figure 2.4: Modeling of uncertainty in Hybrid 2 method

Yield uncertainty: Yield is a static (time-independent) parameter. In Hybrid 2 method, yield uncertainty $\gamma_i(\omega)$ is defined as time-independent scenarios $\gamma_{i,l}$. For this purpose, scenarios are sampled from the yield uncertainty set defined in the Hybrid 1 method and the samples include all vertices of the set. Fig. 2.4b illustrates the yield uncertainty set, which includes 8 samples on the vertices and 17 random samples inside the set. Since the samples cover the vertices of the yield uncertainty set and the constraints are linear with respect to yield uncertainty, the inside samples are redundant.

In this model, adaptive variables are modeled as linear decision rules (affine functions) of demand uncertainty.

$$x_{i,1} = x_{i,1} \qquad \forall i \tag{2.20a}$$

$$x_{i,2}(\xi) = \bar{x}_{i,2} + \tilde{x}_{i,2}\ \xi_1 \qquad \forall i \tag{2.20b}$$

$$y_1(\xi) = \bar{y}_1 + \tilde{y}_1\ \xi_1 \tag{2.20c}$$

$$y_2(\xi) = \bar{y}_2 + \tilde{y}_2\ \xi_1 + \hat{y}_2\ \xi_2 \tag{2.20d}$$

It should be noted that at each time period, only the demand uncertain parameters up to that time period are observed and the future uncertainties are not revealed yet. Therefore, at each time period, variables can only depend on uncertain parameters observed up to that time period. Thus, at time period 1 variables can only depend on ξ_1 and at time period 2 variables can depend on ξ_1 and ξ_2. Following this reasoning, non-anticipativity of uncertainty is enforced in the problem formulation. As illustrated in Fig. 2.2, at time period 1 the installation decision $x_{i,1}$ is decided before the revelation of demand uncertainty ξ_1, therefore it does not depend on ξ_1, but the outsourcing decision $y_1(\xi)$ is decided after the revelation of ξ_1, therefore it depends on ξ_1. The same reasoning follows for second time period variables.

In the appendix, the derivation procedure for the robust counterpart of Hybrid 2 model with respect to demand uncertainty is provided. Eqs. 2.21a to 2.21i present the final Hybrid 2 model.

$$\min \quad \sum_i \alpha_i x_{i,1} + \left[\sum_i \alpha_i \bar{x}_{i,2} + \beta\,(\bar{y}_1 + \bar{y}_2)\right] \tag{2.21a}$$

$$s.t. \quad \sum_{i=1}^{2} x_{i,1} \leq 100 \tag{2.21b}$$

$$V^\top \theta \leq 100 \tag{2.21c}$$

$$W^\top \theta = \sum_{i=1}^{2} [\bar{x}_{i,2} \quad \tilde{x}_{i,2} \quad 0]^\top$$

$$\theta \geq 0$$

$$\sum_i \gamma_{i,l}\; x_{i,1} - V^\top \mu_l \geq \bar{d}_1 \qquad \forall l \tag{2.21d}$$

$$-W^\top \mu_l = [\bar{y}_1 \quad \tilde{y}_1 - \bar{d}_1 \quad 0]^\top \qquad \forall l$$

$$\mu_l \geq 0 \qquad \forall l$$

$$\sum_i \gamma_{i,l}\; x_{i,1} - V^\top \phi_l \geq \bar{d}_2 \qquad \forall l \tag{2.21e}$$

$$-W^\top \phi_l = \sum_i \gamma_{i,l} [\bar{x}_{i,2} \quad \tilde{x}_{i,2} \quad 0]^\top + [\bar{y}_2 \quad \tilde{y}_2 \quad \hat{y}_2 - \bar{d}_2]^\top \qquad \forall l$$

$$\phi_l \geq 0 \qquad \forall l$$

$$x_{i,1} \geq 0 \qquad \forall i \tag{2.21f}$$

$$-V^\top \psi_i \geq 0 \qquad \forall i \tag{2.21g}$$

$$-W^\top \psi_i = [\bar{x}_{i,2} \quad \tilde{x}_{i,2} \quad 0]^\top \qquad \forall i$$

$$\psi_i \geq 0 \qquad \forall i$$

$$-V^\top \eta \geq 0 \tag{2.21h}$$

$$-W^\top \eta = [\bar{y}_1 \quad \tilde{y}_1 \quad 0]^\top$$

$$
\begin{aligned}
&\eta \geq 0 \\
&-V^{\top}\delta \geq 0 \\
&-W^{\top}\delta = [\bar{y}_2 \quad \tilde{y}_2 \quad \hat{y}_2]^{\top} \\
&\delta \geq 0
\end{aligned}
\tag{2.21i}
$$

Results

For this illustrating example, the original deterministic model and the prosed hybrid models are all linear programming problems. They are solved in GAMS 26.1.0 using Cplex solver 12.8.0.0. First, the deterministic model solution for installation ($x_{i,t}$) and outsourcing (y_t) decisions is studied. Since in this problem, the yield factor for second technology ($\gamma_2 = 3.4$) is greater than the first one ($\gamma_1 = 1$), it is more economical to satisfy all the demand requirement using the second technology (P_2). This is the reason that zero installation is observed for technology 1 at both time periods $t = 1$ and $t = 2$ ($x_{1,1} = 0, x_{1,2} = 0$). At $t = 1$, in order to satisfy the demand constraint, full capacity of installation decision is used ($x_{2,1} = 100$, $x_{2,2} = 82.35$) and the remaining demand is satisfied using outsourcing ($y_1 = 200$). At $t = 2$, since the installation capacities of both time periods are accumulated, all the demand is satisfied using installed capacity and the outsourcing decision is zero ($y_2 = 0$). Objective value of 1942.64 is obtained for deterministic formulation.

Results of the hybrid 1 model are studied next. Similar to the discussion presented for deterministic solution, since the yield factor for technology P_2 is greater than technology P_1, installation decisions of P_1 are all zero in all the times (Table 2.3). Greater yield factor means that technology P_2 can generate more products at lower cost. At node $k = 1$ ($t = 0$), only installation decision is made and no outsourcing is required ($y_1 = 0$), but at the remaining nodes, after demand uncertainty is revealed in time periods 1 and 2, some outsourcing decisions are observed (Table 2.3). The objective value for Hybrid 1 formulation is 2426.97.

Table 2.3: Hybrid 1 model solution

k	1	2	3	4	5	6	7	8	9	10	11	12	13
$x_{1,k}$	0	0	0	0	-	-	-	-	-	-	-	-	-
$x_{2,k}$	100	100	100	100	-	-	-	-	-	-	-	-	-
y_k	-	214	234	254	0	3	78	0	3	78	0	3	78

The optimal objective for the Hybrid 2 model is 2510.99. The initial installation decision is $x_{1,1} = 0, x_{2,1} = 100$. The second stage installation decision is $x_{1,2}(\xi) = 0, x_{2,2}(\xi) = 100$. The outpouring decision is $y_1(\xi) = 234+540\xi_1, y_2(\xi) = 39+319.8\xi_2$. Hybrid 1 method provides a better objective value (3.46% lower cost) compared to Hybrid method 2. Similar to Hybrid 1 solution, since the yield factor for technology P2 is greater than P1, all the installation decisions are made

by technology P2 at both time periods 1 and 2 and the installation decisions of technology P1 are zero ($x_{1,1} = 0$, $x_{1,2}$=0). Similarly, we observe some outsourcing decisions at both time periods.

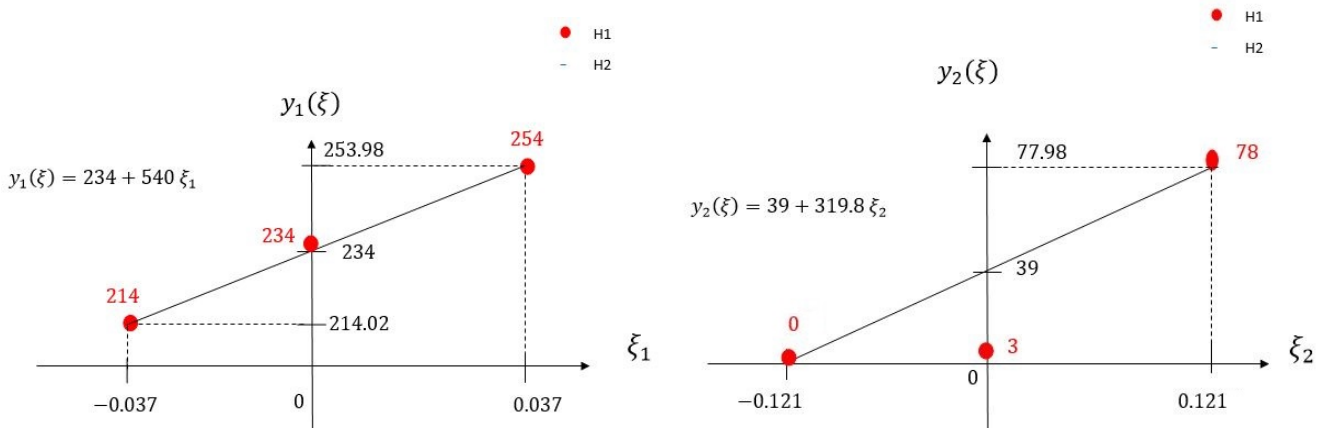


Figure 2.5: Outsourcing decisions for $t = 1$ (left) and $t = 2$ (right)

Figures 2.5 illustrates the outsourcing decisions for both hybrid methods in $t = 1$ and $t = 2$ time periods respectively. The black line represents the Hybrid 2 solution obtained using the LDR technique and the red dots represents the Hybrid 1 solution obtained from the node-tree method. Both figures show that at each time period, if demand increases, the outsourcing decision will increase as well.

Table 2.4: Solution, model size, and run time

Solution Method	Objective	Variables	Constraints	Run time (sec)
Deterministic	1942.64	7	11	0.103
Hybrid 1	2426.97	277	381	0.112
Hybrid 2	2510.99	96	151	0.109

Table 2.4 summarizes the objective value, model size and computation time. Notice that only 8 samples for static uncertainty are considered for Hybrid 2 solution. In this problem, the number of variables and constraints is in the same order for both Hybrid 1 and 2 methods and the run time is under 0.15 seconds while Hybrid 1 method results in slightly better objective function (3.46% lower cost) compared to Hybrid 2. It should be reminded that in both hybrid methods, the solution is feasible with respect to demand and yield uncertainty scenarios or sets, while the deterministic solution may result in constraint violations. While there is a minor difference in runtime for this small problem, for problems with larger number of variables and uncertain parameters, there can be a significant difference in run time. The difference in solution quality and run time between hybrid methods is investigated in the case study presented in section 2.5.

2.4 MSAP with both continuous and integer adaptive variables

In this section, we consider the following MSAP problem with both adaptive continuous and integer variables.

$$\min \quad \rho\left(\sum_t c_t(\xi_{[t]})^\top x_t(\xi_{[t]}) + \sum_t d_t(\xi_{[t]})^\top y_t(\xi_{[t]})\right) \tag{2.22a}$$

$$s.t. \quad \sum_{\tau=1}^{t} E_{t,\tau}(\omega)^\top x_\tau(\xi_{[\tau]}) + \sum_{\tau=1}^{t} F_{t,\tau}(\omega)^\top y_\tau(\xi_{[\tau]}) \le q_t(\xi_{[t]}) \quad \forall t, \xi \in \Xi, \omega \in \Omega \tag{2.22b}$$

$$\sum_{\tau=1}^{t} A_{t,\tau}(\omega)^\top x_\tau(\xi_{[\tau]}) \le b_t(\xi_{[t]}) \quad \forall t, \xi \in \Xi, \omega \in \Omega \tag{2.22c}$$

$$y_t(\xi_{[t]}) \in \{0,1\} \quad \forall t, \xi \in \Xi \tag{2.22d}$$

where binary variable y_t is also adjustable with respect to observed uncertainty ξ. The constraints are classified into those with only continuous variables and those with both continuous and binary or only binary variables. We further assume the following linear model of uncertainty:

$$c_t(\xi_{[t]}) = C_t\xi_{[t]}, d_t(\xi_{[t]}) = D_t\xi_{[t]}, q_t(\xi_{[t]}) = Q_t\xi_{[t]}, b_t(\xi_{[t]}) = B_t\xi_{[t]}$$

$$A_{t,\tau}(\omega) = A_{t,\tau}\omega, E_{t,\tau}(\omega) = E_{t,\tau}\omega, F_{t,\tau}(\omega) = F_{t,\tau}\omega$$

Then, the model is formulated as:

$$\min \quad \rho\left(\sum_t \xi_{[t]}^\top C_t^\top x_t(\xi_{[t]}) + \sum_t \xi_{[t]}^\top D_t^\top y_t(\xi_{[t]})\right) \tag{2.23a}$$

$$s.t. \quad \sum_{\tau=1}^{t} \omega^\top E_{t,\tau}^\top x_\tau(\xi_{[\tau]}) + \sum_{\tau=1}^{t} \omega^\top F_{t,\tau}^\top y_\tau(\xi_{[\tau]}) \le Q_t\xi_{[t]} \quad \forall t, \xi \in \Xi, \omega \in \Omega \tag{2.23b}$$

$$\sum_{\tau=1}^{t} \omega^\top A_{t,\tau}^\top x_\tau(\xi_{[\tau]}) \le B_t\xi_{[t]} \quad \forall t, \xi \in \Xi, \omega \in \Omega \tag{2.23c}$$

$$y_t(\xi_t) \in \{0,1\} \quad \forall t, \xi \in \Xi \tag{2.23d}$$

2.4.1 MSSP formulation

With the expectation based objective, the corresponding MSSP is as follows:

$$\min \quad \mathbb{E}\left(\sum_t \xi_{[t]}^\top C_t^\top x_t(\xi_{[t]}) + \sum_t \xi_{[t]}^\top D_t^\top y_t(\xi_{[t]})\right) \tag{2.24a}$$

$$s.t. \quad \sum_{\tau=1}^{t} \omega^\top E_{t,\tau}^\top x_\tau(\xi_{[\tau]}) + \sum_{\tau=1}^{t} \omega^\top F_{t,\tau}^\top y_\tau(\xi_{[\tau]}) \leq Q_t \xi_{[t]} \quad \forall t, \xi \in \Xi, \omega \in \Omega \tag{2.24b}$$

$$\sum_{\tau=1}^{t} \omega^\top A_{t,\tau}^\top x_\tau(\xi_{[\tau]}) \leq B_t \xi_{[t]} \quad \forall t, \xi \in \Xi, \omega \in \Omega \tag{2.24c}$$

$$y_t(\xi_t) \in \{0,1\} \quad \forall t, \xi \in \Xi \tag{2.24d}$$

In Hybrid 1 method, time dependent uncertainty is modeled with scenario tree and variables (both binary and continuous) are scenario-dependent and the corresponding formulation is provided below:

$$\min \quad \sum_{s \in S} p_s \left(\sum_t c_{t,s} x_{t,s} + \sum_t d_{t,s} y_{t,s} \right) \tag{2.25a}$$

$$s.t. \quad \sum_{\tau=1}^{t} \omega^\top E_{t,\tau}^\top x_{\tau,s} + \sum_{\tau=1}^{t} \omega^\top F_{t,\tau}^\top y_{\tau,s} \leq q_{t,s} \quad \forall t, s, \omega \in \Omega \tag{2.25b}$$

$$\sum_{\tau=1}^{t} \omega^\top A_{t,\tau}^\top x_{\tau,s} \leq b_{t,s} \quad \forall t, s, \omega \in \Omega \tag{2.25c}$$

$$x_{t,s} = x_{t,s'} \quad \forall \{t, s, s'\} \in SP \tag{2.25d}$$

$$y_{t,s} = y_{t,s'} \quad \forall \{t, s, s'\} \in SP \tag{2.25e}$$

where $c_{t,s}$, $d_{t,s}$, $q_{t,s}$ are $b_{t,s}$ are parameter values under scenario s, set SP defines the set where non-anticipativity constraint should be applied.

In Hybrid 2 method, the binary variables are scenario-dependent and the continuous variables are linear decision rules of time-dependent uncertainty. For constraints with only continuous variables (Eq. 2.26c), the polyhedral uncertainty set is employed to model the time-dependent uncertainty and the linear decision rule is applied to the continuous variables. For constraints with both binary and continuous variables (Eq. 2.26b), the binary variables are scenario-dependent and LDR is applied to continuous variables. In this case, the time-dependent uncertainty is modeled using scenario-tree. If the constraint includes only binary variables, it is a especial case of constraint 2.26b where all $E_{t,\tau} = 0$).

$$\min \quad \mathbb{E}\left(\sum_t \xi^\top P_t^\top C_t^\top X_t P_t \xi \right) + \sum_{s \in S} p_s \left(\sum_t d_{t,s} y_{t,s} \right) \tag{2.26a}$$

$$s.t. \quad \sum_{\tau=1}^{t} \omega^\top E_{t,\tau}^\top X_\tau P_\tau \xi_s + \sum_{\tau=1}^{t} \omega^\top F_{t,\tau}^\top y_{\tau,s} \leq q_{t,s} \quad \forall t, s, \omega \in \Omega \tag{2.26b}$$

$$\sum_{\tau=1}^{t} \omega_l^\top A_{t,\tau}^\top X_\tau P_\tau \xi \leq B_t P_t \xi \quad \forall t, l, \xi \in \Xi \tag{2.26c}$$

$$y_{t,s} = y_{t,s'} \qquad \forall \{t, s, s'\} \in SP \tag{2.26d}$$

In the above formulation, P_t is the observation matrix defined in section 2.2.1. Tables 2.5 summarizes the formulation characteristics of Hybrid 1 and 2 methods.

Table 2.5: Summary of the proposed hybrid formulation characteristics

Method	Uncertainty model	Variable model
Hybrid 1	Static: Set Time-dependent: Scenario tree	Integer: Scenario-dependent Continuous: Scenario dependent
Hybrid 2	Static: Samples Time-dependent: Scenario tree and set	Integer: Scenario-dependent Continuous: Decision Rule

2.4.2 MSRO formulation

With worst-case objective, the Hybrid 1 formulation is given below:

$$\begin{aligned}
\min \quad & z && && (2.27a)\\
s.t. \quad & \sum_{\tau=1}^{t} \omega^\top E_{t,\tau}^\top x_{\tau,s} + \sum_{\tau=1}^{t} \omega^\top F_{t,\tau}^\top y_{\tau,s} \leq q_{t,s} && \forall t, s, \omega \in \Omega && (2.27b)\\
& \sum_{\tau=1}^{t} \omega^\top A_{t,\tau}^\top x_{\tau,s} \leq b_{t,s} && \forall t, s, \omega \in \Omega && (2.27c)\\
& \sum_{t} c_{t,s} x_{t,s} + \sum_{t} d_{t,s} y_{t,s} \leq z && \forall s && (2.27d)\\
& x_{t,s} = x_{t,s'} && \forall \{t, s, s'\} \in SP && (2.27e)\\
& y_{t,s} = y_{t,s'} && \forall \{t, s, s'\} \in SP && (2.27f)
\end{aligned}$$

where the worst-case objective is evaluated over the scenarios. On the other hand, for Hybrid 2 method, the worst case objective is evaluated through both uncertainty set and scenarios.

$$\begin{aligned}
\min \quad & z_1 + z_2 && && (2.28a)\\
s.t. \quad & \sum_{\tau=1}^{t} \omega^\top E_{t,\tau}^\top X_\tau P_\tau \xi_s + \sum_{\tau=1}^{t} \omega^\top F_{t,\tau}^\top y_{\tau,s} \leq q_{t,s} && \forall t, s, \omega \in \Omega && (2.28b)\\
& \sum_{\tau=1}^{t} \omega_l^\top A_{t,\tau}^\top X_\tau P_\tau \xi \leq B_t P_t \xi && \forall t, l, \xi \in \Xi && (2.28c)\\
& \sum_{t} \xi^\top P_t^\top C_t^\top X_t P_t \xi \leq z_1 && \forall \xi \in \Xi && (2.28d)\\
& \sum_{t} d_{t,s} y_{t,s} \leq z_2 && \forall s && (2.28e)
\end{aligned}$$

$$y_{t,s} = y_{t,s'} \qquad \forall\{t, s, s'\} \in SP \qquad (2.28f)$$

In the subsequent section, an example is studied to demonstrate the MSSP with both adaptive integer and adaptive continuous variables.

2.5 Capacity expansion planning example

This section investigates a multistage process capacity planning problem with both continuous and integer adaptive variables. Fig. 2.6 illustrates the process network superstructure [65]. Chemicals 1, 2, 3 and 4 are purchased from external resources or produced in the network. Chemical 5 is the final product sold to the market. The objective is to optimize the process planning and maximize the profit of the overall process over a given number of time periods.

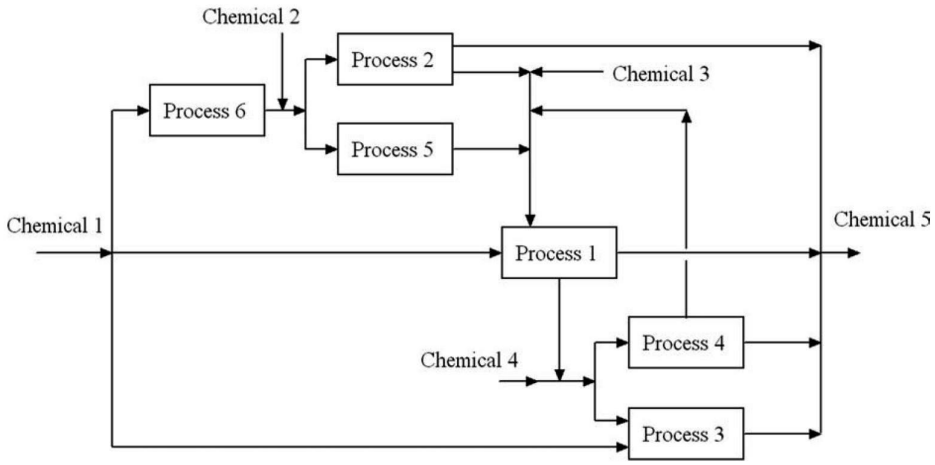


Figure 2.6: Process network superstructure

Multiple sources of uncertainty are considered in this process network optimization. Presence of several sources of uncertainty introduce challenges in the solution of the stochastic optimization problem. The uncertainties include: demand($d_{j,t}(\xi)$), price ($\gamma_{j,t}(\xi)$) and yield ($\eta_{i,j}(\omega)$) uncertainty of chemical 5 ($j = 5$). Demand and price uncertainties are time-dependent while yield uncertainty is time-independent. Fig. 2.7 illustrates decision making and uncertainty revelation sequence. At the beginning of each time period, demand and price uncertainties for chemical 5 are revealed ($d_{j,t}(\xi)$, $\gamma_{j,t}(\xi)$, $j = 5$); subsequently some capacity installation, operation and purchase decisions for processes 1 to 6 and chemicals 1 to 4 ($X_{i,t}(\xi), W_{i,t}(\xi), P_{j,t}(\xi)$) are made. Chemical 5 ($S_{j,t}(\xi, \omega), j = 5$) is sold to the market at the end of each time period. It is assumed that no inventory is available and no product can be stored; therefore all purchased or produced amounts of chemicals 1 to 4 have to be consumed and all the produced chemical 5 has to be sold. It should be mentioned that the sales amount of chemical 5 depends on yield uncertainty of the process network for chemical 5.

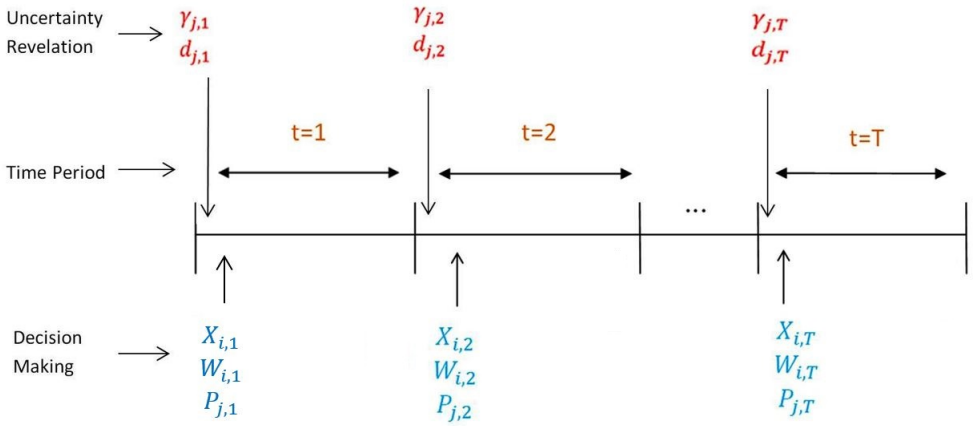


Figure 2.7: Decision timing and uncertainty revelation sequence

In this case study, it is assumed that at each time period, demand and price uncertainties are independent from the other time periods. For instance, demand and price values at time period 1 are independent of time period 2 and there is no correlation between them. In Hybrid 2, Eqs. 2.48 and 2.49 state the independency of uncertainty parameters at different time periods. In order to have equivalent independency of uncertainty at Hybrid 1, in each time period, each cluster of the scenario tree is independent of the other clusters (i.e., in each time period, at each cluster, the middle node accounts for the nominal value of the corresponding time period, and the upper and lower nodes account for +20% and -20% perturbation of the nominal value, respectively.)

The general multistage stochastic programming formulation for this problem is presented as in the following:

$$\max \ \mathbb{E}_{\xi,\omega}[\sum_t \sum_j (\gamma_{j,t}(\xi) S_{j,t}(\xi,\omega) - \Gamma_{j,t} P_{j,t}(\xi)) - \sum_t \sum_i (\delta_{i,t} W_{i,t}(\xi) + \alpha_{i,t} X_{i,t}(\xi) + \beta_{i,t} y_{i,t}(\xi))] \tag{2.29a}$$

$$s.t. \quad y_{i,t}(\xi) X^L_{i,t} \le X_{i,t}(\xi) \le y_{i,t}(\xi) X^U_{i,t} \qquad \forall i,t,\xi \in \Xi \tag{2.29b}$$

$$Q_{i,t}(\xi) = Q_{i,t-1}(\xi) + X_{i,t}(\xi) \qquad \forall i,t,\xi \in \Xi \tag{2.29c}$$

$$\sum_t y_{i,t}(\xi) \le N^{EXP}_i \qquad \forall i,\xi \in \Xi \tag{2.29d}$$

$$\sum_i (\alpha_{i,t} X_{i,t}(\xi) + \beta_{i,t} y_{i,t}(\xi)) \le CI_t \qquad \forall t,\xi \in \Xi \tag{2.29e}$$

$$W_{i,t}(\xi) \le Q_{i,t}(\xi) \qquad \forall i,t,\xi \in \Xi \tag{2.29f}$$

$$P_{j,t}(\xi) + \sum_i \eta_{i,j}(\omega) W_{i,t}(\xi) = S_{j,t}(\xi,\omega) + \sum_i \mu_{i,j} W_{i,t}(\xi) \qquad \forall j,t,\xi \in \Xi, \omega \in \Omega \tag{2.29g}$$

$$P_{j,t}(\xi) \le a_{j,t} \qquad \forall j,t,\xi \in \Xi \tag{2.29h}$$

$$S_{j,t}(\xi,\omega) \ge d_{j,t}(\xi) \qquad \forall j = 5,t,\xi \in \Xi \tag{2.29i}$$

$$X_{i,t}(\xi), W_{i,t}(\xi), P_{i,t}(\xi) \geq 0 \qquad \forall i, t, \xi \in \Xi \tag{2.29j}$$

$$y_{i,t}(\xi) \in \{0, 1\} \qquad \forall i, t, \xi \in \Xi \tag{2.29k}$$

In this formulation, indices i, j and t indicate the process, chemical and time period respectively and ξ is the vector of uncertain parameters. The variable $P_{j,t}(\xi)$ is the amount of chemical j purchased at the beginning of period t, $Q_{i,t}(\xi)$ is the total capacity of process i in period t, $S_{j,t}(\xi, \omega)$ is the amount of chemical j sold at period t, $X_{i,t}(\xi)$ is the units of expansion of process i at period t, and $y_{i,t}(\xi)$ is a binary variable that indicates process i is expanded or not. The parameter $\alpha_{i,t}$ is unit expansion cost of process i at period t, $\beta_{i,t}$ is fixed cost of establishing process i at period t, $\gamma_{i,t}(\xi)$ and $\Gamma_{i,t}$ are selling and buying prices of chemical j in period t, $\delta_{i,t}$ is the unit operation cost for process i in period t, $\eta_{i,j}(\omega)$ is the consumption ratio for chemical j in process i, $\mu_{i,j}$ is the production ratio for chemical j in process i, CI_t is the investment budget for period t, $a_{j,t}$ is the availability of chemical j in period t, $d_{j,t}(\xi)$ is the demand of chemical j in period t, N_i^{EXP} is the allowable number of expansions for process i during all the project time horizon, $X_{i,t}^L$ and $X_{i,t}^U$ are the lower and upper bounds for capacity expansion of process i in period t which are set as 0 and 10^5 tons respectively.

The objective of the stochastic formulation is to maximize the profit obtained from the process network. Eq. 4.32a is the objective function which calculates the revenue obtained from selling chemical 5 and deducts operation, installation and purchase costs. Constraint 4.32b means that expansion capacity ($X_{i,t}(\xi)$) of process i at the beginning of period t is bounded between an upper and lower bound. $y_{i,t}(\xi)$ is a binary variable which indicates process i is expanded at period t or not. Eq. 4.32c calculates the total expanded capacity for each process up to time t. Constraint 4.32d means that the total number of expansions for each process i is restricted during the project horizon. Constraint 4.32e indicates that at each year, there is a capital restriction CI_t for expansion ($\alpha_{i,t}X_{i,t}(\xi)$) and fixed investment ($\beta_{i,t}y_{i,t}(\xi)$) costs. Constraint 4.32f means that the operation level of each process at time t, can not be more than the total installed capacity for that process up to time t. Eq. 4.32g is a material balance equation and it means the total amount of purchased ($P_{j,t}(\xi)$) and produced ($\eta_{i,j}(\omega)W_{i,t}(\xi)$) chemicals must be equal to the total amount of consumed ($\mu_{i,j}W_{i,t}(\xi)$) and sold ($S_{j,t}(\xi, \omega)$) chemicals. Constraint 4.32h indicates that the amount of purchased chemical j at time t, should be less than or equal to the available amount of that chemical ($a_{j,t}$). Constraint 4.32i describes that the amount of sold chemical j at time t ($S_{j,t}(\xi, \omega)$), should be greater than its demand ($d_{j,t}(\xi)$) and constraint 4.32k indicates that problem variables can not be negative.

2.5.1 Hybrid 1 method

In Hybrid 1 method, time-dependent uncertainties are modeled as scenario trees and time-independent uncertainties as static sets. Subsequently, the constraints are reformulated to their robust counterparts with respect to static uncertainty. The solution is feasible for all the uncertainty scenarios defined in the scenario tree and any realization of uncertainty within the static uncertainty set.

Time-independent uncertainty. Yield uncertainty of chemical 5 ($\eta_{i,5}(\omega)$) is static uncertainty modeled as an uncertainty set and it is assumed to be uniformly distributed within 20% perturbation range of nominal yield values (Eqs. 2.30, 2.31). It should be noted that since only the fifth product is sold, demand, selling price, and selling amount parameters and variables are only defined for the fifth product; also yield of chemicals 1 to 4 is assumed to be known and there is no yield uncertainty for chemicals 1 to 4 ($\omega_{i,j} = 0, d_{j,t} = 0, \gamma_{j,t} = 0, S_{j,t} = 0,\ \forall j = 1, ..., 4$).

$$\eta_{i,5}(\omega) = \bar{\eta}_{i,5}(1 + \omega_{i,5}) \qquad \forall i \tag{2.30}$$

$$|\omega_{i,5}| \leq 0.2 \qquad \forall i \tag{2.31}$$

Eq. 2.31 can be written as an inequality on the vector of yield uncertainties $\omega = [\omega_{1,5},\ \omega_{2,5}, \cdots,\ \omega_{6,5}]^\top$ which represents a polyhedral uncertainty set.

$$\Omega = \{\omega : M\omega \leq V\} \tag{2.32}$$

Time-dependent uncertainty. In this problem, demand and price of chemical 5 ($d_{5,t}(\xi), \gamma_{5,t}(\xi)$) are time-dependent uncertainties modeled as scenario trees (Fig. 2.8). Fig. 2.8 illustrates the scenario tree for seven time periods where each node is divided into three branches. In the scenario tree, each path from the first node ($k = 1$) to the nodes at the last time period ($k = 1093$ to 3280) represents an uncertainty scenario. The problem is solved for seven time periods where the tree includes 3280 nodes and the probability of occurrence of each branch is $\frac{1}{3}$ at each node.

Eqs. 2.33a to 2.33n represent the general stochastic Hybrid 1 model for the current problem where node formulation is used for time-dependent uncertainties (demand and price of chemical 5) and time-dependent parameters and variables.

$$\max \sum_k \sum_j p_k \gamma_{j,k} \mathbb{E}_\omega[S_{j,k}(\omega)] - \sum_k \sum_j p_k \Gamma_{j,k} P_{j,k} - \sum_k \sum_i p_k \left(\delta_{i,k} W_{i,k} + \alpha_{i,k} X_{i,k} + \beta_{i,k} y_{i,k}\right) \tag{2.33a}$$

$$s.t. \quad y_{i,k} X^L_{i,k} \leq X_{i,k} \leq y_{i,k} X^U_{i,k} \qquad \forall i, k \in \mathcal{K}_{-1} \tag{2.33b}$$

$$Q_{i,k} = Q_{i,a(k)} + X_{i,k} \qquad \forall i, k \in \mathcal{K}_{-1} \tag{2.33c}$$

$$Q_{i,k} = X_{i,k} \qquad \forall i, k \in \mathcal{K}_2 \tag{2.33d}$$

$$U_{i,k} \leq N^{EXP}_i \qquad \forall i, k \in \mathcal{K}_T \tag{2.33e}$$

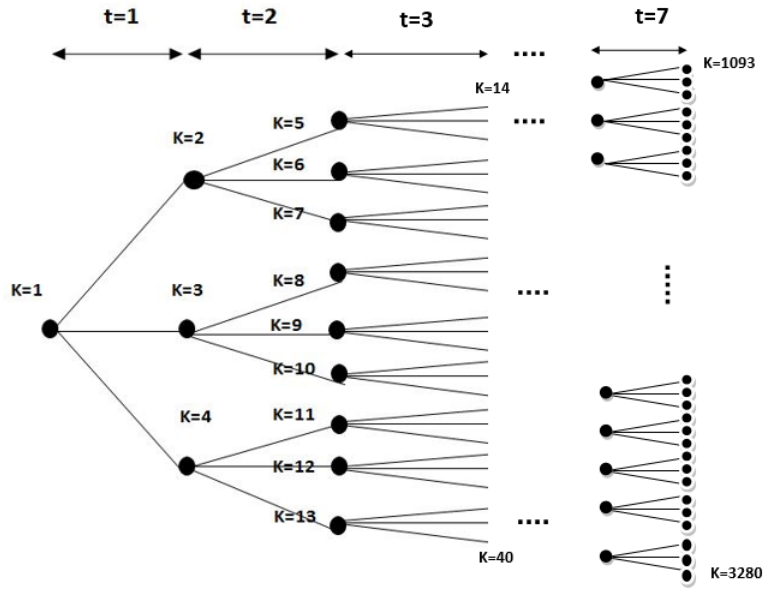


Figure 2.8: Scenario tree for 7 time periods

$$U_{i,k} = y_{i,k} + U_{i,a(k)} \qquad \forall i, k \in \mathcal{K}_{-1} \qquad (2.33f)$$

$$U_{i,1} = 0 \qquad \forall i \qquad (2.33g)$$

$$\sum_i (\alpha_{i,k} X_{i,k} + \beta_{i,k} y_{i,k}) \leq CI_k \qquad \forall k \in \mathcal{K}_{-1} \qquad (2.33h)$$

$$W_{i,k} \leq Q_{i,k} \qquad \forall i, k \in \mathcal{K}_{-1} \qquad (2.33i)$$

$$P_{j,k} + \sum_i \eta_{i,j}(\omega) W_{i,k} = S_{j,k}(\omega) + \sum_i \mu_{i,j} W_{i,k} \qquad \forall j, k \in \mathcal{K}_{-1}, \omega \in \Omega \qquad (2.33j)$$

$$P_{j,k} \leq a_{j,k} \qquad \forall j, k \in \mathcal{K}_{-1} \qquad (2.33k)$$

$$S_{j,k}(\omega) \geq d_{j,k} \qquad \forall j = 5, k \in \mathcal{K}_{-1}, \omega \in \Omega \qquad (2.33l)$$

$$X_{i,k}, W_{i,k}, P_{i,k} \geq 0 \qquad \forall i, k \in \mathcal{K}_{-1} \qquad (2.33m)$$

$$y_{i,k} \in \{0, 1\} \qquad \forall i, k \in \mathcal{K}_{-1} \qquad (2.33n)$$

where p_k represents the probability of each node in the node-tree and Eqs. 2.33e to 2.33g correspond to Eq. 4.32d. In this formulation, $\mathcal{K}_{-1}$ represents the set of all the nodes except the first node, while $\mathcal{K}_2$ and $\mathcal{K}_T$ represent the nodes of the second and last time periods, respectively.

Equation 2.33j is a material balance relation which means all the purchased and produced materials have to consumed and sold. Only chemical 5 is sold, and since the yield of chemical 5 on the left-hand side of the equation is an uncertain parameter ($\eta_{i,5}(\omega)$), in order to satisfy the equality, amount of the sold chemical 5 should depend on yield uncertainty too ($S_{5,k}(\omega)$). For this purpose, the variable $S_{5,k}(\omega)$ is considered to be an affine function of yield uncertainty. Since

each node represents a different scenario for demand and price uncertainty, it means the variable $S_{5,k}(\omega)$ depends on demand, price and yield uncertainties of chemical 5. Notice that the amount of sales is dependent on multiple uncertainties including yield. The sale dependency on demand and price uncertainty is modeled using node index of scenario tree and its dependency on static yield uncertainty is modelled using linear decision rule.

$$S_{j,k}(\omega) = \bar{S}_{j,k} + \hat{S}_{j,k}\omega \qquad \forall j = 5, k \in \mathcal{K}_{-1} \tag{2.34}$$

Since only chemical 5 is sold, the variable $S_{j,k}(\omega)$ is zero for chemicals 1 to 4:

$$\bar{S}_{j,k} = 0, \forall j \in \{1, ..., 4\}, k \in \mathcal{K}_{-1} \tag{2.35}$$

$$\hat{S}_{j,k,i} = 0, \forall i, j \in \{1, ..., 4\}, k \in \mathcal{K}_{-1} \tag{2.36}$$

Under the uniform distribution assumption, the objective can be written as

$$\max \quad \sum_k \sum_j p_k \gamma_k \bar{S}_{j,k} - \sum_k \sum_j p_k \Gamma_{j,k} P_{j,k} - \sum_k \sum_i p_k (\delta_{i,k} W_{i,k} + \alpha_{i,k} X_{i,k} + \beta_{i,k} y_{i,k}) \tag{2.37}$$

Next, the robust form of Hybrid 1 formulation with respect to the time-independent yield uncertainty is derived. Only equation 2.33j and inequality 2.33l depend on yield uncertainty. By substituting the uncertainty relation for yield $\eta_{i,j} = \bar{\eta}_{i,j}(1 + \omega_{i,j})$, the derivation procedure for robust form of equality 2.33j $P_{j,k} + \sum_i \eta_{i,j}(\omega) W_{i,k} = S_{j,k}(\omega) + \sum_i \mu_{i,j} W_{i,k}, \forall j, k \in \mathcal{K}_{-1}, \omega \in \Omega$ is presented below:

$$P_{j,k} + \sum_i \bar{\eta}_{i,j}(1 + \omega_{i,j}) W_{i,k} = \bar{S}_{j,k} + \sum_i \hat{S}_{j,k,i}\omega_{i,j} + \sum_i \mu_{i,j} W_{i,k} \quad \forall j, k \in \mathcal{K}_{-1}, \omega \in \Omega \tag{2.38}$$

In order to satisfy the equality, the coefficients of the uncertain parameter $\xi_{i,j}$ and the terms without parameter $\xi_{i,j}$ are set to be equal on both sides of the equation:

$$P_{j,k} + \sum_i \bar{\eta}_{i,j} W_{i,k} = \bar{S}_{j,k} + \sum_i \mu_{i,j} W_{i,k} \qquad \forall j, k \in \mathcal{K}_{-1} \tag{2.39}$$

$$\bar{\eta}_{i,j} W_{i,k} = \hat{S}_{j,k,i} \qquad \forall i, j = 5, k \in \mathcal{K}_{-1} \tag{2.40}$$

Duality theorem is used to obtain the robust deterministic counterpart of inequality 2.33l. Inequality 2.33l $S_{j,k}(\omega) \geq d_k, \forall j = 5, k \in \mathcal{K}_{-1}, \omega \in \Omega$ is written as:

$$\bar{S}_{j,k} + \hat{S}_{j,k}\omega \geq d_{j,k} \qquad \forall j = 5, k \in \mathcal{K}_{-1}, \omega \in \Omega \tag{2.41}$$

Using uncertainty set $\Omega = \{\omega : M\omega \leq V\}$, the deterministic dual counterpart is obtained:

$$\bar{S}_{j,k} - V^{\top}\phi_{j,k} \geq d_{j,k} \qquad \forall j = 5, k \in \mathcal{K}_{-1} \tag{2.42}$$

$$-M^{T}\phi_{j,k} = \hat{S}_{j,k}^{\top} \qquad \forall j = 5, k \in \mathcal{K}_{-1} \tag{2.43}$$

$$\phi_{j,k} \geq 0 \qquad \forall j = 5, k \in \mathcal{K}_{-1} \tag{2.44}$$

In summary, the Hybrid 1 model for process network of figure 2.6 is represented by objective 2.37, subject to constraints 2.33b-2.33i, 2.33k,2.33m,2.33n, 2.35-2.36, 2.39-2.40, 2.42-2.44. In this formulation the nodal representation of dynamic uncertainty is employed, therefore non-anticipativity constraints are not required.

2.5.2 Hybrid 2 method

In Hybrid 2 formulation, time-dependent uncertainties are modeled as uncertainty sets and the corresponding variables as affine functions of uncertainty sets and time-independent uncertainties are modeled as scenarios. Using duality theorem, constraints can be reformulated to robust deterministic counterparts with respect to time-dependent uncertainties and constraints are enforced with respect to time-independent uncertainty scenarios. The solution is feasible for any realization of uncertainty within the time-dependent uncertainty sets and any scenario of the static uncertainty. For this case study, uncertainty modeling for Hybrid 2 method can be summarized as:

Time-independent uncertainty. Uncertain yield parameter is modeled using samples where $\eta_{i,5}(\omega)$ is the uncertain yield and $\eta_{i,5,l}$ is the l-th yield scenario with probability of p_l.

$$\eta_{i,5}(\omega) \rightarrow \eta_{i,5,l} \qquad \forall i, l \in \mathcal{L} \tag{2.45}$$

Note that only the yield of the fifth product($j = 5$) is uncertain and yield of products 1 to 4 is assumed to be known.

Time-dependent uncertainty: Uncertain demand and price are modeled as perturbations on nominal values:

$$d_{5,t}(\xi) = \bar{d}_{5,t}(1 + \xi_{d_t}) \qquad \forall t \tag{2.46}$$

$$\gamma_{5,t}(\xi) = \bar{\gamma}_{5,t}(1 + \xi_{p_t}) \qquad \forall t \tag{2.47}$$

where $\bar{d}_{5,t}$, $\bar{\gamma}_{5,t}$ are nominal demand and price values for the fifth product. Only the fifth product is sold, therefore there is no demand and price for chemicals 1 to 4 ($d_{j,t} = 0, \gamma_{j,t} = 0 \quad \forall t, j = 1, ..., 4$). Uncertainty is defined for demand and price uncertainty as in the following:

$$|\xi_{p_t}| \leq b_{p_t} \qquad \forall t \tag{2.48}$$

$$|\xi_{d_t}| \leq b_{d_t} \qquad \forall t \tag{2.49}$$

where ξ_{d_t}, ξ_{p_t} are demand and price uncertainty parameters at period t and b_{p_t}, b_{d_t} are the price and demand perturbation values which are set as 20% in this problem. This bounded uncertainty set, can be written in the following inequality form:

$$\Xi = \{\xi : M\ \xi \leq V\} \tag{2.50}$$

Uncertainty vector ξ is defined as the vector that includes all the uncertain demand and price parameters: $\xi = [1, \xi_{p_1}, ..., \xi_{p_T}, \xi_{d_1}, ..., \xi_{d_T}]^\top$. The uncertain demand and price can be further written as:

$$d_{5,t} = \bar{d}_{5,t}(1 + \xi_{d_t}) = \bar{d}_{5,t}(1 + O_{d_t}\xi) \qquad \forall t \tag{2.51}$$

$$\gamma_{5,t} = \bar{\gamma}_{5,t}(1 + \xi_{p_t}) = \bar{\gamma}_{5,t}(1 + O_{p_t}\xi) \qquad \forall t \tag{2.52}$$

where O_{d_t} and O_{p_t} are truncate operators such that: $O_{p_t}\xi = \xi_{p_t}$ and $O_{d_t}\xi = \xi_{d_t}$. All the elements in O_{d_t} and O_{p_t} vectors are zero except the elements corresponding to ξ_{d_t} and ξ_{p_t} in ξ vector; for instance $O_{p_1} = [0\ 1\ 0\ ...\ 0\ 0\ 0\ 0]$.
Continuous variables are modeled as linear decision rules (affine functions) of uncertainty vector.

$$X_{i,t}(\xi) = \bar{X}_{i,t}\xi \qquad \forall i, t \tag{2.53a}$$

$$W_{i,t}(\xi) = \bar{W}_{i,t}\xi \qquad \forall i, t \tag{2.53b}$$

$$P_{j,t}(\xi) = \bar{P}_{j,t}\xi \qquad \forall j, t \tag{2.53c}$$

$$Q_{i,t}(\xi) = \bar{Q}_{i,t}\xi \qquad \forall i, t \tag{2.53d}$$

$$S_{j,t}(\xi, \omega) = \bar{S}_{j,t,l}\xi \qquad \forall j, t, l \tag{2.53e}$$

Non-anticipativity. At each time period t, uncertain parameters up to that time period are revealed and future uncertain parameters are not realized yet. Therefore at each time period, adaptive decision variables can only depend on revealed uncertain parameters. For instance, at the fist time period, the expansion variable $X_{i,t}$ can only depend on ξ_{d_1} and ξ_{p_1}. The following equations represent the non-anticipativity constraints for the continuous decision variables.

$$\bar{X}_{i,t,c} = 0 \qquad \forall \quad i, t < T, t + 2 \leq c \leq T,\ c \geq t + T + 2 \tag{2.54a}$$

$$\bar{P}_{j,t,c} = 0 \qquad \forall \quad j, t < T, t + 2 \leq c \leq T + 1,\ c \geq t + T + 2 \tag{2.54b}$$

$$\bar{W}_{i,t,c} = 0 \qquad \forall \quad i, t < T, t + 2 \leq c \leq T + 1,\ c \geq t + T + 2 \tag{2.54c}$$

$$\bar{Q}_{i,t,c} = 0 \qquad \forall \quad i, t < T, t + 2 \leq c \leq T + 1,\ c \geq t + T + 2 \tag{2.54d}$$

$$S_{j,t,l,c} = 0, \qquad \forall j = 5, l, t < T, t + 2 \leq c \leq T + 1,\ c \geq t + T + 2 \tag{2.54e}$$

The binary variables are scenario-dependent and continuous variables are linear decision rules of dynamic demand and price uncertainty. For those constraints where only continuous variables exist(Eqs. 4.32c, 4.32f to 4.32k), the set definition of dynamic uncertainty is used for LDR (Eqs 2.53a to 2.53e). For constraints where only binary variables exist(e.g., Eq.4.32d), the binary variables are scenario dependent (Eq. 2.55).

$$y_{i,t}(\xi) = y_{i,k} \qquad \forall i,k \tag{2.55}$$

At constraints where both binary and continuous variables exist(Eqs. 4.32b, 4.32e), the node-tree representation of dynamic uncertainty is used in LDR where the uncertainty vector corresponding to each node of the tree is substituted for ξ_k in LDR (Eqs. 2.56).

$$X_{i,t(k)} = \bar{X}_{i,t(k)}\xi_k \qquad \forall i,k \tag{2.56}$$

where the subscript $t(k)$ is the time period corresponding to node k. In constraints where only continuous variables exist, duality theorem is used to obtain the robust counterpart of the constraints. Eqs. 2.57a to 2.57m represent general Hybrid 2 formulation before applying duality to the constraints.

$$\max \quad \mathbb{E}_\xi[\sum_t \sum_j (\gamma_{j,t}(\xi) \sum_l p_l \bar{S}_{j,t,l}\xi - \Gamma_{j,t}\bar{P}_{j,t}\xi) - \sum_t \sum_i (\delta_{i,t}\bar{W}_{i,t}\xi \ + \alpha_{i,t}\bar{X}_{i,t}\xi)] - \sum_k \sum_i p_k \beta_{i,t(k)} y_{i,k} \tag{2.57a}$$

$$s.t. \quad y_{i,k}X^L_{i,t(k)} \le \bar{X}_{i,t(k)}\xi_k \le y_{i,k}X^U_{i,t(k)} \qquad \forall i,k \tag{2.57b}$$

$$\bar{Q}_{i,t}\xi = \bar{Q}_{i,t-1}\xi + \bar{X}_{i,t}\xi \qquad \forall i,t,\xi \in \Xi \tag{2.57c}$$

$$U_{i,k} \le N_i^{EXP} \qquad \forall i,k \in \mathcal{K}_T \tag{2.57d}$$

$$U_{i,k} = y_{i,k} + U_{i,a(k)} \qquad \forall i,k \in \mathcal{K}_{-1} \tag{2.57e}$$

$$U_{i,1} = 0 \qquad \forall i \tag{2.57f}$$

$$\sum_i [\alpha_{i,t(k)}\bar{X}_{i,t(k)}\xi_k + \beta_{i,t(k)}y_{i,k}] \le CI_{t(k)} \qquad \forall k \tag{2.57g}$$

$$\bar{W}_{i,t}\xi \le \bar{Q}_{i,t}\xi \qquad \forall i,t,\xi \in \Xi \tag{2.57h}$$

$$\bar{P}_{j,t}\xi + \sum_i \eta_{i,j,l}\bar{W}_{i,t}\xi = \bar{S}_{j,t,l}\xi + \sum_i \mu_{i,j}\bar{W}_{i,t}\xi \qquad \forall j,t,l,\xi \in \Xi \tag{2.57i}$$

$$\bar{P}_{j,t}\xi \le a_{j,t} \qquad \forall j,t,\xi \in \Xi \tag{2.57j}$$

$$\bar{S}_{j,t}\xi \ge d_{j,t}(\xi) \qquad \forall j,t,\xi \in \Xi \tag{2.57k}$$

$$\bar{X}_{i,t}\xi \ge 0 \qquad \forall i,t,\xi \in \Xi \tag{2.57l}$$

$$\bar{W}_{i,t}\xi \ge 0 \qquad \forall i,t,\xi \in \Xi \tag{2.57m}$$

$$\bar{P}_{i,t}\xi \ge 0 \qquad \forall i,t,\xi \in \Xi \tag{2.57n}$$

$$y_{i,k} \in \{0,1\} \qquad \forall i,k,\xi \in \Xi \qquad (2.57o)$$

The above definition of continuous and binary variables is substituted into the general stochastic model (Eqs. 4.32a to 4.32l) to obtain the general Hybrid 2 model. The derivation procedure for the objective is presented here:

First term:

$$\begin{aligned}
&\mathbb{E}_\xi\left[\sum_t\sum_j\left(\bar{\gamma}_{j,t}(1+O_{p_t}\xi)\sum_l p_l\bar{S}_{j,t,l}\ \xi-\Gamma_{j,t}\bar{P}_{j,t}\xi\right)\right]\\
&=\mathbb{E}_\xi\left[\sum_t\sum_j\bar{\gamma}_{j,t}\sum_l p_l\ \bar{S}_{j,t,l}\ \xi\right]+\mathbb{E}_\xi\left[\sum_t\sum_j\bar{\gamma}_{j,t}(O_{p_t}\xi)^\top\sum_l p_l\bar{S}_{j,t,l}\ \xi\right]-\mathbb{E}_\xi\left[\sum_t\sum_j\Gamma_{j,t}\bar{P}_{j,t}\xi\right]\\
&=\mathbb{E}_\xi\left[\sum_t\sum_j\bar{\gamma}_{j,t}\sum_l p_l\bar{S}_{j,t,l}\ \xi\right]+\mathbb{E}_\xi\left[\sum_t\sum_j\bar{\gamma}_{j,t}\xi^\top(O_{p_t}\sum_l p_l\bar{S}_{j,t,l})\ \xi\right]-\mathbb{E}_\xi\left[\sum_t\sum_j\Gamma_{j,t}\bar{P}_{j,t}\xi\right]\\
&=\sum_t\sum_j\bar{\gamma}_{j,t}\sum_l p_l\bar{S}_{j,t,l}\mathbb{E}(\xi)+\sum_t\sum_j\bar{\gamma}_{j,t}\ \mathrm{Tr}(O_{p_t}\sum_l p_l\bar{S}_{j,t,l}\mathbb{E}(\xi\xi^\top)\ )-\sum_t\sum_j\Gamma_{j,t}\bar{P}_{j,t}\mathbb{E}(\xi)
\end{aligned}$$

Second and third term:

$$\begin{aligned}
&\mathbb{E}_\xi\left[-\sum_t\sum_i(\delta_{i,t}\bar{W}_{i,t}\xi+\alpha_{i,t}\bar{X}_{i,t}\xi)\right]-\sum_k\sum_i p_k\beta_{i,t(k)}y_{i,k}\\
&=-\sum_t\sum_i(\delta_{i,t}\bar{W}_{i,t}+\alpha_{i,t}\bar{X}_{i,t})E(\xi)-\sum_k\sum_i p_k\beta_{i,t(k)}y_{i,k}
\end{aligned}$$

The derivation procedure for the semi-infinite constraints is skipped here. The final robust counterpart for each semi-infinite constraint can be checked through the corresponding equation number. Finally, the overall Hybrid 2 formulation is:

$$\begin{aligned}
\max\ & \sum_t\sum_j\bar{\gamma}_{j,t}\sum_l p_l\bar{S}_{j,t,l}\mathbb{E}(\xi)+\sum_t\sum_j\bar{\gamma}_{j,t}\ \mathrm{Tr}(O_{p_t}\sum_l p_l\bar{S}_{j,t,l}\ \mathbb{E}(\xi\xi^\top)\ ) && && (2.58a)\\
& -\sum_t\sum_j\Gamma_{j,t}\bar{P}_{j,t}\mathbb{E}(\xi)-\sum_t\sum_i(\delta_{i,t}\bar{W}_{i,t}+\alpha_{i,t}\bar{X}_{i,t})\mathbb{E}(\xi)-\ \sum_k\sum_i p_k\beta_{i,t(k)}y_{i,k} && &&\\
s.t.\ & y_{i,k}X^L_{i,t(k)}\le\bar{X}_{i,t(k)}\xi_k\le y_{i,k}X^U_{i,t(k)} && \forall i,k && (2.58b)\\
& \bar{Q}_{i,t}-\bar{Q}_{i,t-1}=\bar{X}_{i,t} && \forall i,t\ge 2 && (2.58c)\\
& \bar{Q}_{i,1}=\bar{X}_{i,1} && \forall i && \\
& U_{i,k}\le N_i^{EXP} && \forall i,k\in\mathcal{K}_T && (2.58d)\\
& U_{i,k}=y_{i,k}+U_{i,a(k)} && \forall i,k\in\mathcal{K}_{-1} && (2.58e)\\
& U_{i,1}=0 && \forall i && (2.58f)\\
& \sum_i[\alpha_{i,t(k)}\bar{X}_{i,t(k)}\xi_k+\beta_{i,t(k)}y_{i,k}]\le CI_{t(k)} && \forall k && (2.58g)\\
& V^\top\theta^4_{i,t}\le 0 && \forall i,t && (2.58h)\\
& M^\top\theta^4_{i,t}=(\bar{W}_{i,t}-\bar{Q}_{i,t})^\top && \forall i,t && \\
& \theta^4_{i,t}\ge 0 && \forall i,t && \\
& \bar{P}_{j,t}+\sum_i\eta_{i,j,l}\bar{W}_{i,t}=\bar{S}_{j,t,l}+\sum_i\mu_{i,j}\bar{W}_{i,t} && \forall j,t,l && (2.58i)\\
& \bar{S}_{j,t,l}=0 && \forall t,l,j=\{1,...,4\} && \\
& \bar{P}_{j,t}=0 && \forall t,j=5 &&
\end{aligned}$$

$$V^{\top}\theta^{5}_{j,t} \leq a_{j,t} \qquad \forall t, j = \{1, ..., 4\} \qquad (2.58j)$$

$$M^{\top}\theta^{5}_{j,t} = \bar{P}^{\top}_{j,t} \qquad \forall t, j = \{1, ..., 4\}$$

$$\theta^{5}_{j,t} \geq 0 \qquad \forall t, j = \{1, ..., 4\}$$

$$(-V)^{\top}\theta^{6}_{t,l,j} \geq \bar{d}_{j,t} \qquad \forall t, l, j = 5 \qquad (2.58k)$$

$$(-M)^{\top}\theta^{6}_{t,l,j} = \bar{S}_{j,t,l} - \bar{d}_{j,t}O_{d_t} \qquad \forall t, l, j = 5$$

$$\theta^{6}_{t,l,j} \geq 0 \qquad \forall t, l, j = 5$$

$$(-V)^{\top}\theta^{7}_{i,t} \geq 0 \qquad \forall i, t \qquad (2.58l)$$

$$(-M)^{\top}\theta^{7}_{i,t} = \bar{X}^{\top}_{i,t} \qquad \forall i, t$$

$$\theta^{7}_{i,t} \geq 0 \qquad \forall i, t$$

$$(-V)^{\top}\theta^{8}_{j,t} \geq 0 \qquad \forall j, t \qquad (2.58m)$$

$$(-M)^{\top}\theta^{8}_{j,t} = \bar{W}^{\top}_{j,t} \qquad \forall j, t$$

$$\theta^{8}_{j,t} \geq 0 \qquad \forall j, t$$

$$(-V)^{\top}\theta^{9}_{i,t} \geq 0 \qquad \forall i, t \qquad (2.58n)$$

$$(-M)^{\top}\theta^{9}_{i,t} = \bar{P}^{\top}_{i,t} \qquad \forall i, t$$

$$\theta^{9}_{i,t} \geq 0 \qquad \forall i, t$$

$$y_{i,k} \in \{0, 1\} \qquad \forall i, k \qquad (2.58o)$$

Eqs. 2.54a-2.54e

2.5.3 Results and discussion

In this study, the capacity expansion problem is formulated under demand, price and yield uncertainty of chemical 5. The problem is solved for $T = 7$ time steps. Table 2.6 summarizes the objective function, run time and problem size for the capacity expansion problem. Hybrid 1 method resulted in a very large problem size where the total number of variables and constraints is greater compared to the Hybrid 2 method. The problem was submitted to a workstation (Intel Xeon Dual 20 Core 2.0 GHz Processor, 128 GB DDR4 ECC RAM) and after 10 hours runtime, the problem reached 1.8% optimality gap while the objective function was still slightly worse than Hybrid 2 method (-0.43%). The Hybrid 2 method is solved using a desktop computer with much less computational power (3.5 GHz CPU, 8 GB RAM), but resulted in a better objective and better run time (7.38 hrs) compared to the Hybrid 1. While it is expected that Hybrid 1 to result in a better objective value due to higher flexibility in continuous variables, its objective value is slightly worse than Hybrid 2 since the Hybrid 1 solution did not reach 0 optimality gap within 10 hrs run time. In this case study, the problem is mixed integer linear optimization (MILP) problem. GAMS 26.1.0 and Cplex solver 12.8.0.0 are employed to solve the problem.

Table 2.6: Objective and run time for capacity expansion problem

Method	Deterministic	Hybrid 1	Hybrid 2
Objective (M$)	7173.36	7160.74	7191.78
Run time*	0.03 sec	10 hrs	7.38 hrs
Optimality gap(%)	0	1.8%	0

* Hybrid 1 method solved using Intel Xeon Dual 20 Core 2.0 GHZ Processor, 128 GB DDR4 ECC RAM. Hybrid 2 model and deterministic model solved using desktop computer 3.5 GHz CPU, 8 GB RAM.

Table 2.7: Number of constraints and variables

Method	Deterministic	Hybrid 1	Hybrid 2
Binary	42	19674	19674
Continuous	197	426376	57107
Constraints	245	692030	79850

Figs. 2.9 to 2.11 compare Hybrid 1 and 2 solutions versus demand uncertainty. For instance, Fig. 2.9 represents the capacity expansion of process 1 with respect to highest, nominal and lowest demand paths. In these figures, Hybrid 2 solution is fixed at most of the time steps for all the demand scenarios (LDR coefficients are zero); however Hybrid 1 solution (red circles) vary with respect to demand uncertainty. For instance, for capacity expansion of process 1 (Fig.2.9), it can be observed that Hybrid 1 solution is different at the third and fourth time steps for the highest demand path compared to nominal and the lowest demand paths but the Hybrid 2 solution is fixed or barely changed with respect to demand scenarios. Similarly for sale of chemical 5 (Fig. 2.11) and purchase of chemical 1 (Fig.2.10), it's observed that Hybrid 1 solution, at the third time step, is affected by demand uncertainty at the highest demand scenario compared to nominal and the lowest demand scenarios while Hybrid 2 solution is almost insensitive to demand variations. Generally, it is observed that Hybrid 1 method is more sensitive to demand uncertainty compared to Hybrid 2 method. This observation is as expected since Hybrid 1 method can accommodate the continuous variables at each node of the tree where each node corresponds to a different uncertainty scenario; therefore, it is more sensitive to different uncertainty scenarios compared to Hybrid 2 that uses LDR for continuous variables.

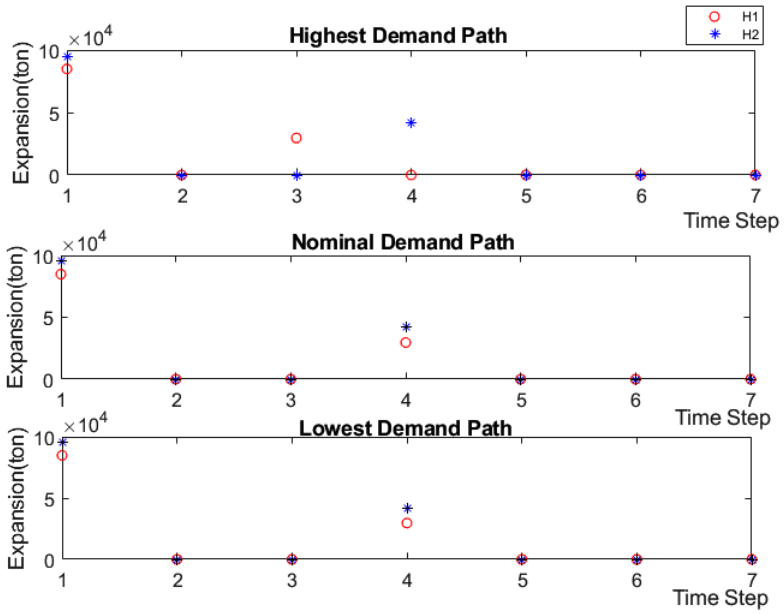


Figure 2.9: Solution comparison for capacity expansion of process 1

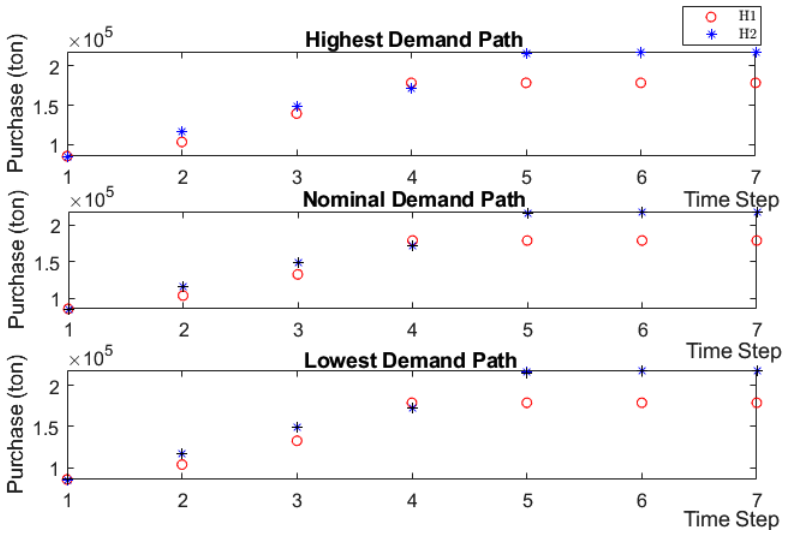


Figure 2.10: Solution comparison for purchase of chemical 1

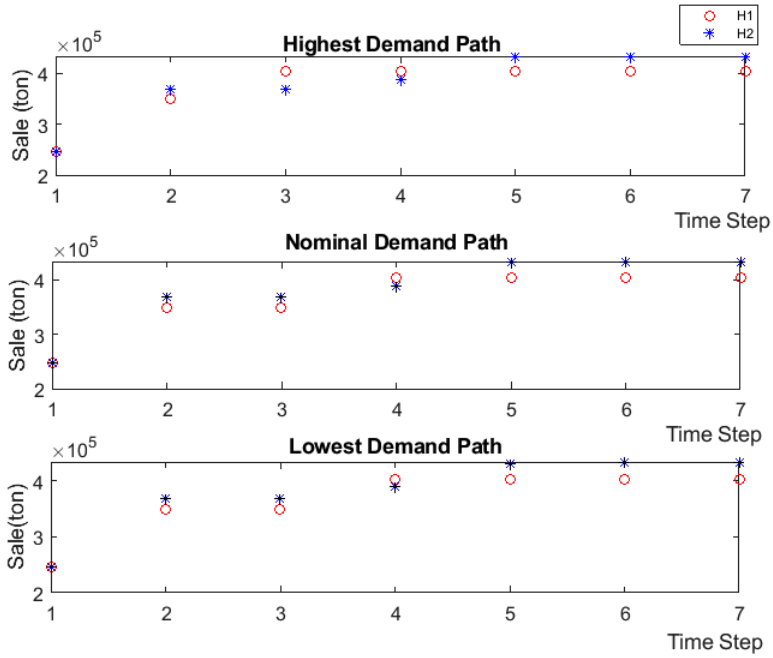


Figure 2.11: Solution comparison for sale of chemical 5

2.6 Conclusion

This work presents novel hybrid formulations to handle adaptive decision making with simultaneous consideration of static and time-dependent uncertainties while scenario-based method may result in computationally intractable formulation and LDR method results in non-convex constraints where strong duality cannot be applied to obtain the robust counterparts.

Comparing run time and computational expense of Hybrid 1 and 2 methods, it can be concluded that Hybrid 2 method results in a more efficient problem formulation requiring less computational effort while the solution quality is preserved (0.4% difference in objective value in this case study). Hybrid 1 method has the ability to accommodate binary and continuous variables at each node of the scenario tree. Therefore, it can capture nonlinear dependency of the binary and continuous variables with respect to uncertain parameters while Hybrid 2 method is restricted to a linear dependency of continuous variables on uncertain parameters. Thus, Hybrid 1 method is more flexible and can result in a better objective value. However, for large scale problems, the problem size can grow exponentially and the solution can be very time consuming and computationally prohibitive.

Therefore, for solving multistage adaptive problems with multiple sources of time-dependent and time-independent uncertainty where uncertain parameters and adaptive variables are multiplied together, Hybrid 2 method is recommended as a time-efficient method while the solution quality is preserved reliably. For applications where runtime and computational expense are not

restrictive and it is desired to obtain better objective value and to observe the solution dependency on uncertainty scenarios, Hybrid 1 method can be considered as a possible choice.

Finally, notice that the Hybrid 2 approach still depends on the scenario tree modeling of time-dependent uncertainty, which suffers from the dimensionality issue while by the time stages and the number of uncertain parameters increase. Future work will investigate decision rule approach for both continuous and binary variables, which is a promising direction to avoid the limitation of dimensionality issue.

Nomenclature

Indices

t	time period
s	scenario
k	node in scenario tree
l	sample
ω	static uncertainty
ξ	time dependent uncertainty

Parameters of capacity planning example

$\alpha_{i,t}$	Unit expansion cost of process i at the beginning of period t
$\beta_{i,t}$	Fixed cost of establishing or expanding process i at the beginning of period t
$\gamma_{i,t}, \Gamma_{i,t}$	Selling and buying prices of chemical j in period t
$\delta_{i,t}$	Unit operating cost for process i during period t
$\eta_{i,j}$	Input proportionality constant for chemical j in process i
$\mu_{i,j}$	Ouput proportionality constant for chemical j in process i
CI_t	Investment budget for period t
$a_{j,t}$	Availability of chemical j in period t
$d_{j,t}$	demand of chemical j in period t
N_i^{EXP}	Allowable number of expansions for process i
$X_{i,t}^L, X_{i,t}^U$	Lower and upper bounds for capacity expansion of process i in period t

Variables of capacity planning example

$P_{j,t}$	Units of chemical j purchased at the beginning of period t
$Q_{i,t}$	Total capacity of process i in period t
$S_{j,t}$	Units of chemical j sold at the end of period t
$W_{i,t}$	Operating level of process i in period t
$X_{i,t}$	Units of expansion of process i at the beginning of period t
$y_{i,t}$	Binary variable; if process i is expanded during period t, then $y_{i,t} = 1$, otherwise $y_{i,t} = 0$
$U_{i,k}$	Number of expansions up to time $t(k)$ for process i

Chapter 3

Multistage Adaptive Binary Optimization: Uncertainty Set Lifting versus Partitioning through Breakpoints Optimization

Abstract

Two methods for multistage adaptive binary optimization are investigated in this work. These methods referred to as binary decision rule and finite adaptability inherently share similarities in dividing the uncertainty set into subsets. In the binary decision rule method, the uncertainty is lifted using indicator functions which result in a nonconvex lifted uncertainty set. Linear decision rule is further applied to a convexified version of the lifted uncertainty set. In the finite adaptability method, the uncertainty set is divided into partitions and a constant decision is applied for each partition. In both methods, breakpoints are utilized either to define the indicator functions in the lifting method or to partition the uncertainty set in the finite adaptability method. In this work, we propose variable breakpoint location optimization for both methods. Extensive computational study on an illustrating example and a larger size case study is conducted. Performance of binary decision rule and finite adaptability methods under fixed and variable breakpoint approaches is compared.

3.1 Introduction

Multistage decision making under uncertainty has practical applications in many areas such as finance, engineering, and operations management, etc. To mention a few examples, Goulart et al. [6] applied stochastic optimization for robust control of linear discrete-time systems. Skaf and Boyd [66] designed an affine controller for linear dynamic systems. Ben Tal et al. [4] addressed the problem of minimizing the overall cost of a supply chain under demand uncertainty. Chuen-Teck and Melvyn [5] proposed a robust optimization method to tackle an inventory control problem where only limited information of demand is available. Chrysanthos et al. [8] studied the robust vehicle routing problem to minimize the delivery costs of a product to geographically dispersed customers using capacity-constrained vehicles. Calafoire [9] presented an affine control method for dynamic asset allocation.

Solving multistage adaptive optimization problem faces challenges. As Alexander and Arkadi [55] pointed out, multistage adaptive optimization problems including real-valued and binary decision variables are computationally intractable in general. Dyer and Stougie [67] have demonstrated that obtaining the optimal solution for the class of single stage uncertain problems involving only real-valued decisions is already P-hard. One popular solution approach to this problem is to use decision rules where variables are modeled as functions of uncertain parameters. Application of decision rules for real-valued functions in stochastic programming goes back to 1972 [68]. However, only recently, the decision rule-based approach has received major attention with the research advances made in robust optimization [69, 70]. Ben-Tal et al. [22] formulated the real-valued functions as affine functions of uncertain parameters. The simple structure of linear decision rules may result in some optimality loss. However, it has the advantage of providing the required scalability to deal with multistage stochastic adaptive problems. It should be mentioned that linear decision rules are shown to be optimal in some problem instances. For instance, Iancu and Parrilo [71] have shown the optimality of affine control policies in one-dimensional uncertainty within robust optimization context. The reader can refer to [72, 66] for other cases.

In order to reduce the loss of optimality due to linear decision rules, various nonlinear decision rule structures have been proposed. Motivated by the success of linear decision rules in providing the favorable scalability features for multi-stage problems, the nonlinear decision rules are formulated as $x(\xi) = x^\top O(\xi)$ where $O : \mathbb{R}^{m'} \to \mathbb{R}^m$ is a nonlinear operator that defines the structure of the decision rule. This structure improves the solution optimality considerably while retaining the scalability property. To mention some instances, different authors have suggested various decision rule structures such as linear[22, 6, 56, 7], piecewise linear [27, 28, 29, 30], multilinear [29], quadratic [73, 29] and polynomial. Kuhn et al. [56] proposed a method to estimate the approximation error introduced by linear decision rules. They have argued the method is applicable to

problems with random recourse and multiple decision stages. Chen et al. [27] addressed uncertain problems where only limited information of underlying stochastic parameters are available. The authors discussed that linear decision rules are inadequate for this type of problems and can result in infeasible solutions. They suggested an alternative second order conic optimization model that can be solved efficiently. Chen and Zhang [28] presented an extended affinely adjustable robust counterpart method to solve multistage uncertain linear problems and illustrated the potential of their proposed method is beyond what is presented in the literature. Georghiou et al. [29] proposed a lifting technique that provides tighter upper and lower bounds compared to the case where linear decision rule is directly applied to the original stochastic problem. They proposed a structured lifting method that gives rise to flexible piecewise linear and nonlinear decision rules. Goh and Sim [30] developed new piecewise linear decision rules that provide a more flexible formulation of the original uncertain problem and results in better bounds on the objective. Bertsimas et al. [32] proposed a framework for tackling linear dynamic systems under uncertainty. They introduced a hierarchy of polynomial control policies that exhibited strong numerical performance at a moderate computational expense.

Although there is a wealth of literature available for real-valued decision rules, the available literature for binary decision rules is relatively scarce. Bertsimas and Georghiou[35] proposed a structure for binary decision rules that models binary variables as piecewise constant functions and can provide high-quality solutions. However, the computational expense is significant. Bertsimas and Caramanis[34] proposed a structure for integer decision rules formulated as $y(\xi) = y^\top \lceil \xi \rceil$, where $y \in \mathbb{Z}^k$ and $\lceil \cdot \rceil$ is the ceiling function. In their work, the resulting problem is approximated and solved using a randomized algorithm [74] that provides only a limited guarantee on solution feasibility. Hanasusanto et al. [36] proposed a so-called k-adaptable structure that can only be applied to two-stage uncertain binary problems where the decision maker pre-commits to k second stage policies and implements the best one once the uncertain parameters are revealed. Recently, Bertsimas and Georghiou [39] proposed a systematic lifting method for binary decision rule that trades off scalability and optimality. This method can be applied to large multistage problems. They demonstrated that the method is highly scalable and provides high-quality solutions and can readily be used along with real-valued decision rules with the general structure of $x(\xi) = x^\top O(\xi), O : \mathbb{R}^{m'} \to \mathbb{R}^m$. Postek and Den Hertog [38] proposed a method to iteratively split the uncertainty set into subsets based on some heuristics. The method keeps the computational complexity at the same level as the static robust optimization problem. Bertsimas and Dunning [63] extended the work of finite adaptability and presented a partition-and-bound method for multistage adaptive mixed integer optimization problem.

While there are many possible ways for uncertainty lifting and uncertainty set partition, we focus on the grid partition based method in this work considering its simplicity in implementa-

tion. This means that for lifting method, we lift each uncertain parameter individually instead of any aggregated version of them. Whereas for the uncertainty set partition method, we divide the uncertainty set using hyper-rectangles. Under this assumption, the binary decision rule (lifting) method and the finite adaptability method (with grid partitioning) for addressing multistage adaptive optimization problems inherently share similarities. The goal of this study is to compare the solution complexity of these two methods in order to obtain insight on advantages and limitations of each method. The contribution of this work is summarized below:

1. Point out the similarity and difference between the lifting method and the finite adaptability (partitioning) method. We demonstrate that under equivalent setting, the partition method in general leads to better solution quality since lifting method has less flexibility due to the restriction of linear decision rule.

2. Propose novel breakpoint optimization formulations for both lifting and partitioning methods. It is shown that breakpoint optimization leads to improved solution quality with the cost of solving mixed integer nonlinear problems (MINLP) compared to mixed integer linear problems (MILP) under a fixed breakpoint setting.

3. Conduct an extensive computational study to investigate the advantages of lifting and finite adaptability methods under fixed breakpoint and variable breakpoint setting. We highlight the tractability of the lifting method for handling large cases with many time stages, and the limitation of finite adaptability method caused by the quick increase of model size.

The paper is organized as follows. Section 3.2 presents the general multistage adaptive binary optimization problem formulation and the traditional scenario tree-based method is applied to an illustrating example. Section 3.3 provides detailed formulations of the binary decision rule method based on lifting technique and the variable breakpoint technique is explained. Section 3.4 provides the detailed formulation for the finite adaptability method based on uncertainty set partitioning and the variable breakpoint based formulation is presented. Section 3.5 presents an extensive computational study using an inventory control problem and finally, Section 3.6 concludes the paper.

3.2 Multistage adaptive binary optimization

The multistage adaptive integer optimization problem to be addressed is as follows:

$$\min \ \mathbb{E}_{\xi}\left(\sum_{t}(\boldsymbol{D}_t\boldsymbol{\xi}_{[t]})^{\top}\ \boldsymbol{y}_t(\boldsymbol{\xi}_{[t]})\right) \tag{3.1a}$$

$$s.t. \ \sum_{\tau=1}^{t}\boldsymbol{B}_{t,\tau}\boldsymbol{y}_{\tau}(\boldsymbol{\xi}_{[\tau]}) \leq \boldsymbol{E}_t\boldsymbol{\xi}_{[t]} \qquad \forall \boldsymbol{\xi} \in \boldsymbol{\Xi}, \quad t = 1, ..., T \tag{3.1b}$$

$$\boldsymbol{y}_t(\boldsymbol{\xi}_{[t]}) \in \{0,1\}^{N_{\mathbf{y}_t}} \tag{3.1c}$$

where $\boldsymbol{D}_t$, $\boldsymbol{B}_{t,\tau}$ and $\boldsymbol{E}_t$ are constant vectors/matrices, $\boldsymbol{y}_t(\boldsymbol{\xi}_{[t]})$ is the adaptive binary decision variable expressed as a function of uncertainty vector $\boldsymbol{\xi}_{[t]}$, based on the following notation:

- $\xi_{q,t}$ scalar, q-th uncertain parameter of stage t
- $\boldsymbol{\xi}_t$ vector of uncertain parameters of stage t: $[\xi_{1,t}, \cdots, \xi_{\bar{q}_t,t}]^\top$, where $\bar{q}_t$ is the number of uncertain parameters in stage t
- $\boldsymbol{\xi}_{[t]}$ vector of uncertain parameters from stage 1 to t: $\left[1, \boldsymbol{\xi}_1^\top, ..., \boldsymbol{\xi}_t^\top\right]^\top$
- $\boldsymbol{\xi}$ vector of all uncertain parameters (from stage 1 to T): $\left[1, \boldsymbol{\xi}_1^\top, ..., \boldsymbol{\xi}_T^\top\right]^\top$, that is, $\boldsymbol{\xi}_{[T]}$

In this work, we consider polyhedral uncertainty set for $\boldsymbol{\xi} \in \mathbb{R}^{|\boldsymbol{\xi}_{[T]}|}$ ($|\boldsymbol{\xi}_{[T]}|$ is the size of $\boldsymbol{\xi}$):

$$\Xi = \{\boldsymbol{\xi} : \boldsymbol{J}\boldsymbol{\xi} \geq \boldsymbol{h}\} \tag{3.2}$$

where $\boldsymbol{J}$ and $\boldsymbol{h}$ are constant matrix and vector, respectively. For each uncertain parameter, we assume that it is subject to an interval $\Xi_{q,t} = \{\underline{\xi}_{q,t} \leq \xi_{q,t} \leq \overline{\xi}_{q,t}\}$.

Illustrating example

Throughout this paper, we will study the following illustrating example while presenting various solution methods. The problem has 2 stages (t=1, 2). Each stage includes only one uncertain parameter. First stage decision y_1 could depend on ξ_1, and the second stage decision y_2 could depend on both ξ_1 and ξ_2. The problem is formulated as:

$$\begin{aligned}
\min \quad & \mathbb{E}_\xi(-y_1(\boldsymbol{\xi}) - y_2(\boldsymbol{\xi})) && && \text{(3.3a)}\\
s.t. \quad & 2y_1(\boldsymbol{\xi}) \leq 1 + 2\xi_1 && \forall \boldsymbol{\xi} \in \Xi && \text{(3.3b)}\\
& 3y_1(\boldsymbol{\xi}) + 2y_2(\boldsymbol{\xi}) \leq 1 + 2\xi_1 + \xi_2 && \forall \boldsymbol{\xi} \in \Xi && \text{(3.3c)}\\
& y_t(\boldsymbol{\xi}) \in \{0,1\} && \forall \boldsymbol{\xi} \in \Xi, t = 1,2 && \text{(3.3d)}
\end{aligned}$$

Assume that each uncertain parameter follows an independent uniform distribution in a certain interval and the uncertainty set Ξ is given as:

$$\Xi = \{\boldsymbol{\xi} : \xi_1 \in [0,3],\ \xi_2 = [0,6]\}$$

The above numeric example can be casted into the formulation 3.1a-3.1c with $B_{11} = 2$, $B_{21} = 3$, $B_{22} = 2$, $E_1 = [1,2]$, $E_2 = [1,2,1]$, $D_1 = [-1,0]$, $D_2 = [-1,0,0]$, $\xi_{[1]} = [1,\xi_1]^\top$, $\xi_{[2]} = [1,\xi_1,\xi_2]^\top$.

Before presenting the lifting and the finite adaptability methods, traditional scenario tree-based method is applied to solve the illustrating problem in order to investigate the problem's true

optimal solution. While the number of scenarios is reasonably large, we can find the approximate optimal solution of the adaptive optimization problem. Figure 3.1 illustrates the scenario tree for 2 time stages where each node includes 3 branches. In the following scenario tree based model, y_k

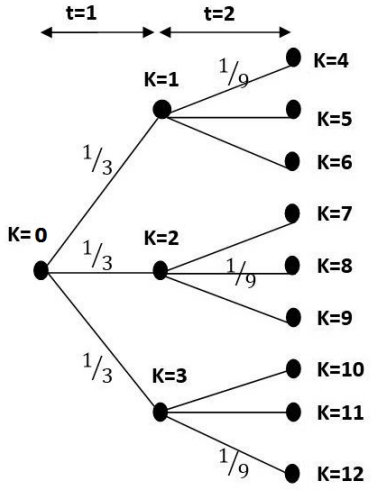


Figure 3.1: Scenario tree for 2 time stages and 3 branches for each node

represents the decision made at node k, $a(k)$ denotes the parent node of k, p_k is the probability of node k. Equations 3.4a to 3.4d present the nodal formulation. For the above scenario tree shown in Figure 3.1, $K_1 = \{1, 2, 3\}$, $K_2 = \{4, \cdots, 12\}$, which indicate the set of nodes at the first and second time steps, respectively.

$$\min \quad -\sum_{k \in K_1 \cup K_2} p_k y_k \tag{3.4a}$$

$$s.t. \quad 2y_k \leq 1 + 2\xi_k \qquad \forall k \in K_1 \tag{3.4b}$$

$$3y_{a(k)} + 2y_k \leq 1 + 2\xi_{a(k)} + \xi_k \qquad \forall k \in K_2 \tag{3.4c}$$

$$y_k \in \{0, 1\} \qquad \forall\ k \in K_1 \cup K_2 \tag{3.4d}$$

In this work, all the mixed integer linear optimization problems were modelled in GAMS and solved using CPLEX solver on a workstation (Intel Xeon Dual 20 Core 2.0 GHz Processor, 128 GB DDR4 ECC RAM). Table 3.1 present the results of the above model for 4, 11, 31 and 99 branches per node. Figures 3.2 and 3.3 illustrate the binary solution for 4 and 31 branches per node, respectively. In theses figures, the black squares indicate value 1 and the white squares indicate 0 value. Table 3.1 shows that by increasing the number of scenarios, the optimal objective value converges to a value around -1.6. We use this as a benchmark for comparing the different methods to be discussed.

Regarding the recourse decision, Figure 3.3 illustrates that for $y_1(\xi_1)$, there is a change point around 1 for ξ_1. For $y_2(\xi_1, \xi_2)$, there are three change points at around 0.5, 1 and 2 for ξ_1, and two change points at around 1 and 2 for ξ_2. Those values can be used as a basis for comparison while

Table 3.1: Results of scenario tree method for the illustrating example

Branches	Objective	Run time	Continuous Variables	Binary Variables
4	-1.625	0.032 s	3	20
11	-1.562	0.041 s	3	132
31	-1.605	0.078 s	3	993
99	-1.594	0.433 s	3	9900

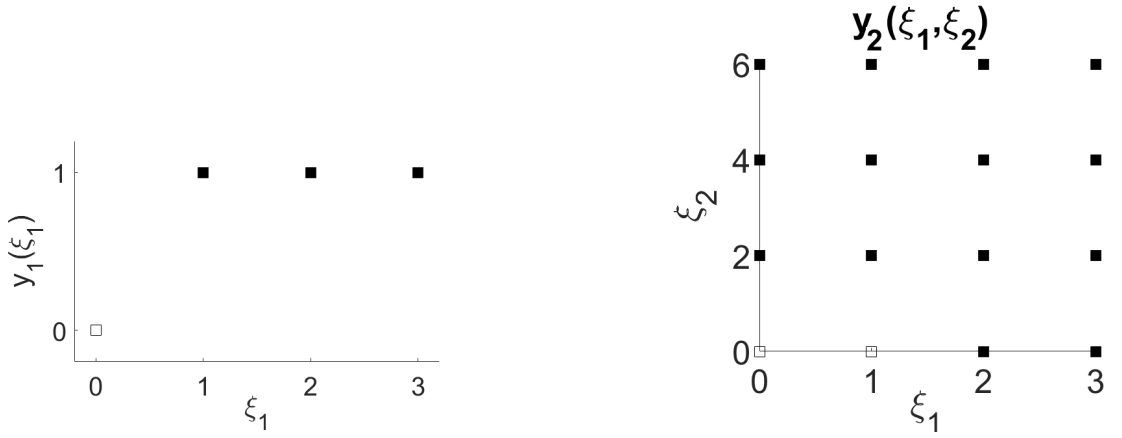


Figure 3.2: Solution under the scenario tree with 4 branches per node

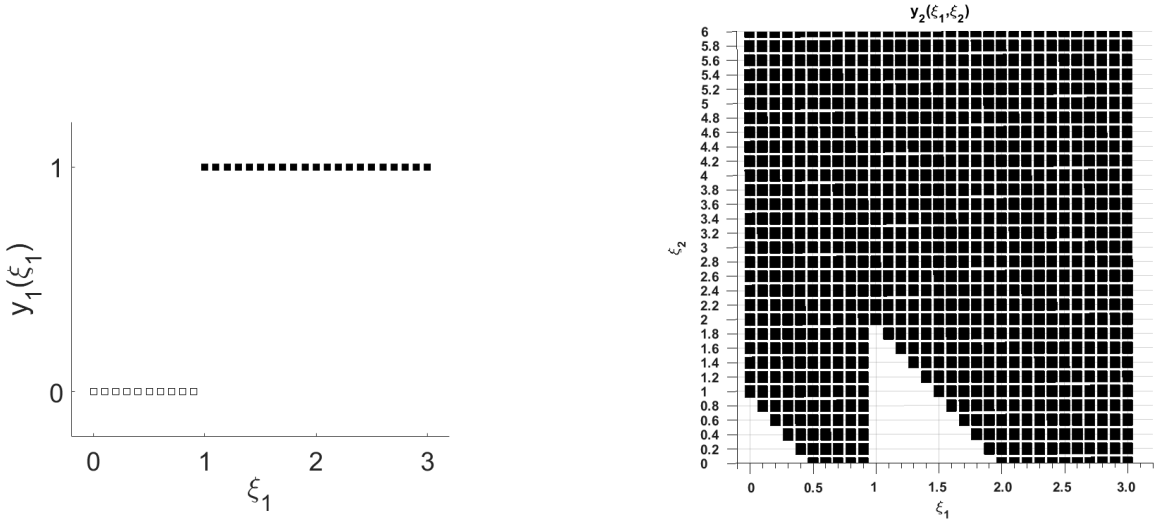


Figure 3.3: Solution under the scenario tree with 31 branches per node

the lifting and partitioning methods are implemented.

3.3 Binary decision rule with lifted uncertainty

3.3.1 Uncertainty lifting

Bertsimas and Georghiou [39] proposed a decision rule method for adaptive binary variables. This method enforces a linear relation between the binary variable and the lifted uncertain parameters. In this method, 0-1 indicator functions are defined based on a set of breakpoints for each uncertain parameter. The utilization of the indicator functions results in a nonconvex lifted uncertainty set. To resolve this problem, convex overestimation is applied to the lifted nonconvex set in order to obtain a convex set. The accuracy of the solution can be improved by increasing the number of breakpoints in the lifted set. While traditional scenario-tree methods result in an exponential growth of model size which restricts the application in large scale problems, the lifting method provides the scalability and tractability required for large scale problems.

Consider a single uncertain parameter $\xi_{q,t}$ subject to the interval $\Xi_{q,t} = \{\underline{\xi}_{q,t} \leq \xi_{q,t} \leq \overline{\xi}_{q,t}\}$, and assume the interval is divided into $(\bar{r}_{q,t}+1)$ subintervals using $\bar{r}_{q,t}$ breakpoints: $\alpha_{r,q,t}, r = 1, ..., \bar{r}_{q,t}$. To lift the uncertainty, the indicator functions $Q_{r,q,t}(\cdot)$ of uncertain parameters are used. The following list summarizes the notation used for lifting a single uncertain parameter $\xi_{q,t}$:

$\bar{r}_{q,t}$ scalar, number of breakpoints applied for $\xi_{q,t}$

$\alpha_{r,q,t}$ scalar, value of the r-th breakpoint for $\xi_{q,t}$, $r = 1, \cdots, \bar{r}_{q,t}$
as a generalization, we denote the bounds as: $\alpha_{0,q,t} \equiv \underline{\xi}_{q,t}$, $\alpha_{\bar{r}_{q,t}+1,q,t} \equiv \overline{\xi}_{q,t}$
$\alpha_{0,q,t} < \alpha_{1,q,t} < ... < \alpha_{\bar{r}_{q,t},q,t} < \alpha_{\bar{r}_{q,t}+1,q,t}$

$Q_{r,q,t}$ the r-th lifted uncertain parameter for $\xi_{q,t}$ (an indicator function of $\xi_{q,t}$),
$$Q_{r,q,t}(\xi_{q,t}) = \begin{cases} 1, & \text{if} \quad \xi_{q,t} \geq \alpha_{r,q,t} \\ 0, & \text{if} \quad \xi_{q,t} < \alpha_{r,q,t} \end{cases}$$

$\mathbf{Q}_{q,t}$ vector of all lifted parameters for $\xi_{q,t}$: $[Q_{1,q,t}, \cdots, Q_{\bar{r}_{q,t},q,t}]^\top$

$\mathbf{Q}_t$ vector of all lifted parameters for $\boldsymbol{\xi}_t$: $= [\mathbf{Q}_{1,t}^\top, \cdots, \mathbf{Q}_{\bar{q}_t,t}^\top]^\top$

$\mathbf{Q}_{[t]}$ vector, all lifted parameters for $\boldsymbol{\xi}_{[t]}$: $[1, \mathbf{Q}_1^\top, ..., \mathbf{Q}_t^\top]^\top$
note: the first element 1 is used for intercept term while implementing linear decision rule

$\boldsymbol{\xi}'_{q,t}$ vector of overall (original + lifted) uncertainty related to parameter $\xi_{q,t}$: $\left[\xi_{q,t}, \mathbf{Q}_{q,t}^\top\right]^\top$

$\boldsymbol{\xi}'_t$ vector of overall uncertainty for stage t: $\left[\boldsymbol{\xi}_{1,t}'^\top, \cdots, \boldsymbol{\xi}_{\bar{q}_t,t}'^\top\right]^\top$

$\boldsymbol{\xi}'_{[t]}$ vector of overall uncertainty from stage 1 to t: $\left[1, \boldsymbol{\xi}_1'^\top, ..., \boldsymbol{\xi}_t'^\top\right]^\top$

$\boldsymbol{\xi}'$ vector of overall uncertainty from stage 1 to T: $\left[1, \boldsymbol{\xi}_1'^\top, ..., \boldsymbol{\xi}_T'^\top\right]^\top$

$\Xi'_{p,q,t}$ the p-th line segment of the lifted uncertainty set for $\boldsymbol{\xi}'_{q,t}$, $p = 1, \cdots, \bar{r}_{q,t}+1$

$\Xi'_{q,t}$ the lifted uncertainty set (nonconvex) for $\boldsymbol{\xi}'_{q,t}$

Ξ' the lifted uncertainty set (nonconvex) for $\boldsymbol{\xi}'$

$\boldsymbol{\nu}_{i,p,q,t}$ the two extreme points of $\Xi'_{p,q,t}$, $i = 1, 2$

$$\boldsymbol{\nu}_{1,p,q,t} = \left[\alpha_{p-1,q,t}, \underbrace{1,\cdots,1}_{p-1 \text{ times}}, \underbrace{0,\cdots,0}_{\bar{r}_{q,t}-p+1 \text{ times}}\right]^\top, \quad p = 1,\cdots,\bar{r}_{q,t}+1$$

$$\boldsymbol{\nu}_{2,p,q,t} = \left[\alpha_{p,q,t}, \underbrace{1,\cdots,1}_{p-1 \text{ times}}, \underbrace{0,\cdots,0}_{\bar{r}_{q,t}-p+1 \text{ times}}\right]^\top, \quad p = 1,\cdots,\bar{r}_{q,t}+1$$

$\eta_{i,p,q,t}$ scalar, coefficient of extreme points in convex full formulation

Figure 3.4 illustrates the lifted uncertainty set ($\Xi'_{q,t}$) for a single uncertain parameter $\xi_{q,t}$ based on 1 and 2 breakpoints. It is clear that the lifted uncertainty set is nonconvex since it contains disconnected pieces ($\Xi'_{1,q,t},\cdots,\Xi'_{R_{q,t}+1,q,t}$) in a higher dimensional space.

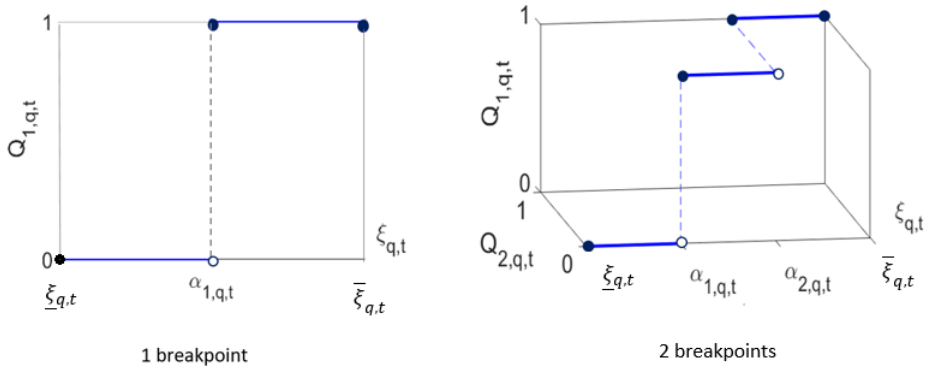


Figure 3.4: Lifting scheme for 1 breakpoint (left) and 2 breakpoints (right) on a single uncertain parameter $\xi_{q,t}$

In addition, projection matrices $\mathbf{P}_{\boldsymbol{\xi}_{[t]}} \in \mathbb{R}^{|\boldsymbol{\xi}_{[t]}|\times|\boldsymbol{\xi}'|}$ and $\mathbf{P}_{\mathbf{Q}_{[t]}} \in \mathbb{R}^{|\mathbf{Q}_{[t]}|\times|\boldsymbol{\xi}'|}$ are used in order to obtain $\boldsymbol{\xi}_{[t]}$ and $\mathbf{Q}_{[t]}(\boldsymbol{\xi})$ from the overall uncertainty vector, as follows:

$$\boldsymbol{\xi}_{[t]} = \mathbf{P}_{\boldsymbol{\xi}_{[t]}}\boldsymbol{\xi}' \qquad t = 1,...,T \tag{3.5a}$$

$$\mathbf{Q}_{[t]}(\boldsymbol{\xi}) = \mathbf{P}_{\mathbf{Q}_{[t]}}\boldsymbol{\xi}' \qquad t = 1,...,T \tag{3.5b}$$

Based on the above notation and the original uncertainty set definition in equation 3.2, the lifted nonconvex uncertainty set for $\boldsymbol{\xi}'$ can be written as:

$$\Xi' = \{\boldsymbol{\xi}' : \mathbf{P}_{\boldsymbol{\xi}_{[T]}}\boldsymbol{\xi}' \in \Xi; \boldsymbol{\xi}'_{q,t} \in \Xi'_{q,t}, t = 1,\cdots,T, q = 1,\cdots,\bar{q}_t\} \tag{3.6}$$

The nonconvexity of the lifted uncertainty set poses challenges to the optimization problem. The lifted set is convexified such that the semi-infinite constraints can be addressed. Given the $2(\bar{r}_{q,t}+1)$ extreme points, the convex hull of $\Xi'_{q,t}$ can be constructed as:

$$conv(\Xi'_{q,t}) = \{\boldsymbol{\xi}'_{q,t} = [\xi_{q,t}, \mathbf{Q}_{q,t}^\top]^\top : \qquad \exists\ \eta_{i,p,q,t} \tag{3.7}$$

$$\boldsymbol{\xi}'_{q,t} = \sum_{p=1}^{\bar{r}_{q,t}+1} \sum_{i=1}^{2} \eta_{i,p,q,t} \boldsymbol{\nu}_{i,p,q,t}$$

$$\sum_{p=1}^{\bar{r}_{q,t}+1} \sum_{i=1}^{2} \eta_{i,p,q,t} = 1$$

$$\eta_{i,p,q,t} > 0, \;\; i = 1, 2; p = 1, \cdots, \bar{r}_{q,t} + 1\}$$

Based on the above convex hull, we define the following convex overestimation for the overall uncertainty set Ξ' defined in equation 3.6:

$$\hat{\Xi}' = \{(\boldsymbol{\xi}', \boldsymbol{\eta}) : \mathbf{P}_{\boldsymbol{\xi}_{[T]}} \boldsymbol{\xi}' \in \Xi; (\boldsymbol{\xi}'_{q,t}, \boldsymbol{\eta}_{q,t}) \in conv(\Xi'_{q,t}), t = 1, \cdots, T, q = 1, \cdots, \bar{q}_t\} \tag{3.8}$$

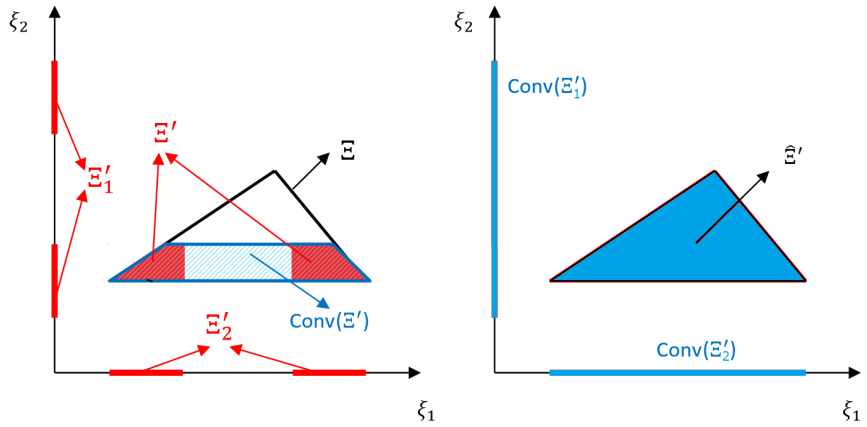


Figure 3.5: Illustration of the relation between $conv(\Xi')$ and its convex overestimation $\hat{\Xi}'$

Notice that the convex hull of $\Xi', conv(\Xi')$, is a subset of the above overestimation. Figure 3.5 illustrates the relation between $conv(\Xi')$ and the overestimation $\hat{\Xi}'$. The left figure illustrates Ξ' according to equation 3.6. Ξ' is obtained by intersecting the original uncertainty set Ξ (the black triangle) and the nonconvex sets Ξ'_1 and Ξ'_2 denoted by discontinuous red lines segments. Notice that the sets Ξ'_1 and Ξ'_2 are at least two dimensional where the dimension depends on the number of breakpoints used in the definition of the lifted set. However, for illustration purpose, the dimension of Ξ'_1 and Ξ'_2 is assumed to be one since using two dimensions on each axis will make the visualization impossible. The blue shaded area demonstrates $conv(\Xi')$. In the right figure, the blue triangle corresponds to the overestimation set $\hat{\Xi}'$ (equation 3.8), which is obtained by intersecting the original uncertainty set Ξ and $conv(\Xi'_1) \times conv(\Xi'_2)$. As the figure shows, $\hat{\Xi}'$ is an overestimation of $conv(\Xi')$.

Finally, for simplicity in derivation, we project the polyhedral uncertainty set $\hat{\Xi}'$ onto the space of $\boldsymbol{\xi}'$, and compactly write it as:

$$\hat{\Xi}' = \{\boldsymbol{\xi}' : \quad \boldsymbol{J}'\boldsymbol{\xi}' \geq \boldsymbol{h}'\} \tag{3.9}$$

where $\boldsymbol{J}'$ and $\boldsymbol{h}'$ are constant matrix calculated from the breakpoints value and the original uncertainty set parameters $\boldsymbol{J}, \boldsymbol{h}$.

Illustrating example (cont.)

In this section, the lifting method is applied to the illustrating example. Two breakpoints are considered for each uncertain parameter (r=1, 2):

$$\begin{aligned} \text{for } \xi_1: &\quad \alpha_{1,1,1} = 1, \alpha_{2,1,1} = 2 \\ \text{for } \xi_2: &\quad \alpha_{1,1,2} = 2, \alpha_{2,1,2} = 4 \end{aligned}$$

The lifted uncertainty vector is formulated as:

$$\boldsymbol{\xi}' = [1, \xi_1, Q_{1,1,1}(\xi_1), Q_{2,1,1}(\xi_1), \xi_2, Q_{1,1,2}(\xi_2), Q_{2,1,2}(\xi_2)]^\top$$

The 0-1 indicator functions are defined as:

$$Q_{1,1,1}(\xi_1) = \begin{cases} 1, & \text{if} \quad \xi_1 \geq 1 \\ 0, & \text{if} \quad \xi_1 < 1 \end{cases}, \qquad Q_{2,1,1}(\xi_1) = \begin{cases} 1, & \text{if} \quad \xi_1 \geq 2 \\ 0, & \text{if} \quad \xi_1 < 2 \end{cases}$$

$$Q_{1,1,2}(\xi_2) = \begin{cases} 1, & \text{if} \quad \xi_2 \geq 2 \\ 0, & \text{if} \quad \xi_2 < 2 \end{cases}, \qquad Q_{2,1,2}(\xi_2) = \begin{cases} 1, & \text{if} \quad \xi_2 \geq 4 \\ 0, & \text{if} \quad \xi_2 < 4 \end{cases}$$

The associated projection matrices and the corresponding (original or lifted) uncertainty vectors are:

$$\boldsymbol{P}_{\xi_{[1]}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{\xi}_{[1]} = \boldsymbol{P}_{\xi_{[1]}} \xi' = [1, \xi_1]^\top$$

$$\boldsymbol{P}_{\xi_{[2]}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}, \quad \boldsymbol{\xi}_{[2]} = \boldsymbol{P}_{\xi_{[2]}} \xi' = [1, \xi_1, \xi_2]^\top$$

$$\boldsymbol{P}_{Q_{[1]}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{Q}_{[1]}(\xi) = \boldsymbol{P}_{Q_{[1]}} \xi' = [1, Q_{1,1,1}, Q_{2,1,1}]^\top$$

$$\boldsymbol{P}_{Q_{[2]}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{Q}_{[2]}(\xi) = \boldsymbol{P}_{Q_{[2]}} \xi' = [1, Q_{1,1,1}, Q_{2,1,1}, Q_{1,1,2}, Q_{2,1,2}]^\top$$

The lifted uncertainty set is defined as: $\Xi' = \{\boldsymbol{\xi}' \in \mathbb{R}^2 : \mathbf{P}_{\boldsymbol{\xi}_{[2]}}\boldsymbol{\xi}' \in \Xi, \boldsymbol{\xi}'_{q,t} \in \Xi'_{q,t}\}$, where $\Xi = \{\xi : \xi_1 \in [0,3],\ \xi_2 = [0,6]\}$. Next, The convex hull for each single lifted uncertain parameter $conv(\Xi'_{q,t})$ is constructed. For simplicity in presentation, $\eta_{i,p,1,1}$ and $\eta_{i,p,1,2}$ are represented by $\eta_{p,i}$ and $\eta'_{p,i}$ in the following equations, respectively.

$conv(\Xi'_{1,1}) =$

$$\left\{ \boldsymbol{\xi}'_{1,1} : \begin{array}{l} \exists\ \eta_{1,1}, \eta_{1,2}, \eta_{2,1}, \eta_{2,2}, \eta_{3,1}, \eta_{3,2} \geq 0, \\ \eta_{1,1} + \eta_{1,2} + \eta_{2,1} + \eta_{2,2} + \eta_{3,1} + \eta_{3,2} = 1 \\ \begin{bmatrix} \xi_1 \\ Q_{1,1,1} \\ Q_{2,1,1} \end{bmatrix} = \eta_{1,1} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \eta_{1,2} \begin{bmatrix} \alpha_{1,1,1} \\ 0 \\ 0 \end{bmatrix} + \eta_{2,1} \begin{bmatrix} \alpha_{1,1,1} \\ 1 \\ 0 \end{bmatrix} + \eta_{2,2} \begin{bmatrix} \alpha_{2,1,1} \\ 1 \\ 0 \end{bmatrix} + \eta_{3,1} \begin{bmatrix} \alpha_{2,1,1} \\ 1 \\ 1 \end{bmatrix} + \eta_{3,2} \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} \end{array} \right\}$$

$conv(\Xi'_{1,2}) =$

$$\left\{ \boldsymbol{\xi}'_{1,2} : \begin{array}{l} \exists\ \eta'_{1,1}, \eta'_{1,2}, \eta'_{2,1}, \eta'_{2,2}, \eta'_{3,1}, \eta'_{3,2} \geq 0, \\ \eta'_{1,1} + \eta'_{1,2} + \eta'_{2,1} + \eta'_{2,2} + \eta'_{3,1} + \eta'_{3,2} = 1 \\ \begin{bmatrix} \xi_2 \\ Q_{1,1,2} \\ Q_{2,1,2} \end{bmatrix} = \eta'_{1,1} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \eta'_{1,2} \begin{bmatrix} \alpha_{1,1,2} \\ 0 \\ 0 \end{bmatrix} + \eta'_{2,1} \begin{bmatrix} \alpha_{1,1,2} \\ 1 \\ 0 \end{bmatrix} + \eta'_{2,2} \begin{bmatrix} \alpha_{2,1,2} \\ 1 \\ 0 \end{bmatrix} + \eta'_{3,1} \begin{bmatrix} \alpha_{2,1,2} \\ 1 \\ 1 \end{bmatrix} + \eta'_{3,2} \begin{bmatrix} 6 \\ 1 \\ 1 \end{bmatrix} \end{array} \right\}$$

The overall convex hull $conv(\Xi')$ is a subset of the following overestimated set $\hat{\Xi}' = \{\boldsymbol{\xi}' \in \mathbb{R}^2 : \mathbf{P}_{\boldsymbol{\xi}_{[2]}}\boldsymbol{\xi}' \in \Xi, \boldsymbol{\xi}'_{q,t} \in conv(\Xi'_{q,t}), q = 1, t = 1, 2\}$.

3.3.2 Binary decision rule

To approximate the optimal adaptive binary solution, binary decision rule is employed. It enforces a linear relation with respect to the lifted uncertainty (indicator function). The binary decision $\boldsymbol{y}_t$ depends on uncertainty up to time stage t. In binary decision rule, $\boldsymbol{y}_t$ is approximated by a linear combination of indicator functions from stage 1 to stage t, $\boldsymbol{Q}_{[t]}(\boldsymbol{\xi})$:

$$\boldsymbol{y}_t(\boldsymbol{\xi}_{[t]}) = \boldsymbol{Y}_t \boldsymbol{Q}_{[t]}(\boldsymbol{\xi}) \qquad t = 1, ..., T \tag{3.10a}$$

where the coefficients only take the following possible integer values: $\boldsymbol{Y}_t \in \{-1, 0, 1\}^{|\boldsymbol{y}_t| \times |\boldsymbol{Q}_{[t]}|}$. Furthermore, binary restriction on original variables need to be enforced, hence:

$$\mathbf{0} \leq \boldsymbol{Y}_t \boldsymbol{Q}_{[t]}(\boldsymbol{\xi}) \leq \boldsymbol{e} \qquad \forall \boldsymbol{\xi} \in \Xi, \quad t = 1, ..., T \tag{3.11}$$

where $\boldsymbol{e}$ is the vector of all ones. By applying binary decision rule to the general stochastic formulation, a semi-infinite optimization problem with finite number of variables but infinite number of

constraints is obtained:

$$\min \ \mathbb{E}_{\xi}\left(\sum_{t}(\boldsymbol{D}_t\boldsymbol{\xi}_{[t]})^{\top}\ \boldsymbol{Y}_t\boldsymbol{Q}_{[t]}(\boldsymbol{\xi})\right) \tag{3.12a}$$

$$s.t.\ \sum_{\tau=1}^{t}\boldsymbol{B}_{t,\tau}\left(\boldsymbol{Y}_{\tau}\mathbf{Q}_{[\tau]}(\boldsymbol{\xi})\right) \leq \boldsymbol{E}_t\boldsymbol{\xi}_{[t]} \qquad \forall\boldsymbol{\xi}\in\Xi, t=1,...,T \tag{3.12b}$$

$$\mathbf{0} \leq \boldsymbol{Y}_t\boldsymbol{Q}_{[t]}(\boldsymbol{\xi}) \leq \mathbf{e} \qquad \forall\boldsymbol{\xi}\in\Xi, t=1,...,T \tag{3.12c}$$

In the above model, the constraints contain nonconvex indicator functions of original uncertain parameters. Using Eqs. 3.5a and 3.5b, the stochastic model can be written as the following model with linear constraints with respect to the lifted uncertainty $\boldsymbol{\xi}'$:

$$\min \ \mathbb{E}_{\xi'}\left(\sum_{t}(\boldsymbol{D}_t\boldsymbol{P}_{\boldsymbol{\xi}_{[t]}}\boldsymbol{\xi}')^{\top}\ \boldsymbol{Y}_t\boldsymbol{P}_{\boldsymbol{Q}_{[t]}}\boldsymbol{\xi}'\right) \tag{3.13a}$$

$$s.t.\ \sum_{\tau=1}^{t}\boldsymbol{B}_{t,\tau}\left(\boldsymbol{Y}_{\tau}\boldsymbol{P}_{\boldsymbol{Q}_{[\tau]}}\boldsymbol{\xi}'\right) \leq \boldsymbol{E}_t\boldsymbol{P}_{\boldsymbol{\xi}_{[t]}}\boldsymbol{\xi}' \qquad \forall\boldsymbol{\xi}'\in\Xi', \forall t\in 1,...,T \tag{3.13b}$$

$$\mathbf{0} \leq \boldsymbol{Y}_t\boldsymbol{P}_{\boldsymbol{Q}_{[t]}}\boldsymbol{\xi}' \leq \boldsymbol{e} \qquad \forall\boldsymbol{\xi}'\in\Xi', \forall t\in 1,...,T \tag{3.13c}$$

Since the lifted uncertainty set Ξ' is nonconvex and consequently the problem is still intractable. The problem can be conservatively approximated by using convex overestimation $\hat{\Xi}'$ of the lifted uncertainty set as defined in equation 3.9.

$$\min \ \mathbb{E}_{\xi'}\left(\sum_{t}(\boldsymbol{D}_t\boldsymbol{P}_{\boldsymbol{\xi}_{[t]}}\boldsymbol{\xi}')^{\top}\ \boldsymbol{Y}_t\boldsymbol{P}_{\boldsymbol{Q}_{[t]}}\boldsymbol{\xi}'\right) \tag{3.14a}$$

$$s.t.\ \sum_{\tau=1}^{t}\boldsymbol{B}_{t,\tau}\left(\boldsymbol{Y}_{\tau}\boldsymbol{P}_{\boldsymbol{Q}_{[\tau]}}\boldsymbol{\xi}'\right) \leq \boldsymbol{E}_t\boldsymbol{P}_{\boldsymbol{\xi}_{[t]}}\boldsymbol{\xi}' \qquad \forall\boldsymbol{\xi}'\in\hat{\Xi}', t\in 1,...,T \tag{3.14b}$$

$$\mathbf{0} \leq \boldsymbol{Y}_t\boldsymbol{P}_{\boldsymbol{Q}_{[t]}}\boldsymbol{\xi}' \leq \mathbf{e} \qquad \forall\boldsymbol{\xi}'\in\hat{\Xi}', t\in 1,...,T \tag{3.14c}$$

Since constraints 3.14b and 3.14c are linear with respect to $\boldsymbol{\xi}'$, and the uncertainty set $\hat{\Xi}'$ is a polyhedral set, duality theorem can be used to covert this semi-infinite problem into its deterministic robust counterpart. As an example, constraint 3.14b can be written as

$$\left(\sum_{\tau=1}^{t}\boldsymbol{B}_{t,\tau}\boldsymbol{Y}_{\tau}\boldsymbol{P}_{\boldsymbol{Q}_{[\tau]}} - \boldsymbol{E}_t\boldsymbol{P}_{\boldsymbol{\xi}_{[t]}}\right)\boldsymbol{\xi}' \leq 0 \quad \forall\boldsymbol{\xi}'\in\hat{\Xi}', t\in 1,...,T$$

or equivalently as in the following, based on the uncertainty set definition in equation 3.9

$$\max_{\boldsymbol{J}'\boldsymbol{\xi}'\geq\boldsymbol{h}'}\left(\sum_{\tau=1}^{t}\boldsymbol{B}_{t,\tau}\boldsymbol{Y}_{\tau}\boldsymbol{P}_{\boldsymbol{Q}_{[\tau]}} - \boldsymbol{E}_t\boldsymbol{P}_{\boldsymbol{\xi}_{[t]}}\right)\boldsymbol{\xi}' \leq 0 \quad \forall t\in 1,...,T$$

Its deterministic counterpart is obtained as follows after applying duality to the inner maximization problem:

$$\begin{cases} -\boldsymbol{h}'^{\top}\boldsymbol{\theta}_t \leq 0 & \forall t \\ -\boldsymbol{J}'^{\top}\boldsymbol{\theta}_t = (\sum_{\tau=1}^{t} \boldsymbol{B_{t,\tau}} \boldsymbol{Y}_{\tau} \boldsymbol{P_{Q_{[\tau]}}} - \boldsymbol{E_t} \boldsymbol{P_{\xi_{[t]}}})^{\top} & \forall t \\ \boldsymbol{\theta}_t \geq 0 & \forall t \end{cases}$$

Similarly, constraint 3.14c can be divided into two parts: $\mathbf{0} \leq \boldsymbol{Y}_t \boldsymbol{P_{Q_{[t]}}} \boldsymbol{\xi}'$ and $\boldsymbol{Y}_t \boldsymbol{P_{Q_{[t]}}} \boldsymbol{\xi}' \leq \boldsymbol{e}$, and the robust counterpart is derived accordingly. The overall deterministic counterpart formulation is given by

$$\min \ \sum_t \mathrm{Tr}(\mathbb{E}[\boldsymbol{\xi}'\boldsymbol{\xi}'^{\top}] \boldsymbol{P}_{\boldsymbol{\xi}_{[t]}}^{\top} \boldsymbol{D}_t^{\top} \boldsymbol{Y}_t \boldsymbol{P_{Q_{[t]}}}) \tag{3.15a}$$

$$s.t. \ -\boldsymbol{h}'^{\top} \boldsymbol{\theta}_t \leq 0 \qquad \forall t \tag{3.15b}$$

$$-\boldsymbol{J}'^{\top}\boldsymbol{\theta}_t = (\sum_{\tau=1}^{t} \boldsymbol{B_{t,\tau}} \boldsymbol{Y}_{\tau} \boldsymbol{P_{Q_{[\tau]}}} - \boldsymbol{E_t} \boldsymbol{P_{\xi_{[t]}}})^{\top} \quad \forall t \tag{3.15c}$$

$$\boldsymbol{\theta}_t \geq 0 \qquad \forall t \tag{3.15d}$$

$$-\boldsymbol{h}'^{\top}\boldsymbol{\lambda}_t \leq 0 \qquad \forall t \tag{3.15e}$$

$$-\boldsymbol{J}'^{\top}\boldsymbol{\lambda}_t = (-\boldsymbol{Y}_t \boldsymbol{P_{Q_{[t]}}})^{\top} \qquad \forall t \tag{3.15f}$$

$$\boldsymbol{\lambda}_t \geq 0 \qquad \forall t \tag{3.15g}$$

$$-\boldsymbol{h}'^{\top}\boldsymbol{\phi}_t \leq \boldsymbol{e} \qquad \forall t \tag{3.15h}$$

$$-\boldsymbol{J}'^{\top}\boldsymbol{\phi}_t = (\boldsymbol{Y}_t \boldsymbol{P_{Q_{[t]}}})^{\top} \qquad \forall t \tag{3.15i}$$

$$\boldsymbol{\phi}_t \geq 0 \qquad \forall t \tag{3.15j}$$

$$\boldsymbol{Y}_t \in \{-1, 0, 1\}^{|\boldsymbol{y}_t| \times |\boldsymbol{Q}_{[t]}|} \tag{3.15k}$$

where $\mathrm{Tr}(\cdot)$ is trace operator and $\mathbb{E}[\boldsymbol{\xi}'\boldsymbol{\xi}'^{\top}]$ can be derived from the know distribution of the uncertainty.

Illustrating example (cont.)

Table 3.2 presents the results of the above lifting method for different number of breakpoints. For 1 breakpoint case, the breakpoint values are set as 1.5 for ξ_1, and 3 for ξ_2. For 2 breakpoints case, the breakpoint values are set as $(1, 2)$ for ξ_1, and $(2, 4)$ for ξ_2. For 9 breakpoints case, the breakpoint values are set as $(0.3, 0.6, 0.9, \cdots, 2.7)$ for ξ_1, and $(0.6, 1.2, \cdots, 5.4)$ for ξ_2. The objective did not improve beyond 1.333 even for 9 breakpoints. This observation indicates that lifting solution quality may be restricted. Using the definition of the lifting method for the binary variable, we can plot the $y_1(\xi_1)$ and $y_2(\xi_1, \xi_2)$ variables. For the case of 2 breakpoints, the model solution is $\boldsymbol{Y}_1 = [0, 1, 0]$ and $\boldsymbol{Y}_2 = [0, 0, 0, 1, 0]$, and the adaptive binary variables are expressed

as:

$$y_1(\xi_1) = \boldsymbol{Y}_1 \boldsymbol{Q}_{[1]} = [0,1,0] \times [1, Q_{1,1,1}(\xi_1), Q_{2,1,1}(\xi_1)]^\top = Q_{1,1,1}(\xi_1)$$
$$y_2(\xi_1,\xi_2) = \boldsymbol{Y}_2 \boldsymbol{Q}_{[2]} = [0,0,0,1,0] \times [1, Q_{2,1,1}(\xi_1), Q_{1,1,2}(\xi_2), Q_{2,1,2}(\xi_2)]^\top = Q_{1,1,2}(\xi_2)$$

As shown in Figure 3.6, the binary variable y_2 is only a function of parameter ξ_2 which indicates a restricted solution quality.

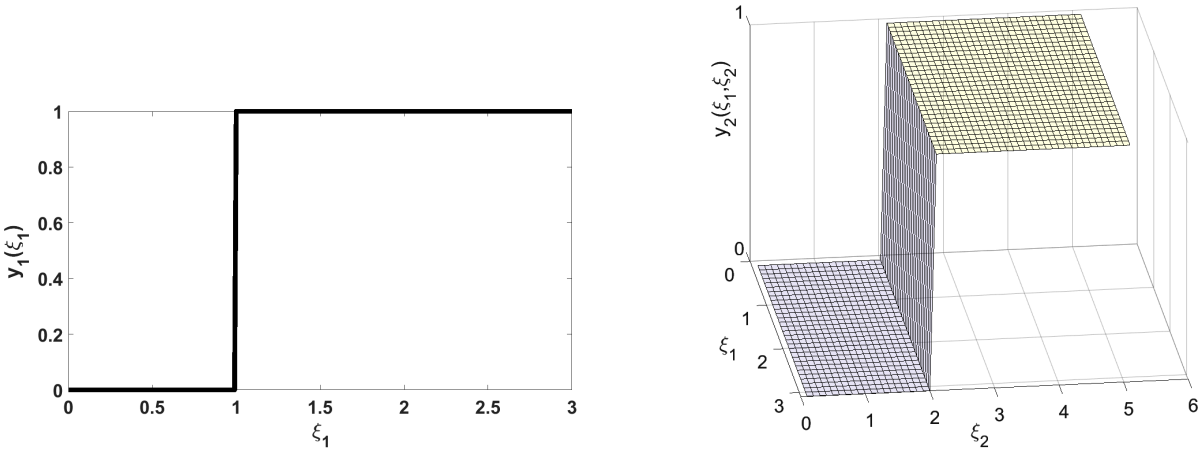


Figure 3.6: $y_1(\xi_1), y_2(\xi_1,\xi_2)$ solution under 2 breakpoints: $(1,2)$ for ξ_1, and $(2,4)$ for ξ_2

Table 3.2: Solution from lifting method with different number of fixed breakpoints

Number of breakpoints	Number of variables	Objective	Run time
1	5 Integer, 159 continuous	-1.000	0.022 s
2	8 Integer, 199 continuous	-1.333	0.078 s
9	29 Integer, 514 continuous	-1.333	0.094 s

3.3.3 Breakpoint optimization for lifting method

In this section, we assume that the location of breakpoints is not pre-fixed and they are optimized instead. Based on equation 3.9, the breakpoints information is contained in the parameters $\boldsymbol{J}', \boldsymbol{h}'$ and can be formulated as:

$$\boldsymbol{J}' = \boldsymbol{J}_0 + \boldsymbol{J}_1 \boldsymbol{\alpha}$$
$$\boldsymbol{h}' = \boldsymbol{h}_0 + \boldsymbol{h}_1 \boldsymbol{\alpha}$$

where $\boldsymbol{J}_0, \boldsymbol{J}_1, \boldsymbol{h}_0, \boldsymbol{h}_1$ are known constant matrices/vectors, $\boldsymbol{\alpha}$ is the vector involving all location information of breakpoints. The above formulated $\boldsymbol{J}'$ and $\boldsymbol{h}'$ should be substituted in the dual counterpart formulation (Eqs 3.15a to 3.15j) to complete the variable breakpoint lifting technique.

$$\min \ \mathbb{E}_{\xi'} \sum_t \mathrm{Tr}(\mathbb{E}[\boldsymbol{\xi}'\boldsymbol{\xi}'^{\top}]\boldsymbol{P}^{\top}_{\boldsymbol{\xi}_{[t]}}\boldsymbol{D}_t^{\top}\boldsymbol{Y}_t\boldsymbol{P}_{\boldsymbol{Q}_{[t]}}) \tag{3.16a}$$

$$s.t. \ -(\boldsymbol{h}_0 + \boldsymbol{h}_1\boldsymbol{\alpha})'^{\top}\boldsymbol{\theta}_t \leq 0 \qquad \forall t \tag{3.16b}$$

$$-(\boldsymbol{J}_0 + \boldsymbol{J}_1\boldsymbol{\alpha})^{\top}\boldsymbol{\theta}_t = (\sum_{s=1}^{t} \boldsymbol{B_{ts}}\boldsymbol{Y}_s\boldsymbol{P}_{\boldsymbol{Q}_{[s]}} - \boldsymbol{E_t}\boldsymbol{P}_{\boldsymbol{\xi}_{[t]}})^{\top} \quad \forall t \tag{3.16c}$$

$$\boldsymbol{\theta}_t \geq 0 \qquad \forall t \tag{3.16d}$$

$$-(\boldsymbol{h}_0 + \boldsymbol{h}_1\boldsymbol{\alpha})'^{\top}\boldsymbol{\lambda}_t \leq 0 \qquad \forall t \tag{3.16e}$$

$$-(\boldsymbol{J}_0 + \boldsymbol{J}_1\boldsymbol{\alpha})'^{\top}\boldsymbol{\lambda}_t = (-\boldsymbol{Y}_t\boldsymbol{P}_{\boldsymbol{Q}_{[t]}})^{\top} \qquad \forall t \tag{3.16f}$$

$$\boldsymbol{\lambda}_t \geq 0 \qquad \forall t \tag{3.16g}$$

$$-(\boldsymbol{h}_0 + \boldsymbol{h}_1\boldsymbol{\alpha})'^{\top}\boldsymbol{\phi}_t \leq \boldsymbol{e} \qquad \forall t \tag{3.16h}$$

$$-(\boldsymbol{J}_0 + \boldsymbol{J}_1\boldsymbol{\alpha})^{\top}\boldsymbol{\phi}_t = (\boldsymbol{Y}_t\boldsymbol{P}_{\boldsymbol{Q}_{[t]}})^{\top} \qquad \forall t \tag{3.16i}$$

$$\boldsymbol{\phi}_t \geq 0 \qquad \forall t \tag{3.16j}$$

$$\boldsymbol{Y}_t \in \{-1, 0, 1\}^{|\boldsymbol{y}_t| \times |\boldsymbol{Q}_{[t]}|} \tag{3.16k}$$

Notice that expectation of the lifted uncertainty $\mathbb{E}(\boldsymbol{\xi}')$ also depends on the location of the variable breakpoints, so $\mathbb{E}[\boldsymbol{\xi}'\boldsymbol{\xi}'^{\top}]$ is also a function of $\boldsymbol{\alpha}$ (which can be evaluated analytically based on the distribution information). The resulting model will be a mixed integer nonlinear optimization problem, where the integer variables are the binary decision rule coefficients $\boldsymbol{Y}_t$, and the continuous variables include the dual variables $\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{\phi}$ and the breakpoints locations $\boldsymbol{\alpha}$.

In this work, all the MINLP problems were solved using the ANTIGONE solver [75] in GAMS platform on a workstation (Intel Xeon Dual 20 Core 2.0 GHz Processor, 128 GB DDR4 ECC RAM) using 10 hours time resource limit. The reported solution are with zero optimality gap, otherwise it is reported. Using a single variable breakpoint in the illustrating example, the obtained objective (-1.333) is better than using a single fixed breakpoint(-1.000). This shows the advantage of variable breakpoint lifting compared to the fixed breakpoint method. However, in this example using more than 1 variable breakpoints did not further improve the objective as shown in Table 3.3. This observation shows that lifting method's solution quality is restricted.

Table 3.3: Solution statistics for variable breakpoint lifting

Number of breakpoints	Number of variables	Objective	Run time
1	5 Integer, 513 continuous	-1.333	1.075 s
2	8 Integer, 876 continuous	-1.333	2.375 s
9	29 Integer, 6112 continuous	-1.333	2 min, 14 s

The optimized breakpoint values are summarized below:

- for 1 breakpoint case, 1 for ξ_1, 2 for ξ_2
- for 2 breakpoints case, (0, 1) for ξ_1, (0, 2) for ξ_2
- for 9 breakpoints case, (1.5, 1.5, 1.5, 1.5, 1.5, 1.5, 3, 3, 3) for ξ_1, (1, 1, 1, 1, 2.52, 3.6, 4.2, 4.8, 5.4) for ξ_2

Notice that the variable breakpoints are not forced to be different to each other, so the solution has overlapped breakpoints while the number of breakpoints increases.

3.4 Finite adaptability with uncertainty set partitioning

3.4.1 Uncertainty set partitioning

In this method, the uncertainty set is divided into small partitions and constant binary decision is assumed for each partition. Similar to lifting method, we define breakpoints for each uncertain parameter and each subset is a box type of uncertainty set. Figure 4.2 illustrates the partitioning of a two dimensional rectangular uncertainty set and the corresponding scenario tree with each branch representing a subinterval for each parameter. In this figure, the interval of each uncertain parameter is divided into 3 segments such that there are 3 nodes in the first stage and 9 nodes in the second stage. The following notations were used in this method:

s	a scenario (each scenario is represented by a matrix structure (allowing empty elements), its element $s_{q,t}$ gives the subinterval index of uncertain parameter $\xi_{q,t}$) under this scenario
$\bar{r}_{q,t}$	scalar, number of breakpoints for $\xi_{q,t}$
$\alpha_{s_{q,t},q,t}$	scalar, upper bound value for element $\xi_{q,t}$ under scenario s

As an example, consider a two-stage problem: the first stage has uncertain parameters $\xi_{1,1}, \xi_{2,1}$ and the second stage has one uncertain parameter $\xi_{1,2}$. Assume no breakpoint is applied to $\xi_{1,1}$, 1 breakpoint is applied to $\xi_{2,1}$, and 2 breakpoints are applied to $\xi_{1,2}$. Then the set $\mathcal{S}$ is:

$$s \in \mathcal{S} = \left\{ \begin{bmatrix} 1 & 1 \\ 1 & * \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 2 & * \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 1 & * \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 2 & * \end{bmatrix}, \begin{bmatrix} 1 & 3 \\ 1 & * \end{bmatrix}, \begin{bmatrix} 1 & 3 \\ 2 & * \end{bmatrix} \right\}$$

where "*" denotes that the corresponding element is not existing.

For each scenario $s \in S$, the uncertainty set is a hyper-rectangular (subset of original set Ξ)

$$\Xi_s = \{\boldsymbol{\xi} : \alpha_{s_{q,t}-1,q,t} \leq \xi_{q,t} \leq \alpha_{s_{q,t},q,t};\ t = 1, ..., T,\ q = 1, ..., \bar{q}_t\}$$

which can be compactly written as

$$\Xi_s = \{\boldsymbol{\xi} : \boldsymbol{W}_s \boldsymbol{\xi} \geq \boldsymbol{V}_s\} \tag{3.17}$$

3.4.2 Finite adaptability

Next, the finite adaptability method (also denoted as "partitioning method" in this work) is applied to the original problem 3.1a-3.1c. The idea is a combination of scenario tree based stochastic formulation and robust optimization. For each scenario, we enforce constraint satisfaction over a uncertainty set as defined in equation 4.4 instead of a single point in the uncertainty space. The resulting model can be cast as:

$$\min_{y} \quad \sum_{s\in S} p_s \sum_{t=1}^{T} \boldsymbol{D}_t \mathbb{E}_{\Xi_s}(\xi) \boldsymbol{y}_{t,s} \tag{3.18a}$$

$$s.t. \quad \sum_{\tau=1}^{t} \boldsymbol{A}_{t\tau} \boldsymbol{y}_{\tau,s} \leq \boldsymbol{E}_t \boldsymbol{\xi} \qquad \forall t, s, \boldsymbol{\xi} \in \Xi_s \tag{3.18b}$$

$$\boldsymbol{y}_{t,s} = \boldsymbol{y}_{t,s'} \qquad \forall (t, s, s') \in SP \tag{3.18c}$$

where the last constraint is enforcing non-anticipativity and SP is the set of all scenarios with the same path up to time t:

$$SP = \{(t, s, s') : \; s_{q,\tau} = s'_{q,\tau}, \quad \forall \tau = 1, ..., t, q = 1, ..., \bar{q}_\tau\}$$

Reformulating the semi-infinite constraint using linear programming duality, the corresponding robust scenario based formulation is

$$\min_{y} \quad \sum_{s\in S} p_s \sum_{t=1}^{T} \boldsymbol{D}_t \mathbb{E}_{\Xi_s}(\xi) \boldsymbol{y}_{t,s} \tag{3.19a}$$

$$s.t. \quad \boldsymbol{V}_s^\top \boldsymbol{\theta}_{t,s} \geq \sum_{\tau}^{t} \boldsymbol{A}_{t\tau} \boldsymbol{y}_{\tau,s} \quad \forall s, t \tag{3.19b}$$

$$\boldsymbol{W}_s^\top \boldsymbol{\theta}_{t,s} = \boldsymbol{E}_t^\top \qquad \forall s, t \tag{3.19c}$$

$$\boldsymbol{\theta}_{t,s} \geq 0 \qquad \forall s, t \tag{3.19d}$$

$$\boldsymbol{y}_{t,s} = \boldsymbol{y}_{t,s'} \qquad \forall (t, s, s') \in SP \tag{3.19e}$$

Illustrating example (cont.)

In this section, the partitioning method is applied to the illustrating example 3.3a-3.3d and the results are presented. Figure 4.2 illustrates the partitioning and the corresponding scenario tree for the illustrating example.

The formulation for partitioning of the uncertainty set is based on 2 equally-spaced breakpoints for each uncertainty element: for $\xi_{1,1} \in [0, 3]$ the breakpoints are set as $0 < 1 < 2 < 3$; for $\xi_{1,2} \in [0, 6]$ the breakpoints are $0 < 2 < 4 < 6$. Hence, $\mathcal{S} = \Big\{[1, 1], [1, 2], [1, 3], [2, 1], [2, 2], [2, 3], [3, 1], [3, 2], [3, 3]\Big\}$.

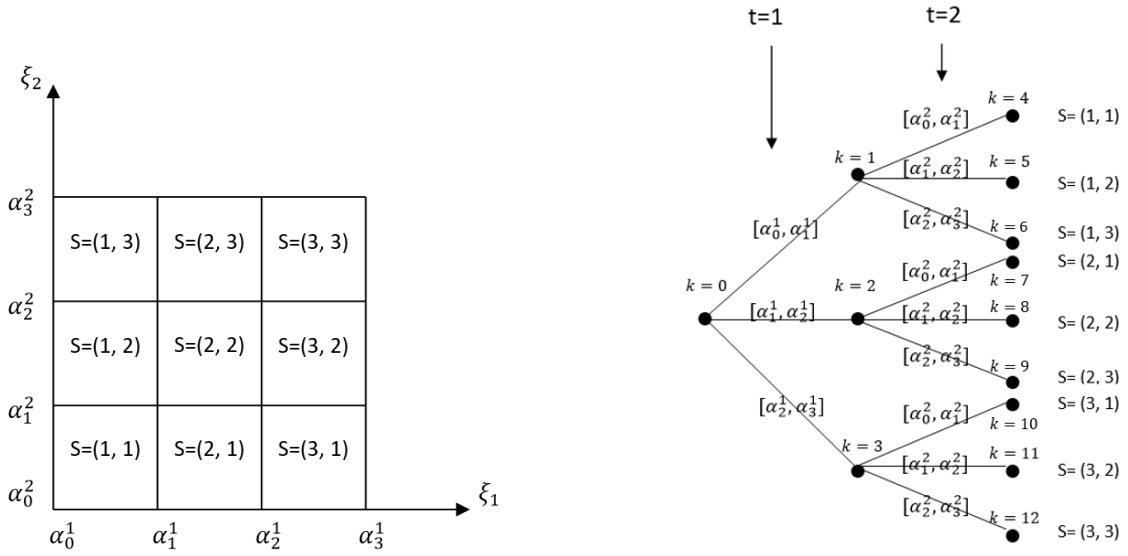


Figure 3.7: Partitioning of the 2-dimensional uncertainty set (left figure) and scenario tree representation (right figure)

The non-anticipativity condition set SP contains the following elements:

$$\Big(1,[1,1],[1,2]\Big),\Big(1,[1,1],[1,3]\Big),\Big(1,[1,2],[1,3]\Big)$$
$$\Big(1,[2,1],[2,2]\Big),\Big(1,[2,1],[2,3]\Big),\Big(1,[2,2],[2,3]\Big)$$
$$\Big(1,[3,1],[3,2]\Big),\Big(1,[3,1],[3,3]\Big),\Big(1,[3,2],[3,3]\Big)$$

The partitioning based model is MILP problem. Table 3.4 presents the results from finite adaptability method. As the number of partitions increases, the objective improves until it reaches close to the optimal solution obtained from scenario method (-1.594 for 99 branches per node). For comparison, the branches and breakpoints are evenly distributed in scenario and finite adaptability method, respectively. For 29 breakpoints in partitioning method and 31 branches in scenario method, the optimal objectives are -1.589 and -1.605, respectively. Figures 3.8 and 3.9 illustrates the partitioning solution for 2 and 29 breakpoints. As the figures demonstrates, there is a close match between the partitioning and scenario solution.

Table 3.4: Solution of finite adaptability method

Number of breakpoints	Number of variables	Objective	Run time
2	18 binary, 109 continuous	-1.444	0.015 s
9	200 binary, 1201 continuous	-1.510	0.016 s
29	1800 binary, 10801 continuous	-1.589	0.078 s

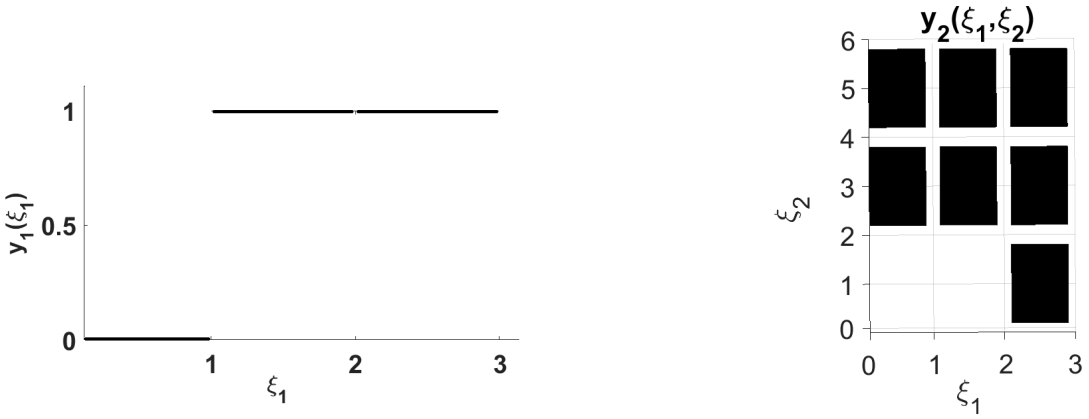


Figure 3.8: Solution from finite adaptability method using 2 breakpoints for each parameter

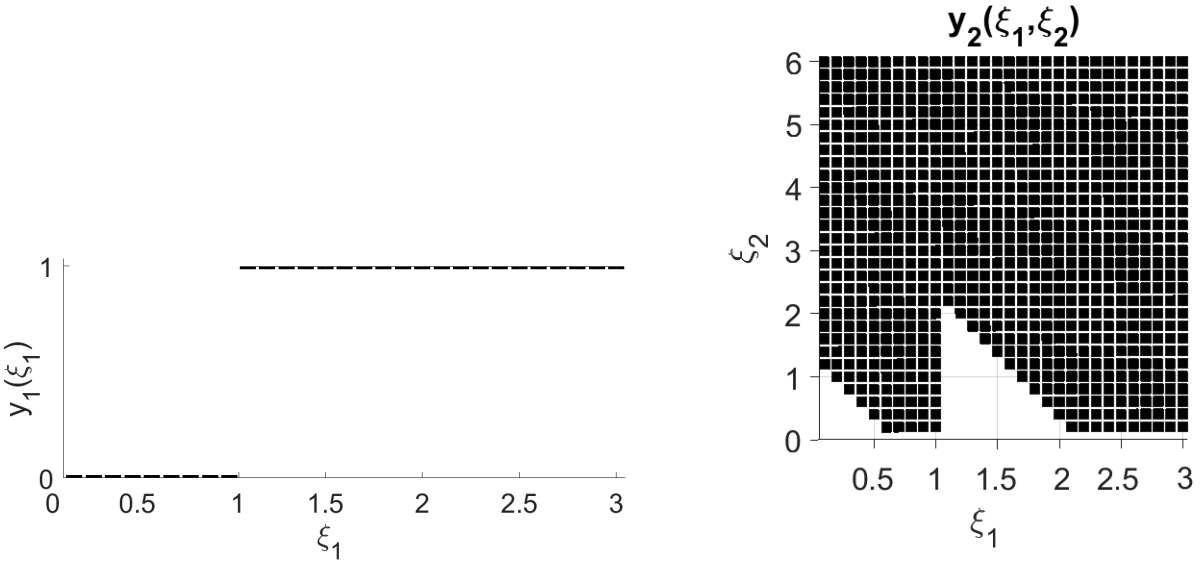


Figure 3.9: Solution from finite adaptability method using 29 breakpoints for each parameter

3.4.3 Breakpoint optimization in partitioning method

In the variable breakpoint partitioning technique, the location of breakpoints is unknown a priori and it is optimized instead. The location of breakpoints is reflected in the $\boldsymbol{V}_s$ matrix of equation 4.4 and it is formulated as follows:

$$\boldsymbol{V}_s = \boldsymbol{V}_s^0 + \boldsymbol{V}_s^1 \boldsymbol{\alpha} \qquad \forall s \in \mathcal{S}$$

where $\boldsymbol{V}_s^0$ and $\boldsymbol{V}_s^1$ are constant matrix. The corresponding robust scenario based formulation is

$$\min_{y} \quad \sum_{s \in S} p_s \sum_{t=1}^{T} \boldsymbol{D}_t \mathbb{E}_{\Xi_s}(\boldsymbol{\xi}) \boldsymbol{y}_{t,s} \tag{3.20a}$$

$$s.t. \quad (\boldsymbol{V}_s^0 + \boldsymbol{V}_s^1\boldsymbol{\alpha})^\top \boldsymbol{\theta}_{t,s} \geq \sum_{\tau=1}^{t} \boldsymbol{A}_{t\tau}\boldsymbol{y}_{\tau,s} \forall s,t \tag{3.20b}$$

$$\boldsymbol{W}_s^\top \boldsymbol{\theta}_{t,s} = \boldsymbol{E}_t^\top \qquad \forall s,t \tag{3.20c}$$

$$\boldsymbol{\theta}_{t,s} \geq 0 \qquad \forall s,t \tag{3.20d}$$

$$\boldsymbol{y}_{t,s} = \boldsymbol{y}_{t,s'} \qquad \forall (t,s,s') \in SP \tag{3.20e}$$

In this formulation, the probability of occurrence of each scenario depends on the location of breakpoints since the length of each dimension in each scenario depends on the location of breakpoints. The probability p_s for each scenario and $\mathbb{E}_{\Xi_s}(\boldsymbol{\xi})$ are both functions of $\boldsymbol{\alpha}$ and they can be evaluated based on the distribution of the uncertainty.

Table 3.5 summarized the results of the variable partitioning technique. It can be observed that for the same number of partitions, the variable partitioning method provided a better objective compared to the fixed breakpoint method. For instance, for 2 breakpoints per uncertain parameter, the variable and fixed methods' objectives are -1.500 and -1.444, respectively. The variable method even provided a better objective with just 3 breakpoints compared to the fixed method with 9 breakpoints. However, for large number of partitions, the fixed breakpoint method is much faster in terms of solution and it could provide a better objective using 29 breakpoints in less than 1s compared to the variable method using 5 breakpoints after about 18 min run time.

Table 3.5: Variable breakpoint partitioning applied to the illustrating example

Number of breakpoints	Number of variables	Objective	Run time
1	8 binary, 83 continuous	-1.333	0.682 s
2	18 binary, 180 continuous	-1.500	0.604 s
3	32 binary, 315 continuous	-1.528	0.928s
5	72 binary, 699 continuous	-1.562	18 min 54 s

3.4.4 Flexibility comparison

We can observe from Table 3.2 that for the illustrating problem, lifting method's objective does not improve beyond -1.333 even for 9 breakpoints, while the partitioning method could provide a better objective of -1.444 with 2 breakpoints for each parameter. In this section, we investigate the reason why lifting method leads to restricted solution quality compared to the partitioning method. For this purpose, the solution from partitioning method is substituted into the decision rule solution from lifting method. Note that same breakpoints are applied in the lifting and partitioning method (as shown in Figure 4.2). This will result in a linear system of equations. If the set of equations has no solution, this means that the lifting method has restricted flexibility such

that it cannot cover the solution of the partitioning method.

Figure 3.8 presents the partitioning solution. In this problem, the interval for ξ_1, ξ_2 are divided into 3 equally distributed pieces. Based on the binary decision rule equations:

$$
\begin{aligned}
y_t &= \boldsymbol{Y}_t \boldsymbol{Q}_{[t]} \\
\boldsymbol{Q}_{[1]} &= [1, Q_{1,1,1}, Q_{2,1,1}] \\
\boldsymbol{Q}_{[2]} &= [1, Q_{1,1,1}, Q_{2,1,1}, Q_{1,1,2}, Q_{2,1,2}]
\end{aligned}
$$

For the three nodes at stage 1, the corresponding lifted uncertainty vector $Q_{[1]}$ takes the following three values:

$$
\begin{aligned}
&\text{Node 1} : \boldsymbol{Y}_1 [1,0,0]^\top = 0 \\
&\text{Node 2} : \boldsymbol{Y}_1 [1,1,0]^\top = 1 \\
&\text{Node 3} : \boldsymbol{Y}_1 [1,1,1]^\top = 1
\end{aligned}
$$

There is a solution $\boldsymbol{Y}_1 = [0,1,0]$ for the above equations. For the nine nodes at stage 2, the lifted uncertainty vector takes 9 different values for nodes 4 to 12:

$$
\begin{aligned}
&\text{Node 4} : \boldsymbol{Y}_2 [1,0,0,0,0]^\top = 0 \\
&\text{Node 5} : \boldsymbol{Y}_2 [1,0,0,1,0]^\top = 1 \\
&\text{Node 6} : \boldsymbol{Y}_2 [1,0,0,1,1]^\top = 1 \\
&\text{Node 7} : \boldsymbol{Y}_2 [1,1,0,0,0]^\top = 0 \\
&\text{Node 8} : \boldsymbol{Y}_2 [1,1,0,1,0]^\top = 1 \\
&\text{Node 9} : \boldsymbol{Y}_2 [1,1,0,1,1]^\top = 1 \\
&\text{Node 10} : \boldsymbol{Y}_2 [1,1,1,0,0]^\top = 1 \\
&\text{Node 11} : \boldsymbol{Y}_2 [1,1,1,1,0]^\top = 1 \\
&\text{Node 12} : \boldsymbol{Y}_2 [1,1,1,1,1]^\top = 1
\end{aligned}
$$

There is no solution $\boldsymbol{Y}_2$ satisfying this linear system of equations (there are 9 equations and 5 variables for $\boldsymbol{Y}_2$). This evidence indicates that lifting method's solution is restricted compared to partitioning method. The limited solution flexibility is due to the affine decision rule over the lifted uncertainty.

3.5 Case study: Inventory Control Problem

In this section, an inventory control problem with discrete ordering decisions is studied where the problem is adapted from [39]. This problem can be formulated as an multistage adaptive optimization problem with fixed recourse and can be explained as follows. At the beginning of each time period $t \in \mathcal{T} = \{1, ..., T\}$, the product demand ξ_t is observed. This demand can be satisfied in two ways: (1) by pre-ordering at period t using maximum N lots (binary variable $z_{n,t}$ is introduced to express whether we order from n lots in time period t or not), each delivering a fixed quantity q^z at the beginning of period t for a unit cost of c^z; (2) by placing a recourse order during the period t using maximum N lots (binary variable $y_{n,t}$ is introduced for this decision), each delivering immediately a fixed quantity q^y for a unit cost of c^y. The pre-ordering cost is always less than the immediate ordering cost ($c^z < c^y$). If the ordered quantity is greater than the demand, the excess units are stored in a warehouse that incurs a unit holding cost of c^h and can be used to satisfy future demand. Furthermore, the cumulative volume of pre-orders $\sum_{\tau=1}^{t}\sum_{n=1}^{N} q^z z_{n,\tau}$ can not exceed the ordering budget $\bar{B}_t$. The objective is to minimize the total ordering and holding costs over the planning horizon by determining the optimal decision of $z_{n,t}$ and $y_{n,t}(\xi_{[t]})$. Eqs 4.3a to 4.3f express the problem formulation.

$$\min \quad \mathbb{E}_{\xi}\left(\sum_{t=1}^{T}\sum_{n=1}^{N} c^z q^z z_{n,t} + c^y q^y y_{n,t}(\boldsymbol{\xi}_{[t]}) + c^h I_t(\boldsymbol{\xi}_{[t]})\right) \tag{3.21a}$$

$$s.t. \quad I_t(\boldsymbol{\xi}_{[t]}) = I_0 + \sum_{\tau=1}^{t}\left(\sum_{n=1}^{N} q^z z_{n,\tau} + q^y y_{n,\tau}(\boldsymbol{\xi}_{[\tau]}) - \xi_\tau\right) \quad \forall t \in \mathcal{T}, \boldsymbol{\xi} \in \Xi \tag{3.21b}$$

$$I_t(\boldsymbol{\xi}_{[t]}) \geq 0 \quad \forall t \in \mathcal{T}, \boldsymbol{\xi} \in \Xi \tag{3.21c}$$

$$\sum_{\tau=1}^{t}\sum_{n=1}^{N} q^z z_{n,\tau} \leq \bar{B}_t \quad \forall t \in \mathcal{T} \tag{3.21d}$$

$$z_{n,t} \in \{0,1\}, y_{n,t}(\boldsymbol{\xi}_{[t]}) \in \{0,1\} \quad \forall n, t \tag{3.21e}$$

Notice that $z_{n,t}$ is static binary decision variable, while $y_{n,t}$ is adaptive binary decision variable dependent on the realized uncertainty $\boldsymbol{\xi}_{[t]}$. The uncertainty set is represented as: $\Xi = \{\boldsymbol{\xi} \in \mathbb{R}^T : \boldsymbol{l} \leq \boldsymbol{\xi} \leq \boldsymbol{u}\}$. Uniform distribution is assumed for all of them. The bounds for each random parameter are chosen from $l_t \in [0, 5]$ and $u_t \in [10, 15]$ for $t = 1, ..., T$. The cumulative ordering budget equals to $\bar{B}_t = 10t$ for $t = 1, ...T$. It is also assumed that $q^z = q^y = 15/N$ and the initial inventory is zero ($I_0 = 0$).

Two different configurations of the inventory problem are studied in this work (Table 3.6). In the second configuration, the unit cost for immediate ordering (c^y), unit cost for storage at warehouse (c^h) and the maximum number of delivery lots (N) are greater than the first config-

uration. Since at the second configuration, the costs for immediate ordering and warehouse are higher, therefore it is more economical to satisfy demand using first stage decisions. Lifting and partitioning method using both fixed and variable breakpoint are applied to the case study and the obtained results are discussed.

Table 3.6: Inventory problem parameters

First Configuration

N	c^z	c^y	c^h	q^y	q^z
2	2	3	4	7.5	7.5

Second Configuration

N	c^z	c^y	c^h	q^y	q^z
3	2	5	6	5	5

First, the comparison is made between lifting and partitioning methods using same fixed breakpoint setting. The following results are obtained:

1. The number of variables grows exponentially in the partitioning method while it grows linearly in the lifting. For a large number of time stages (T) and number of breakpoints used (Br), the number of variables in the partitioning method can be prohibitively large and it may not be possible to run the model (Tables A.13 and A.14). Therefore, considering time and computational resource limitations, lifting method is suggested for large models (experiments next to $T = 10$, $Br = 2$ in this case study, Tables A.11 to A.12). Thus, the ability to run large models can be considered the main advantage of the lifting method.

2. It is observed that for experiments with small and medium model size (up to $T = 10$, $Br = 1$ in this case study), the partitioning method provides a better objective value in shorter run time compared to the lifting method. For instance, for experiments with 5 time steps, the partitioning method provides a better solution in 9.5 seconds (Table A.13, $T = 5$, $Br = 2$) compared to lifting method in 10 hours run time (Table A.11, $T = 5$, $Br = 15$). Figures 3.10 and 3.11 compare the fixed breakpoint lifting and finite adaptability (partitioning) methods for first and second configurations, respectively.

3. In general, in both partitioning and lifting methods, the objective improves by increasing the number of breakpoints except for large problems where the optimality gap is still quite large after 10 hrs run time limitation (Tables A.11 and A.12). Thus, for large models, in order to obtain the best objective in limited run time, there is no need to consider large number of breakpoints.

Next, the comparison is made between fixed and variable breakpoint methods. The following observations are made:

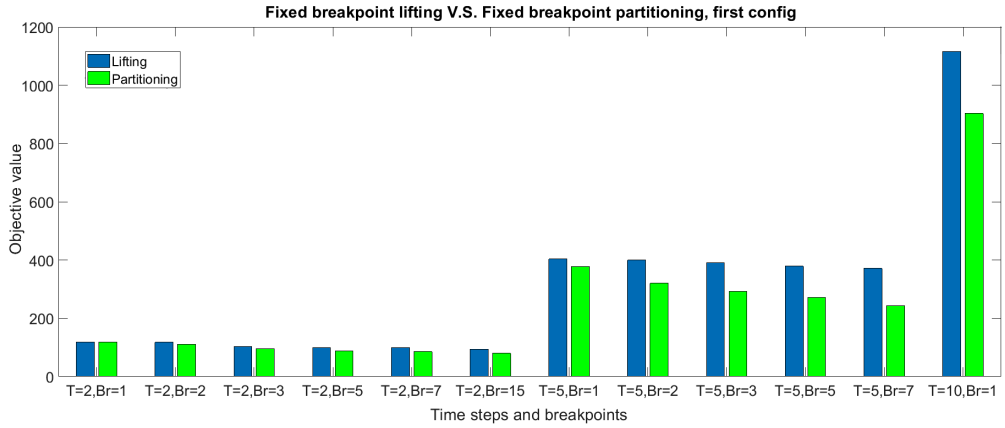


Figure 3.10: Comparison of lifting and finite adaptability method using same fixed breakpoints setting (first configuration).

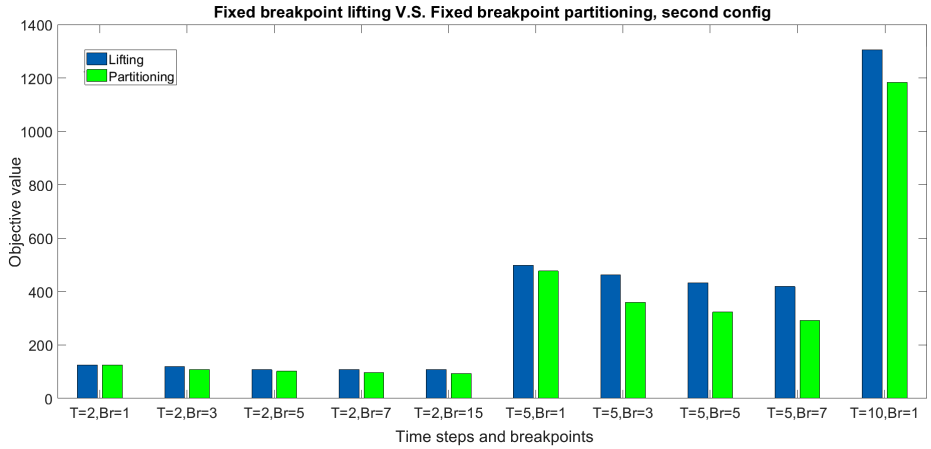


Figure 3.11: Comparison of lifting and finite adaptability method using same fixed breakpoints setting (second configuration).

1. Variable breakpoint lifting and partitioning techniques introduce additional variables and require mixed integer nonlinear optimization compared to mixed integer linear optimization for fixed breakpoint case. As shown in Tables A.11, A.13, A.15, A.16, and Figures 3.12 and 3.13, the variable breakpoint techniques can provide a better objective compared to fixed methods for small to medium size models ($T = 2$ and $T = 5$). However, the run time is longer. For a large number of time steps and breakpoints, the problem size is too large such that it is impossible to run the model under computational resource restrictions.

2. For small to medium size experiments ($T = 2$, $T = 5$), the fixed breakpoints partitioning method is recommended since it provides the best objective within the shortest run time considering computational resource restrictions.

3. For large size problems ($T = 10$, $T = 20$), the lifting method under fixed breakpoints is the only method that that results in a feasible solution considering computational resource

limitation. The partitioning technique leads to large size problems such that the solver could not find a solution in 10 hours time limit.

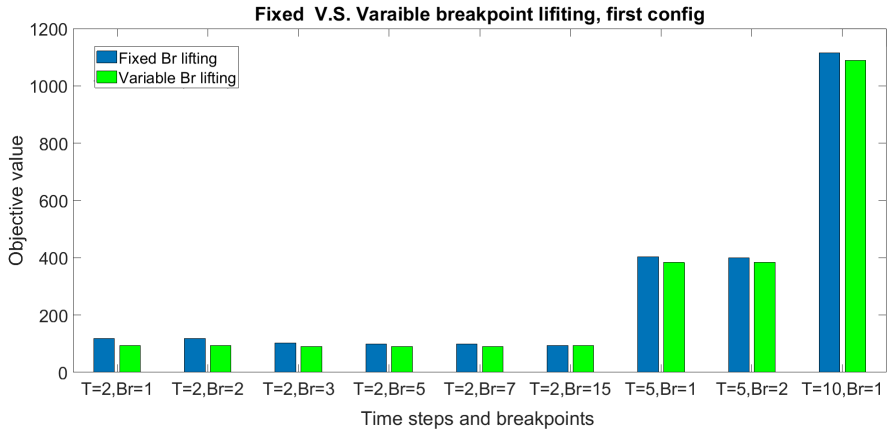


Figure 3.12: Comparison of solution from lifting method using fixed breakpoints and optimized breakpoints

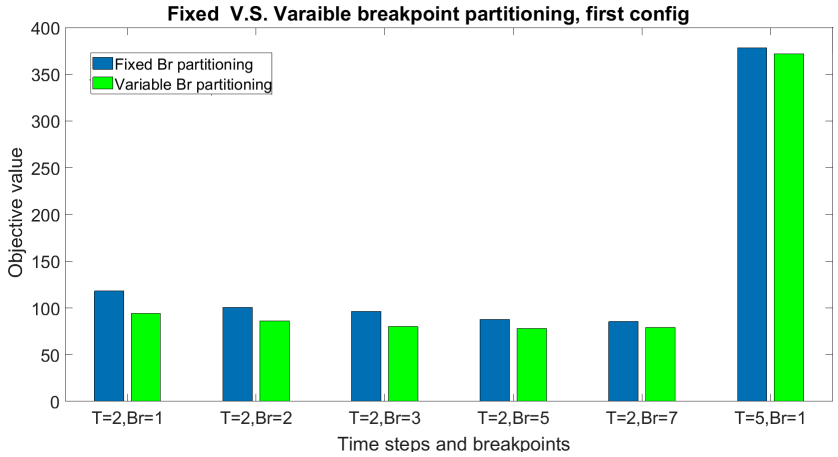


Figure 3.13: Comparison of solution from finite adaptability method using fixed breakpoints and optimized breakpoints

3.6 Conclusions

In this work, the lifting and partitioning methods for multistage adaptive binary optimization problems were studied. Different formulations based on fixed or variable breakpoints setting for each method were presented. Computational studies were made to compare the computational performance and the solution quality. The following conclusions can be made from this study. First, the binary decision rule (lifting) based method and the finite adaptability (partitioning) based method share the similar idea of uncertainty set splitting using breakpoints for each uncertain

parameter. While lifting based method has less solution flexibility than partitioning method (under the same breakpoint setting), the lifting method has the advantage of superior computational tractability and scalability for large problems. While computational resource restrictions is a major concern (especially for problems with large number of stages), the lifting method with fixed breakpoints is suggested and the number of breakpoints can be moderately large to avoid inferior solution quality. Otherwise, the partitioning method is suggested. Variable breakpoints can be implemented for partitioning method especially for small number of stages with the usage of small number of breakpoints. As a future research direction, the number of breakpoints can be optimized to avoid unnecessary model complexity.

Chapter 4

Multistage Adaptive Stochastic Mixed Integer Optimization through Piecewise Decision Rule Approximation

Abstract

This work studies the multistage adaptive stochastic mixed integer optimization problem, where the aim is to find adaptive continuous and integer decision policies that optimize the expected objective. While searching for the exact optimal policy is challenging, piecewise decision rule-based solution framework is studied and two types of strategies are compared: uncertainty set partitioning and uncertainty lifting. Adaptive binary and continuous decisions are approximated through piecewise constant and affine decision rule, respectively. The original optimization problem is solved through robust counterpart optimization technique. The proposed methods are applied to an inventory control problem and a chemical process capacity planning problem. Results show that the two methods provide flexible options with the trade-off between solution quality and computational efficiency. The uncertainty set partitioning based method leads to better solution quality and is appropriate to small-scale problems. On the other hand, the lifting method provides significant computational efficiency especially for large problems.

4.1 Introduction

Stochastic programming is an important technique for decision making under uncertain circumstances. It has many applications in different fields such as operations management [4, 5], control [6, 7], finance [8, 9] and process systems engineering [10, 11, 62, 76]. In the recourse based stochastic programming problem, recourse/adaptive/adjustable decisions can be made based on observed uncertainty. In multistage stochastic programming, the uncertainty is revealed gradually over time stages and the recourse decision can be made sequentially. It is widely known that obtaining the exact optimal adaptive decision policy of multistage stochastic problems including binary and continuous variables is computationally intractable [67, 55].

Two major solution approaches for multistage adaptive stochastic mixed integer problem include the scenario tree based method and the decision rule approximation method. The first method is the traditional approach. The idea is to discretize the uncertain parameters and represent the uncertainty through scenario trees. Then a deterministic counterpart problem can be formulated and solved accordingly. This approach results in exponential growth of the model size which makes the resulting deterministic problem computationally challenging and restricts its application to small size problems. Many literature work have applied the scenario-based method method in different applications. To name a few, Ahmed et al. [12] addressed a multi-period investment model for capacity expansion of a chemical plant under uncertain demand and cost. They used a heuristic scheme and a branch and bound algorithm to solve the problem to optimality. Birge and Rosa [13] studied the problem of greenhouse gas policy decision-making under economic uncertainty. Escudero et al. [14] investigated production and capacity planning problems to assist in raw material supply sourcing decision under demand uncertainty. Takriti et al. [15] developed a model and a solution method for power generation decision making when demand is uncertain. For a review on stochastic optimization, the reader can refer to [16, 17, 18] .

Considering the difficulty and the computational burden in solving stochastic problems using traditional scenario tree methods, there is a need for computationally efficient solution methods. As an alternative approach, the idea of the decision rule based method is to formulate the adaptive variables as special functions of uncertain parameters. These functions are referred to as decision rules. Recent advances in robust optimization enabled researchers to reformulate the problem under decision rule approximation as tractable optimization problem [70, 69]. Ben-Tal et al. applied linear decision rules in the context of adaptive robust optimization [22]. In their work, real valued variables are assumed to be linear functions of uncertain parameters. The simple structure of linear decision rules provides the scalability required for multistage stochastic problems. Some application instances for linear decision rules is reviewed here. Skaf and Boyd[7] designed a general affine controller in which the control input is an affine function of all previous measurements to minimize

a convex objective. They illustrated the method with applications in supply chain management optimization and dynamic portfolio optimization. Calafiore [8] addressed a portfolio optimization problem using an affine parametrization of the recourse policy that provides a sub-optimal but explicit formulation that can be used to solve multistage problems with many constraints and periods. Atamtürk and Zhang [23] presented a two-stage robust optimization method for network flow with uncertain demand and provide applications for lot-sizing and location transportation problems and compared the results with single-stage robust optimization and two-stage scenario-based optimization. Goh et al. [77] developed robust optimization procedures for controlling total completion time in projects with uncertain activity time. Rocha and Kuhn[78] used the linear decision rule method to manage the portfolio of electricity contracts using a multistage stochastic mean-variance optimization model and emphasized the value of adaptability and scalability inherent in this method. Ben-Tal [58] used the affinely adjustable robust counterpart (AARC) method to solve a multi-period supply chain problem. Decision rule based method has also received attention in the process systems engineering community. Lappas and Gounaris [24] developed a multistage adjustable robust optimization framework that accounts for inherent endogenous uncertainty in process scheduling by employing decision-dependent uncertainty sets. Ning and You [25] presented a data-driven framework for adaptive optimization using data and applied the proposed framework on two industrial applications of process scheduling and process network planning. Zhang et al. [26] developed a scheduling model for continuous industrial processes that provide interruptible load. The uncertainty in the timing of the load reduction request is modeled by an adjustable robust optimization approach that integrates recourse decisions using linear decision rules.

In order to address the optimality issue in decision rule methods, various nonlinear decision rules for adaptive real-valued variables have been proposed. The rationale behind the nonlinear decision rules is to improve the optimality while preserving the scalability of linear decision rules. Chen et al. [27] proposed new decision rule structures, segregated and deflected linear decision rules and demonstrated that these proposed methods can outperform sampling approaches when limited information about the underlying probability distributions is available. Chen and Zhang [28] presented a splitting-based extended affinely adjustable formulation and showed that their method is tractable and scalable to multistage problems. Georghiou et al. [29] proposed piecewise linear continuous decision rule based on axial segmentation and lifting of the uncertain parameters. They also proposed a method for piecewise linear continuous decision rule based on general segmentation. The authors recommended ideas for nonlinear continuous decision rules such as quadratic, power, monomial, inverse monomial and multilinear liftings and compared different methods on a dynamic inventory control problem. Goh and Sim [30] studied distributionally robust optimization problem and applied the linear decision rule method to get tractable approximation. Bampou and Kuhn [31] presented a polynomial decision rule to solve multistage stochastic problems. They estimated the suboptimality of the decision rule by solving a dual version of the problem in polynomial

decision rule. Bertsimas et al. [32] introduced a hierarchy of near-optimal polynomial policies and showed that these policies can be computed by solving a single semidefinite programming problem. They evaluated the framework using three classical applications in the context of inventory management and robust regulation of a suspension system. Recently, Avraamidou and Pistikopoulos [33] proposed a method based on generalized affine decision rules for linear mixed integer robust optimization problems using multi-parametric programming and showed that the method can find the exact global solution.

Although many decision rule structures are proposed for real-valued variables, the available literature for adaptive integer variables is scarce. Bertsimas and Caramanis [34] proposed a linear decision rule and approximated the semi-infinite optimization problem using a sampling algorithm. Bertsimas and Georghiou [35] suggested a structure that can provide near-optimal solutions but restricted scalability. Hanasusanto et al. [36] presented a binary decision rule based on the previous work done by Bertsimans and Caramanis [37] that can only be applied to two-stage problems. Postek and den Hertog [38] presented a method for iterative splitting of uncertainty set that can be used for adaptive integer variables and demonstrated the advantage of their method on a capital budgeting and a lot sizing problem. Recently, Bertsimas and Georghiou [39] proposed a binary decision rule that lifts the original uncertain parameters using 0-1 indicator functions. The trade-off between the solution optimality and scalability can be adjusted based on the segmentation resolution over the uncertainty set. The authors suggested a scalable formulation that can be used for large-scale multistage problems. For an extensive review on adaptive optimization, the reader can refer to Bertsimas et al., [40], Gabrel et al. [41] and Yanıkoğlu et al. [79].

While most of the existing work study adaptive continuous decisions and binary decisions separately, this study presents a novel piecewise decision rule solution framework that integrates both adaptive continuous and adaptive binary variables. To evaluate the efficiency of the piecewise decision rule method, two strategies are presented and compared: an uncertainty lifting based method and an uncertainty set partitioning based method. The first method combines the uncertainty lifting based method for decision rule approximation for continuous and binary variables. This combination provides a robust scalable solution for multistage stochastic problems. It can be adjusted to result in a continuous or discontinuous solution for the continuous variables. The discontinuous solution provides more flexibility and consequently a better objective value. The second method combines the scenario tree representation and uncertainty set based robust counterpart optimization technique. In the proposed partitioning method, we partition the uncertainty set using fixed breakpoints at the initial stage. Such a partition is used since we try to compare the performance with the lifting method, which essentially also partitions the uncertainty set through axial segmentation. Overall, the new contribution of the work is: 1) combination of the continuous lifting and binary lifting strategies, 2) a predefined uncertainty set partitioning method and the

associated piecewise decision rule, 3) comparative study of the two piecewise decision rule framework and discussion on the advantages and limitations.

The rest of the paper is organized as follows: Section 4.2 presents the general formulation for multistage stochastic problems under exogenous uncertainty and describes an inventory planning case study. In section 4.3, the uncertainty set partitioning based method is described. Section 4.4 presents the mathematical framework for adaptive continuous and adaptive binary variables and shows how to integrate both continuous and binary variables in a single framework. In Section 4.5, the developed framework is applied to an inventory control problem and the benefits of the proposed method are discussed. Section 4.6 presents a case study about expansion planning of a chemical process network and section 5.8 concludes the study.

4.2 Multistage stochastic mixed integer optimization

Formulation 4.1 presents the multistage stochastic problem that is addressed in this study.

$$\begin{aligned} \min \quad & \mathbb{E}_{\boldsymbol{\eta}} \left(\sum_t c_t^\top(\boldsymbol{\eta}) x_t(\boldsymbol{\eta}) + d_t^\top(\boldsymbol{\eta}) y_t(\boldsymbol{\eta}) \right) \\ s.t. \quad & \sum_{\tau=1}^{t} \boldsymbol{A}_{t,\tau}^\top x_\tau(\boldsymbol{\eta}) + \sum_{\tau=1}^{t} \boldsymbol{B}_{t,\tau}^\top y_\tau(\boldsymbol{\eta}) \leq f_t(\boldsymbol{\eta}) \qquad \forall t, \boldsymbol{\eta} \in \Omega \\ & y_t(\boldsymbol{\eta}) \in \{0,1\} \qquad \forall t, \boldsymbol{\eta} \in \Omega \end{aligned} \tag{4.1}$$

where $\boldsymbol{\eta}$ is the primitive exogenous uncertainty vector parameter, the objective coefficients $c_t(\boldsymbol{\eta})$, $d_t(\boldsymbol{\eta})$, and constraint right hand side $f_t(\boldsymbol{\eta})$ are assumed to be dependent on the primitive uncertainty. Furthermore, we assume they are linearly dependent on the vector of uncertain parameters up to time t, $\boldsymbol{\eta}_{[t]}$:

$$c_t(\boldsymbol{\eta}) = \boldsymbol{C}_t \boldsymbol{\eta}_{[t]}, d_t(\boldsymbol{\eta}) = \boldsymbol{D}_t \boldsymbol{\eta}_{[t]}, f_t(\boldsymbol{\eta}) = \boldsymbol{F}_t \boldsymbol{\eta}_{[t]}$$

Constraints coefficients $\boldsymbol{A}_{t,\tau}$ and $\boldsymbol{B}_{t,\tau}$ are assumed to be independent of uncertainty.

This formulation aims at finding the optimal adaptive decision (policy) $x_t(\boldsymbol{\eta})$ and $y_t(\boldsymbol{\eta})$ that are feasible for all uncertainty realizations in the uncertainty set Ω and minimizes the expected objective performance ($\mathbb{E}(\cdot)$ is the expectation operator). Due to non-anticipativity requirements (i.e., the decision can not be made based on unknown future information), the adaptive continuous decision $x_t(\boldsymbol{\eta})$ and adaptive binary decision $y_t(\boldsymbol{\eta})$ are functions of uncertain parameters up to time t, $\boldsymbol{\eta}_{[t]}$:

$$x_t(\boldsymbol{\eta}) = x_t(\boldsymbol{\eta}_{[t]}), y_t(\boldsymbol{\eta}) = y_t(\boldsymbol{\eta}_{[t]})$$

Under the above assumption, the problem under investigation can be re-written as

$$\begin{aligned} \min \quad & \mathbb{E}_{\boldsymbol{\eta}} \left(\sum_t {\boldsymbol{\eta}_{[t]}}^\top \boldsymbol{C}_t^\top x_t(\boldsymbol{\eta}_{[t]}) + {\boldsymbol{\eta}_{[t]}}^\top \boldsymbol{D}_t^\top y_t(\boldsymbol{\eta}_{[t]}) \right) & & \\ s.t. \quad & \sum_{\tau=1}^{t} (\boldsymbol{A}_{t,\tau}^\top x_\tau(\boldsymbol{\eta}_{[\tau]}) + \boldsymbol{B}_{t,\tau}^\top y_\tau(\boldsymbol{\eta}_{[\tau]})) \leq \boldsymbol{F}_t \boldsymbol{\eta}_{[t]} & \forall t, \boldsymbol{\eta} \in \Omega & \quad (4.2) \\ & y_t(\boldsymbol{\eta}_{[t]}) \in \{0,1\} & \forall t, \boldsymbol{\eta} \in \Omega & \end{aligned}$$

While it is in general hard to get the optimal decision policy $x_t(\boldsymbol{\eta})$ and $y_t(\boldsymbol{\eta})$, we study piecewise decision rule based approximation method in this paper. The unknown true optimal policy will be approximated by piecewise linear/binary decision rule. In subsequent section 3, we present the uncertainty set partitioning based method. For each partition, linear and constant binary decision rules are applied to continuous and binary variables, respectively. In Section 4, we present the uncertainty lifting method. Applying linear decision rule over the lifted uncertainty will also lead to piecewise linear decision rule for continuous variable and piecewise constant decision rule for binary variables. Then the two methods will be compared in terms of solution quality and computational efficiency.

4.2.1 Example: Inventory planning problem

As a motivating example, the problem of inventory planning of a single product is presented here. As Figure 4.1 shows, at the beginning of each time period, the demand uncertainty η_t is revealed. In order to satisfy the demand, two types of decisions are made. The continuous ordering decision x_t can only be used to satisfy the demand for the next time period. The binary ordering decision $y_{n,t}$ (related to fixed order amount q_n from lot n) can be used to satisfy the demand immediately at the same time period but at a higher ordering cost ($c_x < c_{y_n}$). At each time step, the demand should be satisfied and the remaining products will be held at the inventory and are used to satisfy the demand for the next time period. Holding the products at the inventory will incur the holding cost of c_h per unit product. At $t = 1$, only the binary decision $y_{n,1}$ is available to satisfy the demand at this period, since x_1 decision can only be used to satisfy the demand at the next time period. The initial inventory I_0 is assumed to be zero. This decision making sequence is shown in Figure 4.1.

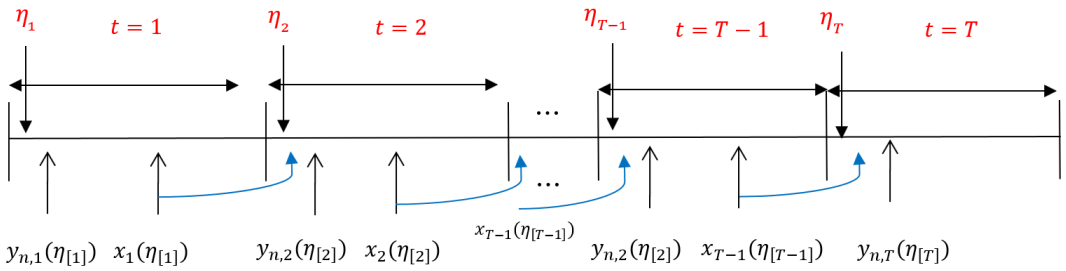


Figure 4.1: Decision making and uncertainty realization sequence

Equations 4.3a-4.3f present the multistage stochastic programming formulation for this inventory planning problem under demand uncertainty.

$$\min \quad \mathbb{E}_{\boldsymbol{\eta}} \left(\sum_{t=1}^{T-1} c_x x_t(\boldsymbol{\eta}_{[t]}) + \sum_{t=1}^{T} \sum_{n=1}^{N} c_{y_n} q_n y_{n,t}(\boldsymbol{\eta}_{[t]}) + \sum_{t=1}^{T} c_h I_t(\boldsymbol{\eta}_{[t]}) \right) \tag{4.3a}$$

$$s.t. \quad I_t(\boldsymbol{\eta}_{[t]}) = I_0 + \sum_{\tau=1}^{t} \left[x_{\tau-1}(\boldsymbol{\eta}_{[\tau-1]}) + \sum_{n=1}^{N} q_n y_{n,\tau}(\boldsymbol{\eta}_{[\tau]}) - \eta_\tau \right] \quad \forall t \in \{1, ..., T\}, \boldsymbol{\eta} \in \Omega \tag{4.3b}$$

$$I_t(\boldsymbol{\eta}_{[t]}) \geq 0 \quad \forall t \in \{1, ..., T\}, \boldsymbol{\eta} \in \Omega \tag{4.3c}$$

$$x_t(\boldsymbol{\eta}_{[t]}) \geq 0 \quad \forall t \in \{1, ..., T-1\}, \boldsymbol{\eta} \in \Omega \tag{4.3d}$$

$$\sum_{\tau=1}^{t} x_\tau(\boldsymbol{\eta}_{[\tau]}) \leq CCAP_t \quad \forall t \in \{1, ..., T-1\}, \boldsymbol{\eta} \in \Omega \tag{4.3e}$$

$$y_{n,t}(\boldsymbol{\eta}_{[t]}) \in \{0, 1\} \quad \forall n, t, \boldsymbol{\eta} \in \Omega \tag{4.3f}$$

The problem parameters and variables are described as follows. η_t is the uncertain demand parameter of time period t. c_{y_n} is the cost for buying a unit of product from lot n for immediate delivery. c_x is the cost per product unit for purchases that are delivered in the subsequent period. c_h is the holding cost for a unit product in the inventory. q_n is the fixed quantity delivered when the purchase decision $y_{n,t}$ is made. N is the number of lots. CAP_t is the order limit at time step t and $CCAP_t$ is the cumulative order limit up to period t, $CCAP_t = \sum_{\tau=1}^{t} CAP_\tau$. I_t is the level of available inventory at period t. x_t is the product amount ordered in period t.

Table 4.1: Problem parameters

c_x	c_{y_1}	c_{y_2}	c_h	N	q_n
2	7	9	5	2	50

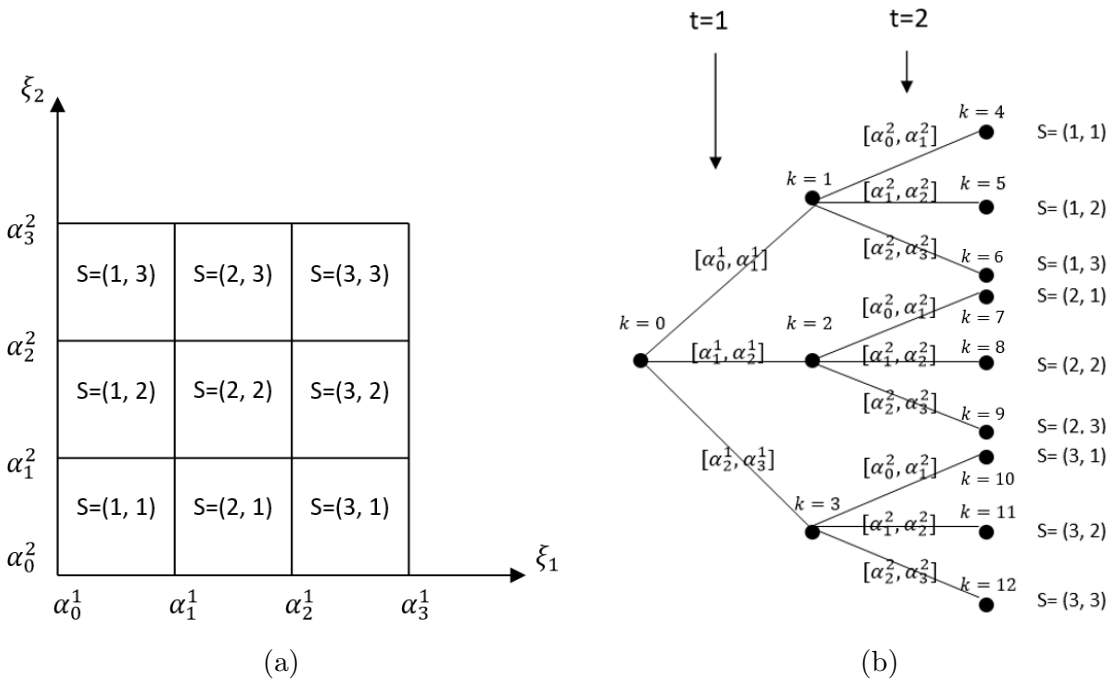
Table 4.2: Order limit for each time period

t	1	2	3	4	5	6	7	8	9
CAP_t	50	80	60	70	50	90	40	70	80

In this problem, the demand uncertainty is assumed to be uniformly distributed within the uncertainty set $\{\boldsymbol{\eta} : l_t \leq \eta_t \leq u_t, t = 1, \cdots, T\}$ where l_t and u_t are the lowest and highest possible demand at each time period t that are considered to be 20 and 100, respectively.

4.3 Uncertainty set partitioning based method

In this method, the uncertainty set is divided into partitions. Adaptive continuous variables are modeled as affine functions of uncertain parameters and adaptive binary decisions are assumed be constant (0 or 1) for each partition. To generate the partitions, a set of breakpoints are pre-defined for each uncertain parameter. Figure 4.2 illustrates the partitioning of a two dimensional rectangular uncertainty set using two breakpoints for each uncertain parameter. With this partition, a corresponding scenario tree can be constructed with each branch representing a subinterval for each parameter. In this figure, each uncertain parameter is divided into 3 segments such that there are 3 nodes in the first stage and 9 nodes in the second stage.



(a) (b)

Figure 4.2: Partitioning of a 2-dimensional uncertainty set (left) and scenario tree representation (right)

In order to clarify how each scenario is defined, consider a two-stage problem: the first stage

has uncertain parameters η_1 and the second stage has one uncertain parameter η_2. Assume 1 breakpoint is applied to η_1 and two breakpoints are applied to η_2. Therefore, the set $\mathcal{S}$ is:

$$s \in \mathcal{S} = \left\{ \begin{bmatrix} 1, & 1 \end{bmatrix}, \begin{bmatrix} 1, & 2 \end{bmatrix}, \begin{bmatrix} 1, & 3 \end{bmatrix}, \begin{bmatrix} 2, & 1 \end{bmatrix}, \begin{bmatrix} 2, & 2 \end{bmatrix}, \begin{bmatrix} 2, & 3 \end{bmatrix} \right\}$$

For each scenario $s \in S$, the corresponding uncertainty set is a subset of original set Ω. The probability of uncertainty realization occurs in subset Ω_s is p_s. Its value can be evaluated based on the known probability distribution of the uncertainty. The subset is compactly written as

$$\Omega^s = \{\boldsymbol{\eta} : \boldsymbol{W}_s \boldsymbol{\eta} \geq \boldsymbol{V}_s\} \tag{4.4}$$

The uncertain set partitioning method (also described as finite adaptability method) is applied to the original problem 4.1. This method is a combination of scenario tree based stochastic formulation and robust optimization. For each scenario, the constraints are satisfied over an uncertainty set as defined in equation 4.4 instead of a single point in the uncertainty set. For each scenario, the binary variables take constant value and continuous variables are assumed to be affine functions of uncertain parameters.

$$y_t(\boldsymbol{\eta}) = y_t^s \qquad \forall t, s, \boldsymbol{\eta} \in \Omega^s \tag{4.5}$$

$$x_t(\boldsymbol{\eta}) = \boldsymbol{X}_t^s \boldsymbol{\eta}_{[t]} \qquad \forall t, s, \boldsymbol{\eta} \in \Omega^s \tag{4.6}$$

where $\boldsymbol{\eta}_{[t]}$ denotes the vector of uncertain parameters up to time step t. A truncate matrix $\boldsymbol{R}_t$ can be applied to the vector of all uncertain parameters $\boldsymbol{\eta}$ to obtain $\boldsymbol{\eta}_{[t]}$.

$$\boldsymbol{R}_t \boldsymbol{\eta} = \boldsymbol{\eta}_{[t]} \qquad \forall t \tag{4.7}$$

The original problem 4.1 can be formulated as:

$$\begin{aligned}
\min \quad & \sum_{s \in \mathcal{S}} p_s \sum_t \mathbb{E}_{\boldsymbol{\eta} \in \Omega^s} \left(\boldsymbol{\eta}^\top \boldsymbol{R}_t^\top \boldsymbol{C}_t^\top \boldsymbol{X}_t^s \boldsymbol{R}_t \boldsymbol{\eta} + \boldsymbol{\eta}^\top \boldsymbol{R}_t^\top \boldsymbol{D}_t^\top y_t^s \right) & & \\
s.t. \quad & \sum_{\tau=1}^{t} (\boldsymbol{A}_{t,\tau}^\top \boldsymbol{X}_\tau^s \boldsymbol{R}_t \boldsymbol{\eta} + \boldsymbol{B}_{t,\tau}^\top y_\tau^s) \leq \boldsymbol{F}_t \boldsymbol{R}_t \boldsymbol{\eta} & & \forall t, s, \boldsymbol{\eta} \in \Omega^s \\
& y_t^s \in \{0, 1\} & & \forall t, s \\
& y_t^s = y_t^{s'} & & \forall (t, s, s') \in SP \\
& \boldsymbol{X}_t^s = \boldsymbol{X}_t^{s'} & & \forall (t, s, s') \in SP
\end{aligned} \tag{4.8}$$

where the last two constraints enforce non-anticipativity; at each time step, the variables can only depend on the revealed uncertain parameters up to time step t. Therefore, scenarios with the same

path up to time step t should have the same solution. This set of scenarios is represented by SP:

$$SP = \{(t, s, s') : \ s_\tau = s'_\tau, \quad \forall \tau = 1, ..., t\} \tag{4.9}$$

where s_t gives the subinterval index of uncertain parameter η_t under this scenario. For instance, the set SP corresponding to Figure 4.2 contains the following elements:

$$\Big(1,[1,1],[1,2]\Big), \Big(1,[1,1],[1,3]\Big), \Big(1,[1,2],[1,3]\Big)$$
$$\Big(1,[2,1],[2,2]\Big), \Big(1,[2,1],[2,3]\Big), \Big(1,[2,2],[2,3]\Big)$$
$$\Big(1,[3,1],[3,2]\Big), \Big(1,[3,1],[3,3]\Big), \Big(1,[3,2],[3,3]\Big)$$

Next, the semi-infinite constraint is reformulated using duality theorem and the corresponding deterministic counterpart formulation is obtained:

$$\begin{aligned}
\min \quad & \sum_{s\in\mathcal{S}} p_s \sum_t \left(\mathrm{Tr}(\boldsymbol{R}_t^\top \boldsymbol{C}_t^\top \boldsymbol{X}_t^s \boldsymbol{R}_t \mathbb{E}_{\Omega^s}(\boldsymbol{\eta}\boldsymbol{\eta}^\top)) + \mathbb{E}_{\Omega^s}(\boldsymbol{\eta})^\top \boldsymbol{R}_t^\top \boldsymbol{D}_t^\top y_t^s\right) & & \\
s.t. \quad & {\boldsymbol{V}_s}^\top \boldsymbol{\theta}_{t,s} \geq \sum_{\tau=1}^{t} \boldsymbol{B}_{t,\tau}^\top y_\tau^s & \forall t, s & \text{(4.10a)} \\
& {\boldsymbol{W}_s}^\top \boldsymbol{\theta}_{t,s} = -\sum_{\tau}^{t} \boldsymbol{A}_{t,\tau}^\top \boldsymbol{X}_\tau^s \boldsymbol{R}_\tau + \boldsymbol{F}_t \boldsymbol{R}_t & \forall t, s & \text{(4.10b)} \\
& \boldsymbol{\theta}_{t,s} \geq 0 & \forall t, s & \text{(4.10c)} \\
& y_t^s \in \{0,1\} & \forall t, s & \text{(4.10d)} \\
& y_t^s = y_t^{s'} & \forall (t, s, s') \in SP & \text{(4.10e)} \\
& \boldsymbol{X}_t^s = \boldsymbol{X}_t^{s'} & \forall (t, s, s') \in SP & \text{(4.10f)}
\end{aligned}$$

where $\mathrm{Tr}(\cdot)$ denotes the trace operator, $\boldsymbol{\theta}_{t,s}$ represents the dual variables. The final model 10 is a mixed integer linear optimization (MILP) problem that can be solved using off-the-shelf optimization solver such as CPLEX. Notice that the second order moment matrix in the objective can be evaluated using formula $\mathbb{E}[\boldsymbol{\eta}\boldsymbol{\eta}^\top] = \mathrm{Cov}(\boldsymbol{\eta}) + \mathbb{E}[\boldsymbol{\eta}]\mathbb{E}[\boldsymbol{\eta}]^\top$ with $\mathrm{Cov}(\boldsymbol{\eta})$ being the covariance matrix of vector $\boldsymbol{\eta}$. Specifically, based on matrix calculation property $\mathbb{E}[\mathrm{Tr}(M_{n\times n})] = \mathrm{Tr}(\mathbb{E}[M_{n\times n}])$ and $\mathrm{Tr}(M_{m\times n}N_{n\times m}) = \mathrm{Tr}(N_{n\times m}M_{m\times n})$, the first term is derived from the following property

$$\mathbb{E}[\boldsymbol{\eta}^\top M \boldsymbol{\eta}] = \mathrm{Tr}(\mathbb{E}[\boldsymbol{\eta}^\top M \boldsymbol{\eta}]) = \mathbb{E}[\mathrm{Tr}(\boldsymbol{\eta}^\top M \boldsymbol{\eta})] = \mathbb{E}[\mathrm{Tr}(M\boldsymbol{\eta}\boldsymbol{\eta}^\top)] = \mathrm{Tr}(\mathbb{E}[M\boldsymbol{\eta}\boldsymbol{\eta}^\top]) = \mathrm{Tr}(M\mathbb{E}[\boldsymbol{\eta}\boldsymbol{\eta}^\top])$$

4.4 Uncertainty lifting based method

This section explains the mathematical framework for lifting uncertain parameters. The traditional scenario tree method results in exponential growth of the problem size that makes the problem computationally expensive while the lifting method provides the scalability and tractability required for large-scale problems. The lifted parameters are used to build piecewise linear/binary decision rules for continuous and binary variables. Construction of convex overestimation set from the original uncertainty set is also presented.

4.4.1 Piecewise linear lifting of uncertainty

The concept of lifting for piecewise linear continuous decision rule was proposed by Georghiou et al. [29]. In this framework, a set of breakpoints are selected for each uncertain parameter. A lifting operator is defined based on the selected breakpoints that result in a nonconvex lifted set. The nonconvex set is convexified using its extreme points. The adaptive continuous variable will be approximated by an affine function of the lifted uncertain parameters.

Consider a single uncertain parameter η_i defined within the interval $\Omega_i \in [l_i, u_i]$. The uncertain parameter is axially segmented using $r_i - 1$ number of breakpoints. Figure 4.3 illustrates that two breakpoints ($r_i = 3$) are selected for the uncertainty interval.

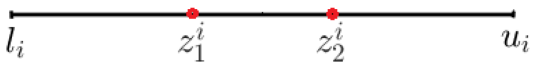


Figure 4.3: Two breakpoint locations are selected to segment the uncertainty interval

The location of breakpoints is denoted by $z_j^i, j = 1, \cdots, r_i - 1$. To generalize the notation, we define $z_0^i \equiv l_i, z_{r_i}^i \equiv u_i$. Hence $z_0^i < z_1^i < ... < z_{r_i-1}^i < z_{r_i}^i$. The lifted uncertainty for a single uncertain parameter η_i is defined as:

$$G_{i,j} = \begin{cases} \eta_i & r_i = 1 \\ \min\{\eta_i, z_1^i\} & r_i \geq 2,\ j = 1 \\ \max\{\min\{\eta_i, z_j^i\} - z_{j-1}^i, 0\} & r_i \geq 2,\ j = 2, ..., r_i - 1 \\ \max\{\eta_i - z_{j-1}^i, 0\} & r_i \geq 2,\ j = r_i \end{cases} \tag{4.11}$$

Notice that if $r_i = 1$, it means that there are no breakpoints for i-th uncertain parameter and the lifting operator will reduce to the identity mapping. The original uncertain parameter can be retrieved through the following equation: $\eta_i = \sum_{i=1}^{r_i} G_{i,j}$ where $\boldsymbol{G_i} = (G_{i,1}, \cdots, G_{i,r_i})$. Figure 4.4 illustrates the lifted uncertain parameter for two breakpoints.

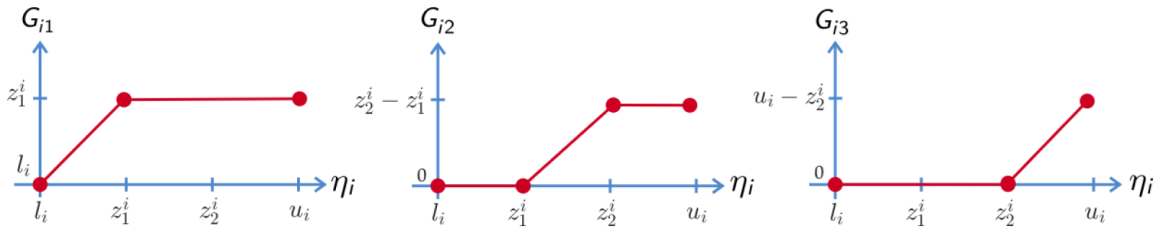


Figure 4.4: Illustration of the lifted uncertain parameter for two-breakpoint case

Figure A.4 illustrates the sequence for convexifying the lifted set. The left figure shows that two breakpoints are selected for lifting the uncertain parameter. The middle figure presents the lifted nonconvex set (Ω_i^{liftC}) in the space of the lifted uncertain parameters $G_{i,j},\ j = 1, \cdots, r_i$ and the right figure is obtained by convexifying the lifted nonconvex set (Ω_i^{liftC}) using its extreme points.

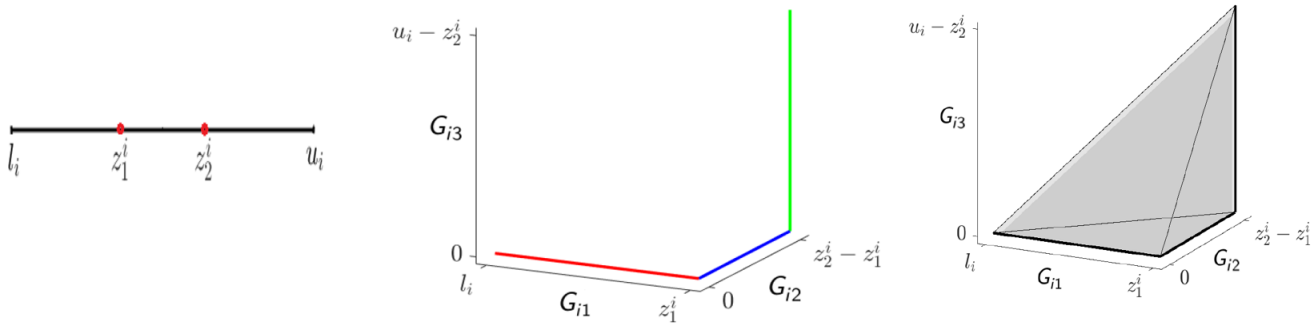


Figure 4.5: Lifted uncertain parameter and convexified uncertainty set

The nonconvex set Ω_i^{liftC} is a union of r_i connected finite line segments. Its extreme points are $\boldsymbol{\nu_0} = (l_i, 0, \cdots, 0)^\top$, $\boldsymbol{\nu_1} = (z_1^i, 0, \cdots, 0)^\top$, $\boldsymbol{\nu_2} = (z_1^i, z_2^i - z_1^i, 0, \cdots, 0)^\top$,$\cdots$, $\boldsymbol{\nu_{r_i}} = (z_1^i, z_2^i - z_1^i, \cdots, z_{r_i-1}^i - z_{r_i-2}^i, u_i - z_{r_i-1}^i)^\top \in \mathbb{R}^{r_i}$. Therefore, the convex hull of Ω_i^{liftC} is generated by the simplex with vertices $\{\nu_j\}_{j=0}^{r_i}$: $\{\mathbf{G}_i \in \mathbb{R}^{r_i} : \mathbf{G}_i = \sum_{j=0}^{r_i} \lambda_j \boldsymbol{\nu_j}, \sum_{j=0}^{r_i} \lambda_j = 1, \lambda_0, \cdots, \lambda_{r_i} \geq 0\}$. This can be rewritten as:

$$conv(\Omega_i^{liftC}) = \{\mathbf{G}_i \in \mathbb{R}^{r_i} : \exists \boldsymbol{\lambda} \in R^{r_i+1}, \mathbf{A}_i \boldsymbol{\lambda} = (1, \mathbf{G}_i^\top)^\top, \boldsymbol{\lambda} \geq 0\} \tag{4.12}$$

$$= \{\mathbf{G}_i \in \mathbb{R}^{r_i} : \mathbf{A}_i^{-1}(1, \mathbf{G}_i^\top)^\top \geq 0\} \tag{4.13}$$

where

$$\mathbf{A}_i = \begin{bmatrix} 1 & \cdots & 1 \\ \boldsymbol{\nu_0} & \cdots & \boldsymbol{\nu_{r_i}} \end{bmatrix}, \mathbf{A}_i^{-1} = \begin{bmatrix} \frac{z_1^i}{z_1^i - l_i} & -\frac{1}{z_1^i - l_i} & & & & \\ -\frac{l_i}{z_1^i - l_i} & \frac{1}{z_1^i - l_i} & -\frac{1}{z_2^i - z_1^i} & & & \\ & & \frac{1}{z_2^i - z_1^i} & \ddots & & \\ & & & \ddots & -\frac{1}{z_{r_i-1}^i - z_{r_i-2}^i} & \\ & & & & \frac{1}{z_{r_i-1}^i - z_{r_i-2}^i} & -\frac{1}{u_i - z_{r_i-1}^i} \\ & & & & & \frac{1}{u_i - z_{r_i-1}^i} \end{bmatrix}$$

Note that $\mathbf{A}_i \in \mathbb{R}^{(r_i+1)\times(r_i+1)}$. In reference [29], the continuous variable is approximated by an affine function of lifted uncertain parameters $x(\eta) = \boldsymbol{X}[1, \boldsymbol{G}]^\top$. For instance, for a single uncertain parameter with two breakpoints, the continuous variable is formulated as: $x(\eta) = x_0 + x_1 G_{i,1} + x_2 G_{i,2} + x_3 G_{i,3} = \boldsymbol{X}[1, G_{i,1}, G_{i,2}, G_{i,3}]^\top$. It should be noted that the lifted uncertain parameters $G_{i,j}$ are continuous with respect to the original uncertain parameters η_i, therefore the continuous variable will always be continuous with respect to η. This property restricts the flexibility of the decision rule. In section 5.5.4, a discontinuous decision rule for the continuous variable is presented that is inherently more flexible.

4.4.2 Piecewise binary lifting of uncertainty

The idea of adaptive binary variables based on decision rule method was proposed by Bertsimas and Georghiou [39]. In this method, each uncertain parameter is lifted using the indicator functions. The lifted set is non-convex and it is convexified which results in a convex over-estimation of the original uncertainty set.

Consider a single uncertain parameter η_i subject to the interval $l_i \leq \eta_i \leq u_i$, and assume the interval is divided into r_i subintervals using $r_i - 1$ breakpoints $z_j^i, j = 1, ..., r_i - 1$, (Figure 4.3). To lift the uncertainty, the indicator functions $Q_{i,j}(\cdot)$ of uncertain parameters are employed.

$$Q_{i,j}(\eta_i) = \begin{cases} 1, & \text{if} \quad \eta_i \geq z_j^i \\ 0, & \text{if} \quad \eta_i < z_j^i \end{cases} \tag{4.14}$$

While two breakpoints are applied on a single uncertain parameter η_i, the corresponding indicator functions are illustrated in Figure 4.6.

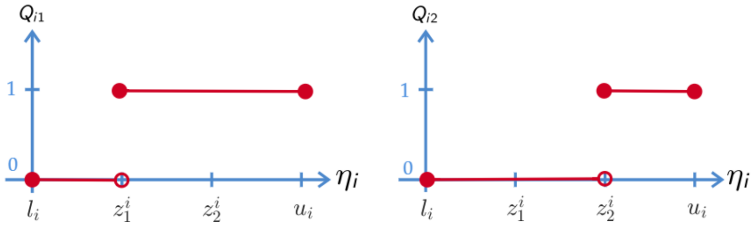


Figure 4.6: Indicator functions for two breakpoints on uncertain parameter η_i

Figure 4.7 illustrates the lifted uncertainty set for two breakpoints. The lifted set is nonconvex since it consists of disconnected pieces. Next, using the extreme points of the lifted nonconvex set, a convex polyhedron is formulated.

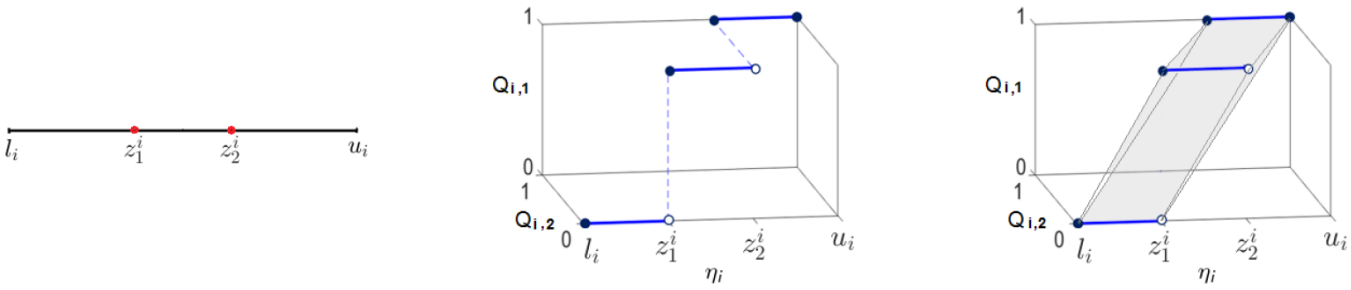


Figure 4.7: Lifted uncertain parameter and convexified uncertainty set to be used for adaptive binary decision rule

Convex hull of lifted uncertainty set for a single uncertain parameter is formulated as:

$$conv(\Omega_i^{liftB}) = \{\boldsymbol{\eta_i'} = (\eta_i, \boldsymbol{Q_i}) : \exists\ \lambda_{k,p,i} > 0, \boldsymbol{\eta_i'} = \sum_{p=1}^{r_i}\sum_{k=1}^{2} \lambda_{k,p,i}\boldsymbol{\nu}_{k,p,i}, \sum_{p=1}^{r_i}\sum_{k=1}^{2} \lambda_{k,p,i} = 1\} \tag{4.15}$$

where $\boldsymbol{Q_i} = (Q_{i,1}, \cdots, Q_{i,r_i-1})$ is the vector of indicator functions for uncertain parameter η_i, $\lambda_{k,p,i}$ is the coefficient of extreme points in the convex hull and $\boldsymbol{\nu}_{1,p,i}$ and $\boldsymbol{\nu}_{2,p,i}$ are extreme points of the p-th segment of the lifted non-convex set (Ω_i^{liftB}) related to parameter η_i:

$$\boldsymbol{\nu}_{1,p,i} = [z_{p-1}^i, \underbrace{1, \cdots, 1}_{p-1 \text{ times}}, \underbrace{0, \cdots, 0}_{r_i - p \text{ times}}]^\top, \quad \boldsymbol{\nu}_{2,p,i} = [z_p^i, \underbrace{1, \cdots, 1}_{p-1 \text{ times}}, \underbrace{0, \cdots, 0}_{r_i - p \text{ times}}]^\top \in \mathbb{R}^{r_i}$$

therefore,

$$conv(\Omega_i^{liftB}) = \{\boldsymbol{\eta_i'} = (\eta_i, \boldsymbol{Q_i}) \in \mathbb{R}^{r_i} : \exists\ \boldsymbol{\lambda} \in \mathbb{R}^{2r_i}, \mathbf{A}_i\boldsymbol{\lambda} = (1, \boldsymbol{\eta_i'}^\top)^\top, \boldsymbol{\lambda} > 0\} \tag{4.16}$$

where

$$\mathbf{A}_i = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ \boldsymbol{\nu}_{1,1,i} & \boldsymbol{\nu}_{2,1,i} & \cdots & \boldsymbol{\nu}_{1,r_i,i} & \boldsymbol{\nu}_{2,r_i,i} \end{bmatrix} \in \mathbb{R}^{(r_i+1)\times 2r_i}$$

The binary variable will be a linear combination of indicator functions that can generate a piecewise constant solution. For instance, for a single uncertain parameter and two breakpoints, the decision rule for the binary variable will be: $y(\eta_i) = y_0 + y_1 Q_{i,1} + y_2 Q_{i,2} = Y[1, Q_{i,1}, Q_{i,2}]^\top$. As an illustration, Table 4.3 demonstrates that for the case with 2 breakpoints, the binary decision rule can generate all the possible values of the binary variable at different uncertainty intervals. The integer variable $Y \in \{-1, 0, +1\}$.

Table 4.3: Piecewise adaptive binary variable for 2 break points case

$y_{\eta_i \in [l^i, z_1^i)}$	$y_{\eta_i \in [z_1^i, z_2^i)}$	$y_{\eta_i \in [z_2^i, u^i]}$	y_0	y_1	y_2
1	0	0	1	-1	0
0	1	0	0	1	-1
0	0	1	0	0	1
1	1	0	1	0	-1
1	0	1	1	-1	1
0	1	1	0	1	0
1	1	1	1	0	0
0	0	0	0	0	0

4.4.3 Overall lifted uncertainty

In this study, both adaptive continuous and adaptive binary variables are considered in the problem formulation. Therefore, the uncertain parameter includes the lifting elements for both binary and continuous variables. In the following, the notation used for the overall uncertainty vector for all time steps is summarized here.

- $\boldsymbol{\eta'_i} = [\eta_i, \underbrace{Q_{i,1}, \cdots, Q_{i,r_i-1}}_{\text{binary}}, \underbrace{G_{i,1}, \cdots, G_{i,r_i}}_{\text{continuous}}]^\top = [\eta_i, \boldsymbol{Q_i}^\top, \boldsymbol{G_i}^\top]^\top$
- $\boldsymbol{\eta'_{[t]}}$ vector of overall uncertainty from stage 1 to t: $[1, \boldsymbol{\eta'_1}^\top, \cdots, \boldsymbol{\eta'_t}^\top]^\top$
- $\boldsymbol{\eta'} \equiv \boldsymbol{\eta'_{[T]}}$ vector of overall uncertainty from stage 1 to T: $[1, \boldsymbol{\eta'_1}^\top, \cdots, \boldsymbol{\eta'_T}^\top]^\top$

Assume the original uncertainty set is a general polyhedral set:

$$\Omega = \{\boldsymbol{\eta} : \boldsymbol{J\eta} \geq \boldsymbol{h}\} \tag{4.17}$$

The convex overestimation for the overall lifted uncertainty set is defined as follows. It includes constraints from the original uncertainty set, the convex hull for the continuous lifted set, the convex hull for the binary lifted set and the correlation between the lifted parameters.

$$\hat{\Omega}' = \{\boldsymbol{\eta'} : \boldsymbol{J\eta} \geq \boldsymbol{h}, \boldsymbol{G_i} \in conv(\Omega_i^{liftC}), (\eta_i, \boldsymbol{Q_i}) \in conv(\Omega_i^{liftB}), \eta_i = \mathbf{1}^\top \boldsymbol{G_i}, i = 1, \cdots, T\} \quad (4.18)$$

The convex overestimated set after projection to the space of $\boldsymbol{\eta'}$ can be compactly written as:

$$\hat{\Omega}' = \{\boldsymbol{\eta'} : \boldsymbol{J'\eta'} \geq \boldsymbol{h'}\} \quad (4.19)$$

Projection (truncation) matrices can be used to retrieve the specific uncertainty to be used in the decision rule approximation model:

$$\boldsymbol{\phi_{[t]}} = [1, \boldsymbol{G_1}^\top, \cdots, \boldsymbol{G_t}^\top, \boldsymbol{Q_1}^\top, \cdots, \boldsymbol{Q_t}^\top]^\top = \boldsymbol{P}_{\boldsymbol{\eta_{[t]}}} \boldsymbol{\eta'} \quad (4.20)$$

$$\boldsymbol{Q_{[t]}} = [1, \boldsymbol{Q_1}^\top, \cdots, \boldsymbol{Q_t}^\top]^\top = \boldsymbol{P}_{\boldsymbol{Q_{[t]}}} \boldsymbol{\eta'} \quad (4.21)$$

$$\boldsymbol{G_{[t]}} = [1, \boldsymbol{G_1}^\top, \cdots, \boldsymbol{G_t}^\top]^\top = \boldsymbol{P}_{\boldsymbol{G_{[t]}}} \boldsymbol{\eta'} \quad (4.22)$$

4.4.4 Decision rule approximation

The decision rule for continuous variables is based on defining an affine function of the overall lifted uncertainty as presented in section 5.5.3. The adaptive continuous variable is approximated as:

$$x_t(\eta_{[t]}) = \boldsymbol{X_t}[1, \boldsymbol{G_1}^\top, \cdots, \boldsymbol{G_t}^\top, \boldsymbol{Q_1}^\top, \cdots, \boldsymbol{Q_t}^\top]^\top = \boldsymbol{X_t} \boldsymbol{P}_{\boldsymbol{\eta_{[t]}}} \boldsymbol{\eta'} \quad (4.23)$$

This formulation provides the flexibility to achieve a discontinuous solution for the continuous variable. Notice that if a piecewise continuous decision rule is necessary to be enforced, we can express the continuous variables as the following affine function of the lifted uncertain parameter:

$$x_t(\eta_{[t]}) = \boldsymbol{X_t}[1, \boldsymbol{G_1}^\top, \cdots, \boldsymbol{G_t}^\top]^\top = \boldsymbol{X_t} \boldsymbol{P}_{\boldsymbol{G_{[t]}}} \boldsymbol{\eta'} \quad (4.24)$$

Adaptive binary variables are approximated using the following affine function of the lifted uncertain parameter:

$$y_t(\eta_{[t]}) = \boldsymbol{Y_t}[1, \boldsymbol{Q_1}^\top, \cdots, \boldsymbol{Q_t}^\top]^\top = \boldsymbol{Y_t} \boldsymbol{P}_{\boldsymbol{Q_{[t]}}} \boldsymbol{\eta'} \quad (4.25)$$

Using the above linear decision rules, the original problem formulation (equation 4.2) will become a semi-infinite optimization problem since the parameter $\boldsymbol{\eta'}$ belongs to convex set $\hat{\Omega}'$ and can take

infinite number of values (equation 4.26b).

$$\min \quad \mathbb{E}(\sum_{t}(\boldsymbol{\eta'}^{\top}\boldsymbol{P_{\eta_{[t]}}}^{\top}\boldsymbol{C_t}^{\top}\boldsymbol{X_t}\boldsymbol{P_{\eta_{[t]}}}\boldsymbol{\eta'} + \boldsymbol{\eta'}^{\top}\boldsymbol{P_{\eta_{[t]}}}^{\top}\boldsymbol{D_t}^{\top}\boldsymbol{Y_t}\boldsymbol{P_{Q_{[t]}}}\ \boldsymbol{\eta'})) \tag{4.26a}$$

$$s.t. \quad \sum_{\tau=1}^{t}(\boldsymbol{A}_{t,\tau}^{\top}\boldsymbol{X_t}\boldsymbol{P_{\eta_{[t]}}}\boldsymbol{\eta'} + \boldsymbol{B}_{t,\tau}^{\top}\boldsymbol{Y_t}\boldsymbol{P_{Q_{[t]}}}\boldsymbol{\eta'}) \leq F_t\boldsymbol{P_{\eta_{[t]}}}\boldsymbol{\eta'} \qquad \forall t, \boldsymbol{\eta'} \in \hat{\Omega}' \tag{4.26b}$$

$$\boldsymbol{Y_t} \in \{-1, 0, 1\} \qquad \forall t \tag{4.26c}$$

To solve this problem, duality based standard robust linear optimization formulation is applied to the constraints and the resulting deterministic counterpart is obtained.

$$\min \quad \mathrm{Tr}\sum_{t}(\boldsymbol{P_{\eta_{[t]}}}^{\top}\boldsymbol{C_t}^{\top}\boldsymbol{X_t}\boldsymbol{P_{\eta_{[t]}}}\mathbb{E}[\boldsymbol{\eta'}\boldsymbol{\eta'}^{\top}] + \boldsymbol{P_{\eta_{[t]}}}^{\top}\boldsymbol{D_t}^{\top}\boldsymbol{Y_t}\boldsymbol{P_{Q_{[t]}}}\ \mathbb{E}[\boldsymbol{\eta'}\boldsymbol{\eta'}^{\top}]) \tag{4.27a}$$

$$s.t. \quad \boldsymbol{h'\theta_t} \geq \boldsymbol{0} \qquad \forall t \tag{4.27b}$$

$$\boldsymbol{J'}^{\top}\boldsymbol{\theta_t} = -\sum_{\tau=1}^{t}(\mathbf{A}_{t,\tau}^{\top}\boldsymbol{X_t}\boldsymbol{P_{\eta_{[t]}}} + \mathbf{B}_{t,\tau}^{\top}\boldsymbol{Y_t}\boldsymbol{P_{Q_{[t]}}}) + \mathbf{F}_t\boldsymbol{P_{\eta_{[t]}}} \quad \forall t \tag{4.27c}$$

$$\boldsymbol{\theta_t} \geq \boldsymbol{0} \qquad \forall t \tag{4.27d}$$

$$\boldsymbol{Y_t} \in \{-1, 0, 1\} \qquad \forall t \tag{4.27e}$$

where $\boldsymbol{\theta}_t$ is the dual variable. The final model is a MILP problem.

Remark: Limitation of the binary decision rule

In the binary decision rule method, the binary variable is a linear combination of 0-1 indicator functions and each indicator function is a function of only one uncertain parameter. Therefore, the final binary variable can only be a function of one of the uncertain parameters. This restriction limits the solution quality. To illustrate this limitation, assume a problem with two time steps and one uncertain parameter at each time step. For simplicity, consider one breakpoint for each uncertain parameter. Under this setting, for $t = 1$ and $t = 2$, the binary variable is formulated as:

$$t = 1,\ y_1(\eta_{[1]}) = Y_0 + Y_{11}Q_{11}(\eta_1)$$

$$t = 2,\ y_2(\eta_{[2]}) = Y_0 + Y_{11}Q_{11}(\eta_1) + Y_{21}Q_{21}(\eta_2)$$

At $t = 2$, there are two indicator functions ($Q_{11}(\eta_1)$ and $Q_{21}(\eta_2)$) involved in the decision rule. Considering the range of coefficients $Y \in \{-1, 0, +1\}$, linear combination of independent 0-1 indicator functions will not necessarily be binary. Therefore, at each time step, the binary variable can only be a function of one uncertain parameter. Table 4.4 shows all the possible combinations of the indicator functions at $t = 2$. Among all the 18 combinations, 6 are feasible that result in a binary variable. In each feasible case, the indicator function related to only one of the uncertain

parameters is involved, hence the solution quality is restricted.

Table 4.4: Combinations of indicator functions at $t = 2$

$y_2(\eta_{[2]})$	NF	NF	NF	NF	F	F	NF	F	NF	NF	F	NF	F	F	NF	NF	NF	NF
Y_0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1
Y_{11}	-1	-1	-1	0	0	0	1	1	1	-1	-1	-1	0	0	0	1	1	1
Y_{21}	-1	0	1	-1	0	1	-1	0	1	-1	0	1	-1	0	1	-1	0	1

NF: Not Feasible, F: Feasible

4.5 Results for inventory planning problem

The example problem introduced in Section 2 is studied here using the different piecewise decision rule approximation methods presented in Section 3 and 4. All the numerical experiments are run on a desktop computer with 8 Gb RAM and Intel(R) Core(TM) i5-7500 CPU @ 3.40GHz processor.

4.5.1 Lifting method results

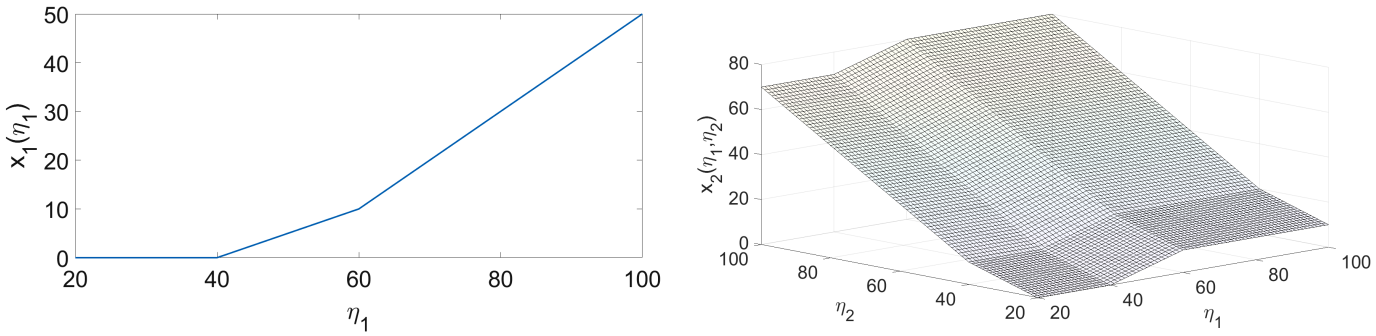
The inventory problem is investigated for cases with different time horizon $T = 2, 3, 5, 10$. In order to demonstrate the effectiveness of the proposed framework, 4 combinations of variables are examined for each case: adaptive, static continuous, static binary, all static. In the adaptive setting, both continuous and binary variables are adaptive. In the static continuous setting, the continuous variables are forced to be static and the binary variables are adaptive. In the static binary setting, the continuous variables are adaptive and the binary variables are static. As a basis for comparison, when continuous and binary variables are all static, the optimal objective for the four cases $T = 2, 3, 5, 10$ are 1850, 2900, 5350, 15350, respectively.

Tables 4.5 to 4.8 present the results for different cases using various number of breakpoints. Note that if a problem is not solved to optimality, then the relative optimality gap is also reported along with the objective in the table. It is evident that in all the time stages, the adaptive and all static settings result in the best and worst objective, respectively. According to the results of the static continuous and static binary setting, in most of the cases, the contribution of the adaptive continuous variables is greater than the adaptive binary variables. For static continuous setting, the objective deteriorates more significantly compared to the static binary setting.

Another important observation is that increasing the number of breakpoints improves the objective as expected. For time horizons $T = 2, 3, 5$, even using a large number of breakpoints, the run time for the adaptive settings is under 1 min, therefore it justifies to increase the number of breakpoints

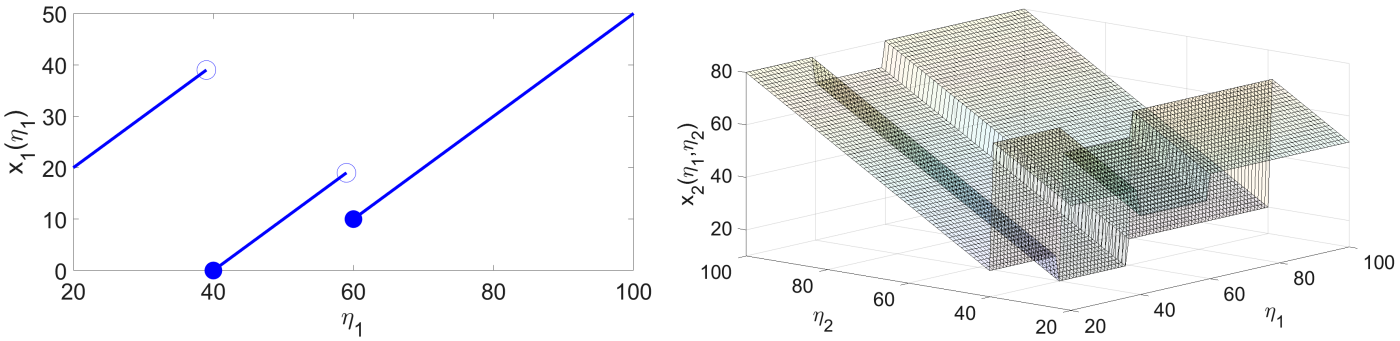
to obtain a better objective. For $T = 2, 3, 5$, increasing the number of breakpoints beyond 5 has not improved the objective. For $T = 10$, the model size is very large and increasing the number of breakpoints beyond 3 results in notable increase of run time. Therefore, depending on the problem size, a reasonable number of breakpoints should be selected to obtain the best objective and the least run time.

Figures 4.8a, 4.8b show the continuous decision rule at $t = 1$ and $t = 2$ for the case of $T = 3$ using 3 breakpoints. As these figures demonstrate, the continuous decision rule results in a continuous solution that restricts the solution flexibility. Figures 4.9a and 4.9b illustrate the same figures for discontinuous decision rule. It can be observed that the continuous solution is piecewise discontinuous linear with respect to demand uncertain parameter. This results in an improved objective compared to static or simply linear solutions. The continuous solution show that the ordering amount increases when the demand increases, which is reasonable. The binary solutions can be observed from Figure 4.10. The figure demonstrates that the obtained binary solution is adaptive and therefore the solution quality is enhanced.



(a) (b)

Figure 4.8: Adaptive solution under piecewise continuous decision rule, for $T = 3$, 3 breakpoints



(a) (b)

Figure 4.9: Adaptive solution under piecewise discontinuous decision rule $T = 3$, 3 breakpoints applied

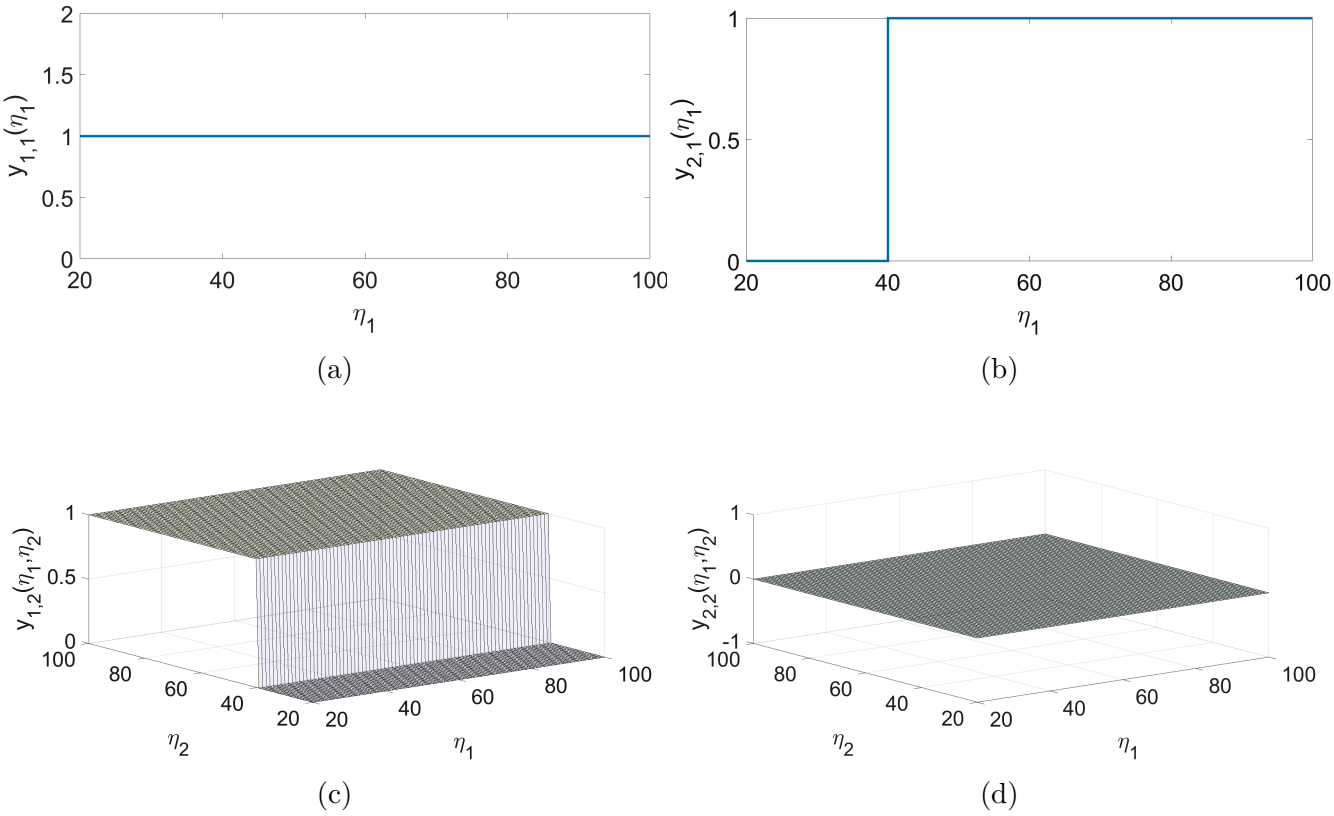


Figure 4.10: Adaptive binary solution for $T = 3$; 3 breakpoints applied for each uncertain parameter

Table 4.5: Results for the inventory problem, T=2

Break points[1]	Adaptive [2]	Static Cont[3]	Static Bin[4]	All Static[5]	Variables[6]	Run Time[7]
Br=1	1620.00	1690.00	1622.50	1850.00	466, 10	0.3 s
Br=3	1287.50	1462.50	1613.75	1850.00	710, 22	0.3 s
Br=5	1120.00	1268.75	1609.37	1850.00	954, 34	0.4 s
Br=7	1120.00	1268.75	1268.75	1850.00	1138, 46	0.7 s

[1] Number of breakpoints used in uncertainty lifting.

[2] Optimal objective value by allowing both continuous and integer variables be adaptive.

[3] Optimal objective value by allowing only adaptive continuous be adaptive.

[4] Optimal objective value by allowing only integer variables be adaptive.

[5] Optimal objective value by forcing all continuous and integer variables to be non-adaptive.

[6] Number of continuous and binary variables in the model with both adaptive continuous and integer variables.

Table 4.6: Results for the inventory problem, T=3

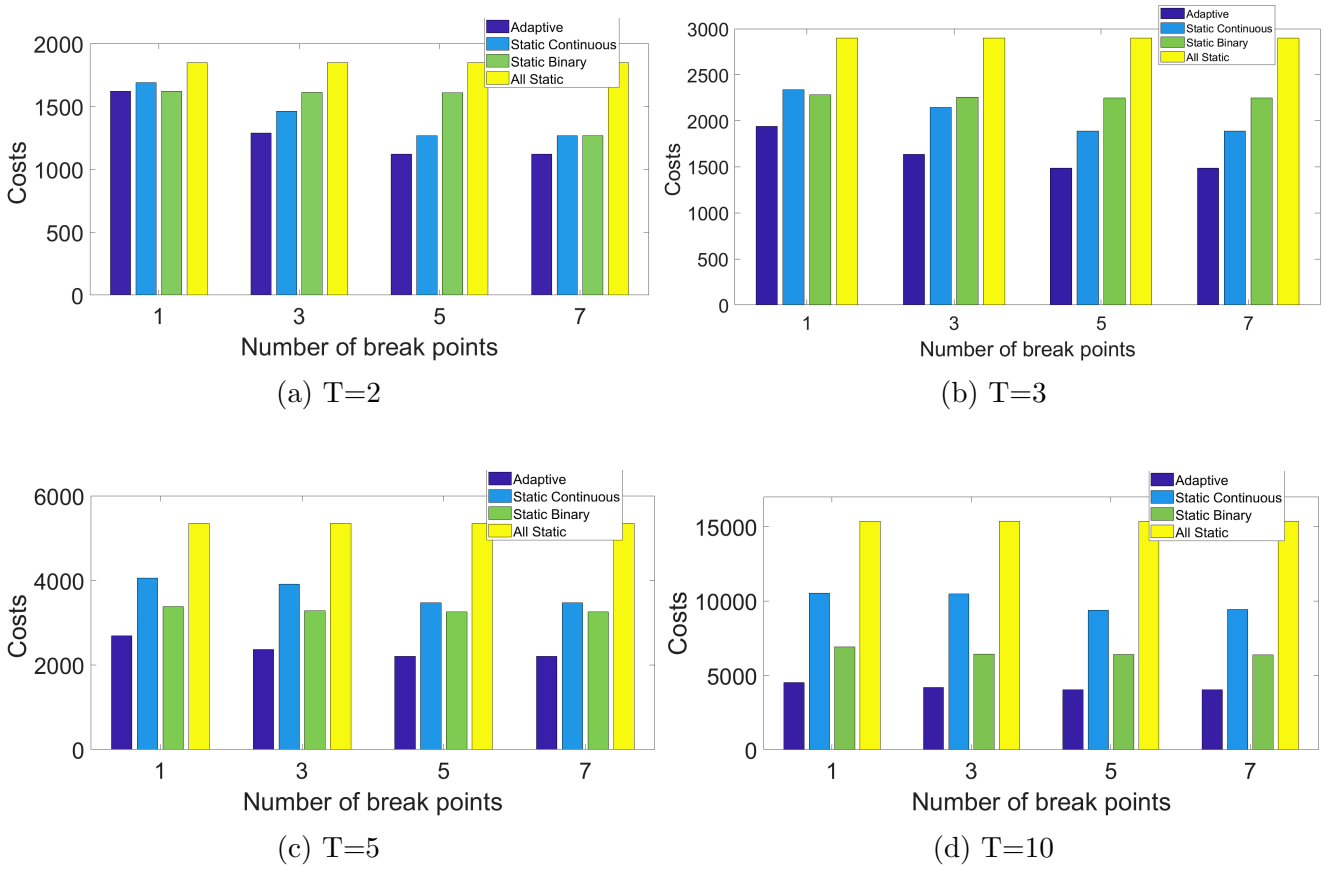
Break points	Adaptive	Static Cont	Static Bin	All Static	Variables	Run time
Br=1	1940.00	2340.00	2282.50	2900.00	1064, 18	0.2 s
Br=3	1637.50	2147.50	2258.75	2900.00	1646, 42	0.4 s
Br=5	1487.50	1891.25	2246.87	2900.00	2228, 66	1.0 s
Br=7	1487.50	1891.25	2246.87	2900.00	2810, 90	1.5 s

Table 4.7: Results for the inventory problem, T=5

Break Points	Adaptive	Static Cont	Static Bin	All Static	Variables	Run time
Br=1	2692.50	4055.00	3377.50	5350.00	2983, 40	0.4 s
Br=3	2362.50	3912.50	3278.50	5350.00	4673, 100	2.3 s
Br=5	2207.50	3475.00	3260.00	5350.00	6363, 160	5.1 s
Br=7	2207.50	3475.00	3255.62	5350.00	8053, 220	41.1 s

Table 4.8: Results for the inventory problem, T=10

Break points	Adaptive	Static Cont	Static Bin	All Static	Variables	Run time
Br=1	4517.50	10530.00	6902.50	15350.00	11998, 130	1.6 s
Br=3	4177.50	10442.50	6433.75	15350.00	18978, 350	87 s
Br=5	4032.50, 4.8% gap	9372.50	6391.87	15350.00	25958, 570	15 min
Br=7	4027.50, 8.1% gap	9417.50, 7.1% gap	6364.37	15350.00	32938, 790	30 min



(a) T=2 (b) T=3

(c) T=5 (d) T=10

Figure 4.11: At all the time steps, the model that includes both adaptive binary and adaptive continuous variables results in the best objective (minimum cost) compared to models that include static continuous, static binary or both static continuous and static binary variables.

Figure 4.11 summarizes the solution for different cases. It gives an intuitive illustration on the effect of the adaptability of different type of decisions. It is also clear to see the trend that as the number of breakpoints increases, the solution quality increases. Note that in the study, the set of smaller number of breakpoints is always the subset of the large number of breakpoints in the same case. This is to ensure the location of the breakpoints is not making impact. However, we point out the location of the breakpoints is also important and may affect the quality of the solution significantly.

4.5.2 Partitioning method results

Tables 4.9 to 4.12 present the results for the partitioning method applied to the inventory problem. For small cases with $T = 2, 3$, the run time is close to the lifting method. In terms of the objective, partitioning method provides slightly better objective values. However, for $T = 5, 10$, as the model size becomes larger, the lifting method becomes much more efficient from run time point of view and the objective value is similar or even better than the partitioning method. For instance for the case with $T = 10$, the partitioning method results in the objective of 4259.19 in 30 min while

the lifting could provide the better objective of 4177.50 in just 87 s or 4032.50 in 15 min. It is also important to notice that for $T = 10$, Br=3/5/7, the partitioning method results in such a large model size that the CPLEX solver could not return a feasible solution after 3 hours while the lifting method was solved very efficiently in 87s. Thus, the lifting method with the flexible discontinuous decision rule is efficient for large model cases.

Table 4.9: Results for the inventory problem, T=2

Break points	Objective	Variables	Run Time
Br=1	1490.00	110, 16	0.3 s
Br=3	1207.50	422, 64	0.4 s
Br=5	1113.75	942, 144	0.4 s
Br=7	1111.87	1670, 256	0.4 s

Table 4.10: Results for the inventory problem, T=3

Break points	Objective	Variables	Run Time
Br=1	1820.00	495, 48	0.2 s
Br=3	1556.25	3911, 384	1.3 s
Br=5	1466.25	13183, 1296	4.8 s
Br=7	1459.77, 2.3%	31239, 3072	30 min

Table 4.11: Results for the inventory problem, T=5

Break points	Objective	Variables	Run Time
Br=1	2508.75	5449, 320	0.7 s
Br=3	2235.95, 2.1%	174089, 10240	30 min
Br=5	2199.54, 8.7%	1321929, 77760	30 min
Br=7 **	—	—	—

** Problem size is too large. No feasible solution reported after 3 hrs.

Table 4.12: Results for the inventory problem, T=10

Break points	Objective	Variables	Run Time
Br=1	4259.19, 0.91%	686094, 20480	30 min
Br=3 **	—	—	—

** Problem size is too large. No feasible solution reported after 3 hrs.

We can use the partitioning method as a reference to check the sub-optimality of the results of the lifting method. While the number of breakpoints is large, we can get a high-resolution partition of the uncertainty set. In other words, we divide the uncertainty set into very small sub-regions and each sub-region is associated with a decision rule. This way, the obtained solution will be close to the optimal solution. The partitioning method results reported in Table 9 with 5 and 7 breakpoints shows that the optimal solution is around 1111 for $T = 2$. This provides a basis for comparison of the solution quality from lifting method as reported in Table 5, where the best solution 1120 is not far from this reference. Similar evaluation of solution quality can be made for $T = 3, 5$ for the inventory planning problem.

4.6 Capacity Expansion Planning

This section investigates a case study for multistage capacity planning of a chemical plant. Figure 4.12 illustrates the process network superstructure. In this network, chemicals $j = 1, \cdots, 4$ can be purchased from external resources and chemical $j = 5$ is produced and sold to the market. It is assumed that there is no inventory in this problem, therefore all the purchased or produced chemicals are consumed or sold in the process. The demand at each time step must be satisfied and the objective is to maximize the overall revenue over a number of time stages.

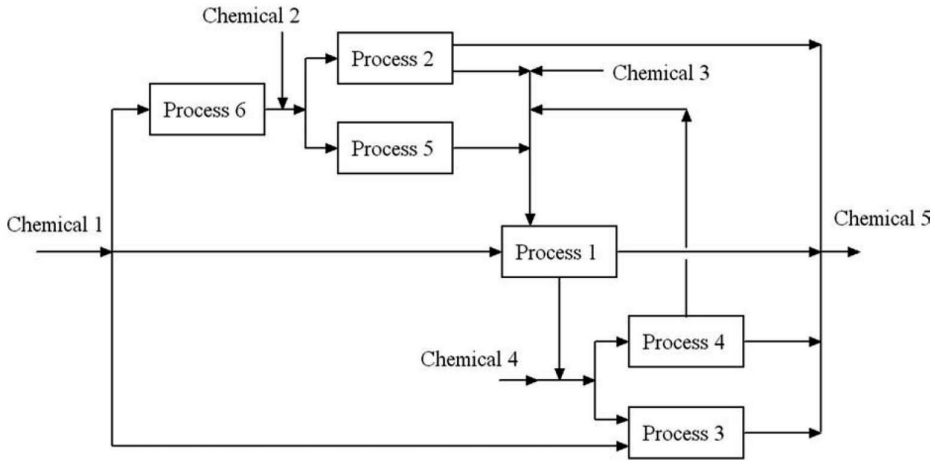


Figure 4.12: Chemical plant superstructure

There are two sources of uncertainty in this problem: demand $d_{5,t}$ and price $\gamma_{5,t}$ of the sold product (chemical 5). Figure 4.13 illustrates the decision making and uncertainty revelation sequence. At the beginning of the first time period, there is an initial pre-built capacity ($Q_{i,1} = 100000$) for each process $i = 1, \cdots, 6$. During each time period t and after demand and price uncertainties are revealed, some capacity expansion decisions $(y_{i,t}(\eta_{[t]}), X_{i,t}(\eta_{[t]}))$ are made based on the revealed uncertainty up to time period t, $\eta_{[t]}$. The process capacity at each time period is the summation

of the existing capacity and the expanded capacity in the previous time period. $Q_{i,t}(\eta_{[t-1]})$ is used to satisfy the demand at the same time period. At each time step, after uncertainty revelation, operation $W_{i,t}(\eta_{[t]})$ and purchase decisions $P_{j,t}(\eta_{[t]})$ are also made.

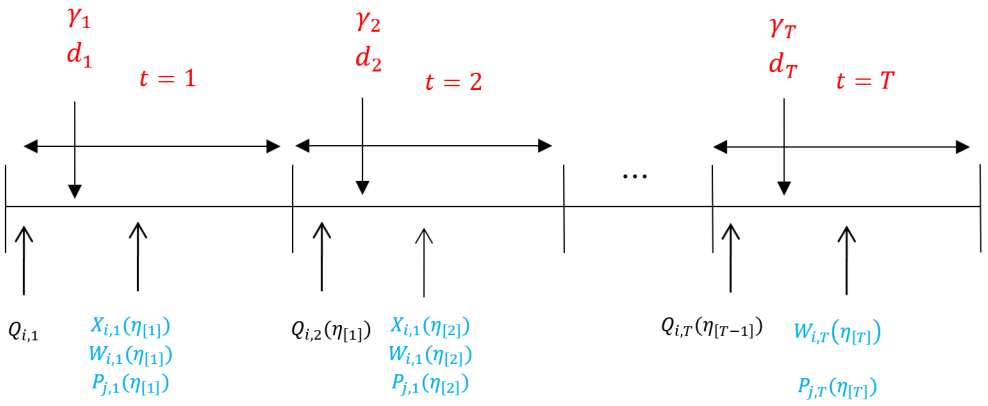


Figure 4.13: Decision making and uncertainty realization sequence

In this problem, demand and price uncertainties at each time step are defined as:

$$\gamma_{j,t}(\eta_t^p) = \bar{\gamma}_{j,t}(1 + \eta_t^p) \qquad \forall t, j = 5 \tag{4.28}$$

$$d_{j,t}(\eta_t^d) = \bar{d}_{j,t}(1 + \eta_t^d) \qquad \forall t, j = 5 \tag{4.29}$$

where $\bar{\gamma}_{j,t}$ and $\bar{d}_{j,t}$ are the minimum price and demand for the fifth product respectively. Primitive uncertain parameters $\eta_t^d \in [0, 1.5]$, $\eta_t^p \in [0, 2]$ are assumed to be uniformly distributed and independent at each time step.

The capital budget restriction CI_t at each year t and the minimum number of expansions for each process during the project time horizon N_i^{min} are considered to be linear functions of uncertainty.

$$CI_t(\eta) = \bar{CI}_t(1 + \eta_t^p + \eta_t^d) \qquad \forall t \tag{4.30}$$

$$N_i^{min}(\eta) = \frac{1}{T-1} \sum_{t=1}^{T-1} \eta_t^d - \bar{\alpha} \qquad \forall i \tag{4.31}$$

where $\bar{CI}_t$ is the minimum available budget at each year t and $\bar{\alpha}$ is a constant parameter (0.75) to adjust the minimum required number of process expansions considering the average range of demand uncertainty. Equation 4.30 means when the demand or price of the sold product increases, it justifies to allocate a larger budget for process expansion in order to meet the demand and increase the revenue. Equation 4.31 indicates that when demand increases, the minimum number of process expansions will also increase to meet the demand.

The general multistage stochastic programming formulation for this problem is presented as follows:

$$\max\ \mathbb{E}_{\eta}\Big[\sum_{t=1}^{T}\sum_{j=1}^{5}(\gamma_{j,t}(\eta_t^p)S_{j,t}(\eta_{[t]})-\Gamma_{j,t}P_{j,t}(\eta_{[t]}))-\sum_{t=1}^{T}\sum_{i=1}^{6}\delta_{i,t}W_{i,t}(\eta_{[t]})-\sum_{t=1}^{T-1}\sum_{i=1}^{6}(\alpha_{i,t}X_{i,t}(\eta_{[t]})+\beta_{i,t}y_{i,t}(\eta_{[t]}))\Big] \tag{4.32a}$$

s.t.

$$0\le X_{i,t}(\eta_{[t]})\le y_{i,t}(\eta_{[t]})X_{i,t}^{U} \qquad \forall i,t\le T-1,\eta\in\Omega \tag{4.32b}$$

$$Q_{i,t}(\eta_{[t-1]})=Q_{i,t-1}(\eta_{[t-2]})+X_{i,t-1}(\eta_{[t-1]}) \qquad \forall i,t\ge 2,\eta\in\Omega \tag{4.32c}$$

$$N_i^{min}(\eta)\le\sum_{t=1}^{T-1}y_{i,t}(\eta_{[t]})\le N_i^{max} \qquad \forall i,\eta\in\Omega \tag{4.32d}$$

$$\sum_{i=1}^{6}(\alpha_{i,t}X_{i,t}(\eta_{[t]})+\beta_{i,t}y_{i,t}(\eta_{[t]}))\le CI_t(\eta) \qquad \forall t\le T-1,\eta\in\Omega \tag{4.32e}$$

$$W_{i,t}(\eta_{[t]})\le Q_{i,t}(\eta_{[t-1]}) \qquad \forall i,t,\eta\in\Omega \tag{4.32f}$$

$$P_{j,t}(\eta_{[t]})+\sum_{i=1}^{6}\nu_{i,j}W_{i,t}(\eta_{[t]})=S_{j,t}(\eta_{[t]})+\sum_{i=1}^{6}\mu_{i,j}W_{i,t}(\eta_{[t]}) \qquad \forall j,t,\eta\in\Omega \tag{4.32g}$$

$$P_{j,t}(\eta_{[t]})\le a_{j,t} \qquad \forall j,t,\eta\in\Omega \tag{4.32h}$$

$$S_{j,t}(\eta_{[t]})\ge d_{j,t}(\eta_t^d) \qquad \forall j=5,t,\eta\in\Omega \tag{4.32i}$$

$$X_{i,t}(\eta_{[t]})\ge 0 \qquad \forall i,t\le T-1,\eta\in\Omega \tag{4.32j}$$

$$W_{i,t}(\eta_{[t]}),P_{j,t}(\eta_{[t]})\ge 0 \qquad \forall i,j,t,\eta\in\Omega \tag{4.32k}$$

$$y_{i,t}(\eta_{[t]})\in\{0,1\} \qquad \forall i,t\le T-1,\eta\in\Omega \tag{4.32l}$$

$$P_{j,t}(\eta_{[t]})=0 \qquad \forall j=5,t,\eta\in\Omega \tag{4.32m}$$

$$S_{j,t}(\eta_{[t]})=0 \qquad \forall j=1,2,3,4,t,\eta\in\Omega \tag{4.32n}$$

In this problem, the objective is to maximize the profit over a given time horizon. Eq. 4.32a is the objective function that accounts for the revenue from selling chemical 5 and the costs of operation, capacity installation and material purchase. Constraint 4.32b enforces bounds on the expansion capacity ($X_{i,t}(\eta_{[t]})$) of process i. $y_{i,t}(\eta_{[t]})$ is a binary variable which indicates whether process i is expanded or not at period t. Eq. 4.32c calculates the total capacity for each process up to time t. Constraint 4.32d means that the total number of expansions for each process i is restricted between a lower and upper bound during the project horizon. Constraint 4.32e indicates that, at each year, there is a capital restriction $CI_t(\eta_t)$ for expansion ($\alpha_{i,t}X_{i,t}(\eta)$) and fixed investment ($\beta_{i,t}y_{i,t}(\eta)$) costs. Constraint 4.32f means that the operation level of each process at time t can not be more than the total installed capacity for that process up to time t. Eq. 4.32g is

a material balance equation and it means the total amount of purchased ($P_{j,t}(\eta_{[t]})$) and produced ($\eta_{i,j}W_{i,t}(\eta_{[t]})$) chemicals must be equal to the total amount of consumed ($\mu_{i,j}W_{i,t}(\eta_{[t]})$) and sold ($S_{j,t}(\eta_{[t]})$) chemicals. Constraint 4.32h indicates that the amount of purchased chemical j at time t should be less than or equal to the available amount of that chemical ($a_{j,t}$). Constraint 4.32i describes that the amount of sold chemical j at time t ($S_{j,t}(\eta_{[t]})$) should satisfy the demand ($d_{j,t}(\eta_t^d)$). Constraint 4.32k indicates that problem variables can not be negative. Constraint 4.32m indicates that chemical 5 is sold and constraint 4.32n means that chemicals 1 to 4 are purchased and not sold. Following the methodology explained in Section 4, the binary and continuous variables are formulated as affine functions of lifted uncertain parameters and the original problem is converted to its deterministic counterpart and solved using MILP solver CPLEX. All the parameter values are listed in the appendix.

The capacity expansion problem is studied for $T = 3, 4$ and 5 time stages. Note that for the case that all variables are static, the optimal objective for $T = 3, 4, 5$ are 10111.69, 15194.75 and 20953.12, respectively. Tables 4.13 to 4.15 present the results for lifting method under discontinuous decision rules for adaptive continuous variables. The contribution of the adaptive continuous variables is greater than the adaptive binary variables. For instance, in T=4 with 3 breakpoints, the objective reduced by 5.7% for the static continuous case while it was only reduced by 0.2% for the static binary. In all the time horizons, increasing the break points have improved the objective value. However, for larger settings, the run time increases significantly. For example, in T=4 setting, when the number of break points is increased from 1 to 3, the run time has increased from 20 s to 5 hrs which may not be desirable. Thus, for larger settings, selecting a reasonable number of breakpoints can save a substantial amount of time and computational effort.

Figures 4.14a, 4.14b provide a graphical illustration of the results for the continuous decision rule for $T = 3$ (the same figures for the discontinuous decision rule are similar). Both figures show that when the demand increases, the installed capacity and total capacity increase accordingly. At each time step, the continuous variables depend on all the revealed uncertain parameters in the previous time steps. For instance, the amount of sold product at $t = 5$ discontinuous decision rule (1 breakpoint) is provided below:
$S_{5,5}(\eta) = 425490+55306.41G^d_{2,1}+72644.72G^d_{3,1}+1192.07G^p_{3,1}+178000+23863.62Q^p_{2,1}+48401.03Q^p_{3,1}$
where G^d_{tj}, G^p_{tj} and $Q^d_{t,j}, Q^p_{t,j}$ are the j-th component of the lifted continuous and binary demand and price uncertain parameters at time step t respectively.

Table 4.13: Results for chemical plant based on discontinuous decision rule, T=3

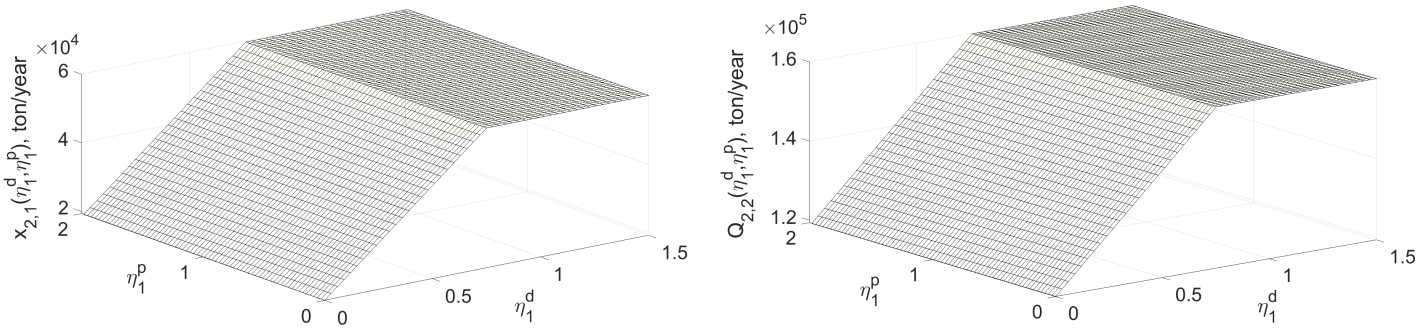
Break points	Adaptive	Static Cont	Static Bin	Variables	Run time
1	10784.58	10533.73	10653.70	14204, 48	1.6 s
3	10872.66	10539.83	10769.44	22750, 120	50.7 s

Table 4.14: Results for the chemical plant based on discontinuous decision rule,T=4

Break points	Adaptive	Static Cont	Static Bin	Variables	Run time
1	16405.54	15512.61	16225.68	25550, 90	20.0 s
3	16467.41	15528.53	16427.24	41086, 234	5hr

Table 4.15: Results for the chemical plant based on discontinuous decision rule, T=5

Break points	Adaptive	Static Cont	Static Bin	Variables	Run time
1	22710.20	20961.42	22520.43	40192, 144	4 min
3	23054.45	21098.23	23015.95	62478, 384	38hr



(a) Capacity expansion of process $i = 2$ at $t = 1$

(b) Total capacity of process $i = 2$ at $t = 2$

Figure 4.14: Capacity expansion $x_{2,1}$ and total capacity $Q_{2,2}$ of process $i = 2$ at $t = 1$ and $t = 2$ for $T = 3$

4.7 Conclusion

In this study, a novel solution framework is proposed for solving multistage mixed integer stochastic problems that contains both adaptive binary and adaptive continuous variables, based on uncertainty set partitioning and uncertainty lifting ideas. While the traditional uncertainty discretization and scenario tree representation based method results in exponential growth of problem size that makes the problem computationally prohibitive, the proposed piecewise decision rule based

methods provide flexibility for seeking solution quality and computational efficiency. The uncertainty lifting method is compared with the uncertainty partitioning method in terms of solution quality and computational efficiency. While the uncertainty set partitioning method leads to better solution quality and is appropriate for small scale problems. It is demonstrated that the lifting method can provide the scalability and tractability required for large scale problems with significant computational advantage.

Breakpoints are used in both methods of the proposed framework. Increasing the number of breakpoints leads to more flexible adaptive solution and may improve the solution. However, beyond a certain point, the run time increases considerably while the objective remains the same or improves negligibly. Depending on the problem size, a reasonable number of breakpoints need to be selected. In this paper, we didn't investigate the impact of the location of the breakpoints. However, it is worth pointing out that the location is also as important as the number of breakpoints. Furthermore, this work addresses the exogenous uncertainty only. Future work will be extended to investigate the problem with endogenous type of uncertainty.

Nomenclature for capacity planning example

Indices

t time period

i process

j chemical

Parameters

η_t^d	Demand uncertainty at period t
η_t^p	Price uncertainty at period t
$\alpha_{i,t}$	Unit expansion cost of process i at the beginning of period t
$\beta_{i,t}$	Fixed cost of establishing or expanding process i at the beginning of period t
$\gamma_{i,t}, \Gamma_{i,t}$	Selling and buying prices of chemical j in period t
$\delta_{i,t}$	Unit operating cost for process i during period t
$\nu_{i,j}$	Input proportionality constant for chemical j in process i
$\mu_{i,j}$	Output proportionality constant for chemical j in process i
CI_t	Investment budget for period t
$a_{j,t}$	Availability of chemical j in period t
$d_{j,t}$	Demand of chemical j in period t
N_i^{EXP}	Allowable number of expansions for process i
$X_{i,t}^L, X_{i,t}^U$	Lower and upper bounds for capacity expansion of process i in period t

Variables

$P_{j,t}$	Units of chemical j purchased at period t
$Q_{i,t}$	Total capacity of process i in period t
$S_{j,t}$	Units of chemical j sold at the end of period t
$W_{i,t}$	Operating level of process i in period t
$X_{i,t}$	Units of expansion of process i at period t
$y_{i,t}$	Binary variable; if process i is expanded during period t, then $y_{i,t} = 1$, otherwise $y_{i,t} = 0$

Chapter 5

Multistage Stochastic Mixed Integer Optimization Under Endogenous and Exogenous Uncertainty

Abstract

A novel mathematical framework is proposed for multistage adaptive optimization that incorporates endogenous and exogenous uncertainty based on robust optimization technique. The endogenous uncertainty is formulated based on the k-adaptability (partitioning) method where at each scenario, the binary variable is constant and the continuous variable is an affine function of endogenous uncertain parameters. The exogenous uncertainty is modeled using the lifting method where the binary variables are affine functions of 0-1 indicator functions and the continuous variables are approximated using affine functions of lifted uncertain parameters. In order to demonstrate the applicability of the proposed framework, a systematic approach is followed with respect to problem comlexity beginning from the simplest case of only binary variables under endogenous uncertainty to the more inclusive case of both binary and continuous variables under endogenous and exogenous uncertainty. A case study for infrastructure and production planning of a gas filed is employed. The results show that the proposed framework can successfully be applied to large problems and the solution effectively incorporates both types of uncertainty for binary and continuous variables.

5.1 Introduction

Optimization under uncertainty aims to make the best decisions under uncertain conditions. Uncertainty is categorized as exogenous and endogenous uncertainty. If the uncertainty is not affected by the problem decisions, it is called exogenous uncertainty, otherwise it is referred to as endogenous uncertainty. Endogenous uncertainty is typically classified into two types. In type 1, the decisions affect the underlying distribution of uncertainty. For instance, a company may decide to lower the price of its products in order to change the probability distribution of its sale toward higher values and to increase its market share. In type 2, the decisions determine the time of uncertainty revelation . For example, the true amount of recoverable gas in a gas field is usually unknown until after a costly well drilling operation is completed. Two main approaches are used in this field: scenario-tree based optimization and robust duality based optimization. Different studies that address scenario-tree based method or robust optimization based method under exogenous or endogenous uncertainty are reviewed in the following.

In the context of scenario-tree based optimization under exogenous uncertainty, many studies are carried out and a few are reviewed in the following. Takamatsu et al.,1974[80] addressed the parameter uncertainty involved in design of unit operations in a process system and applied the method on a simple chemical reaction process. Grossmann and Sargent, 1978 [81] proposed a strategy for optimum design of chemical plants where the uncertain parameters are bounded. The method ensures that the plant specifications are met for any feasible realization of uncertain parameters. The method is illustrated on a reactor-separator system and a heat exchanger network. Gupta and Maranas, 2000 [82] suggested a two-stage stochastic programming approach to address demand uncertainty in supply-chain planning problems. The production decisions are first stage and the supply-chain decisions are second stage decisions. It was shown that the computational requirements are much smaller than the Monte Carlo sampling. Applequist, 2000[83] presented a metric for evaluating supply chain design and planning projects that accounts for the effects of demand uncertainties on revenue while considering uncertainty in inventory over time. The method provides a rational balance between expected value of investment performance and its variance. We refer to Sahinidis, 2004 [17] for a review on stochastic programming under exogenous uncertainty.

Scenario-tree based optimization under type 2 endogenous uncertainty is addressed by different authors. Goel and Grossmann, 2004 [54] proposed a stochastic programming model that incorporated the dependency of scenario tree on decisions. They addressed an optimal investment and operational planning of gas field development under uncertainty in gas reserves. Goel and Grossmann, 2006 [20] extend the stochastic programming framework by presenting a mixed-integer formulation for stochastic problems where decisions influence the time of information discovery.

Boland et al., 2008, [84] formulated a mixed integer multistage stochastic method to address the scheduling of open pit mine production. In this framework, the decision made in later time periods can depend on observations of geological properties in earlier periods. The authors extended the formulation to general stochastic problems under endogenous uncertainty. They discussed that using the structure of the proposed method, a significant number of non-anticipativity constraints can be omitted in some cases. Colvin and Maravelias, 2008 [85] presented a multi-stage stochastic formulation for planning of clinical trials in pharmaceutical research. Given a portfolio of potential drugs and limited resources, the model determines the trials to be performed in each period. Solak et al., 2010, [86] developed a detailed description of research and development project portfolio management and modeled the problem as a multistage stochastic problem with endogenous uncertainty. They proposed a solution technique than can be used along with scenario decomposition methods.

Scenario-based optimization under type 1 endogenous uncertainty has received less attention. Peeta et al., 2010 [87] addressed a pre-disaster planning problem for strengthening a highway network where links are subject to failure due to disaster. The goal is to select links to invest in considering budget restrictions with the objective of maximizing the post disaster connectivity and minimizing travel costs. Escudero et al., 2018 [88] addressed a three stage problem for managing natural disaster mitigation. The endogenous uncertainty is based on the decision-maker's investment to obtain greater accuracy in regard to the occurrence of disaster and reinforcing the network infrastructure. Hellemo at al., 2018 [89] formulated two-stage stochastic problems with recourse where the prior probabilities are distorted by an affine transformation or combined using a convex combination of several probability distributions through decision variables.

Since scenario-based multistage stochastic programming problems usually suffer from the curse of dimensionality (exponential growth of model size at multistage settings), researches have tried to tackle this issue using different methods. Apap and Grossmann, 2017 [19] proposed a sequential scenario decomposition heuristic and a Lagrangean decomposition method. Goel and Grossmann, 2006 [20] presented a Lagrangean branch and bound algorithm to reduce the model size. Gupta and Grossmann, 2014 [21] developed a new Lagrangean decomposition algorithm for solving large-scale stochastic problems with endogenous uncertainty that can reduce the computational expense by reducing the number of non-anticipativity constraints. Colvin and Maravelias, 2010 [11] developed novel branch and cut algorithms where non-anticipativity constraints that are unlikely to be active are removed from the formulation and added only if they are violated within the search tree.

Robust optimization under exogenous uncertainty is investigated in various studies. Mitra et al., 2012 [90] applied robust optimization to uncertain electricity prices using an uncertainty set with multiple features that can account for correlated data. They showed the differences

between a deterministic and robust scheduling solution for an air separation plant. Vujanic et al ., 2012 [91] tackled the problem of flexible electricity consumption using robust optimization. They applied their proposed method on scheduling of cement milling machines. Amaran et al., 2016 [92]addressed the problem of maintenance planning for integrated chemical sites under uncertainty of maintenance duration. A combined robust optimization and stochastic programming approach is proposed in their study.

Studies that addressed robust optimization under endogenous uncertainty are rare in the literature. Vayanos et al.,2011 [57] approximated the true solution by partitioning the uncertainty set and in each partition, the binary variables are constant and continuous variables are affine functions of uncertain parameters. They derived the non-anticipativity constraints for decision dependent uncertain parameters for their partitioning method. Lappas and Gounaris., 2016 [60] proposed a adjustable robust optimization approach for process scheduling under uncertainty. They showed that adjusting decisions to past parameter realizations leads to significant improvements in the objective. A decision-dependent uncertainty set is to model the endogenous uncertainty inherent in process scheduling is also formulated in their study. Nohadani and Sharma, 2018 [93] proposed a decision dependent uncertainty set whose size depends on binary decisions and reduces conservatism of robust optimization approaches. They discussed a shortest path problem where the uncertain arc lengths are affected by decisions. Vayanos et al., 2019 [94] proposed a method inspired by k-adaptability approximation where k candidate strategies for each decision stage are chosen here-and-now and at the beginning of each period, the best strategy is selected after the uncertain parameters are revealed. They generalized their approach to minimize piecewise linear convex functions.

Studies that address both endogenous and exogenous uncertainty are limited in the literature. In the context of stochastic programming, two notable studies deal with this type of problems. Goel and Grossmann, 2006 [20] presented a mixed integer programming formulation and applied Lagrangean duality based branch and bound algorithm to reduce to number of non-anticipativity constraints. Apap and Grossmann, 2017 [19] proposed a composite scenario tree that captures both types of uncertainty and discussed two solution approaches to reduce the number of non-anticipativity constraints and solve the problem. A sequential scenario decomposition heuristic and a Lagrangean decomposition method. Dupačová (2006) [42] provided only a general description of problems under both types of uncertainty but did not present a solution framework or numerical results.

This study presents a novel mathematical framework that addresses both exogenous and endogenous uncertainty in a multistage mixed integer optimization problem structure. The partitioning method is employed to model the endogenous uncertainty and the lifting method is used to model the exogenous uncertainty. In this framework, both binary and continuous variables

are adaptive with respect to endogenous and exogenous uncertain parameters and the problem is solved based on robust optimization technique.

The rest of the paper is organized as follows. Section 5.2 presents the general problem formulation and states the required assumptions. Sections 5.3 explains the partitioning method for enogenous uncertainty and presents a simple problem structure with only binary variables under endogenous uncertainty. Section 5.4 describes problem structures that include continuous and binary variables under endogenous uncertainty. Section 5.5 explains the lifting method for exogenous uncertainty for binary and continuous variables and describes a problem structre where continuous variables are functions of both endogenous and exogenous uncertainty and binary variables are only a function of endogenous uncertainty. Section 5.6 presents the most comprehensive problem structure where both binary and contiuous variables are functions of both types of uncertainty. Section 5.7 illustrates a numerically demanding case study about infastructe and production planning of a gas field and section 5.8 concludes the paper.

5.2 Problem statement

The general formulation of a multistage stochastic problem addressed in this study is presented below.

$$\min_{x,y,z} \quad \mathbb{E}\Big(\sum_t c_t^\top(\boldsymbol{\eta})x_t(\boldsymbol{\xi},\boldsymbol{\eta}) + d_t^\top(\boldsymbol{\eta})y_t(\boldsymbol{\xi},\boldsymbol{\eta}) + h_t^\top(\boldsymbol{\eta})z_t(\boldsymbol{\xi})\Big) \tag{5.1a}$$

$$s.t. \quad \sum_{\tau=1}^{t} A_{t,\tau}^\top x_\tau(\boldsymbol{\xi},\boldsymbol{\eta}) + \sum_{\tau=1}^{t} B_{t,\tau}^\top y_\tau(\boldsymbol{\xi},\boldsymbol{\eta}) + \sum_{\tau=1}^{t} M_{t,\tau}^\top z_\tau(\boldsymbol{\xi}) \le e_t(\boldsymbol{\xi}) + f_t(\boldsymbol{\eta}) \;\; \forall t, \xi \in \Xi, \eta \in \Omega \tag{5.1b}$$

$$x_t(\boldsymbol{\xi},\boldsymbol{\eta}) = x_t(z_{t-1} \circ \boldsymbol{\xi}, \boldsymbol{\eta}_{[t]}) \qquad \forall t, \xi \in \Xi, \eta \in \Omega \tag{5.1c}$$

$$y_t(\boldsymbol{\xi},\boldsymbol{\eta}) = y_t(z_{t-1} \circ \boldsymbol{\xi}, \boldsymbol{\eta}_{[t]}) \qquad \forall t, \xi \in \Xi, \eta \in \Omega \tag{5.1d}$$

$$z_t(\boldsymbol{\xi}) = z_t(z_{t-1} \circ \boldsymbol{\xi}) \qquad \forall t, \xi \in \Xi \tag{5.1e}$$

$$z_t(\boldsymbol{\xi}) \in \{0,1\}, y_t(\boldsymbol{\xi},\boldsymbol{\eta}) \in \{0,1\}, x_t(\boldsymbol{\xi},\boldsymbol{\eta}) \in \mathbb{R} \qquad \forall t, \xi \in \Xi, \eta \in \Omega \tag{5.1f}$$

where ξ, η are the vectors of primitive endogenous and exogenous uncertain parameter, respectively. $x_t(\boldsymbol{\xi},\boldsymbol{\eta}), y_t(\boldsymbol{\xi},\boldsymbol{\eta})$ are continuous and binary variables and $z_t(\boldsymbol{\xi})$ is the indicator binary variable that indicates the revelation of endogenous uncertainty. Notice that we assume scalar variables x, y, z here but the variables can be easily extended to vector case. The notation $\circ$ indicates the Hadamard product which is the element-wise product. The product $z_{t-1} \circ \boldsymbol{\xi}$ means revealed endogenous parameters up to time step $t-1$ and $\boldsymbol{\eta}_{[t]}$ is the vector of exogenous uncertain parameters up to time step t, $\boldsymbol{\eta}_{[t]} = [\eta_1, \cdots, \eta_t]$. The objective coefficients $c_t(\boldsymbol{\eta})$, $d_t(\boldsymbol{\eta})$, $h_t(\boldsymbol{\eta})$, and constraint right hand side parameters $e_t(\boldsymbol{\xi}), f_t(\boldsymbol{\eta})$ are assumed to be linearly dependent on the primitive

uncertainty.

$$c_t(\boldsymbol{\eta}) = C_t\boldsymbol{\eta}_{[t]}, d_t(\boldsymbol{\eta}) = D_t\boldsymbol{\eta}_{[t]}, h_t(\boldsymbol{\eta}) = H_t\boldsymbol{\eta}_{[t]}, f_t(\boldsymbol{\eta}) = F_t\boldsymbol{\eta}_{[t]}, e_t(\boldsymbol{\xi}) = E_t\boldsymbol{\xi}$$

Constraints coefficients $A_{t,\tau}$, $B_{t,\tau}$ and $M_{t,\tau}$ are assumed to be independent of uncertainty. Ξ and Ω are the uncertainty set for endogenous and exogenous uncertainty, respectively. In this paper, they are both assumed to take polyhedral representation and both types of uncertainty are suumed to follow uniform distribution. To address this problem, we start by a special case and gradually increase the problem complexity.

5.3 Binary variables under endogenous uncertainty

We first consider the following problem with only endogenous uncertainty and binary variables indicating the revealing status

$$\min_{z} \quad \mathbb{E}\big(\sum_{t} c_t z_t(\boldsymbol{\xi})\big) \tag{5.2a}$$

$$s.t. \quad \sum_{\tau=1}^{t} A_{t,\tau}^{\top} z_\tau(\boldsymbol{\xi}) \leq E_t\boldsymbol{\xi} \qquad \forall t, \xi \in \Xi \tag{5.2b}$$

$$z_t(\boldsymbol{\xi}) = z_t(z_{t-1} \circ \boldsymbol{\xi}) \qquad \forall t, \xi \in \Xi \tag{5.2c}$$

$$z_t(\boldsymbol{\xi}) \in \{0,1\}^k \qquad \forall t, \xi \in \Xi \tag{5.2d}$$

where ξ is the vector of primitive endogenous uncertain parameters. $z_t(\boldsymbol{\xi})$ is the indicator binary variable that indicates the revelation of endogenous uncertainty.

5.3.1 Uncertainty set partitioning

In this method, the interval for each endogenous uncertain parameter is divided into segments. The overall uncertainty set is partitioned into scenarios where each scenario is a hyper-rectangle of the following form:

$$\Xi_s = \{\boldsymbol{\xi} \in \Xi : v^i_{s_{i-1}} \leq \xi_i < v^i_{s_i}, i = 1, ..., k\}$$

where $s \in \mathbb{S} = \bigtimes_{i=1}^{k}\{1, ..., r_i\} \subseteq \mathbb{N}^k$. $v^i_0 < v^i_1 < v^i_2 < \cdots < v^i_{r_i-1} < v^i_{r_i}$ represents $r_i - 1$ breakpoints along the ξ_i axis and v^i_0, $v^i_{r_i}$ indicate the initial and the end parameter values for ξ_i. For simplicity in the subsequent derivation, we re-write the above uncertainty set into compact form as

$$\Xi_s = \{\boldsymbol{\xi} : W_s\xi \geq V_s\}$$

5.3.2 Solution method

Using the above partitioning method, the binary decisions are approximated by piecewise constant decision rules of the following form, where $z_{t,s} \in \{0,1\}^k$:

$$z_t(\xi) = z_{t,s} \qquad \forall t \in \mathbb{T}, \xi \in \Xi_s, s \in \mathbb{S} \tag{5.3}$$

At each time step, the continuous and binary variables can only depend on uncertain parameters revealed up to the previous time step. This feature is referred to as non-anticipativity in the literature. The revelation of endogenous uncertainty is indicated by the binary variable $z(\xi)$. To simplify the notation, the domain of the following parameters is removed: $t \in \mathbb{T}, s, s' \in \mathbb{S}, j, j' \in \{1, ..., k\}$. The non-anticipativity for endogenous uncertainty can be expressed in the following form:

$$|z_{t,j',s} - z_{t,j',s'}| \leq z_{t-1,j,s} \qquad \forall j, j', t, s, s' : s_{-j} = s'_{-j} \tag{5.4}$$

where $s_{-j} = s'_{-j}$ means all the components in the scenarios$s$ and s' are the same, except the j-th component.

With the above approximation to the adaptive binary variables, we obtain a robust optimization formulation. Furthermore, we use duality based robust linear optimization and obtain the following deterministic counterpart optimization model

$$\min_{z,\theta} \quad \sum_s p_s \sum_t \sum_j c_t z_{t,j,s} \tag{5.5a}$$

$$s.t. \quad V_s^\top \theta_{t,s} \geq \sum_\tau^t A_{t,\tau}^\top z_{\tau,s} \qquad \forall t, s \tag{5.5b}$$

$$W_s^\top \theta_{t,s} = E_t \qquad \forall t, s \tag{5.5c}$$

$$\theta_{t,s} \geq 0 \qquad \forall t, s \tag{5.5d}$$

$$|z_{t,j',s} - z_{t,j',s'}| \leq z_{t-1,j,s} \qquad \forall j, j', t, s, s' : s_{-j} = s'_{-j} \tag{5.5e}$$

$$z_{t,j,s} \in \{0,1\} \qquad \forall j, t, s \tag{5.5f}$$

5.3.3 Example

In this section, a numeric example that includes binary variables under endogenous uncertainty is presented.

$$\max \quad \mathbb{E}_\xi(z_{1,1}(\xi) + z_{1,2}(\xi) + z_{2,1}(\xi) + z_{2,2}(\xi)) \tag{5.6a}$$

$$s.t. \quad z_{1,1}(\xi) + z_{1,2}(\xi) \leq D_1 \xi \qquad \forall \xi \in \Xi \tag{5.6b}$$

$$3z_{2,1}(\xi) + 2z_{2,2}(\xi) \leq D_2 \xi \qquad \forall \xi \in \Xi \tag{5.6c}$$

$$z_{2,j}(\xi) \geq z_{1,j}(\xi) \qquad \forall j, \xi \in \Xi \tag{5.6d}$$

$$z_{2,j}(\xi) = z_{2,j}(z_1(\xi) \circ \xi) \qquad \forall j, \xi \in \Xi \tag{5.6e}$$

$$z_{1,j}(\xi) = z_{1,j} \qquad \forall j, \xi \in \Xi \tag{5.6f}$$

$$z_{t,j}(\xi) \in \{0,1\} \qquad \forall t, j, \xi \in \Xi \tag{5.6g}$$

where ξ_j is the j-th endogenous uncertain parameter which is independent of time. $z_{t,j}$ is the binary variable that indicates the revelation of j-th uncertain parameter at time step t. In this illustration problem, $t = 1, 2$ and $j = 1, 2$. ξ is defined as $\xi = [1, \xi_1, \xi_2]$. Constraint 5.6d means once an uncertain parameter is revealed, it remains revealed at the next time period. Equation 5.6e is the non-anticipativity constraint which means the binary variables at the second time step can only be a function of revealed uncertainty parameters at the first time step. Constraints 5.6f indicates that at the first time step, the binary variables are non-adaptive with respect to uncertain parameters since at the first step no uncertain parameter is revealed yet. Based on the uncertainty set partitioning method, the following robust optimization problem is obtained and its deterministic counterpart can be derived accordingly:

$$\max \quad \sum_s p_s(z_{1,1,s} + z_{1,2,s} + z_{2,1,s} + z_{2,2,s}) \tag{5.7a}$$

$$s.t. \quad z_{1,1,s} + z_{1,2,s} \leq D_1\xi \qquad \forall s, \xi \in \Xi^s \tag{5.7b}$$

$$3z_{2,1,s} + 2z_{2,2,s} \leq D_2\xi \qquad \forall s, \xi \in \Xi^s \tag{5.7c}$$

$$z_{2,j,s} \geq z_{1,j,s} \qquad \forall j, s \tag{5.7d}$$

$$z_{1,j,s} = z_{1,j,s'} \qquad \forall j, s, s' \tag{5.7e}$$

$$\mid z_{2,j,s} - z_{2,j',s'} \mid \leq z_{1,j,s} \qquad \forall j, j', s, s' : s_{-j} = s'_{-j} \tag{5.7f}$$

$$z_{t,j,s} \in \{0,1\} \qquad \forall j, t, s \tag{5.7g}$$

where $D_1 = [1, 1, 3]$, $D_2 = [2, 2, 1]$. Breakpoint locations for different number of partitions is provided in the following table:

Table 5.1: Breakpoint locations

Number of partitions	Number of scenarios	$\xi_1 \in [0, 3]$	$\xi_2 \in [0, 6]$
2	4	1.8	3.6
5	25	0.6, 1.2, 1.8, 2.4	1.2, 2.4, 3.6, 4.8
10	100	0.3 : 0.3 : 2.7	0.6 : 0.6 : 5.4

Table 5.2: Solution statistics for different number of partitions

Number of partitions	Objective	Binary, Continuous Variables	Run time
2	2.4	16, 49	0.235 s
5	2.4	100, 301	0.236 s
10	2.5	400, 1201	1.339 s

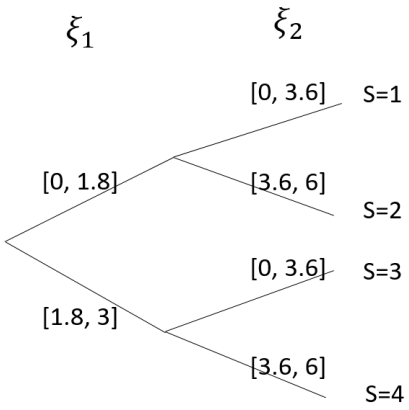


Figure 5.1: Scenarios 1 to 4 when each uncertain parameter is segmented into 2 partitions using 1 breakpoint

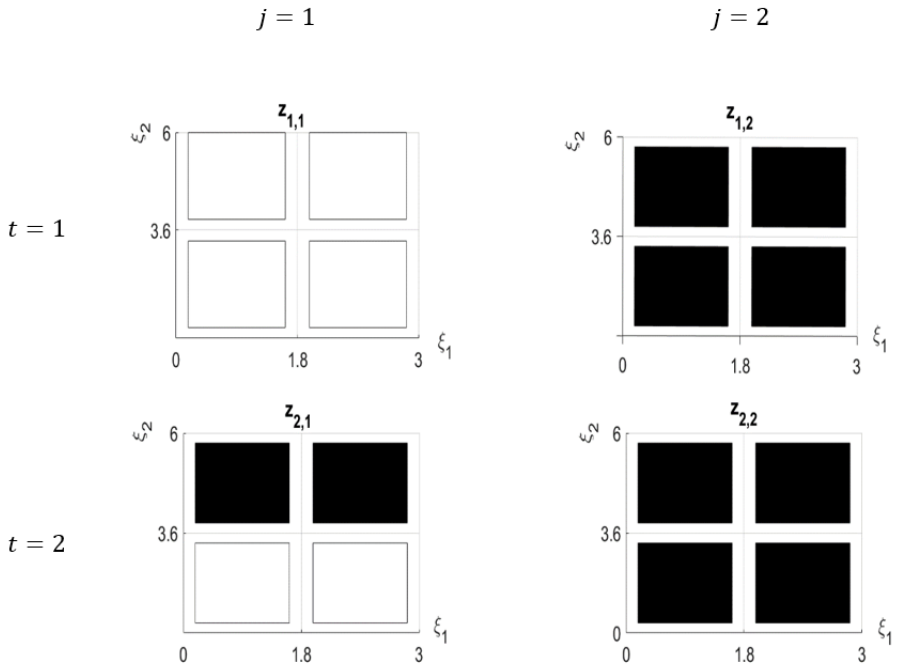


Figure 5.2: $z_{t,j}$ binary solution at 4 scenarios

Figure 5.1 illustrates how scenarios are generated when the range of each uncertain parameter

Table 5.3: Binary solution for 4 scenarios $z(\xi) = z_{t,j,s}$

t	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2
j	1	1	1	1	2	2	2	2	1	1	1	1	2	2	2	2
s	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
$z_{t,j,s}$	0	0	0	0	1	1	1	1	0	1	0	1	1	1	1	1

is divided into 2 partitions. Figure 5.2 and table 5.3 show that at $t = 1$, ξ_1 is not revealed ($z_{1,1} = 0$) but ξ_2 is revealed ($z_{1,2} = 1$). At $t = 2$, $z_{2,1}$ is adaptive with respect to the revealed parameter (ξ_2) and non-adaptive with respect to the unrevealed parameter (ξ_1). Since the ξ_2 is revealed at $t = 1$, it remains revealed at $t = 2$ which means $z_{2,2} = 1$. Table 5.2 shows when the number of partitions increase, the objective improves. It means that increasing the number of partitions results in a more refined solution space and subsequently a better objective value.

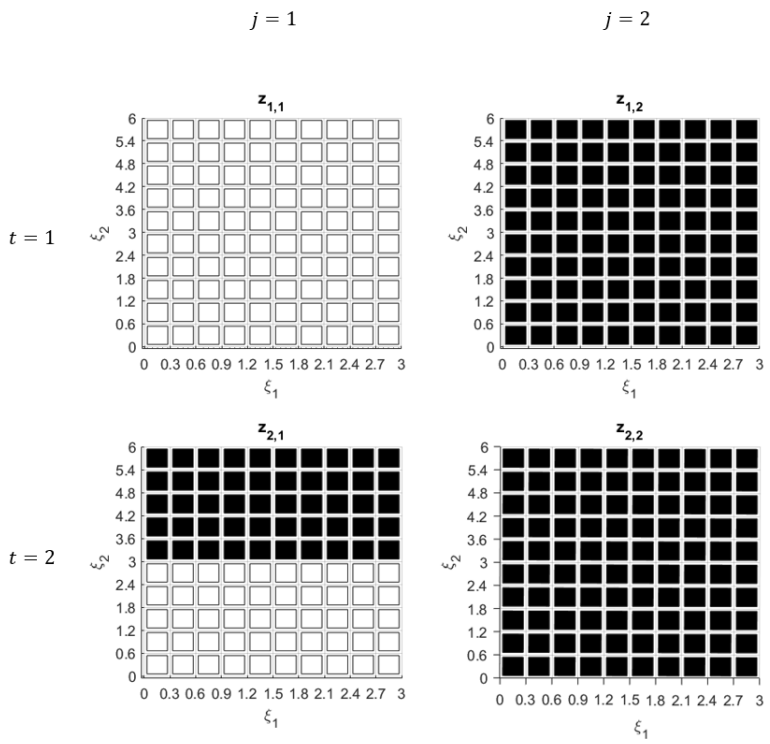


Figure 5.3: Binary solution for 100 scenarios

The explanation for the solution obtained from 10 partitions for each uncertain parameter (100 scenarios) is similar to the solution obtained from 2 partitions (4 scenarios).

5.4 Continuous and binary variables under endogenous uncertainty

In this section, we consider problems that include binary and continuous variables under endogenous uncertainty

$$\min_{x,z} \quad \mathbb{E}\big(\sum_t c_t x_t(\boldsymbol{\xi}) + d_t z_t(\boldsymbol{\xi})\big) \tag{5.8a}$$

$$s.t. \quad \sum_{\tau=1}^{t} A_{t,\tau}^{\top} x_\tau(\boldsymbol{\xi}) + \sum_{\tau=1}^{t} B_{t,\tau}^{\top} z_\tau(\boldsymbol{\xi}) \leq E_t \boldsymbol{\xi} \quad \forall t, \xi \in \Xi \tag{5.8b}$$

$$x_t(\boldsymbol{\xi}) = x_t(z_{t-1} \circ \boldsymbol{\xi}) \quad \forall t, \xi \in \Xi \tag{5.8c}$$

$$z_t(\boldsymbol{\xi}) = z_t(z_{t-1} \circ \boldsymbol{\xi}) \quad \forall t, \xi \in \Xi \tag{5.8d}$$

$$z_t(\boldsymbol{\xi}) \in \{0,1\}, x_t(\boldsymbol{\xi}) \in \mathbb{R} \quad \forall t, \xi \in \Xi \tag{5.8e}$$

where $\boldsymbol{\xi}$ is the vector of the primitive endogenous uncertain parameters. $x_t(\boldsymbol{\xi})$ is the continuous variable and $z_t(\boldsymbol{\xi})$ is the indicator binary variable that indicates the revelation of endogenous uncertainty.

5.4.1 Solution method

Under the same uncertainty set partitioning method, the binary and continuous decisions are approximated by piecewise constant and piecewise linear decision rules of the following form:

$$z_t(\xi) = z_{t,s} \qquad \forall \xi \in \Xi_s, s \in \mathbb{S} \tag{5.9}$$

$$x_t(\xi) = X_{t,s}\boldsymbol{\xi} \qquad \forall \xi \in \Xi_s, s \in \mathbb{S} \tag{5.10}$$

where $z_{t,s} \in \{0,1\}^k$, $X_{t,s} \in \mathbb{R}^{k+1}$. The non-anticipativity for endogenous uncertainty can be expressed in the following form:

$$|z_{t,j',s} - z_{t,j',s'}| \leq z_{t-1,j,s} \qquad \forall j, j', t, s, s' : s_{-j} = s'_{-j} \tag{5.11}$$

$$|X_{t,s,j'} - X_{t,s',j'}| \leq M z_{t-1,j,s} \qquad \forall j, j', s, s' : s_{-j} = s'_{-j} \tag{5.12}$$

$$|X_{t,s,j}| \leq M z_{t-1,j,s} \qquad \forall j, t, s \tag{5.13}$$

where $s_{-j} = s'_{-j}$ means all the components in the scenarios s and s' are the same, except the j-th component. M is a sufficiently large constant. With the above constraints, the deterministic counterpart formulation is then given as following

$$\min_{X,z,\theta} \quad \sum_s p_s \sum_t c_t X_{\tau,s} \mathbb{E}_{\Xi_s}(\xi) + d_t z_{t,s} \tag{5.14a}$$

$$
\begin{aligned}
s.t. \quad & V_s^\top \theta_{t,s} \geq \sum_{\tau}^{t} B_{t,\tau}^\top z_{\tau,s} && \forall t, s && (5.14b)\\
& W_s^\top \theta_{t,s} = \sum_{\tau}^{t} A_{t,\tau}^\top X_{\tau,s} - E_t && \forall t, s && (5.14c)\\
& \theta_{t,s} \geq 0 && \forall t, s && (5.14d)\\
& |z_{t,j',s} - z_{t,j',s'}| \leq z_{t-1,j,s} && \forall j, j', t, s, s' : s_{-j} = s'_{-j} && (5.14e)\\
& |X_{t,s,j'} - X_{t,s',j'}| \leq M z_{t-1,j,s} && \forall j, j', s, s' : s_{-j} = s'_{-j} && (5.14f)\\
& |X_{t,s,j}| \leq M z_{t-1,j,s} && \forall j, t, s && (5.14g)\\
& z_{t,j,s} \in \{0,1\} && \forall j, t, s && (5.14h)
\end{aligned}
$$

5.4.2 Example

Consider a numeric example given below

$$
\begin{aligned}
\max \quad & \mathbb{E}_\xi(z_{1,1}(\xi) + z_{1,2}(\xi) + z_{2,1}(\xi) + z_{2,2}(\xi) + x_2(\xi) + x_3(\xi)) && && (5.15a)\\
s.t. \quad & z_{1,1}(\xi) + z_{1,2}(\xi) \leq D_1 \boldsymbol{\xi} && \forall \xi \in \Xi && (5.15b)\\
& 3z_{2,1}(\xi) + 2z_{2,2}(\xi) \leq D_2 \boldsymbol{\xi} && \forall \xi \in \Xi && (5.15c)\\
& x_1(\xi) = 0 && \forall \xi \in \Xi && (5.15d)\\
& x_2(\xi) \leq D_3 \boldsymbol{\xi} z_{2,1} && \forall \xi \in \Xi && (5.15e)\\
& x_3(\xi) \leq D_4 \boldsymbol{\xi} && \forall \xi \in \Xi && (5.15f)\\
& z_{2,j}(\xi) \geq z_{1,j}(\xi) && \forall j, \xi \in \Xi && (5.15g)\\
& z_{1,j}(\xi) = z_{1,j} && \forall j, \xi \in \Xi && (5.15h)\\
& z_{2,j}(\xi) = z_{2,j}(z_1(\xi) \circ \boldsymbol{\xi}) && \forall j, \xi \in \Xi && (5.15i)\\
& x_2(\xi) = x_2(z_1(\xi) \circ \boldsymbol{\xi}) && \forall \xi \in \Xi && (5.15j)\\
& x_3(\xi) = x_3(z_2(\xi) \circ \boldsymbol{\xi}) && \forall \xi \in \Xi && (5.15k)\\
& z_{t,j}(\xi) \in \{0,1\} && \forall j, t, \xi && (5.15l)
\end{aligned}
$$

In this model, the only new type of constraints are the non-anticipativity constraints 5.15j and 5.15k. There constraints enforce that at each time step, the continuous variable can only be a function of revealed uncertainties up to the previous time step.

$$
\begin{aligned}
\max \quad & \sum_s p_s(z_{1,1,s} + z_{1,2,s} + z_{2,1,s} + z_{2,2,s} + X_{2,s}\mathbb{E}_{\Xi_s}(\xi) + X_{3,s}\mathbb{E}_{\Xi_s}(\xi)) && && (5.16a)\\
s.t. \quad & z_{1,1,s} + z_{1,2,s} \leq D_1 \boldsymbol{\xi} && \forall s, \xi \in \Xi^s && (5.16b)\\
& 3z_{2,1,s} + 2z_{2,2,s} \leq D_2 \boldsymbol{\xi} && \forall s, \xi \in \Xi^s && (5.16c)\\
& x_1 = 0 && && (5.16d)
\end{aligned}
$$

$$X_{2,s}\xi \leq D_3\boldsymbol{\xi} z_{2,1,s} \qquad \forall s, \xi \in \Xi^s \tag{5.16e}$$

$$X_{3,s}\xi \leq D_4\boldsymbol{\xi} \qquad \forall s, \xi \in \Xi^s \tag{5.16f}$$

$$\mid z_{t+1,j',s} - z_{t+1,j',s'} \mid \leq z_{t,j,s} \qquad \forall t = 1, s, s', j, j' : s_{-j} = s'_{-j} \tag{5.16g}$$

$$\mid X_{t+1,j',s} - X_{t+1,j',s'} \mid \leq M z_{t,j,s} \qquad \forall t, s, s', j, j' : s_{-j} = s'_{-j} \tag{5.16h}$$

$$\mid X_{t+1,j,s} \mid \leq M z_{t,j,s} \qquad \forall t, j, s \tag{5.16i}$$

$$z_{t+1,j,s} \geq z_{t,j,s} \qquad \forall j, s, t = 1 \tag{5.16j}$$

$$z_{1,j,s} = z_{1,j,s'} \qquad \forall j, s, s' \tag{5.16k}$$

$$z_{t,j,s} \in \{0, 1\} \qquad \forall j, t, s \tag{5.16l}$$

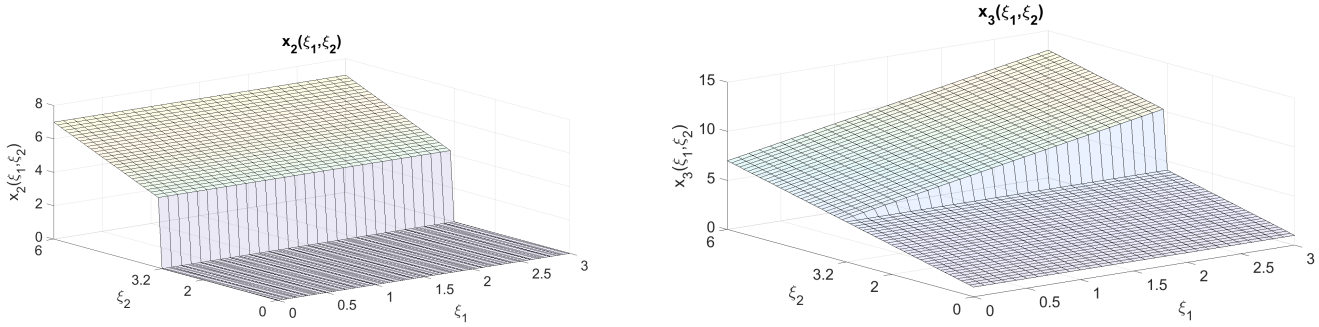
where $t = 1, 2$, $\xi = [1, \xi_1, \xi_2]$, $D_1 = [1, 1, 3]$, $D_2 = [2, 2, 1]$, $D_3 = [1, 3, 1]$ and $D_4 = [1, 2, 1]$. Breakpoint locations are provided in table 5.4. Constraints 5.16g and 5.16h stipulates the non-anticipativity for binary and continuous variables across different scenarios. Constraint 5.16i enforces the non-anticipativity for continuous variables within each scenario.

Table 5.4: Breakpoint locations

Number of partitions	Number of scenarios	$\xi_1 \in [0, 3]$	$\xi_2 \in [0, 6]$
2	4	1.8	3.6
5	25	0.6, 1.2, 1.8, 2.4	1.2, 2.4, 3.6, 4.8
15	225	0.1875 : 01875 : 2.8125	0.3750 : 0.3750 : 5.625

Table 5.5: Solution statistics for different number of partitions

Number of partitions	Objective	Binary, Continuous variables	Run time
2	9.92	16, 122	0.221 s
5	9.92	100, 752	0.489 s
15	10.48	900, 6752	49.932 s



(a) $x_2(\xi_1, \xi_2)$ (b) $x_3(\xi_1, \xi_2)$

Figure 5.6: Continuous solution for 225 scenarios

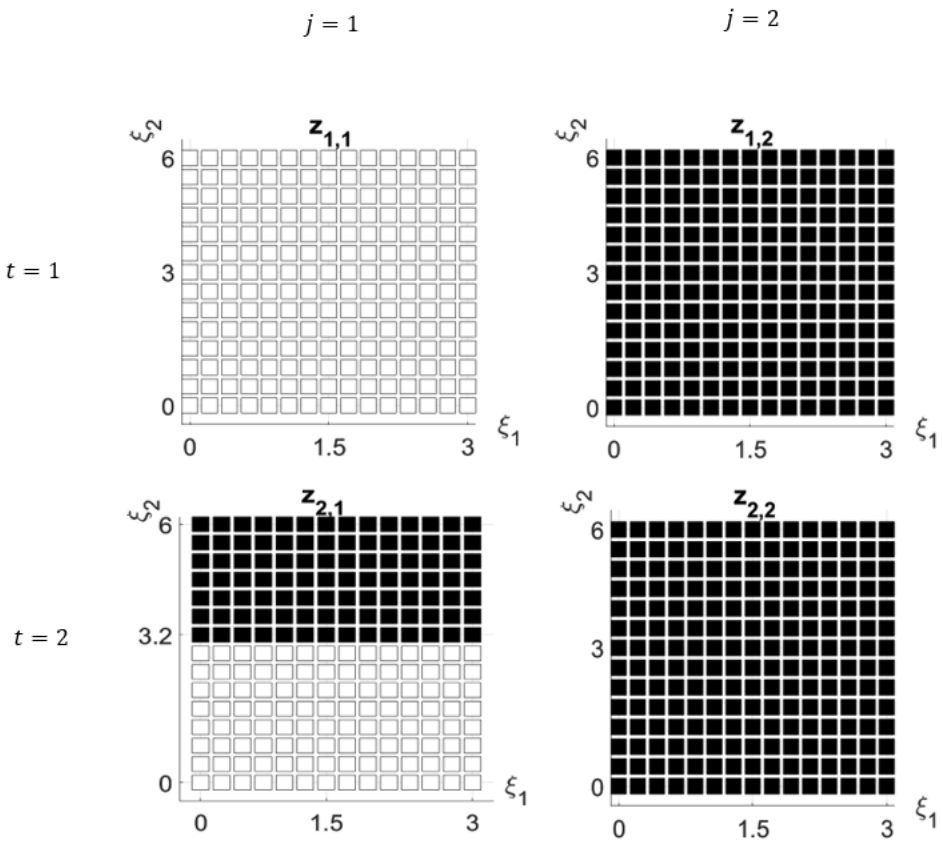
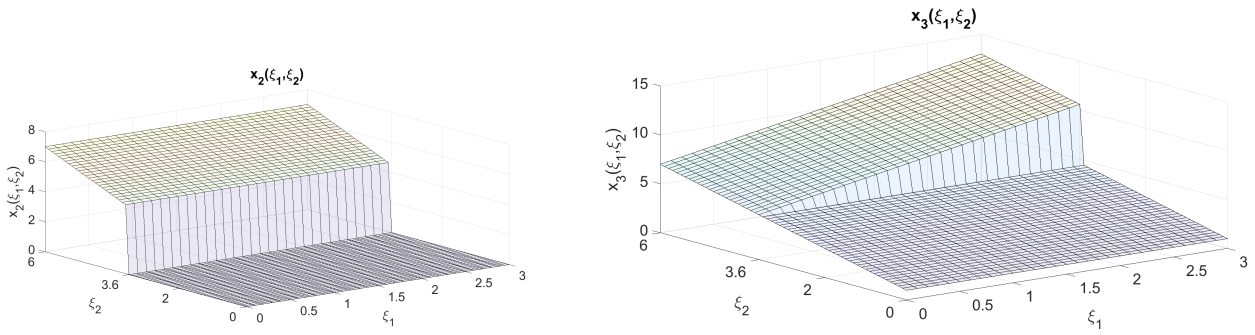


Figure 5.4: $z_{t,j}$ binary solution at 225 scenarios

Table 5.5 shows when the number of partitions increase, the objective improves. Figure 5.4 illustrates the binary solution for 225 scenarios (the binary solution for 4 scenarios is the same as the solution shown in Figure 5.2). Figures 5.5 and 5.6 show that the continuous solution is piecewise discontinuous. Using 225 scenarios, the break point location has shifted from 3.6 to 3.2 and the objective has improved, but the solution shows similar trend. Figures 5.5a and 5.6a show that at $t = 1$, the continuous solution is only a function of revealed uncertainty ξ_2 and constant with respect to unrevealed uncertainty ξ_1. Figures 5.5b and 5.6b show that at $t = 2$, the continuous variable changes with respect to both uncertain parameters since at $t = 2$, both uncertain parameters are revealed at some scenarios.



(a) $x_2(\xi_1, \xi_2)$ (b) $x_3(\xi_1, \xi_2)$

Figure 5.5: Continuous solution for 4 scenarios

5.5 Continuous variables under exogenous and endogenous uncertainty and indicator binary variables under endogenous uncertainty

This section studies problems that involve continuous variables under endogenous and exogenous uncertainty and binary variables under only endogenous uncertainty.

$$\min_{x,z} \quad \mathbb{E}\left(\sum_t c_t^\top(\boldsymbol{\eta}) x_t(\boldsymbol{\xi}, \boldsymbol{\eta}) + d_t^\top(\boldsymbol{\eta}) z_t(\boldsymbol{\xi})\right) \tag{5.17a}$$

$$s.t. \quad \sum_{\tau=1}^{t} A_{t,\tau}^\top x_\tau(\boldsymbol{\xi}, \boldsymbol{\eta}) + \sum_{\tau=1}^{t} B_{t,\tau}^\top z_\tau(\boldsymbol{\xi}) \leq e_t(\boldsymbol{\xi}) + f_t(\boldsymbol{\eta}) \qquad \forall t, \xi \in \Xi, \eta \in \Omega \tag{5.17b}$$

$$x_t(\boldsymbol{\xi}, \boldsymbol{\eta}) = x_t(z_{t-1} \circ \boldsymbol{\xi}, \boldsymbol{\eta}_{[t]}) \qquad \forall t, \xi \in \Xi, \eta \in \Omega \tag{5.17c}$$

$$z_t(\boldsymbol{\xi}) = z_t(z_{t-1} \circ \boldsymbol{\xi}) \qquad \forall t, \xi \in \Xi \tag{5.17d}$$

$$z_t(\boldsymbol{\xi}) \in \{0, 1\}, x_t(\boldsymbol{\xi}, \boldsymbol{\eta}) \in \mathbb{R} \qquad \forall t, \xi \in \Xi, \eta \in \Omega \tag{5.17e}$$

where ξ, η are the primitive endogenous and exogenous uncertain parameter vectors, respectively. $x_t(\boldsymbol{\xi}, \boldsymbol{\eta})$ are continuous variables and $z_t(\boldsymbol{\xi})$ is the indicator binary variable that indicates the revelation of endogenous uncertainty. $\boldsymbol{\eta}_{[t]}$ is the vector of exogenous uncertain parameters up to time step t, $\boldsymbol{\eta}_{[t]} = [\eta_1, \cdots, \eta_t]$. The objective coefficients $c_t(\boldsymbol{\eta})$, $d_t(\boldsymbol{\eta})$, and constraint right hand side coefficients $e_t(\boldsymbol{\xi})$, $f_t(\boldsymbol{\eta})$ are assumed to be linearly dependent on the primitive uncertainty.

$$c_t(\boldsymbol{\eta}) = C_t \boldsymbol{\eta}_{[t]}, d_t(\boldsymbol{\eta}) = D_t \boldsymbol{\eta}_{[t]}, f_t(\boldsymbol{\eta}) = F_t \boldsymbol{\eta}_{[t]}, e_t(\boldsymbol{\xi}) = E_t \boldsymbol{\xi}$$

Partitioning method is used to model endogenous uncertainty and lifting method is used for exogenous uncertainty. Since in the following problems, lifting method is used to formulate adaptive continous and adaptive binary variables under exogenous uncertaitny, it is explained in detail in

the following section and illustrated in the Appendix.

5.5.1 Piecewise linear lifting of exogenous uncertainty

Georghiou et al., 2015, [29] proposed the idea of adaptive continuous variables based on linear decision rule. In this method, the primitive uncertain parameter is mapped to a higher dimension lifted set based on predefined breakpoints and lifted parameters . Assume η_i is the i-th uncertain parameter defined within $\eta_i \in [l_i, u_i]$ and $r_i - 1$ is the number of breakpoints; the lifted parameters are defined as:

$$G_{i,j} = \begin{cases} \eta_i & r_i = 1 \\ \min\{\eta_i, v_1^i\} & r_i \geq 2,\ j = 1 \\ \max\{\min\{\eta_i, v_j^i\} - v_{j-1}^i, 0\} & r_i \geq 2,\ j = 2, ..., r_i - 1 \\ \max\{\eta_i - v_{j-1}^i, 0\} & r_i \geq 2,\ j = r_i \end{cases} \tag{5.18}$$

If $r_i = 1$, it means there is no breakpoint for the i-th uncertain parameter and the lifting operator is reduced to a identity mapping. Notice that for $r_i - 1$ breakpoints, the number of lifted uncertain parameters is r_i. The original uncertain parameter can be retrieved by $\eta_i = \sum_{j=1}^{r_i} G_{i,j}$. The locations of breakpoints for η_i is defined as: $v_0^i < v_1^i < .... < v_{ri}^i$, where $v_0^i = l_i$ and $v_{ri}^i = u_i$. The nonconvex set Ω_i^{liftC} is a union of r_i connected finite line segments. Its extreme points are $\psi_0 = (l_i, 0, \cdots, 0)^\top$, $\psi_1 = (v_1^i, 0, \cdots, 0)^\top$, $\psi_2 = (v_1^i, v_2^i - v_1^i, 0, \cdots, 0)^\top$,$\cdots$, $\psi_{r_i} = (v_1^i, v_2^i - v_1^i, \cdots, v_{r_i-1}^i - v_{r_i-2}^i, u_i - v_{r_i-1}^i)^\top \in \mathbb{R}^{r_i}$. Thus, the convex hull of Ω_i^{liftC} is generated by the simplex with vertices $\{\psi_j\}_{j=0}^{r_i}$: $\{\mathbf{G}_i \in \mathbb{R}^{r_i} : \mathbf{G}_i = \sum_{j=0}^{r_i} \lambda_j \psi_j, \sum_{j=0}^{r_i} \lambda_j = 1, \lambda_0, \cdots, \lambda_{r_i} \geq 0\}$. This can be rewritten as:

$$conv(\Omega_i^{liftC}) = \{\mathbf{G}_i \in \mathbb{R}^{r_i} : \exists \lambda \in R^{r_i+1}, W_i^c \lambda = (1, \mathbf{G}_i^\top)^\top, \lambda \geq 0\} \tag{5.19}$$

where $W_i^c = \begin{bmatrix} 1 & \cdots & 1 \\ \nu_0 & \cdots & \nu_{r_i} \end{bmatrix}$.

5.5.2 Piecewise binary lifting of exogenous uncertainty

Bertsimas and Georghiou, 2015 [39] originally developed the concept of adaptive binary variables based on linear decision rule. In this method, the primitive uncertain parameter is mapped to a higher dimensional uncertainty set using 0-1 indicator functions. For each breakpoint, an indicator function is defined. The space of the original uncertain parameter and the indicator functions generates a non-convex set that is convexified using its extreme points which results in an over-estimated convex polyhedron. The binary variable is a linear combination of 0-1 indicator functions. Using decision rule formulation results in semi-infinite constraints and subsequently duality theorem is applied to convert the semi-infinite constraints to their deterministic counter-

parts. While traditional scenario-tree methods suffers from curse of dimensionality, the decision rule method provides the scalability required to tackle large-scale problems. In this method, the trade-off between the computational burden and the solution quality can be adjusted by the number of breakpoints. Increasing the number of breakpoints results in a better solution quality but increased computational expense.

Assume that η_i belongs to the interval $\eta_i \in [l_i, u_i]$ and the breakpoints are denoted by $v_1^i < v_2^i < ... < v_{r_i-1}^i$. The indicator functions are defined based on the location of the breakpoints.

$$Q_{i,j}(\eta_i) = \begin{cases} 0, & \text{if} \quad \eta_i < v_j^i \\ 1, & \text{if} \quad \eta_i \geq v_j^i \end{cases} \tag{5.20}$$

Convex hull of lifted uncertainty set for a single uncertain parameter is formulated as:

$$conv(\Omega_i^{liftB}) = \{\eta_i' = (\eta_i, Q_i) : \exists\ \lambda_{k,p,i} > 0, \eta_i' = \sum_{p=1}^{r_i}\sum_{k=1}^{2} \lambda_{k,p,i}\psi_{k,p,i}, \sum_{p=1}^{r_i}\sum_{k=1}^{2} \lambda_{k,p,i} = 1\} \tag{5.21}$$

where $Q_i = (Q_{i,1}, \cdots, Q_{i,r_i-1})$ is the vector of indicator functions for uncertain parameter η_i, $\lambda_{k,p,i}$ is the coefficient of extreme points in the convex hull and $\psi_{1,p,i}$ and $\psi_{2,p,i}$ are extreme points of the p-th segment of the lifted non-convex set (Ω_i^{liftB}) related to parameter η_i:

$$\psi_{1,p,i} = [v_{p-1}^i, \underbrace{1, \cdots, 1}_{p-1 \text{ times}}, \underbrace{0, \cdots, 0}_{r_i - p \text{ times}}]^\top, \quad \psi_{2,p,i} = [v_p^i, \underbrace{1, \cdots, 1}_{p-1 \text{ times}}, \underbrace{0, \cdots, 0}_{r_i - p \text{ times}}]^\top \in \mathbb{R}^{r_i}$$

thus,

$$conv(\Omega_i^{liftB}) = \{\eta_i' = (\eta_i, Q_i) \in \mathbb{R}^{r_i} : \exists\ \lambda \in \mathbb{R}^{2r_i}, W_i^b\lambda = (1, {\eta_i'}^\top)^\top, \lambda > 0\} \tag{5.22}$$

where $W_i^b = \begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ \psi_{1,1,i} & \psi_{2,1,i} & \cdots & \psi_{1,r_i,i} & \psi_{2,r_i,i} \end{bmatrix} \in \mathbb{R}^{(r_i+1)\times(2r_i)}$.

5.5.3 Overall lifted exogenous uncertainty

In this study, both adaptive continuous and adaptive binary variables are considered in the problem formulation. Therefore, the uncertain parameter includes the lifting elements for both binary and continuous variables. In the following, the notation used for the overall uncertainty vector for all time steps is summarized here.

- $\eta_i' = [\eta_i, \underbrace{Q_{i,1}, \cdots, Q_{i,r_i-1}}_{\text{binary}}, \underbrace{G_{i,1}, \cdots, G_{i,r_i}}_{\text{continuous}}]^\top = [\eta_i, {Q_i}^\top, {G_i}^\top]^\top$
- $\eta_{[t]}'$ vector of overall uncertainty from stage 1 to t: $[1, {\eta_1'}^\top, \cdots, {\eta_t'}^\top]^\top$

- $\eta' \equiv \eta'_{[T]}$ vector of overall uncertainty from stage 1 to T: $[1, {\eta'_1}^\top, \cdots, {\eta'_T}^\top]^\top$

Assume the original uncertainty set is a general polyhedral set:

$$\Psi = \{\eta : J\eta \geq h\} \tag{5.23}$$

The convex overestimation for the overall lifted uncertainty set is defined as follows. It includes constraints from original uncertainty set, the convex hull for the continuous lifted set, the convex hull for the binary lifted set and the correlation between the lifted parameters.

$$\hat{\Psi}' = \{\eta' : J\eta \geq h, G_i \in conv(\Omega_i^{liftC}), (\eta_i, Q_i) \in conv(\Omega_i^{liftB}), \eta_i = 1^\top G_i, i = 1, \cdots, T\} \tag{5.24}$$

The convex overestimated set after projection to the space of η' can be compactly written as:

$$\hat{\Psi}' = \{\eta' : J'\eta' \geq h'\} \tag{5.25}$$

Projection/truncation matrices can be used to retrieve the specific uncertainty to be used in the decision rule approximation model for continuous and binary variables:

$$\eta_{[t]} = [\eta_1, \cdots, \eta_t] = P_{\eta_{[t]}}\boldsymbol{\eta'} \tag{5.26}$$

$$Q_{[t]} = [1, Q_1^\top, \cdots, Q_t^\top]^\top = P_{Q_{[t]}}\boldsymbol{\eta'} \tag{5.27}$$

$$G_{[t]} = [1, G_1^\top, \cdots, G_t^\top]^\top = P_{G_{[t]}}\eta' \tag{5.28}$$

5.5.4 Decision rule approximation

As an approximation policy, we set the continuous variables as function of the lifted exogenous uncertain parameter and the endogenous uncertainty:

$$x_t(\xi, \eta) = X_{t,s}\boldsymbol{\xi} + X'_{t,s}\boldsymbol{\eta}'_{[t]} \qquad \forall t, s, \xi \in \Xi_s, \eta' \in \hat{\Psi}' \tag{5.29}$$

Adaptive binary variables under endogenous uncertainty are approximated using the partitioning method explained in section 5.3.1.

$$z_{t,j}(\xi) = z_{t,j,s} \qquad \forall t, j, s \tag{5.30}$$

Using the above approximation methods in the original problem formulation will result in a semi-infinite problem since the parameter η' belongs to convex set $\hat{\Omega}'$ and can take infinite number of values. Through standard duality based robust counterpart constraint formulation, the determin-

istic counterpart optimization problem is formulated as

$$\begin{aligned}
\min \quad & \sum_s p_s \sum_t (\mathbb{E}(\boldsymbol{\eta}')^\top P_{\eta_{[t]}}^\top C_t^\top X_{t,s} \mathbb{E}_{\Xi_s}(\boldsymbol{\xi}) + \mathrm{Tr}(P_{\eta_{[t]}}^\top \ {C_t}^\top X'_{t,s} P_{G_{[t]}} \mathbb{E}[\eta' \boldsymbol{\eta}'^\top])) && \text{(5.31a)} \\
& + \sum_s p_s \sum_t \mathbb{E}(\boldsymbol{\eta}')^\top {P_{\eta_{[t]}}}^\top D_t^\top z_{t,s} \\
s.t. \quad & V_s \theta_{t,s} + V'_s \theta'_{t,s} \geq \sum_{\tau=1}^{t} {B_{t,\tau}}^\top z_{\tau,s} && \forall t, s && \text{(5.31b)} \\
& W_s^\top \theta_{t,s} = -\sum_{\tau=1}^{t} (A_{t,\tau}^\top X_{\tau,s}) + E_t && \forall t, s && \text{(5.31c)} \\
& {W'_s}^\top \theta'_{t,s} = -\sum_{\tau=1}^{t} (A_{t,\tau}^\top X'_{\tau,s} P_{G_{[t]}}) + F_t P_{\eta_{[t]}} && \forall t, s && \text{(5.31d)} \\
& \theta_{t,s} \geq 0, \theta'_{t,s} \geq 0 && \forall t, s && \text{(5.31e)} \\
& |z_{t,j',s} - z_{t,j',s'}| \leq z_{t-1,j,s} && \forall j, j', t, s, s' : s_{-j} = s'_{-j} && \text{(5.31f)} \\
& |X_{t,s,j'} X_{-t,s',j'}| \leq M z_{t-1,j,s} && \forall j, j', s, s' : s_{-j} = s'_{-j} && \text{(5.31g)} \\
& |X_{t,s,j}| \leq M z_{t-1,j,s} && \forall j, t, s && \text{(5.31h)} \\
& z_{t,j,s} \in \{0,1\} && \forall t, j, s && \text{(5.31i)}
\end{aligned}$$

5.5.5 Example

An illustrating example is studied in this section.

$$\begin{aligned}
\min \quad & -\mathbb{E}_{\xi,\eta}(z_{1,1}(\xi) + z_{1,2}(\xi) + z_{2,1}(\xi) + z_{2,2}(\xi)) + x_2(\xi,\eta) + x_3(\xi,\eta)) && \text{(5.32a)} \\
s.t. \quad & z_{1,1}(\xi) + z_{1,2}(\xi) \leq D_1 \boldsymbol{\xi} && \forall \xi \in \Xi && \text{(5.32b)} \\
& 3z_{2,1}(\xi) + 2z_{2,2}(\xi) \leq D_2 \boldsymbol{\xi} && \forall \xi \in \Xi \\
& z_{2,j}(\xi) \geq z_{1,j}(\xi) && \forall j, \xi \in \Xi && \text{(5.32c)} \\
& z_{1,j}(\xi) = z_{1,j} && \forall j, \xi \in \Xi && \text{(5.32d)} \\
& x_1 = 0 && && \text{(5.32e)} \\
& x_2(\xi,\eta) \geq D_3 \eta - D_1 \boldsymbol{\xi} && \forall \xi \in \Xi, \eta \in \Psi && \text{(5.32f)} \\
& x_3(\xi,\eta) \geq D_4 \eta - D_2 \boldsymbol{\xi} && \forall \xi \in \Xi, \eta \in \Psi && \text{(5.32g)} \\
& z_{2,j}(\xi) = z_{2,j}(\mathbf{z}_1(\xi) \circ \boldsymbol{\xi}) && \forall j, \xi \in \Xi && \text{(5.32h)} \\
& x_3(\xi,\eta) = x_3(\mathbf{z}_2(\xi) \circ \boldsymbol{\xi}, \eta_{[2]}) && \xi \in \Xi, \eta \in \Psi && \text{(5.32i)} \\
& x_2(\xi,\eta) = x_2(\mathbf{z}_1(\xi) \circ \boldsymbol{\xi}, \eta_{[1]}) && \xi \in \Xi, \eta \in \Psi && \text{(5.32j)} \\
& z_{t,j}(\xi) \in \{0,1\} && \forall t, j, \xi \in \Xi, \eta \in \Psi && \text{(5.32k)}
\end{aligned}$$

Using the approximation methods for continuous and binary variables, the following semi-infinite

problem formulation is obtained that can be converted to its robust deterministic counterpart using dulaity method.

$$\min \quad -\sum_s p_s(\sum_t \sum_j z_{t,j,s} + \sum_t X_{t+1,s} E_{\Xi_s}(\boldsymbol{\xi}) + \sum_t X'_{t+1,s} E(\boldsymbol{\eta}'_{[t]})) \tag{5.33a}$$

$$s.t. \quad z_{1,1,s} + z_{1,2,s} \leq D_1 \boldsymbol{\xi} \qquad \forall s, \xi \in \Xi^s \tag{5.33b}$$

$$3z_{2,1,s} + 2z_{2,2,s} \leq D_2 \boldsymbol{\xi} \qquad \forall s, \xi \in \Xi^s \tag{5.33c}$$

$$z_{t+1,j,s} \geq z_{t,j,s} \qquad \forall j, s, t = 1 \tag{5.33d}$$

$$x_1 = 0 \tag{5.33e}$$

$$X_{2,s}\xi + X'_{2,s}\eta'_{[2]} \geq D_3 \eta_{[2]} + D_1 \boldsymbol{\xi} \qquad \forall s, \xi \in \Xi^s, \eta \in \Psi, \eta' \in \Psi'' \tag{5.33f}$$

$$X_{3,s}\boldsymbol{\xi} + X'_{3,s}\eta'_{[3]} \geq D_4 \eta_{[3]} + D_2 \boldsymbol{\xi} \qquad \forall s, \xi \in \Xi^s, \eta \in \Psi, \eta' \in \Psi'' \tag{5.33g}$$

$$\mid z_{t+1,s,j'} - z_{t+1,s',j'} \mid \leq z_{t,s,j} \qquad \forall t = 1, s, s', j, j' : s_{-j} = s'_{-j} \tag{5.33h}$$

$$\mid X_{t+1,s,j'} - X_{t+1,s',j'} \mid \leq M z_{t,s,j} \qquad \forall t, s, s', j, j' : s_{-j} = s'_{-j} \tag{5.33i}$$

$$\mid X_{t+1,s,j} \mid \leq M z_{t,s,j} \qquad \forall t, j, s \tag{5.33j}$$

$$z_{1,s,j} = z_{1,s',j} \qquad \forall j, s, s' \tag{5.33k}$$

$$z_{t,s,j} \in \{0, 1\} \tag{5.33l}$$

where $D_1 = [1, 1, 3]$, $D_2 = [2, 2, 1]$, $D_3 = [1, 3]$, $D_4 = [1, 2, 1]$. In this problem, there are two time steps $t = 1, 2$, two endogenous ξ_1, ξ_2 and two exogenous η_1, η_2 uncertain parameters. Constraints 5.33h and 5.33i enforce the non-anticipativity for binary and continuous variables across different scenarios and constraint 5.33j stipulate the non-anticipativity for continuous variables within the same scenario. The breakpoint locations for endogenous and exogenous parameters are provided in tables 5.6 and 5.7 respectively.

Table 5.6: Breakpoint locations for endogenous uncertain parameters

Number of breakpoints	Number of scenarios	$\xi_1 \in [0, 3]$	$\xi_2 \in [0, 6]$
1	4	2	4
2	9	1, 2	2, 4
5	36	0.5, 1, 1.5, 2, 2.5	1, 2, 3, 4, 5

Table 5.7: Breakpoint locations for exogenous uncertain parameters

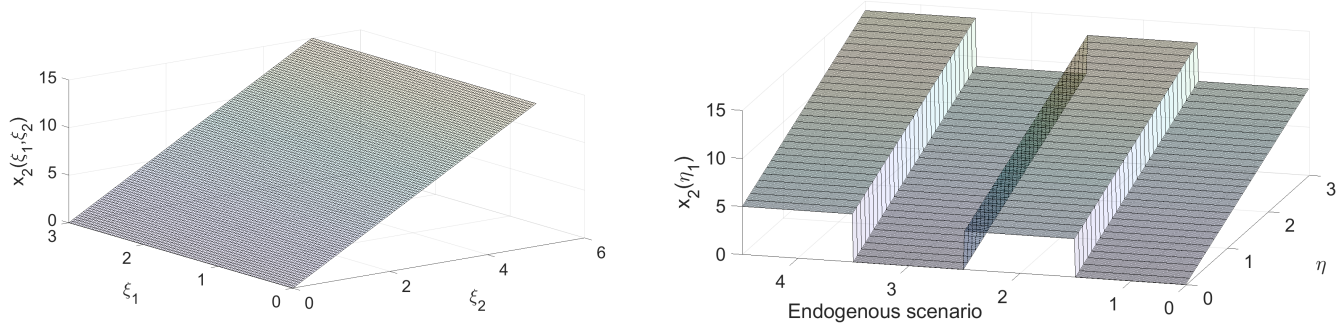
Number of breakpoints	$\eta_1 \in [0, 3]$	$\eta_2 \in [0, 6]$
1	2	4
2	1, 2	2, 4
5	0.5, 1, 1.5, 2, 2.5	1, 2, 3, 4, 5

Table 5.8: Solution statistics for different number of breakpoints

Number of breakpoints	Objective	Discrete, Continuous variables	Run time
1	31.61	16, 150	0.316 s
2	30.83	36, 307	0.467 s
5	30.00	144, 1138	0.907 s

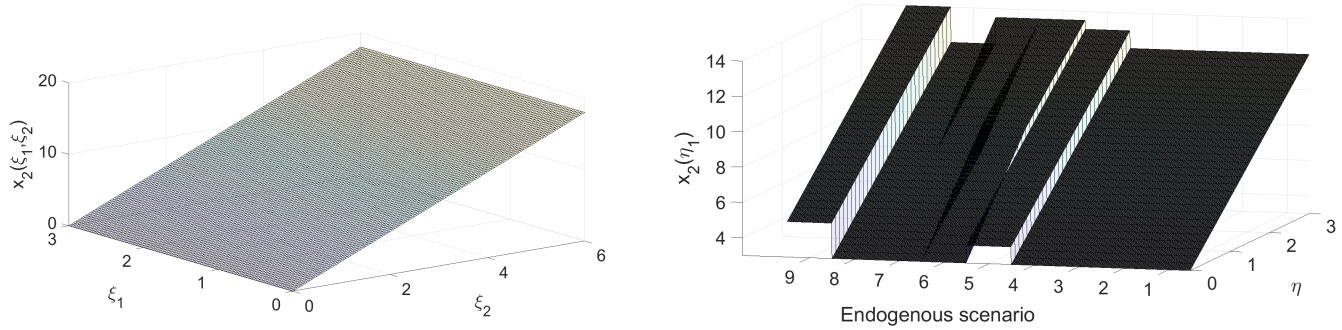
Table 5.9: Binary solution for 4 scenarios $z(\xi) = z_{t,j,s}$

t	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2
j	1	1	1	1	2	2	2	2	1	1	1	1	2	2	2	2
s	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
$z_{t,j,s}$	0	0	0	0	1	1	1	1	0	1	0	1	1	1	1	1



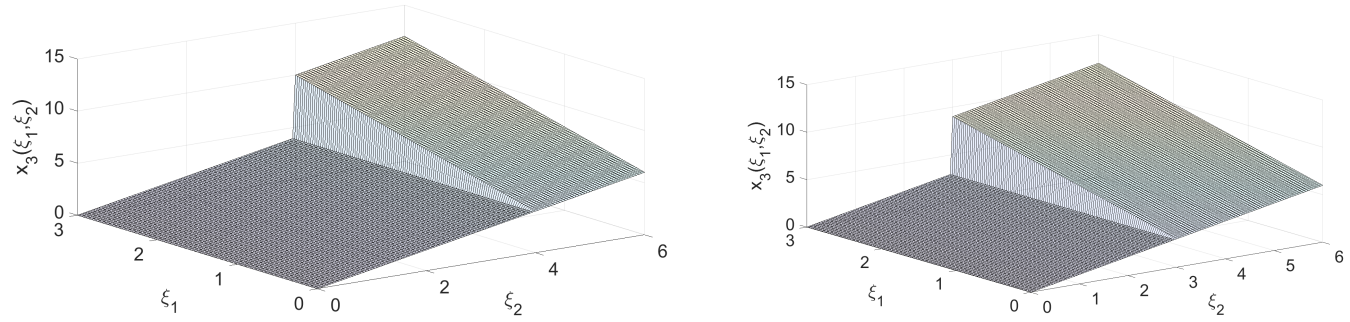
(a) $x_2(\xi_1, \xi_2)$ (b) $x_2(\eta_1)$

Figure 5.7: Endogenous $x_2(\xi_1, \xi_2)$, exogenous $x_2(\eta_1)$ parts of the continuous solution for 4 scenarios at $t = 1$



(a) $x_2(\xi_1, \xi_2)$ (b) $x_2(\eta_1)$

Figure 5.8: Endogenous $x_2(\xi_1, \xi_2)$, exogenous $x_2(\eta_1)$ parts of the continuous solution for 9 scenarios at $t = 1$



(a) 4 scenarios (b) 36 scenarios

Figure 5.9: Endogenous part of the continuous solution $x_3(\xi_1, \xi_2)$ for different number of scenarios at $t = 2$ (the exogenous part is not illustrated since the dimensions exceed 3)

The binary solution in this problem is similar to the previous problems (Table 5.9). Figures 5.7 and 5.8 illustrate the endogenous $x_2(\xi_1, \xi_2)$ and exogenous $x_2(\eta_1)$ parts of the continuous solution at $t = 1$ ($x_2(\xi, \eta) = x_2(\xi_1, \xi_2) + x_2(\eta_1)$). As figure 5.2 and table 5.9 indicate, at t=1, only ξ_2 is revealed and ξ_1 is not realized. It means that at t=1, the continuous variable can only be a linear function of ξ_2 and η_1. Figures 5.7a and 5.8a illustrate that the endogenous part of continuous variable ($x_2(\xi_1, \xi_2)$) is only a linear funtion of ξ_2 and fixed with respect to ξ_1. Figures 5.7b and 5.8b shows that at each scenario, the exogenous part of continuous variable ($x_2(\eta_1)$) is a linear function of exogenous uncertainty.

Figure 5.9 presents the endogenous part of the continuous solution at $t = 2$ ($x_3(\xi_1, \xi_2)$). At the second time step, ξ_2 is revealed at all scenarios while ξ_1 is only revealed at certain scenarios (Table 5.9). Therefore, the continous solution is a linear function of ξ_2 at all scenarios and only a function of ξ_1 at some scenarios. For instance for the case of 4 scenarios, $x_3(\xi_1, \xi_2)$ is only a function of ξ_1 when $\xi_2 \in [4, 6]$ which is equivalent to scenarios 2 and 4. Similar solution is obtained for 9

and 36 scenarios. Figure 5.9 shows that by increasign the number of scenarions from 4 to 36, the breakpoint along the ξ_2 axis has shifted from 4 to 3 which has contirbuted to an improved objective value. As table 5.8 shows, increasing the numer of breakpoints for both types of uncertainty has provided a more refined search space and subsequently has resulted in a better solution quality.

5.6 General problem formulation under exogenous and endogenous uncertainty

Here, we go back to the original problem stated in Section 2, which includes indicator binary variables under endogenous uncertainty, binary and continuous variables under both endogenous and exogenous uncertainty. Under the assumption of the coefficients: $c_t(\boldsymbol{\eta}) = C_t\boldsymbol{\eta}_{[t]}$, $d_t(\boldsymbol{\eta}) = D_t\boldsymbol{\eta}_{[t]}$, $h_t(\boldsymbol{\eta}) = H_t\boldsymbol{\eta}_{[t]}$, $f_t(\boldsymbol{\eta}) = F_t\boldsymbol{\eta}_{[t]}$, $e_t(\boldsymbol{\xi}) = E_t\boldsymbol{\xi}$, the model is

$$\min_{x,y,z} \quad \mathbb{E}\left(\sum_t \boldsymbol{\eta}_{[t]}^\top C_t^\top x_t(\boldsymbol{\xi},\boldsymbol{\eta}) + \boldsymbol{\eta}_{[t]}^\top D_t^\top y_t(\boldsymbol{\xi},\boldsymbol{\eta}) + \boldsymbol{\eta}_{[t]}^\top H_t^\top z_t(\boldsymbol{\xi})\right) \tag{5.34a}$$

$$s.t. \quad \sum_{\tau=1}^{t} A_{t,\tau}^\top x_\tau(\boldsymbol{\xi},\boldsymbol{\eta}) + \sum_{\tau=1}^{t} B_{t,\tau}^\top y_\tau(\boldsymbol{\xi},\boldsymbol{\eta}) + \sum_{\tau=1}^{t} M_{t,\tau}^\top z_\tau(\boldsymbol{\xi}) \le E_t\boldsymbol{\xi} + F_t\boldsymbol{\eta}_{[t]} \quad \forall t, \xi \in \Xi, \eta \in \Omega \tag{5.34b}$$

$$x_t(\boldsymbol{\xi},\boldsymbol{\eta}) = x_t(z_{t-1} \circ \boldsymbol{\xi}, \boldsymbol{\eta}_{[t]}) \qquad \forall t, \xi \in \Xi, \eta \in \Omega \tag{5.34c}$$

$$y_t(\boldsymbol{\xi},\boldsymbol{\eta}) = y_t(z_{t-1} \circ \boldsymbol{\xi}, \boldsymbol{\eta}_{[t]}) \qquad \forall t, \xi \in \Xi, \eta \in \Omega \tag{5.34d}$$

$$z_t(\boldsymbol{\xi}) = z_t(z_{t-1} \circ \boldsymbol{\xi}) \qquad \forall t, \xi \in \Xi \tag{5.34e}$$

$$z_t(\boldsymbol{\xi}) \in \{0,1\}, y_t(\boldsymbol{\xi},\boldsymbol{\eta}) \in \{0,1\}, x_t(\boldsymbol{\xi},\boldsymbol{\eta}) \in \mathbb{R} \qquad \forall t, \xi \in \Xi, \eta \in \Omega \tag{5.34f}$$

5.6.1 Solution method

Both binary $y(\xi,\eta)$ and continuous $x(\xi,\eta)$ variables are functions of endogenous and exogenous uncertainty. $z(\xi)$ is the binary decision that indicates the revelation of endogenous uncertainty. In this general formulation, the continuous variable is formulated as a linear combination of endogenous and exogenous uncertainty. The binary varaible is formulated as the multiplication of endogenous and exogenous binary components which represents the AND logic between two types of uncertainty. In both binary and continuous varibles, the endogenous component is formulated using the partitioning method and the exogenous component is formulated based on the lifting method. In both binary and continuous variables, for each scenario of endogenous uncertainty,

there is a lifting solution for exogenous uncertainty.

$$z_t(\xi) = z_{t,s} \qquad \forall t, s \tag{5.35a}$$
$$y_t(\xi, \eta) = y_{t,s} Y_{t,s} Q_{[t]}(\eta) \qquad \forall s, \eta \in \Psi \tag{5.35b}$$
$$x_t(\xi, \eta) = X_{t,s}\boldsymbol{\xi} + {X'}_{t,s}\eta'_{[t]} \qquad \forall s, \xi \in \Xi^s, \eta' \in \hat{\Psi}' \tag{5.35c}$$

Following the above described variable modeling technique and by applying duality to semi-infinite constraints, the deterministic counterpart optimization model is obtained.

$$\min \quad \sum_s p_s \sum_t (\mathbb{E}(\boldsymbol{\eta'})^\top {P_{\eta_{[t]}}}^\top C_t^\top X_{t,s} \mathbb{E}_{\Xi_s}(\boldsymbol{\xi}) + \mathrm{Tr}({P_{\eta_{[t]}}}^\top C_t^\top X'_{t,s} \;\; P_{G_{[t]}} \mathbb{E}[\eta' \boldsymbol{\eta'}^\top])) \tag{5.36a}$$
$$+ \quad \sum_s p_s \sum_t (\mathrm{Tr}({P_{\eta_{[t]}}}^\top D_t^\top y_{t,s} Y_{t,s} P_{Q_{[t]}} \mathbb{E}[\eta' \boldsymbol{\eta'}^\top]) + \mathbb{E}(\boldsymbol{\eta'})^\top {P_{\eta_{[t]}}}^\top \; H_t^\top z_{t,s})$$
$$s.t. \tag{5.36b}$$
$$V_s \theta_{t,s} + V'_s \theta'_{t,s} \geq \sum_{\tau=1}^{t} {M_{t,\tau}}^\top z_{\tau,s} \qquad \forall t, s \tag{5.36c}$$
$$W_s^\top \theta_{t,s} = -\sum_{\tau=1}^{t} (\mathbf{A}_{t,\tau}^\top X_{\tau,s}) + \mathbf{E}_t \qquad \forall t, s \tag{5.36d}$$
$$W_s'^\top \theta'_{t,s} = -\sum_{\tau=1}^{t} (\mathbf{A}_{t,\tau}^\top X'_{\tau,s} P_{G_{[t]}} - B_{t,\tau}^\top y_{\tau,s} Y_{\tau,s} P_{Q_{[t]}}) + \mathbf{F}_t P_{\eta_{[t]}} \qquad \forall t, s \tag{5.36e}$$
$$\theta_{t,s} \geq 0, \theta'_{t,s} \geq 0 \qquad \forall t, s \tag{5.36f}$$
$$|z_{j',t,s} - z_{j',t,s'}| \leq z_{j,t-1,s} \qquad \forall j, j', t, s, s' : s_{-j} = s'_{-j} \tag{5.36g}$$
$$|y_{t,s} - y_{t,s'}| \leq z_{j,t-1,s} \qquad \forall j, t, s, s' : s_{-j} = s'_{-j} \tag{5.36h}$$
$$|X_{t,s,j'} - X_{t,s',j'}| \leq M z_{j,t-1,s} \qquad \forall j, j', t, s, s' : s_{-j} = s'_{-j} \tag{5.36i}$$
$$|X_{t,s,j}| \leq M z_{j,t-1,s} \qquad \forall j, t, s \tag{5.36j}$$
$$z_{t,s} \in \{0, 1\}, y_{t,s} \in \{0, 1\} \qquad \forall t, s \tag{5.36k}$$
$$Y_{t,s} \in \{-1, 0, +1\} \qquad \forall t, s \tag{5.36l}$$

5.6.2 Example

Consider the following illustrating example

$$\max \quad \mathbb{E}_{\xi,\eta}[(2+\eta)x(\xi,\eta) - 6y(\xi,\eta) - 4z(\xi) - 0.1x(\xi,\eta)] \tag{5.37a}$$
$$s.t. \quad x(\xi,\eta) \leq \xi \qquad \forall \xi, \eta \tag{5.37b}$$
$$8x(\xi,\eta) \leq 16 + \eta \qquad \forall \xi, \eta \tag{5.37c}$$
$$(\eta - 8)z(\xi) \leq 16y(\xi,\eta) \qquad \forall \xi, \eta \tag{5.37d}$$

$$x(\xi,\eta) \le 20z(\xi) \qquad \forall \xi,\eta \qquad (5.37e)$$
$$x(\xi,\eta) \ge 0 \qquad \forall \xi,\eta \qquad (5.37f)$$
$$x(\xi,\eta) = x(z(\xi)\circ\xi,\eta) \qquad \forall \xi,\eta \qquad (5.37g)$$
$$y(\xi,\eta) = y(z(\xi)\circ\xi,\eta) \qquad \forall \xi,\eta \qquad (5.37h)$$
$$y(\xi,\eta), z(\xi) \in \{0,1\} \qquad \forall \xi,\eta \qquad (5.37i)$$

By employing the explained variable modeling technique and using duality method, the deterministic counterpart formulation can be derived.

$$\max \quad \sum_s p_s[2(X_s\mathbb{E}_{\Xi_s}(\xi) + X_s'\mathbb{E}(\boldsymbol{\eta}') + \mathbb{E}(\eta)X_s\mathbb{E}_{\Xi_s}(\boldsymbol{\xi}) + X_s'E(\eta\boldsymbol{\eta}'^{\top})$$
$$- 6y_sY_s\mathbb{E}(Q(\eta)) - 4z - 0.1((X_s\mathbb{E}_{\Xi_s}(\xi) + X_s'\mathbb{E}(\boldsymbol{\eta}'))] \qquad (5.38a)$$
$$s.t. \quad X_s\xi + X_s'\boldsymbol{\eta}' \le \xi \qquad \forall s, \xi \in \Xi_s, \eta' \in \hat{\Psi}' \qquad (5.38b)$$
$$8(X_s\xi + X_s'\boldsymbol{\eta}') \le 16 + \eta \qquad \forall s, \xi \in \Xi_s, \eta \in \Psi, \eta' \in \hat{\Psi}' \qquad (5.38c)$$
$$(\eta - 8)z \le 16(y_sY_sQ(\eta)) \qquad \forall s, \eta \in \Psi \qquad (5.38d)$$
$$X_s\xi + X_s'\boldsymbol{\eta}' \le 20z_s \qquad \forall s, \xi \in \Xi_s, \eta' \in \hat{\Psi}' \qquad (5.38e)$$
$$X_s\xi + X_s'\boldsymbol{\eta}' \ge 0 \qquad \forall s, \xi \in \Xi_s, \eta' \in \hat{\Psi}' \qquad (5.38f)$$
$$|X_{s,j}| \le Mz_s \qquad \forall s, j = 2 \qquad (5.38g)$$
$$z_s = z_{s'} \qquad \forall s, s' \qquad (5.38h)$$
$$0 \le Y_sQ(\eta) \le 1 \qquad \forall s, \eta \in \Psi \qquad (5.38i)$$
$$z_s \in \{0,1\}, y_s \in \{0,1\} \qquad \forall s, t, j \qquad (5.38j)$$

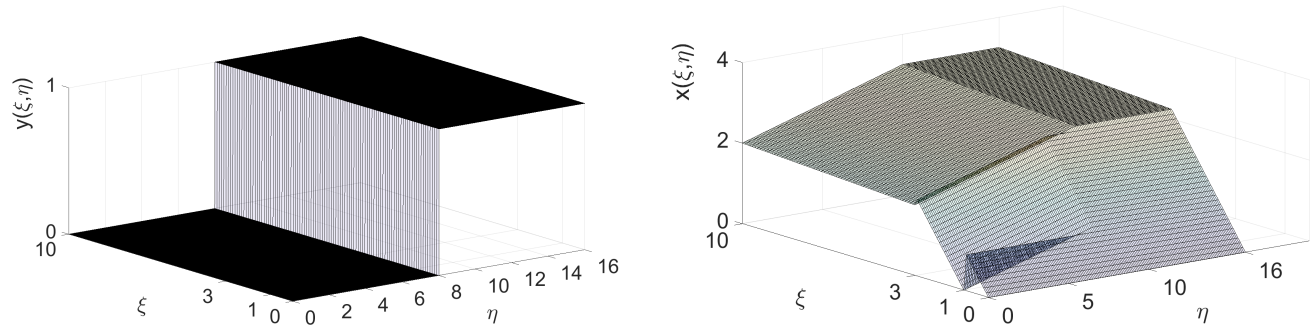
Uncertainty interval and breakpoint locations are presented in Table 5.10.

Table 5.10: Breakpoint locations for endogenous and exogenous uncertain parameters

Endogenous	$\xi \in [0,10]$	Exogenous	$\eta \in [0,16]$
Number of breakpoints	Location	Number of breakpoints	Location
1	5	1	8
2	1, 3	3	4, 8, 12
3	1, 3, 7	7	2, 4, 6, 8, 10, 12, 14, 16

Table 5.11: Solution statistics for different number of breakpoints

Br Endo, Exog	Objective	Disc, Cont variables	Run time
1, 1	14.133	12, 190	0.348 s
1, 3	14.133	20, 262	0.359 s
1, 7	14.133	36, 402	0.414 s
2, 1	17.217	18, 280	0.463 s
2, 3	17.426	30, 388	0.384 s
2, 7	17.657	54, 604	0.514 s
3, 1	18.360	24, 370	0.384 s
3, 3	18.479	40, 514	0.574 s
3, 7	18.493	72, 802	0.635 s



(a) Binary variable

(b) Continuous variable

Figure 5.10: Both binary and continuous variables are piecewise discontinuous functions of ξ and η (illustration for 2 endogenous and 7 exogenous breakpoints)

Figure 5.10 illustrates the binary and continuous solution for the case with 2 and 7 breakpoints for endogenous and exogenous uncertaint parameters, respectively. It can be observed that both binary and continuous variables are piecewise discontinuous functions of endogenous and exogenous uncertainty. Table 5.11 demonstrates that by increasing the number of breakpoints for endogenous or exogenous uncertain parameters, the objective value improves. In this problem, increasing the number of partitions for endogenous uncertainty has a much more substantial effect on improving the objective value than increasing the number of breakpoints for exogenous uncertainty.

5.7 Shale gas problem

In this section, a case study about infrastructure construction and production planning of a gas field is studied. Figure 5.11 illustrates the general network of the gas field. There are 3 platforms and 5 pipelines. The goal is to optimally plan the construction and gas production during a certain

time horizon in order to maximize the revenue. Equatoins 5.39a to 5.39o present the problem objective and constraints. The problem includes both endogenous and exogenous uncertainty. ξ is the gas reservoir size that represents endogenous uncertainty and η is the gas price that indicates the exogenous uncertainty. Explanation of all the problem variables and parameters is provided in the Appendix.

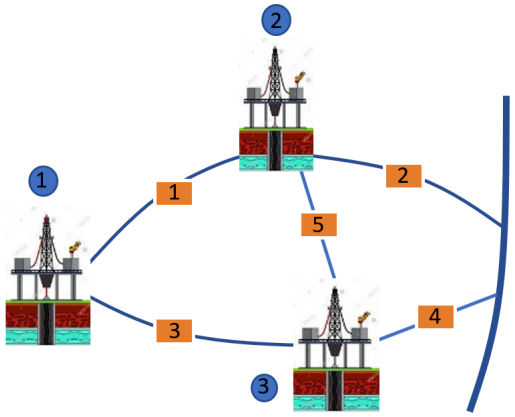


Figure 5.11: Platform, Pipeline Network

$$\max \quad \mathbb{E}_{\xi,\eta} \sum_{t\in T} d_t[c_t^g(\eta_t)(x_{2,t}^f(\xi,\eta)+x_{4,t}^f(\xi,\eta)) - \sum_{i\in L} c_i^l(y_{i,t}^l(\xi,\eta) - y_{i,t-1}^l(\xi,\eta)) \tag{5.39a}$$
$$- \sum_{j\in P} c_j^p(y_{j,t}^p(\xi) - y_{j,t-1}^p(\xi)) - \sum_{j\in P} c_j^c x_{j,t}^c(\xi,\eta) - \sum_{j\in P} c_j^e\, x_{j,t}^e(\xi,\eta)]$$

$$s.t. \quad \sum_{t\in T} x_{j,t}^e(\xi,\eta) \le \xi^j \qquad \forall j, \xi\in\Xi, \eta\in\Psi \tag{5.39b}$$

$$0 \le x_{j,t}^e(\xi,\eta) \le r_j + \eta_t \qquad \forall j,t\in T, \xi\in\Xi,,\eta\in\Psi \tag{5.39c}$$

$$x_{j,t}^e(\xi,\eta) + \sum_{i\in L^+(p)} x_{i,t}^f(\xi,\eta) \ge \sum_{i\in L^-(p)} x_{i,t}^f(\xi,\eta) \qquad \forall j\in P, t\in T, \xi\in\Xi, \eta\in\Psi \tag{5.39d}$$

$$\sum_{i\in L^-(p)} x_{i,t}^f(\xi,\eta) \le \sum_{t'=1}^{t} x_{j,t'}^c(\xi,\eta) \qquad \forall j\in\mathbb{P}, t\in T, \xi\in\Xi, \eta\in\Psi \tag{5.39e}$$

$$0 \le x_{i,t}^f(\xi,\eta) \le M y_{i,t}^l(\xi,\eta) \qquad \forall i\in L, t\in T, \xi\in\Xi, \eta\in\Psi \tag{5.39f}$$

$$0 \le x_{j,t}^c(\xi,\eta) \le M y_{j,t}^p(\xi) \qquad \forall j\in P, t\in T, \xi\in\Xi, \eta\in\Psi \tag{5.39g}$$

$$y_{j,t}^p(\xi) \ge y_{j,t-1}^p(\xi) \qquad j\in P, t\in T, \xi\in\Xi \tag{5.39h}$$

$$y^l_{i,t}(\xi,\eta) \geq y^l_{i,t-1}(\xi,\eta) \qquad i \in L, t \in T, \xi \in \Xi, \eta \in \Psi \quad (5.39i)$$

$$\sum_{j\in P} x^p_{j,t}(\xi) - \sum_{j\in P} x^p_{j,t-1}(\xi) \leq N^{max} \qquad \forall t, \xi \in \Xi \quad (5.39j)$$

$$y^l_t(\xi,\eta), y^p_t(\xi) \in \{0,1\} \qquad \forall t, \xi \in \Xi, \eta \in \Psi \quad (5.39k)$$

$$y^p_{t+1}(\xi) = y^p_{t+1}(y^p_t(\xi) \circ \boldsymbol{\xi}) \qquad \forall \in t, \xi \quad (5.39l)$$

$$x^f_{i,t+1}(\xi,\eta) = x^f_{i,t+1}(y^p_t(\xi) \circ \boldsymbol{\xi}, \eta_{[t+1]}) \qquad \forall i, t, p, \xi, \eta \quad (5.39m)$$

$$x^c_{j,t+1}(\xi,\eta) = x^c_{j,t+1}(y^p_t(\xi) \circ \boldsymbol{\xi}, \eta_{[t+1]}) \qquad \forall j, t, p, \xi, \eta \quad (5.39n)$$

$$x^e_{j,t+1}(\xi,\eta) = x^e_{j,t+1}(y^p_t(\xi) \circ \boldsymbol{\xi}, \eta_{[t+1]}) \qquad \forall j, t, p, \xi, \eta \quad (5.39o)$$

Equation 5.39b means that the total amount of gas extracted over the entire time horizon from a single gas field should not exceed the gas field size. Equation 5.39c indicates that gas production is limited by a maximum production rate at a particular production platform. If the gas price increases at each year, the amount of gas extraction increases proportionally. Equation 5.39d is the flow conservation constraint in the network. Equation 5.39e enforces that gas flow from a particular production platform should not exceed its capacity. Equation 5.39f enforces that gas can not flow through pipeline i if the pipeline has not been built. Equation 5.39g means that no platform expansion can be constructed if platform p has not been built. Equation 5.39h and 5.39i enforces that once a platform or pipelines is built, it does not vanish in the network. Equation 5.39j enforces the maximum number of platforms that can be built at each year. Equation 5.39a is the objective that constitute of the revenue obtaine from selling gas minus the costs for building pipelines, platforms, gas extraction and platform capacity expansion.

The endogenous uncertainty is modeled using partitioning method and the exogenous uncertainty is modeled using lifting method. The binary variable $y^p_{j,t}(\xi)$ indicates the construction of platforms and it depends only on endogenous uncertainty. The binary variable $y^l_{i,t}(\xi,\eta)$ indicates pipeline construction and it depends on both type of uncertainties. It is modeled as a multiplication of an endogenous binary variable and an exogenous binary variable (AND logic between endogenous and exogenous uncertainty). The continuous variables are modeled as linear combination of endogenous and lifted exogenous uncertain parameters. In both binary and continuous variables, for each scenario of endogenous uncertainty, there is a lifting solution for exogenous uncertainty. Under the linear decision rule approximation, the derivation of the deterministic counterpart optimization problem is reported in Appendix.

$$y^p_{j,t}(\xi) = y^p_{j,t,s} \qquad \forall j, s, \xi \in \Xi_s \quad (5.40)$$

$$y^l_{i,t}(\xi,\eta) = y^l_{i,t,s} Y^l_{i,t,s} Q_{[t]}(\eta) \qquad \forall t, i, s, \xi \in \Xi_s, \eta \in \Psi \quad (5.41)$$

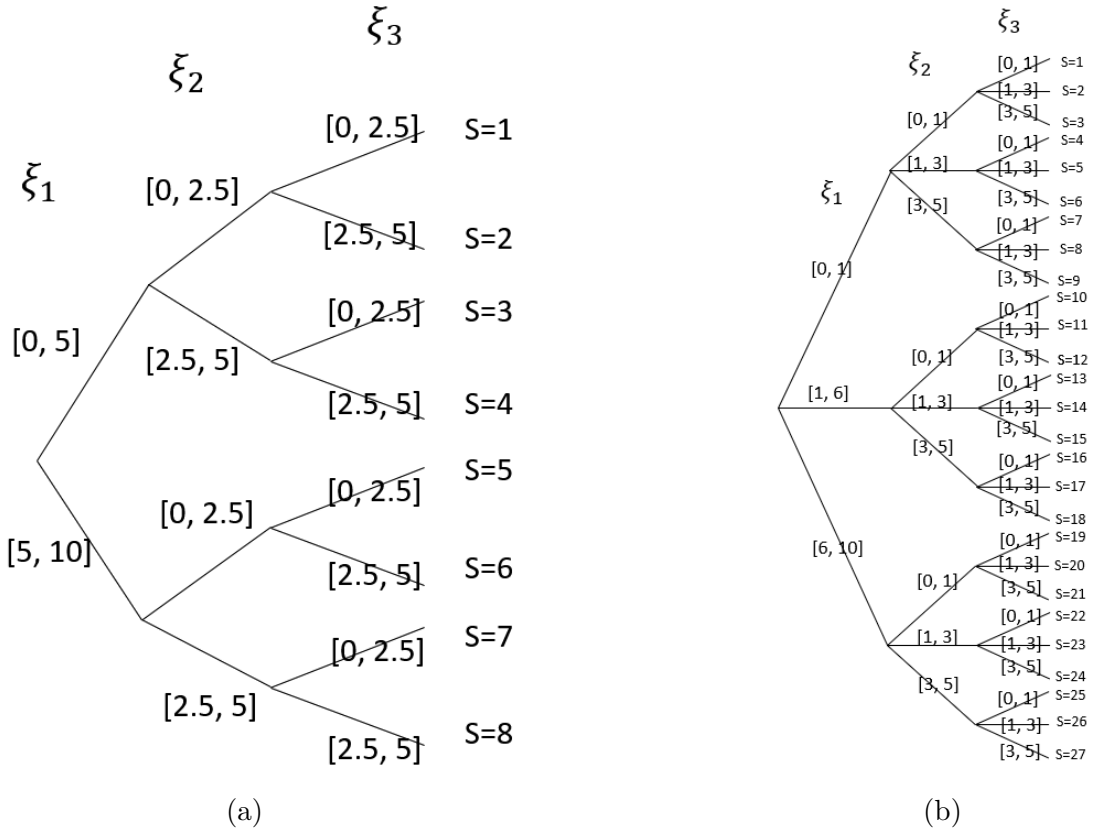
$$x^f_{i,t}(\xi,\eta) = X^f_{i,t,s}\boldsymbol{\xi} + X^{'f}_{i,t,s}\boldsymbol{\eta}'_{[t]} \qquad \forall t, i, s, \xi \in \Xi_s, \eta' \in \hat{\Psi}' \quad (5.42)$$

$$x^c_{j,t}(\xi,\eta) = X^c_{j,t,s}\boldsymbol{\xi} + X^{'c}_{j,t,s}\boldsymbol{\eta}'_{[t]} \qquad \forall t, j, s, \xi \in \Xi_s, \eta' \in \hat{\Psi}' \quad (5.43)$$

$$x^{e}_{j,t}(\xi,\eta) = X^{e}_{j,t,s}\boldsymbol{\xi} + X^{'e}_{j,t,s}\boldsymbol{\eta}'_{[t]} \qquad \forall t, j, s, \xi \in \Xi_s, \eta' \in \hat{\Psi}' \tag{5.44}$$

Table 5.12: Breakpoint locations for endogenous and exogenous uncertain parameters

Endogenous				Exogenous	
Number of breakpoints	$\xi_1 \in [0,10]$	$\xi_2 \in [0,5]$	$\xi_3 \in [0,5]$	Number of breakpoints	$\eta \in [0,1]$
1	5	2.5	2.5	1	0.5
2	1, 6	1, 3	1, 3	3	0.25, 0.5, 0.75



(a) (b)

Figure 5.12: Reservoir size illustration for 8 and 27 scenarios

Table 5.12 demonstrates the uncertainty range and breakpoints locations and figure 5.12 illustrates the partitioning of endogenous uncertainty (gas reservoir size) for 8 and 27 scenarios. The problem solution for $T = 3, 4, 5$ time periods and different number of breakpoints is reported in table 5.13. Increasing the number of exogenous and endogenous breakpoints has improved the objective while the effect of endogenous uncertainty (reservoir size) is more significant. For instance at $T = 3$, increasing the number of endogenous breakpoints from 1 to 2 has improved the objective by 11.45% while increasign the number of exogenous breakpoints from 1 to 3 has resulted in 8.32% improvment in the objective. It should be noted that increasing the breakpoints has substantially

increased the run time. As an example, for $T = 3$, when the number of exogenous breakpoints increases from 1 to 3, the runtime increases from 16 min with zero gap to 17 hrs with 5.8% gap. Therefore, depending on the model size, selecting a reasonable number of breakpoints can save significant run time.

Table 5.13: Solution statistics for different number of partitions

Time step	Br Endo, Exog	Obj	Best possible obj	Opt gap	Disc, Cont variables	Run time
T=3	1, 1	22.69	22.69	0	912, 33965	∼16 min
	1, 3	23.22	24.58	5.87%	1872, 50045	∼ 17 hr
	2, 1	24.29	25.29	3.9%	3078, 114601	∼ 336 hr
	2, 3	23.30	27.65	18.64%	6318, 168871	∼ 26 hr
T=5	1, 1	25.85	25.85	0	1920, 84973	∼ 118 hr
	1, 3	25.86	28.68	10.89%	4320, 30413	∼ 40 hr
	2, 1	25.02	29.28	17.04%	6480, 286753	∼ 56 hr
	2, 3	23.80	30.49	28.09%	14580, 440113	∼ 91 hr
T=10	1, 1	26.31	28.07	6.2 %	5840, 311613	∼ 336 hr

5.7.1 Solution for binary decision rule

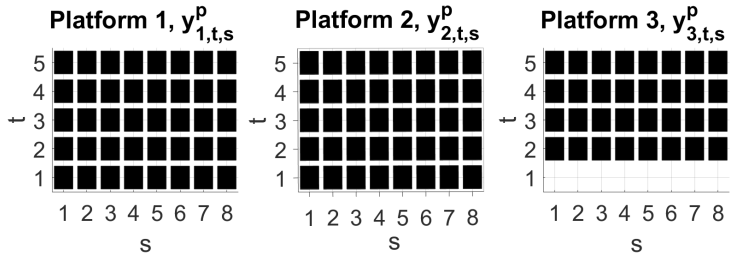


Figure 5.13: Platform construction decision at year t and scenario s

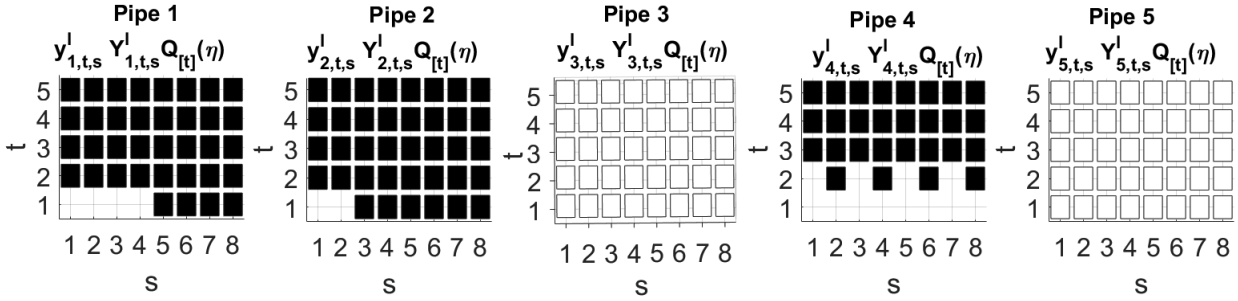


Figure 5.14: Pipe construction decision at year t and scenario s

Figures 5.13 and 5.14 illustrate the platform and pipe construction decisions. platforms 1 and 2 are built from the first time step and platform 3 is built starting from the second time step at all the

scenarios. Considering constraint 5.39j that stipulates at each time step only two platforms can be built and building the platform is required to extract and sell the gas, these results are rational. Figure 5.14 illustrates the pipe construction decision for the 8 scenarios setting (figure 5.12a). Pipes 3 and 5 are never built since these pipes are the most expensive pipes to build and their absence does not prevent gas delivery from any platform to the main pipeline (figure 5.11). Pipe 1 is only built after scenario 4 at $t = 1$. Since pipe 1 delivers the gas from platform 1 and scenarios 1 to 4 and 5 to 8 correspond to the smaller and larger segments of reservoir 1 respectively ([0, 5] and [5, 10]), the solver has decided it's not economical to build pipe 1 for smaller segment of the reservoir.

Pipe 2 is only built after scenario 2 at $t = 1$. Since pipe 2 delivers the gas from platform 2 and scenarios 1 and 2 correspond to the smaller segment of reservoir 2, the solver has decided that it is only profitable to build pipe 2 starting from scenario 3 where scenarios correspond to the larger segments of reservoirs 2 and 1. Pipe 4 delivers the gas from platform 3. Since platform 3 is not built at $t = 1$, subsequently pipe 4 can not be built at $t = 1$ too. At $t = 2$, pipe 4 is only built at scenarios 2, 4, 6, 8 that correspond to the larger segment of reservoir 3 ([2.5, 5]) in order to achieve the most profit.

5.7.2 Solution for continuous decision rule

The continuous varaibles include extraction, flow and capacity expansion decisions ($x^e_{j,t}(\xi, \eta)$, $x^f_{j,t}(\xi, \eta)$, $x^c_{j,t}(\xi, \eta)$) that depend on both price and reservoir size uncertainty. At each time step, the continuous variables can depend on all previous uncertain parameters up to the current time step. Since different scenarios represent different reserves for each of the three reservoirs, the solution corresponding to each scenario can be different. Table 5.14 presents the extraction solution for platform 3 at $t = 5$ under the setting T=5, 1 endogenous and 1 exogenous breakpoint. It can be observed in the table that for platform 3, each group of scenarios 1,3,5,7 and 2,4,6,8 have the same solution since the first group correspond to the smaller segment ([0, 2.5]) and the second group corresponds to the larger segment ([2.5, 5]) of reservoir 3. For the latter group, the extraction amount depends on both reservoir and price uncertainty; as the reservoir size or the price increases, the extraction increases too that follows constraitns 5.39b and 5.39c.

5.8 Conclusion

In this study, a novel framework capable of handling both endogenous and exogenous uncertainty for multistage mixed integer optimization is formulated based on robust optimization method. The endogenous and exogenous uncertainties are modeled using the partitioning and the lifting methods, respectively. The continuous variable is a summation of endogenous and exogenous components. The endogenous component is based on the partitioning method where, at each scenario,

Table 5.14: $x^e_{j,t}(\xi,\eta)$ for platform 3 at t=5 under the setting T=5, 1 endogenous, 1 exogenous breakpoint

Scenario Number	Platform 3
1	$x^e_{3,5}(\xi,\eta) = X^e_{3,5,1}\xi + X'^e_{3,5,1}\eta'_{[5]} = 0.33\xi_3$
2	$x^e_{3,5}(\xi,\eta) = X^e_{3,5,2}\xi + X'^e_{3,5,2}\eta'_{[5]} = 0.4\xi_3 - 1 + 0.17G_{5,2}$
3	$x^e_{3,5}(\xi,\eta) = X^e_{3,5,3}\xi + X'^e_{3,5,3}\eta'_{[5]} = 0.33\xi_3$
4	$x^e_{3,5}(\xi,\eta) = X^e_{3,5,4}\xi + X'^e_{3,5,4}\eta'_{[5]} = 0.4\xi_3 - 1 + 0.17G_{5,2}$
5	$x^e_{3,5}(\xi,\eta) = X^e_{3,5,5}\xi + X'^e_{3,5,5}\eta'_{[5]} = 0.33\xi_3$
6	$x^e_{3,5}(\xi,\eta) = X^e_{3,5,6}\xi + X'^e_{3,5,6}\eta'_{[5]} = 0.4\xi_3 - 1 + 0.17G_{5,2}$
7	$x^e_{3,5}(\xi,\eta) = X^e_{3,5,7}\xi + X'^e_{3,5,7}\eta'_{[5]} = 0.33\xi_3$
8	$x^e_{3,5}(\xi,\eta) = X^e_{3,5,8}\xi + X'^e_{3,5,8}\eta'_{[5]} = 0.4\xi_3 - 1 + 0.17G_{5,2}$

the continuous variable is approximated using an affine function of endogenous uncertain parameters. The exogenous component is based on the lifting method where the continuous variable is approximated using an affine function of lifted uncertain parameters at each scenario of endogenous uncertainty. This formulation results in a flexible piece-wise discontinuous solution for the continuous variable. The binary variable is a multiplication (AND logic) of endogenous and exogenous components. The endogenous component is constant at each scenario. The exogenous component is approximated using an affine function of 0-1 indicator functions at each scenario of endogenous uncertainty. The proposed framework is applied to a numerically demanding case study for infrastructure construction and production planning of a gas filed. The obtained results demonstrate that the proposed framework is effective for large problem settings and it can successfully integrate both endogenous and exogenous uncertainty in the solution.

Chapter 6

Strategic Planning of SAGD Reservoir Development under Reservoir Production and Oil Price Uncertainty

Abstract

During the lifecycle of an oilfield project, well development is a critical phase due to intensive investments required for capital cost. Determining whether it is economic to develop an oilfield or not and finding the best plan for field development are the main concerns in this field of study. Drainage area development and production planning are crucial in a SAGD project of oil sand industry. In this study, an optimization framework for planning the development of SAGD drainage areas under uncertainty is presented. The proposed framework includes the following two major elements. First, a mixed integer optimization model is developed to arrange the multiperiod development plan of the drainage areas with consideration of capital and steam allocation restrictions. Second, uncertainties in crude oil price and reservoir production are investigated based on a multistage stochastic programming model. The results demonstrate that the method can effectively generate an economically optimal development plan that maximizes the profit.

6.1 Introduction

Oil extraction and production is one of the main income streams of Alberta Canada. After Venezuela with 297 billion barrels and Saudi Arabia with 268 billion barrels of oil reserves, Alberta of Canada has 166.3 billion barrels of proven oil reserves [95]. Most of the the oil reserves in Alberta Canada is oil sands bitumen. Since the start of commercial production in 1967 in Alberta, only 5.9% of initial crude bitumen reserves have been produced. Athabasca, Peace River and cold lake are the major areas for bitumen production in Alberta where Athabasca holds the largest reserves. Bitumen is produced using surface mining and in-situ methods: surface mining methods are used in areas where the bitumen reserve is close to the surface and in-situation methods are employed where the bitumen is deep below the ground. It is estimated that 80% of the total bitumen reserves are extractable using in-situ methods and surface mining will account for only 20%. In 2012, production from in-situ methods surpassed mined production. In 2014, 55% of total bitumen production (366.3 thousand m^3 per day) was from in-situ methods. By 2024, total bitumen production is expected to reach 642 thousand m^3 per day and in-situ production is expected to reach 60% [96]. Cyclic-steam stimulation (CSS) and Steam Assisted Gravity Drainage (SAGD) are the commonly used methods for bitumen production.

In SAGD operation, high temperature and high pressure steam is injected into the underground reservoir to reduce the bitumen viscosity so that it can be pumped to the surface. Usually several pairs of horizontal injector and producer wells are installed into a drainage area from a surface pad (SP) (Figure 6.1). The injector is installed around 5 meters above the producer. Hot steam is injected into the injector well. As a result, steam chamber grows and propagates into the reservoir, and liquidizes the bitumen. Subsequently, bitumen is pumped out. The SAGD operation mainly consists of three stages: rising, spread and depletion stages. In rising stage, steam chamber gradually propagates into the reservoir and production builds up. During spread stage, steam chamber has already reached top of the pay zone and it propagates laterally. In the depletion stage, most of the bitumen is already produced and production gradually declines. The produced bitumen is processed in the central processing facility (CPF). CPF is a surface facility that produces steam for injection into the underground oil reservoir, separates oil-water emulsion and recycles the produced water for steam generation.

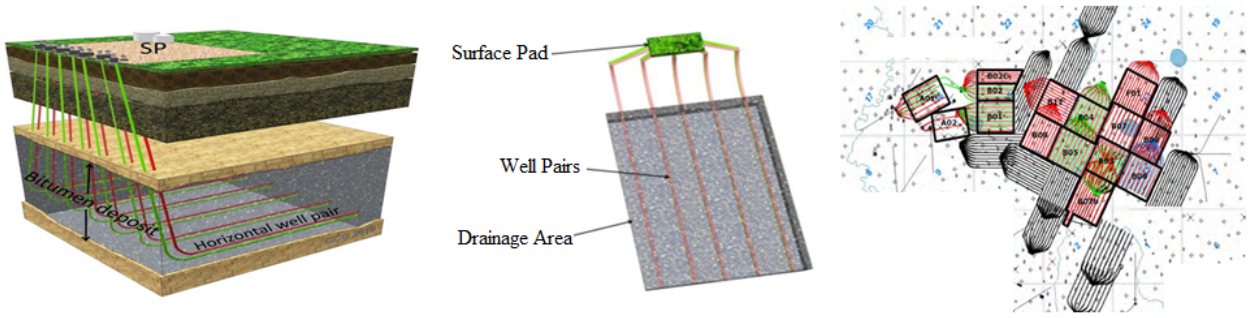


Figure 6.1: Left: 3D view of a DA; Middle: aerial view of a DA; Right: aerial view of multiple DAs

In the context of oil and gas infrastructure investment and production planning, most of the publications use a deterministic approach[97, 98, 99, 100]. Van Den et al. [98] presented a review of available literature for these types of problems. Shahandeh et al. [100] proposed an optimization framework to optimally divide an oil field into several drainage areas, in order to have maximum access to bitumen. They developed a deterministic optimization algorithm to maximize the revenue from SAGD operations while uncertainty is not addressed. Charry et al. [101] proposed an energy optimization model for upgrading of unconventional oil to minimize the cost while meeting environmental regulations and product demands. The proposed model for upgrading plants can determine the energy production costs and identify significant parameters that influence upgrading operations. Betancourt-Torcat et al. [102] presented an energy optimization model with similar objectives. Shahandeh and Li [103] proposed an optimal bitumen upgrading facility with the goal of CO_2 reduction. They discussed optimal configurations for different scenarios including plant capacity, price of natural gas, electricity and crude oil. The obtained model is a large-scale mixed-integer nonlinear problem which is solved using augmented Lagrangian decomposition method.

Uncertainty can be classified into endogenous and exogenous type [53]. If an uncertain parameter's revelation time depends on the decisions, it is referred to as endogenous uncertainty, otherwise it is called exogenous uncertainty. In the stochastic programming context, the reader can refer to the review [16, 17] on stochastic problems, and the following works [104, 105, 51, 106] as examples for addressing exogenous uncertainty. For endogenous uncertainty in optimization problem, distribution of uncertain parameters and the revelation time could both be affected by the optimization decisions. Hence, stochastic optimization problems including decision-dependent uncertainty revelation are considered to be difficult and the solution methods in this context are rare. Jonsbråten [53] developed a stochastic model with decision-dependent uncertainty under the assumption that variables controlling uncertainty revelation are discrete and only affect first stage decisions. Stochastic problems including discrete endogenous uncertainties can be formulated as deterministic mixed-binary programs [54],but the number of non-anticipativity constraints and binary variables can increase exponentially by expanding the size of the problem, which makes the

problem computationally intractable. Therefore, researchers have tried to propose approximation methods to provide feasible but suboptimal solutions to the original problem. In order to reduce the complexity of the problem, measurement decisions (i.e. decisions that reveal uncertain information) can be fixed as the first stage variables. Goel and Grossmann [54] employed decomposition techniques to solve this type of stochastic problems. Colvin and Maravelis [11] proposed branch and cut algorithm that can incorporate the adaptive measurement variables.

Addressing uncertainty in reservoir development planning is critical since it significantly affects the economic performance of a project. Among the various sources of uncertainties, oil price and reservoir production capacity represent the major uncertain factors to be considered. In the past, stochastic programming has received lots of attention in the context of oil and gas investment and production planning under uncertainty. Most of the algorithms for solving stochastic problems require discrete distribution of uncertain parameters. Jonsbråten [53] used a finite scenario tree to model the decision procedure and the binary decisions were used to indicate the uncertainty revelation. He addressed production optimization under reservoir size and price uncertainty. Goel and Grossmann [54] presented a stochastic programming approach for offshore gas filed development under reserve uncertainty. Later on, Goel et al. [107] developed a branch and bound based algorithm to solve the similar problem. Gupta and Grosmann [10] proposed iterative algorithms based on relaxation of non-anticipativity constraints for measurement variables. Betancourt-Torcat et al. [108] extended the deterministic model developed by Betacourt-Torcat et al. [102] to account for uncertainty in steam-oil-ratio (SOR) and natural gas price. It provides optimal arrangement of energy supply and oil producer infrastructures. There are few numbers of publications addressing oil and gas planning problems including endogenous uncertainty. Gupta and Grossmann [109] proposed a new decomposition algorithm based on relaxation of non-anticipativity constraints that results in reduced computational expense while preserving the same optimality gap.

All of the above described solution methods are based on scenario representation (discretization) of uncertain parameters. Discretization is an effective method for small size problems, but for problems with large number of uncertain parameters, discretization can result in very large number of states which makes the problem computationally inefficient. On the contrary, using few scenarios may result in suboptimal or even practically infeasible solutions. As an alternative, decision rule approximation techniques have been applied in various stochastic and robust optimization problems [30, 56, 22]. To handle endogenous uncertainty, Vayanos et al. [57] proposed a method for solving stochastic dynamic problems based on robust optimization techniques. They introduced binary variable for information revealing status. Furthermore, they presented method for approximating binary decisions by piecewise constant functions and adaptive real-valued decisions by piecewise linear functions of uncertainty. This decision rule approximation leads to a mixed-integer linear programming problem (MILP).

In this work, long term development planning of the drainage areas is addressed. This study is complementary to a previous work [100] where a method for configuring the drainage area layout was developed. First, an optimization framework is built to account for rising, spread and depletion stages of SAGD operation. Furthermore, we incorporate the effect of price and oil production uncertainties into the optimization framework under realistic restrictions. Following the idea proposed by Vayanos et al. [57], in this work, binary variables are used to indicate the revelation of endogenous uncertainty. The goal of this optimization framework is to maximize the profit. The problem is finally solved using linear decision rule techniques.

The rest of the paper is organized as follows. Section 2 provides a general description of the problem. In Section 3 a detailed description and mathematical formulation of oil production and steam injection models in SAGD operation is presented. Multistage deterministic and stochastic development models are explained in Sections 4 and 5, respectively. Section 6 discusses the obtained results and compares deterministic and stochastic solutions. Section 7 concludes the paper.

6.2 Problem statement

In this study, the goal is to determine an optimal reservoir development and oil production plan that maximizes the net present value (NPV) over a long term horizon. DAs have different geological properties such as porosity, permeability and oil saturation. SAGD oil production profile consists of three stages [110, 111] where the duration and the amount oil production at each stage depends on the geological properties of each DA.

The steam required by all drainage areas will be provided by the steam generation facility. Since the steam generation capacity is limited, the steam resources have to be optimally allocated to different drainage areas during the lifetime of the project. In addition, the CPF has a limited capacity for processing the produced bitumen, which accordingly restricts the amount of possible oil production. In summary, bitumen production from multiple drainage areas is restricted by the limited steam generation and oil production capacity of the CPF. Finally, the capital investment is constrained by the available budget for each year.

In this work, it is assumed that any DA development and oil production decision takes immediate effect at the beginning of each year and that oil production uncertainty (ζ_i) is revealed at the middle of the year after DA installation decision is made and installation is completed. Since price for each year is considered to be the average price of the whole year, price uncertainty (ϵ_t) for each year is revealed at the end of the year when the oil price for the whole year is known (Figure 6.2).

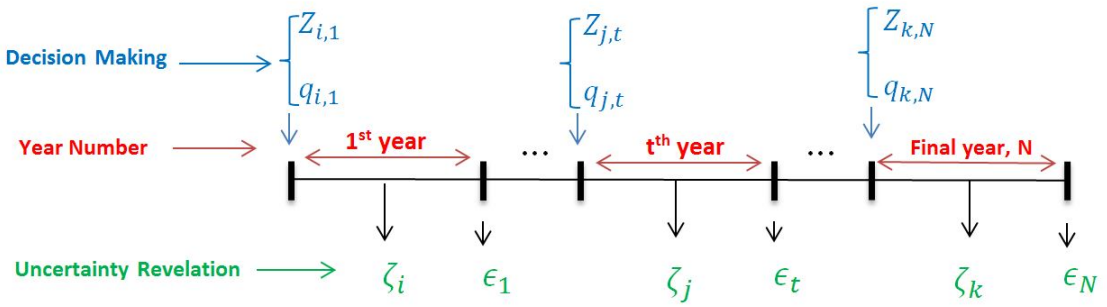


Figure 6.2: Uncertainty revelation and decision making sequence

In the subsequent sections, a deterministic optimization model based on previous work [100] is first developed. Next, a stochastic optimization framework under price and production uncertainty is proposed. In this stochastic model, future oil price and oil production capacity are set as the uncertain parameters. Future oil price does not depend on optimization decisions while oil production capacity depends on reservoir uncertainty which is revealed at each stage based on the well development decisions. Therefore, oil price is an exogenous uncertainty but oil production capacity is an endogenous uncertainty. Using linear decision rule approximation and robust optimization techniques, the stochastic model is converted to a deterministic model. Both deterministic and stochastic solutions provide the time for DA installation, oil production sequencing and production rate for each DA. Deterministic and stochastic models are explained in detail in Section 4 and 6, respectively.

6.3 SAGD process model

The objective of this section is to calculate the oil production capacity and steam injection rate that are required to formulate the deterministic and stochastic optimization problem for DA development. Next, oil production and steam injection models are described and the required parameters are provided and explained. These models are employed to simulate the oil production and steam injection rates for each DA according to its geological properties.

6.3.1 Oil production model

In SAGD operation, the oil production profile mainly consists of three stages: rising, spread and depletion stages (Figure 6.3). In rising stage, steam chamber is gradually expanding in the underground reservoir and oil production is increasing until the steam chamber reaches top of the net pay zone. At this point, steam chamber starts developing laterally and an almost constant oil production profile is observed until most of the bitumen in the pay zone is extracted. At the end of spread stage, oil production rate starts declining and depletion stage starts since most of

the bitumen is produced up to this point. Duration and production rate at each of these three stages vary according to reservoir geological properties. In the following, the mathematical model for each of these three stages [110, 111] is presented and the required parameters are explained.

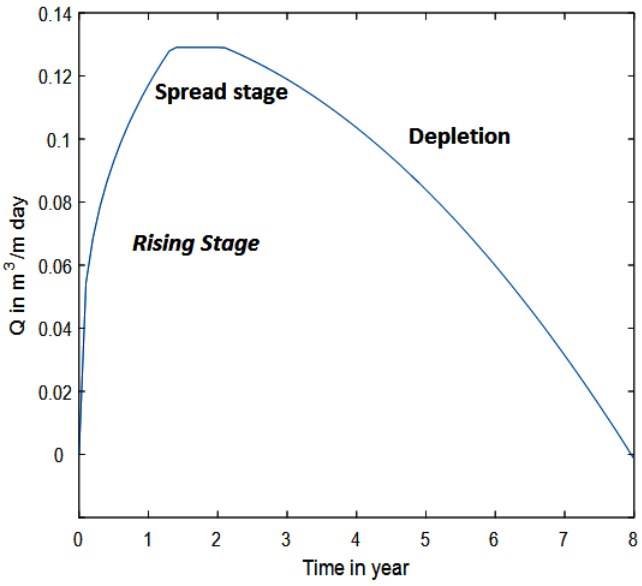


Figure 6.3: Rising, spread and depletion stages of oil production in SAGD operation

The oil flow rate for the rising stage is obtained using the following equation:

$$q_{rising} = 3(\frac{kg\alpha}{m\nu_s})^{\frac{2}{3}}\ (\phi\Delta S_0)^{\frac{1}{3}}\ t^{\frac{1}{3}}CF_1 \tag{6.1}$$

The oil flow rate for the spread stage is calculated as:

$$q_{spread} = 2\sqrt{\frac{\beta kg\alpha\phi\Delta S_0 h}{m\nu_s}}CF_2 \tag{6.2}$$

The oil flow rate for the depletion stage is obtained using Eqs. 6.3 to 6.6:

$$q_{depletion} = 2Q^*/Factor \tag{6.3}$$

$$Q^* = \sqrt{\frac{3}{2}} - t^{*2}\sqrt{\frac{2}{3}} \tag{6.4}$$

$$t^* = \frac{t}{w}\sqrt{\frac{kg\alpha}{\phi\Delta_0 m\nu_s h}} \tag{6.5}$$

$$Factor = \sqrt{\frac{m\nu_s}{kg\alpha h\phi\Delta S_0}}CF_3 \tag{6.6}$$

where q_{rising}, q_{spread} and $q_{depletion}(m^3/m \times day)$ are oil production rates per meter of the well length in rising, spread and depletion stages, respectively. In addition, $k(m^2)$is the reservoir effective permeability, $\alpha(m^2/s)$ is the reservoir thermal diffusivity, $\nu_s(m^2/s)$ is the oil viscosity at steam temperature, $g(m/s^2)$ is the gravitational acceleration, ΔS_o is the displaceable oil saturation, m is the viscosity-temperature correlation parameter, ϕ is the reservoir porosity, $h(m)$ is the effective drainage height, t (second) is time, t^* is dimensionless time, β, γ_1 and γ_2 are empirical constants, CF_1, CF_2 and CF_3 are conversion factors ($CF_1 = (24 \times 3600)^{4/3}$, $CF_2 = 24 \times 3600$, $CF_3 = 1/(24 \times 3600)$) to calculate production rate in $(m^3/m \times day)$.

Eqs. 1 to 6 are used to calculate oil production rates. Using geological data specific to each drainage area which are obtained from Eqs. 14 to 18, the oil production capacity for each drainage area at each year of its lifetime ($q_{i,k}^{maxoil}$) is calculated. The $q_{i,k}^{maxoil}$ values are organized as a numeric table and used in the optimization problem.

6.3.2 Steam consumption model

In this section, an approximate model for steam consumption is presented. The cumulative heat consumption in SAGD well consists of two main components: heat inside the steam chamber H_{inside} and cumulative heat loss H_{loss}. H_{inside} is calculated based on the required energy for expansion of steam chamber (Eq. 6.7). Cumulative heat loss consists of heat loss to overburden, chamber sides and underburden. Heat loss to overburden H_{top} is calculated based on cumulative heat loss to a semi-infinite plane (Eq. 6.8). The heat outside the steam chamber is composed of heat losses from the chamber sides and underburden which can empirically considered to be one-third of the heat losses from the overburden [112]. Therefore, the total cumulative heat consumption H_{total} and total required steam volume V_{steam} is calculated using Eqs. 6.7 to 6.11 [1]:

$$H_{inside} = A\Delta T C_{\nu r} h \eta_s \tag{6.7}$$

$$H_{top} = \frac{4}{3} A\Delta T \sqrt{\frac{k_t C_{\nu o} t}{\pi}} \tag{6.8}$$

$$H_{loss} = H_{top} + \frac{1}{3} H_{top} \tag{6.9}$$

$$H_{total} = H_{inside} + H_{loss} = A\Delta T (C_{\nu r} h \eta_s + \sqrt{k_t C_{\nu o} t}) \tag{6.10}$$

$$V_{steam}(t) = \frac{H_{total}}{H_{l\nu}} = \frac{A\Delta T}{H_{l\nu}} (C_{\nu r} h \eta_s + \sqrt{k_t C_{\nu o} t}) \tag{6.11}$$

Based on the theoretical derivation and experimental work by Butler, [110] the height of the steam chamber can be calculated as a function of time and reservoir geological properties. The steam chamber height is calculated using Eq. 6.12 before it reaches the reservoir top. After reaching the reservoir top, it is replaced by constant effective reservoir height. Finally, the cumulative steam

consumption for each year is obtained using Eq. 6.13.

$$h = 2(\frac{kgk_t}{m\nu_s\phi\Delta S_0 C_{\nu 0}})^{\frac{1}{3}} t^{\frac{2}{3}} \tag{6.12}$$

$$V_{steam}(t) = \frac{H_t}{H_{l\nu}} = \frac{A\Delta t}{H_{l\nu}}(C_{\nu r} h\eta_s + \sqrt{k_t C_{\nu o} t}) \tag{6.13}$$

In the above equations, ΔT is the temperature difference between reservoir temperature and injected steam, A is the planar area of the steam chamber, C_{vr} is the reservoir volumetric heat capacity, C_{vo} is the overburden volumetric heat capacity, h_t is the height of the steam chamber as a function of time, η_s is the effective sweep efficiency, k_t is the overburden thermal conductivity, k is effective reservoir permeability, m is the viscosity-temperature correlation coefficient, H_{lv} is the steam latent heat and t is time since the start of steam injection. Eqs. 7 to 13 are used to calculate the cumulative volume of injected steam $V_{steam}(t)$ for each drainage area up to year t . The formula $V_{steam}(t) - V_{steam}(t-1)$ is used to calculate the volume of injected steam for each drainage area at each year of its life time, $q_{i,k}^{steam}$. These values are organized as a numeric table and used in the optimization problem.

Athabasca reservoir properties [2] were used in order to calculate oil production and steam consumption rates for each DA. Average reservoir porosity, permeability and initial oil saturation are parameters that can be calculated based on the following empirical equations [3]:

$$\phi = \frac{0.25NCB}{DA_{cell}} \tag{6.14}$$

$$A_\nu = \frac{6}{d_p} \tag{6.15}$$

$$k = \frac{1}{2\tau A_\nu^2} \frac{\phi^3}{(1-\phi)^2} \tag{6.16}$$

$$S_w = \frac{100\phi^{2.25}}{\sqrt{k}} \tag{6.17}$$

$$S_o = 1 - S_w \tag{6.18}$$

where NCB is a dimensionless geological parameter that indicates the net continuous bitumen in each DA [100], DA_{cell} is the number of cells in each DA, $A_\nu(m^{-1})$ is the specific area of spherical shape grains, d_p (m) is the average diameter of grains, S_w is water saturation and τ is tortuosity. It is assumed in this work that same number of well pairs were installed in each DA. All the required geologic parameters are provided in supporting information Table S1.

6.4 Deterministic development planning

In this section, a deterministic SAGD drainage area development planning model is proposed. The result of this development plan includes the starting time and oil production rate for each DA during its lifetime. This plan is optimized to maximize the net profit. In the following, constraints and objective function of the optimization model are described. In the model, indices i, t and k refer to DA number, year within the lifetime of the project and year within life time of each DA, respectively.

Production start time is one of the decision variables in the optimization model. It is assumed that any DA installation and production decision can only be made at the start of each year and it is calculated by the following equation:

$$t_i^{start} = \sum_t z_{i,t} t \qquad \forall i \in \mathcal{I} \tag{6.19}$$

where $z_{i,t}$ is a binary variable indicating whether i-th drainage area starts production at year t or not, and t_i^{start} is the start time of i-th DA.

To enforce that each DA production starts only once, the following constraint is applied:

$$\sum_t z_{i,t} \leq 1 \qquad \forall i \in \mathcal{I} \tag{6.20}$$

Since the lifetime of DAs varies, a mathematical mechanism is required to enable shifting lifetime of each DA during the project horizon. The following equation enables shifting lifetime of each DA and ensures that nonnegative oil production rate from each DA ($q_{i,t}^{oil}$) is less than or equal to the capacity ($q_{i,k}^{maxoil} N_w$) calculated from theoretical SAGD model presented above (Eqs. 6.1 to 6.7).

$$q_{i,t}^{oil} \leq \sum_{k \in \mathcal{K}_i} z_{i,t-k+1} q_{i,k}^{maxoil} N_w \qquad \forall t \in \mathcal{T}, i \in \mathcal{I} \tag{6.21}$$

$$q_{i,t}^{oil} \geq 0 \qquad \forall t \in \mathcal{T}, i \in \mathcal{I} \tag{6.22}$$

where $q_{i,t}^{oil}(m^3/year)$ is the oil production rate for each DA during the project time horizon, $q_{i,k}^{maxoil}(m^3/year)$ is the oil production capacity for each well pair during the DA life time.

In addition, CPF capacity is an upper limit for total oil processing volume at each year. Therefore the oil production at each year must always be less than or equal to the CPF capacity.

$$\sum_{i \in \mathcal{I}} q_{i,t}^{oil} \leq Cap_{CPF} \qquad \forall t \in \mathcal{T} \tag{6.23}$$

For each DA, steam injection should start at the same year of oil production. Therefore a similar equation is used for steam injection while a steam consumption restriction (Cap_{steam}) is applied.

$$q_{i,t}^{steam} = \sum_{k \in \mathcal{K}_i} z_{i,t-k+1} q_{i,k}^{steam} N_w \qquad \forall i \in \mathcal{I}, t \in \mathcal{T} \tag{6.24}$$

$$\sum_{i \in \mathcal{I}} q_{i,t}^{steam} \leq Cap_{steam} \qquad \forall t \in \mathcal{T} \tag{6.25}$$

where $q_{i,t}^{steam}(m^3/year)$ is the steam flow rate for each DA during the project time horizon, $q_{i,k}^{steam}(m^3/year)$ represents steam flow rates for each well pair during the DA life time and Cap_{steam} is the yearly steam generation capacity.

Capital cost of the project consists of initial investment for installation of well pairs for all DAs, $DA_{cost}(M\$)$, and initial investment for central processing facility [113]. In this study, it is assumed the CPF is already available in the field. The next constraint imposes a capital restriction for DA installation per year, $CC_{DA}(M\$)$.

$$DA_{cost} \sum_{i \in \mathcal{I}} z_{i,t} \leq CC_{DA} \qquad \forall t \in \mathcal{T} \tag{6.26}$$

Next, the objective function is discussed which is based on the net present value calculation. The first term is the Total Annual Revenue (TAR) which accounts for the profit of selling the oil:

$$TAR_t = (1 - RY) OP_t \sum_{i \in \mathcal{I}} q_{i,t}^{oil} \qquad \forall t \in \mathcal{T} \tag{6.27}$$

where OP_t is the oil price in CAN$\$/m^3$ and RY is royalty parameter. If a company decides to extracts oil from an oil sands field in Alberta, it has to pay a percentage of its revenue as royalty to the government. The royalty value is scaled according to the price of West Texas Intermediate [113, 114].

The main capital cost in the objective function is DA capital investment (CC_{DA}) which is calculated as following:

$$TCC_t = DA_{cost} \sum_{i \in \mathcal{I}} z_{i,t} \qquad \forall t \in \mathcal{T} \tag{6.28}$$

where DA_{cost} is the capital cost for each DA and TCC_t is the required investment for all the DAs per year. As mentioned above, it is assumed that investment for CPF is already been made and the facility is available, so the capital cost of CPF is not counted here.

The cost for steam generation is calculated next. Vaporizing water requires fuel consumption and the cost is obtained by the following equation:

$$TFC_t = FC \sum_{i \in \mathcal{I}} q_{i,t}^{steam} \qquad \forall t \in \mathcal{T} \tag{6.29}$$

where FC is the fuel cost in ($M\$/m^3$) and TFC_t is the total fuel cost in ($M\$/year$). Finally, the objective is to maximize the NPV defined using the following equation:

$$\max NPV = \sum_{t \in \mathcal{T}} \frac{1}{(1+DR)^t} [TAR_t - TCC_t - TFC_t] \tag{6.30}$$

where the term $1/(1+DR)^t$ is the discount cash flow term required to calculate NPV and DR is the discount rate. The overall multiperiod deterministic planning (MSDP) model includes Eqs. 19 to 30. It is a mixed integer linear optimization problem.

6.5 Uncertainty

In this work, it is assumed that for each DA, production capacity follows a uniform distribution around a nominal value. For oil production capacity uncertainty, a perturbation ζ_i is considered based on the theoretical value calculated using geological parameters. This oil production uncertainty reflects uncertainty in different reservoir geologic parameters such as porosity, permeability, saturation, etc.

$$\tilde{q}_{i,k}^{maxoil} = q_{i,k}^{maxoil}(1+\zeta_i), \quad \forall i \in \mathcal{I}, \forall k \in \mathcal{K}_i \tag{6.31}$$

where $q_{i,k}^{maxoil}$ is the nominal value for oil production and $\tilde{q}_{i,k}^{maxoil}$ is the uncertain oil production of i-th DA at the k-th year of its lifetime. It is worth pointing out that the production uncertainty is an endogenous uncertainty in this work. That is, the uncertainty will be revealed once the well is developed.

As a manipulated variable for SAGD well control, the steam injection rate is determined and controlled by the company that manages the SAGD operation. Steam injection rate is known in this sense. Therefore, it is not considered as a source of uncertainty. The formula provided in this paper for steam consumption rate is an estimate of the actual field values. However, the oil production rate is dependent on the oil reserve capacity underground, which is not accurately known to the operator and is considered as an uncertain parameter.

To model the oil price uncertainty, a time series model structure, Auto Regressive Moving

Average $ARMA(p, q)$, is used:

$$OP_t = \{\phi_1 OP_{t-1} + \cdots + \phi_p OP_{t-p}\} + \{\epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}\} \tag{6.32}$$

where the parameters $\phi_1, \cdots, \phi_p, \theta_1, \cdots, \theta_q$ can be estimated from history data [115] as shown in Figure 6.4, ϵ_t represents uncorrelated white noise (with zero mean and constant variance).

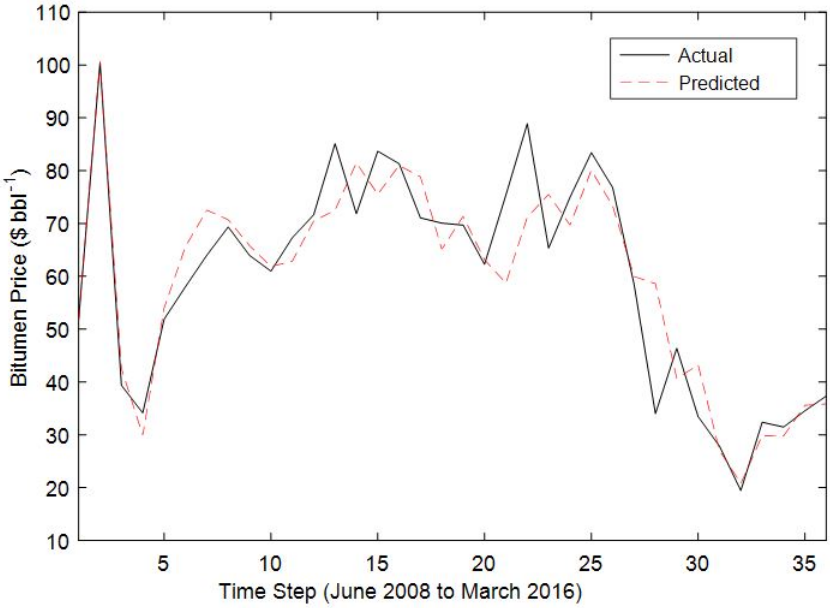


Figure 6.4: ARMA training using historical bitumen price data

$ARMA$ is a practical tool for modeling and predicting time series data. It provides a description of a stationary stochastic process based on two polynomials: autoregression (AR) and moving average (MA). The AR polynomial regresses the variable on its own lagged (past) values $\{\phi_1 OP_{t-1} + \cdots + \phi_p OP_{t-p}\}$ and the MA part models the error term as a linear combination of error terms at the same and previous time periods $\{\epsilon_t + \theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}\}$. The model is commonly represented as $ARMA(p, q)$ where p and q are the orders of the AR and MA polynomials respectively.

To predict the future oil price from time $t + 1$ to $t + T$, we define $\epsilon = [\epsilon_{t+1}, \cdots, \epsilon_{t+T}]^T$. Then the future oil price can be compactly written as:

$$OP_{t+i} = A_{t+i}\epsilon + b_{t+i}, \quad i \in \{1, \cdots, T\} \tag{6.33}$$

where A_{t+i} is a vector, b_{t+i} is a scalar, both calculated using the model parameters, past noise and past oil price data.

In order to demonstrate the derivation procedure for future oil price prediction (Eq. 6.33), the

derivation for a simple case of $p = 1$ and $q = 1$ is presented here. By assuming the process starts at time t=0, oil price for next two time steps 1 and 2 is calculated. According to Eq. 6.32, oil price for the current and the next two time steps can be written as:

$$OP_0 = \phi_1 OP_{-1} + \epsilon_0 + \theta_1 \epsilon_{-1} \tag{6.34}$$

$$OP_1 = \phi_1 OP_0 + \epsilon_1 + \theta_1 \epsilon_0 \tag{6.35}$$

$$OP_2 = \phi_1 OP_1 + \epsilon_2 + \theta_1 \epsilon_1 \tag{6.36}$$

In Eqs. 6.34 to 6.36, the goal is to find oil price for the next two time steps OP_1 and OP_2; parameters θ_1 and ϕ_1 are known parameters obtained from ARMA model training; OP_{-1} and OP_0 are the oil price for the previous and current years which are known; ϵ_{-1} can be set as zero because of zero mean assumption; ϵ_0 can be calculated using Eq. 6.34; ϵ_1 and ϵ_2 are uncertain parameters for future time steps. In order to obtain OP_1, ϵ_0 is calculated from Eq. 6.34 and then it is plugged into Eq. 6.35.

$$\epsilon_0 = OP_0 - \phi_1 OP_{-1} \tag{6.37}$$

$$OP_1 = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} \epsilon_1 \\ 0 \end{bmatrix} + (\phi_1 + \theta_1 OP_0 - \phi_1 \theta_1 OP_{-1}) \tag{6.38}$$

In order to obtain OP_2, OP_1 from Eq. 6.38 is plugged into Eq. 6.36.

$$OP_2 = \begin{bmatrix} \phi_1 + \theta_1 & 1 \end{bmatrix} \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \end{bmatrix} + (\phi_1^2 + \theta_1 \phi_1) OP_0 - \phi_1^2 \theta_1 OP_{-1} \tag{6.39}$$

In Eqs. 6.38 and 6.39, all the parameters are known. For larger values of p and q, a similar procedure is followed and oil price for any future time step can be calculated. The general formula for future price prediction is presented in Eq. 6.33.

The overall uncertainty vector is defined as:

$$\xi = [1, \zeta_1, \cdots, \zeta_N, \epsilon_1, \cdots, \epsilon_T]^T \tag{6.40}$$

where the first element 1 is used to simplify the linear decision rule expression. To use the above uncertainty vector for oil flow rate, the row vector e_i is defined such that $e_i \xi = \zeta_i$ (all elements of e_i are zero except the $i + 1$-th element is 1). Using e_i , the oil production uncertainty can be written as:

$$\tilde{q}_{i,k}^{maxoil}(\xi) = q_{i,k}^{maxoil}(1 + e_i \xi), \quad \forall i \in \mathcal{I} \tag{6.41}$$

For oil price, a truncate operator P is defined such that: $P\xi = [\epsilon_1, \cdots, \epsilon_T]^T$, then oil price uncertainty can be written as:

$$OP_t(\xi) = A_t P\xi + b_t, \quad \forall t = 1, \cdots, T \tag{6.42}$$

Uncertainty set

For the stochastic problem, the following uncertainty set is constructed:

$$\begin{aligned}
\Xi = \{\xi = [1, \zeta_1, \cdots, \zeta_N, \epsilon_1, \cdots, \epsilon_T]^T : \\
\zeta_i \in [\underline{\zeta}_i, \bar{\zeta}_i], \qquad \forall i, & \quad (6.43a)\\
|\epsilon_t| \leq z_{1-\alpha}, \qquad \forall t, & \quad (6.43b)\\
\sum_t |\epsilon_t| \leq \Gamma z_{1-\alpha}, & \quad (6.43c)\\
A_t\epsilon + b_t \geq 10, \qquad \forall t\} & \quad (6.43d)
\end{aligned}$$

The constraints applied in the uncertainty set are explained as following:

- For each drainage area i, the production capacity varies in a interval $\zeta_i \in [\underline{\zeta}_i, \bar{\zeta}_i]$.
- For oil price, the variation range of white noise is controlled by applying conditions on each ϵ_t such that $|\epsilon_t| \leq z_{1-\alpha}$, where $z_{1-\alpha}$ is the $1-\alpha$ confidence level for normal distribution.
- The total change of budget is controlled by: $\sum_t |\epsilon_t| \leq \Gamma z_{1-\alpha}$, where Γ is a scalar to control range of white noise change, t is time in project horizon.
- Considering minimum oil price \$10: $OP_t(\xi) = A_t P\xi + b_t \geq 10, \forall t$.

The above uncertainty set is a convex polyhedron. In compact matrix notation it is written as:

$$\Xi = \{\xi : W\xi \geq \nu\}$$

In order to investigate effect of uncertainty set size on predicted oil price, 3 different combinations of Γ and CI were selected (Γ=25 & CI=99%, Γ=12 & CI=95%, Γ=7, CI=90%) and the average oil price $\pm$ std is plotted. As Γ and CI move toward smaller values, the size of price uncertainty set reduces and considering that a minimum 10\$ price is applied, reduced mean oil price and standard deviation is observed such that the uncertainty set corresponding to $\Gamma = 25$ and CI=99% have the highest mean oil price and standard deviation and $\Gamma = 7$ and CI=90% have the lowest values (Figure 6.5). Figure 6.5 was generated based on the ARMA model trained from history bitumen price data. The mean price represents the average price of 1000 random samples for next 25 years of bitumen price predicted using the ARMA model. Notice that for a single sample of price prediction, the profile is not necessarily declining. This average trend is reflecting

the latest history price declining trend in the bitumen price between June 2008 and March 2016 as shown in Figure 6.4.

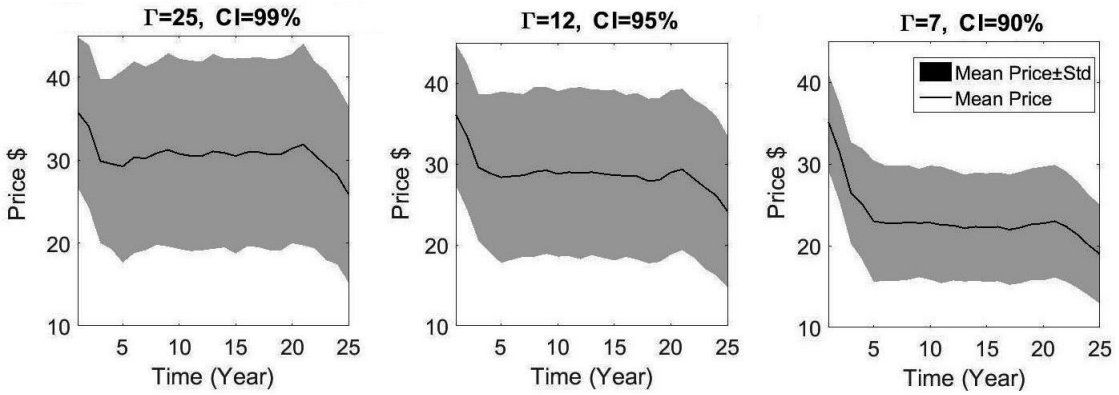


Figure 6.5: Effect of uncertainty set size on predicted bitumen price

6.6 Stochastic planning

In multistage stochastic programming, uncertainty in bitumen price and reservoir capacity is included in the model. Before drilling each oil well in SAGD operation, there is only an initial estimate about the reservoir property and after well completion, the uncertainty is revealed. Future oil price is also uncertain and neglecting price uncertainty can deteriorate planning outcome. Therefore, in long term reservoir development, considering reservoir property and price uncertainty can provide a feasible and optimal development plan that significantly improves the planning results while neglecting uncertainty may lead to infeasible plans. In the following section, multistage stochastic programming (MSSP) formulation are presented to tackle the oil price and reservoir production uncertainty.

6.6.1 Multistage stochastic programming model

In the deterministic planning model, it is assumed that there is no uncertainty in the model and all parameters take deterministic value that is known. However, in stochastic modeling at each stage only the uncertain information up to previous stage are known and the present and future uncertainty are unknown. The constraints that reflect this fact are referred to as non-anticipativity constraints. To summarize, the difference between stochastic and deterministic models is the price and oil production uncertainty and non-anticipativity constraints. Eqs. (6.44a) to (6.44m) present

the stochastic model:

$$\max \quad \sum_{t\in\mathcal{T}} \frac{1}{(1+DR)^t}\mathbb{E}\left[(1-RY)OP_t(\xi)\sum_{i\in\mathcal{I}} q_{i,t}^{oil}(\xi) - DA_{cost}\sum_{i\in\mathcal{I}} z_{i,t}(\xi) - FC\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}_i} z_{i,t-k+1}(\xi)q_{i,k}^{steam}N_w\right] \tag{6.44a}$$

$$s.t. \quad DA_{cost}\sum_{i\in\mathcal{I}} z_{i,t}(\xi) \le CC_{DA} \qquad \forall t\in\mathcal{T}, \xi\in\Xi \tag{6.44b}$$

$$\sum_{t\in\mathcal{T}} z_{i,t}(\xi) \le 1 \qquad \forall i\in\mathcal{I}, \xi\in\Xi \tag{6.44c}$$

$$q_{i,t}^{oil}(\xi) \le \sum_{k\in\mathcal{K}_i} z_{i,t-k+1}(\xi)\,\tilde{q}_{i,k}^{maxoil}(\xi)N_w \qquad \forall i\in\mathcal{I}, t\in\mathcal{T}, \xi\in\Xi \tag{6.44d}$$

$$\sum_{i\in\mathcal{I}} q_{i,t}^{oil}(\xi) \le Cap_{CPF} \qquad \forall t\in\mathcal{T}, \xi\in\Xi \tag{6.44e}$$

$$\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}_i} z_{i,t-k+1}(\xi)\,q_{i,k}^{steam}N_w \le Cap_{steam} \qquad \forall t\in\mathcal{T}, \xi\in\Xi \tag{6.44f}$$

$$q_{i,t}^{oil}(\xi) \ge 0 \qquad \forall i\in\mathcal{I}, t\in\mathcal{T}, \xi\in\Xi \tag{6.44g}$$

$$x_{1,t}(\xi) = 1 \qquad \forall t\in\mathcal{T} \tag{6.44h}$$

$$x_{i+1,t}(\xi) = \sum_{t'\le t} z_{i,t'}(\xi) \qquad \forall i\in\mathcal{I}, t\in\mathcal{T}, \xi\in\Xi \tag{6.44i}$$

$$x_{l,t}(\xi) = 0 \qquad \forall l \ge N_{DA}+3, t < l-(1+N_{DA}), \xi\in\Xi \tag{6.44j}$$

$$x_{l,t}(\xi) = 1 \qquad \forall l \ge N_{DA}+2, t \ge l-(1+N_{DA}), \xi\in\Xi \tag{6.44k}$$

$$q_{i,t}^{oil}(\xi) = q_{i,t}^{oil}(x_{t-1}(\xi)\bullet\xi) \qquad \forall i\in\mathcal{I}, t\in\mathcal{T}, \xi\in\Xi \tag{6.44l}$$

$$z_{i,t}(\xi)\in\{0,1\}, x_{l,t}(\xi)\in\{0,1\} \qquad \forall i\in\mathcal{I}, t\in\mathcal{T}, \xi\in\Xi \tag{6.44m}$$

Here, new binary variable x is introduced, which denotes the information revealing status of the uncertainty. Constraint (6.44h) is for the trivial element 1 in the uncertain vector ξ. The non-anticipativity relation between information revealing status variable x and development decision z is modeled by constraint (6.44i). For price uncertainty, the information revealing status variable is only dependent on the time, this is modeled by constraint (6.44j) and (6.44k). Constraint (6.44l) represents non-anticipativity constraints for production decision variables (it is only decided based on information revealed) and and sign ($\bullet$) in Eq. (6.44l) represents Hadamard product.

6.6.2 Linear decision rule based solution method

Multistage stochastic model includes constraints with the uncertain parameter ξ, and each of these constraints has to be satisfied for infinite number of uncertain ξ values. This stochastic optimiza-

tion model is intractable. One method to simplify the problem is to employ a decision rule and approximate the adjustable variable using affine function of the uncertainty. This leads to the so-called linear decision rule, which is used in this work to solve the stochastic problem. The result of the LDR solution is a set of linear decision rules rather than exact values which approximates the optimal solution. Affine decision rule method has received considerable application in robust optimization where the objective is evaluated based on the worst case performance over the uncertainty set. In this study, the method is applied to multistage stochastic programming model where the objective is based on the expected performance over the uncertainty set. In comparison to traditional scenario (tree) based stochastic programming, LDR method seeks solution feasibility over an uncertainty set instead of finite number of scenarios.

In this work, we do not use the scenario based approach since this problem has relative large number of stages (for example 25 stages in the case study of next section). This will results in a scenario tree with very large size even if two branches are considered for each node. Compared to scenario based method, the LDR method leads to less model complexity. In addition, it provides an efficient way to deal with endogenous uncertainty.

In the proposed linear decision rule method of this paper, the binary variables are considered to be constant function of uncertainty; therefore, the following decision rule is used:

$$z_{i,t}(\xi) = z_{i,t}, \quad \forall \xi \in \Xi, \forall i \in \mathcal{I}, t \in \mathcal{T} \tag{6.45}$$

$$x_{l,t}(\xi) = x_{l,t}, \quad \forall \xi \in \Xi, \forall l \in \mathcal{L}, t \in \mathcal{T} \tag{6.46}$$

where $z_{i,t} \in \{0,1\}, x_{l,t} \in \{0,1\}$.

Continuous variable $q^{oil}_{i,t}(\xi)$ is considered to be a linear function of uncertainty, hence the following linear decision rule is employed:

$$q^{oil}_{i,t}(\xi) = Q_{i,t}\xi, \quad \forall \xi \in \Xi, \forall i \in \mathcal{I}, t \in \mathcal{T} \tag{6.47}$$

where $Q_{i,t}$ is a row vector $[Q_{i,t,1}, \cdots, Q_{i,t,1+N+T}]$.

Non-anticipativity

As described in Section 6.1, non-anticipativity of uncertain information means that at each time step, uncertain information up to previous time step are revealed and therefore decisions depend on uncertain information only up to previous time step. At the first step, no uncertainty is revealed, therefore the coefficient of LDR solution does not depend on any uncertainty and it is set as zero. For time step greater than 1, the decision at each time step, only depends on uncertainties revealed up to previous time step. To enforce non-anticipativity, $Q_{i,t,l}$ must satisfy:

$$Q(i,1,l) = 0 \qquad \forall i \in \mathcal{I}, l \in \mathcal{L} \tag{6.48}$$

$$|Q_{i,t,l}| < M x_{l,t-1}, \quad \forall i \in \mathcal{I}, l \in \mathcal{L}, t \geq 2 \tag{6.49}$$

where M is a big constant (can be set as the maximum production capacity). For the trivial element 1 in the uncertain vector ξ, we have:

$$x_{1,t} = 1, \quad \forall t \in \mathcal{T} \tag{6.50}$$

The non-anticipativity relation between information revealing status variable x and development decision z is defined as:

$$x_{i+1,t} = \sum_{t' \leq t} z_{i,t'}, \quad \forall i \in \mathcal{I}, t \in \mathcal{T} \tag{6.51}$$

For price uncertainty, the information revealing status variable only depends on time.

$$x_{l,t} = 0, \quad \forall l \geq N_{DA} + 3, t < l - (1 + N_{DA}) \tag{6.52}$$

$$x_{l,t} = 1, \quad \forall l \geq N_{DA} + 2, t \geq l - (1 + N_{DA}) \tag{6.53}$$

Constraints 6.52 and 6.53 describe the price uncertainty revelation status for each year. The price uncertainty of each year is revealed at the same year but it is unknown at previous years. The numbers 2 and 3 are required to adjust the price uncertainty revelation for each year. Finally, the following matrix represents the structure of the x variable.

$$[x_{lt}] = \begin{bmatrix} 1 & 1 & ... & 1 \\ z_{1,1} & z_{1,1} + z_{1,2} & ... & \sum_t z_{1,t} \\ z_{2,1} & z_{2,1} + z_{2,2} & ... & \sum_t z_{2,t} \\ : & : & \ddots & : \\ z_{N,1} & z_{N,1} + z_{N,2} & ... & \sum_t z_{N,t} \\ 1 & 1 & ... & 1 \\ 0 & 1 & ... & 1 \\ : & : & \ddots & : \\ 0 & 0 & ... & 1 \end{bmatrix} \tag{6.54}$$

The second half of the matrix in Eq. 6.54 illustrates the price uncertainty revelation at each year. For example the last row of the matrix corresponds to price uncertainty of the final year. The price uncertainty of the final year is only revealed at the last year. Therefore all the elements of the last row are zero except the last column which corresponds to the final year.

Overall deterministic counterpart model

For each semi-infinite constraint, it is converted to deterministic counterpart based on duality

principle. Detailed procedure on robust counterpart derivation of the constraints and the expectation of the objective function with respect to the uncertainty are provided in the supporting information. The overall deterministic counterpart of the stochastic model is expressed in Eqs. 6.55a to 6.55n. The deterministic planning model and the LDR based stochastic planning model are both mixed integer linear optimization problem, which can be solved using standard MILP solver.

$$\max \quad \sum_{t\in\mathcal{T}} \frac{1-RY}{(1+DR)^t}\left[tr(P^T A_t^T \sum_{i\in\mathcal{I}} Q_{i,t}\,\mathbb{E}(\xi\xi^T)\,) + b_t \sum_{i\in\mathcal{I}} Q_{i,t}\mathbb{E}(\xi)\right] - \sum_{t\in\mathcal{T}} \frac{DA_{cost}}{(1+DR)^t}\sum_i z_{i,t}$$
$$- \sum_{t\in\mathcal{T}} \frac{N_w FC}{(1+DR)^t}\sum_i \sum_{k\in\mathcal{K}_i} z_{i,t-k+1} q_{i,k}^{steam} \tag{6.55a}$$

$$s.t. \quad DA_{cost}\sum_{i\in\mathcal{I}} z_{i,t} \le CC_{DA} \qquad \forall t\in\mathcal{T} \tag{6.55b}$$

$$\sum_{t\in\mathcal{T}} z_{i,t} \le 1 \qquad \forall i\in\mathcal{I} \tag{6.55c}$$

$$-\nu\Lambda_{i,t} - \sum_{k\in\mathcal{K}_i} z_{i,t-k+1}\, q_{i,k}^{maxoil} N_w \le 0 \qquad \forall i\in\mathcal{I}, t\in\mathcal{T} \tag{6.55d}$$

$$-W\Lambda_{i,t} = Q_{i,t} - \sum_{k\in\mathcal{K}_i} z_{i,t-k+1}\, q_{i,k}^{maxoil} N_w e_i \qquad \forall i\in\mathcal{I}, t\in\mathcal{T}$$

$$\Lambda_{i,t} \ge 0 \qquad \forall i\in\mathcal{I}, t\in\mathcal{T}$$

$$-\nu\lambda_t \le Cap_{CPF} \qquad \forall t\in\mathcal{T} \tag{6.55e}$$

$$-W\lambda_t = \sum_{i\in\mathcal{I}} Q_{i,t} \qquad \forall t\in\mathcal{T}$$

$$\lambda_t \ge 0 \qquad \forall t\in\mathcal{T}$$

$$\sum_{i\in\mathcal{I}}\sum_{k\in\mathcal{K}_i} z_{i,t-k+1}\, q_{i,k}^{steam} N_w \le Cap_{steam} \qquad \forall t\in\mathcal{T} \tag{6.55f}$$

$$\nu\mu_{i,t} \ge 0 \qquad \forall i\in\mathcal{I}, t\in\mathcal{T} \tag{6.55g}$$

$$W\mu_{i,t} = Q_{i,t} \qquad \forall i\in\mathcal{I}, t\in\mathcal{T}$$

$$\mu_{i,t} \ge 0 \qquad \forall i\in\mathcal{I}, t\in\mathcal{T}$$

$$x_{1,t} = 1 \qquad \forall t\in\mathcal{T} \tag{6.55h}$$

$$x_{i+1,t} = \sum_{t'\le t} z_{i,t'} \qquad \forall i\in\mathcal{I}, t\in\mathcal{T} \tag{6.55i}$$

$$x_{l,t} = 0 \qquad \forall l \ge N_{DA}+3, t < l-(1+N_{DA}) \tag{6.55j}$$

$$x_{l,t} = 1 \qquad \forall l \ge N_{DA}+2, t \ge l-(1+N_{DA}) \tag{6.55k}$$

$$Q(i,1,l) = 0 \qquad \forall i \in \mathcal{I}, l \in \mathcal{L} \tag{6.55l}$$

$$|Q_{i,t,l}| \leq M x_{l,t-1} \qquad \forall i \in \mathcal{I}, l \in \mathcal{L}, t \geq 2 \tag{6.55m}$$

$$z_{i,t} \in \{0,1\}, x_{l,t} \in \{0,1\} \tag{6.55n}$$

6.7 Case study

In this section, an example problem is used to investigate the proposed planning method using deterministic and stochastic optimization models. There are 43 DAs to be developed in total, each DA includes 6 pairs of wells and the project time horizon is considered to be 25 years. Oil production and steam injection tables are calculated in MATLAB and it was found that life time of the DAs vary between 6 to 10 years while the life time of each drainage area depends on the amount of bitumen available in each DA. NCB parameter used in Eq. 6.14 reflect the amount of available bitumen in each DA which varies between 307 to 504 in this problem [100]. The length and width of each DA is 850 m and 600 m respectively. In this study, it is assumed that DA installation and production decisions are made at the beginning of each year, therefore there are 25 stages. All the relevant parameters for the deterministic model are provided in supporting information file Table S2.

Table 6.1: Statistic of the deterministic and stochastic models

Statistics	Stochastic Model	Deterministic Model
Number of constraints	1303962	3580
Number of variables	1067144	4537
Optimality gap	1%	1%
Run time	Max: 4.5 hr; Min: 500 sec	Max: 0.724 sec

All the optimization problems are modeled and solved in GAMS 23.9 using the CPLEX solver 12.71. Table 6.1 reports the models' statistics for deterministic and stochastic models. Relative optimality gap of 1% was considered as the stopping criteria which was met at all the runs. Using Intel(R) Core(TM) i5-3470 CPU @ 3.20 GHZ, 8.00 GB RAM and 64 bit windows operating system, the minimum and maximum run time for the stochastic model are 500 seconds and 4.5 hours, respectively. While the maximum run time for the deterministic model is only 0.724 seconds. As Table 6.1 indicates, number of constraints and variables for the stochastic model is much larger than the deterministic model.

6.7.1 Deterministic planning solution

Deterministic planning model was solved first. Figure 6.6 illustrates DA development sequence, TNPV (NPV up to year t), number of operating wells and total oil production at each year for the specific scenario of CPF= $5.25 \times 10^6 m^3/year$ and steam capacity of $16 \times 10^6 m^3/year$. It

can be observed that generally drainage areas with larger bitumen in place (DAs 35 to 43) are developed at the beginning years (years 1 to 10) and those with smaller bitumen in place (DAs 1 to 10) are developed at the final years which is the expected and reasonable trend and DAs are scattered to maximize the revenue while satisfying the constraints. From the beginning up to year 15, number of operating wells and total oil production per year increases; from year 15 to 25, number of operating wells decline and the slope of TNPV curve testifies this trend.

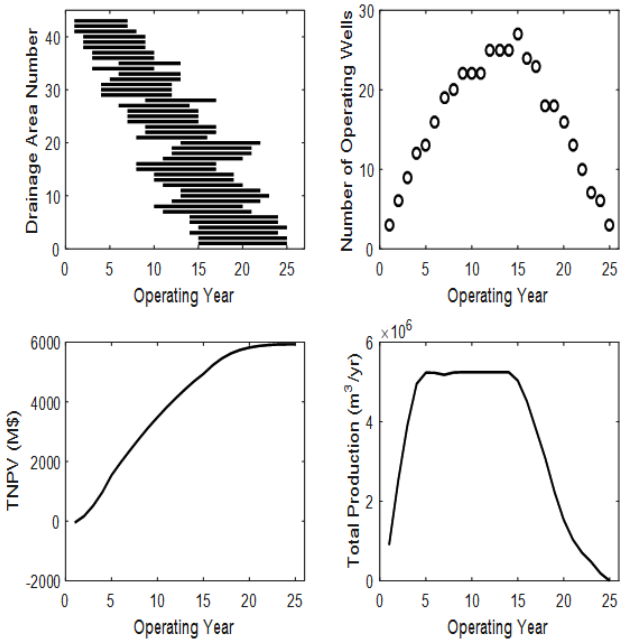


Figure 6.6: Deterministic planning solution

Next, we evaluate the performance of deterministic planning solution through sampling. In all the cases, CPF capcity is 5.25 $\times 10^6$ $m^3/year$ and steam generation capacity is 16×10^6 $m^3/year$. For each Γ value, 1000 samples from the corresponding uncertainty set are generated and sampling method is applied on the results of stochastic and deterministic solutions. Confidence interval and flow perturbation factors are fixed at 95% and 10% accordingly and Γ is varied in the range of 5 to 25.

As Table 6.2 demonstrates, deterministic solution always results in considerable percentage of oil flow constraint violations among the 1000 samples. It should be clarified at this point that in deterministic sampling, for each flow constraint violation, the infeasible flow rate is considered to be zero in NPV calculation. This replacement has resulted in significantly smaller NPV from deterministic solution compared to stochastic solution.

Table 6.2: Sampling evaluation results of deterministic solution

Γ	NPV(M$)	Mean Violations
Γ=25	1711	16.62%
Γ=15	1711	16.62%
Γ=10	1161	16.69%
Γ=7	1142	16.64%
Γ=5	756	16.72%

6.7.2 Stochastic planning solution

In order to investigate the effect of price uncertainty on NPV, 15 experiments were conducted using different Γ and confidence interval (CI) values. Table 6.3 illustrates the obtained results. For greater values of Γ and CI, the price uncertainty set is wider and since a minimum price constraint (bitumen price> 10$/*bbl*) is applied in the uncertainty set, wider price uncertainty set results in greater bitumen price values and the NPV increases accordingly. Figure 6.7 illustrates the stochastic planning solution for extreme and middle cases of uncertainty set parameters.

Table 6.3: Effect of price uncertainty set size on NPV objective of the stochastic model

	CI=99%	CI=95%	CI=90%
Γ=25	6098.88	5688.20	5364.91
Γ=15	6110.96	5690.01	5363.14
Γ=10	6061.87	5562.50	5025.23
Γ=7	5482.12	4678.88	4216.94
Γ=5	4535.58	3902.16	3659.33

Note that the results presented in Table 6.3 is based on stopping criterion of 1% optimality gap. If we reset the optimality gap to zero and solve the problem to optimality, then the objective is expected to decrease as Γ decrease under the same CI value.

To study the effect of reservoir and price uncertainty on stochastic optimization results, 1000 samples from the widest uncertainty set (10% bitumen flow perturbation, Γ=25, CI=99%, Bitumen Price>10 $) were generated and applied on stochastic results. Table 6.4 illustrates the results of sampling method. It shows a decreasing trend of NPV from wider uncertainty set (greater values of Γ and CI) towards narrower uncertainty set. Solution obtained from narrower uncertainty set (smaller Γ and CI values) encloses smaller price uncertainty set and results in smaller NPV. Figure 6.8 illustrates sampling NPV distribution for a specific uncertainty set. In the stochastic solution, changing the size of uncertainty set (Γ,CI) will finally affect the planning of the drainage

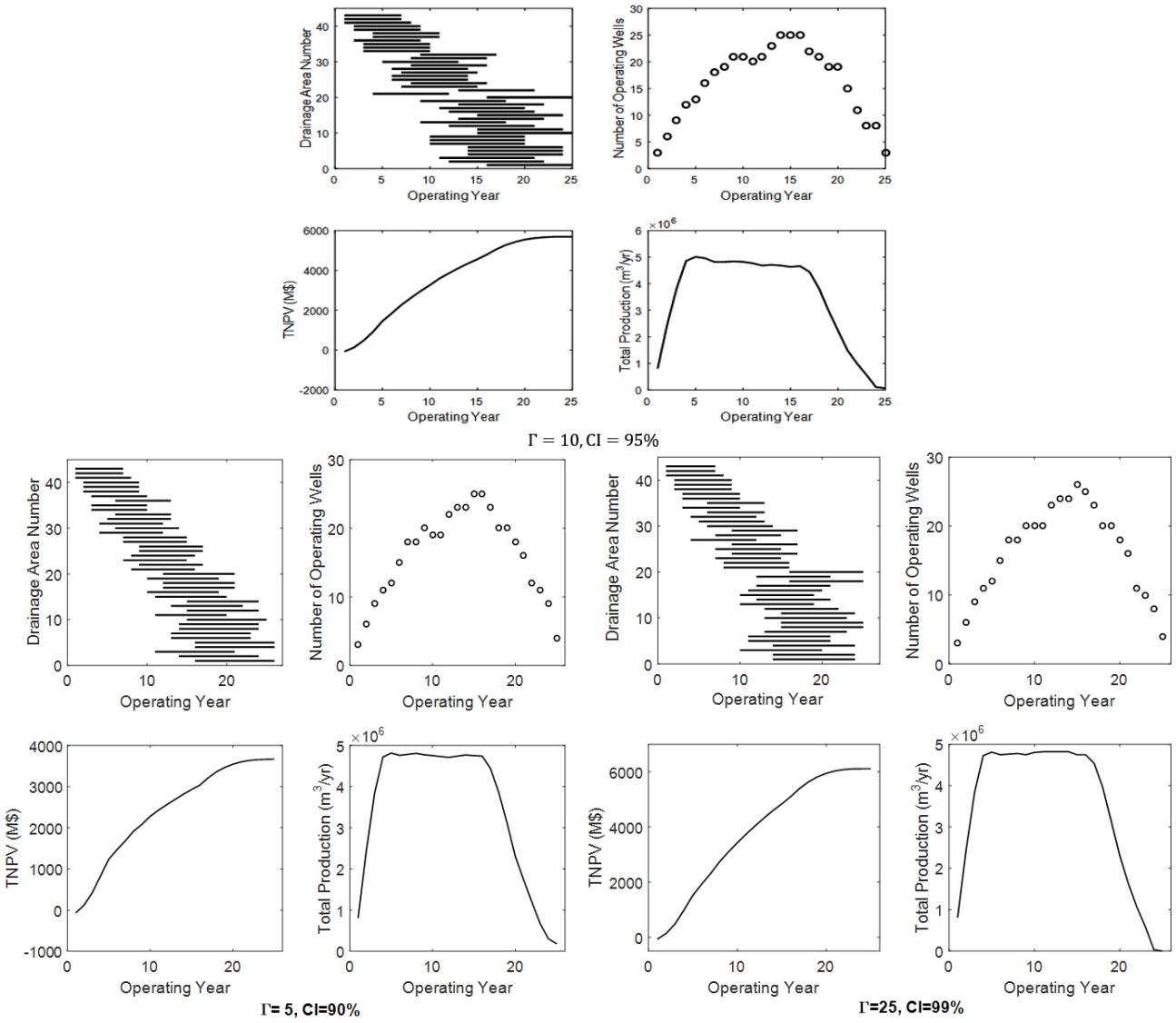


Figure 6.7: Stochastic planning solution

areas in SAGD operation during the project time horizon. Smaller uncertainty sets restrict the price variations which results in lower NPV and vice versa. More description about the effect of uncertainty set on stochastic solution is provided in Section 7.3 and 7.4.

Table 6.4: Sampling based evaluation of stochastic solution (each based on the samples from the corresponding uncertainty set)

	CI=99%	CI=95%	CI=90%
Γ=25	6099	5845	5857
Γ=15	6111	5857	5854
Γ=10	5861	5859	5856
Γ=7	5860	5852	5857
Γ=5	5867	5855	5854

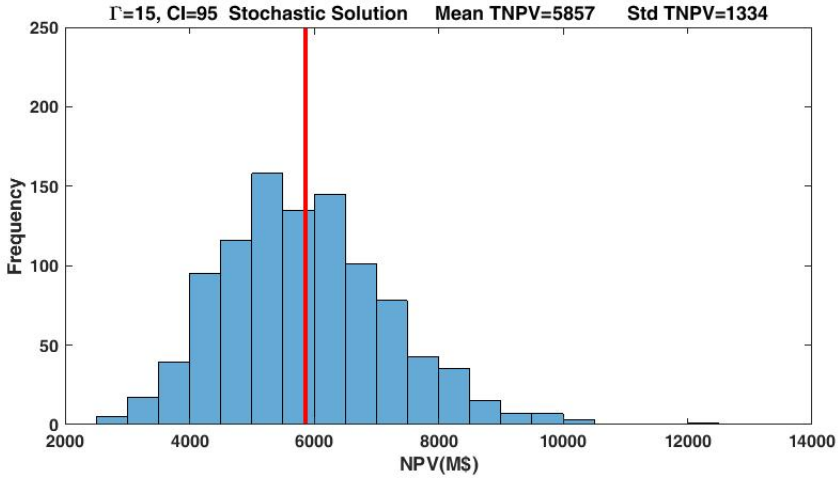


Figure 6.8: NPV distribution for $\Gamma = 15$ and CI=95%

In order to evaluate the performance of the stochastic solutions in terms of feasibility subject to the widest uncertainty set, the sampling method is used. The following factors affect the uncertainty set: Γ and CI affect price uncertainty and bitumen flow perturbation affects production uncertainty. In order to obtain the most comprehensive assessment, 1000 samples from the widest uncertainty set (CI=99% and $\Gamma = 25$, flow perturbation is always kept at 10%) were generated and applied to the solutions obtained from different combinations of CI and Γ. During sampling procedure, the constraint for maximum bitumen production was violated for some of the samples. Figure 6.9 illustrates the mean number of constraint violations and Table 6.5 presents the mean percentage of constraint violations for 1000 samples. Solutions obtained from narrower uncertainty set may not be feasible for uncertainty values outside of their uncertainty set and result in greater number of violations. For instance, solution obtained from $\Gamma = 25$ and CI=99% shows no violations while solution obtained from $\Gamma = 7$ and CI=90% has resulted in the highest percentage of violations.

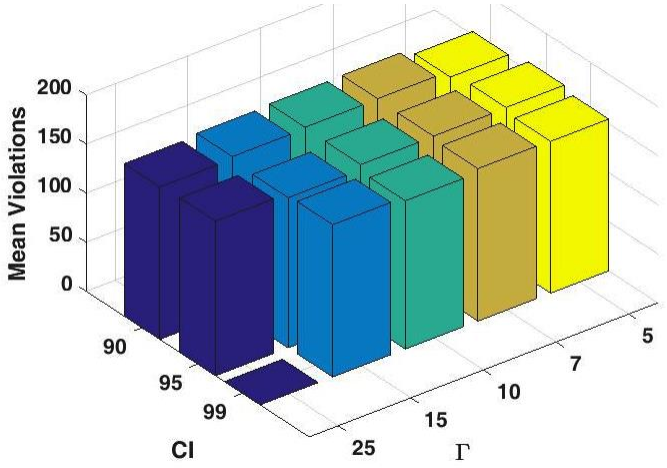


Figure 6.9: Average number of violations

Table 6.5: Mean percentage of constraint violations in sampling (all based on the same set of samples from the largest uncertainty set)

	CI=99%	CI=95%	CI=90%
Γ=25	0	14.64%	14.43%
Γ=15	14.42%	14.22%	14.66%
Γ=10	14.04%	14.72%	14.80%
Γ=7	14.44%	14.76%	14.86%

In the stochastic solution, changing CI and Γ values will finally affect the planning of the drainage areas. While smaller uncertainty sets restrict the price variations, it results in lower NPV and more constraint valuations. As it is described in Section 7.2, the stochastic solution obtained from the largest uncertainty set provides the highest NPV and least constraint violations subject to samples obtained from the widest uncertainty set which is preferable from profit and safety point of views.

6.7.3 Sub-optimality of LDR solution

Sub-optimality of LDR solution is investigated based on continuous and binary components of the solution. In order to study suboptimality of continuous LDR solution, two different solutions are compared: complete MSSP solution and partial MSSP solution. Stochastic solution consists of two parts: continuous variable $q(\xi)$ and binary variable z. In *complete MSSP solution*, the decision rule solution (z and $q(\xi)$) are fixed. Then using sample data from the uncertainty set (Γ=10,CI=95%), the obtained objective function value is 5562 M$. For *partial MSSP solution*, the binary variables are fixed to get a linear programming (LP) problem. For 1000 uncertainty samples (from CI=95% and $\Gamma = 10$ set), the resulting LP problem is called 1000 times from MATLAB and solved using

GAMS for each uncertainty sample. The final objective function value is 5655 M$. Hence, a gap of 93 M$ is observed for continuous LDR solution.

Sub-optimality of binary LDR Solution is examined by comparing two solutions: Partial MSSP solution and Partial MSDP Solution. For *Partial MSSP Solution*, the MSSP binary decision rule is fixed and for each known uncertainty sample (perfect Information), the LP problem is solved. With this, the objective is 5655 M$. For *Partial MSDP Solution*, the MSDP binary decision rule is fixed and for each known uncertainty sample (perfect Information), the LP problem is solved. The objective value is 5770 M$. As a result, the gap for binary LDR solution is 115 M$.

The above calculation shows that LDR solution is suboptimal and there is room to improve the solution. It should be noted that the partial MSDP and partial MSSP solution are not applicable in real applications since they are based on known realizations of the uncertainty.

6.8 Conclusion

In this work, multistage deterministic and stochastic optimization models for a SAGD reservoir development planning were built. The mathematical modeling procedure resulted in mixed integer linear programming for both deterministic and stochastic models. The objective of the models was to maximize the revenue obtained from the SAGD operation during the project horizon considering restrictions on capital investment, steam generation and oil processing capacities. The stochastic model was developed to incorporate oil price and oil production uncertainty. The multistage stochastic programming model handles both exogenous and endogenous uncertainties simultaneously. Using a linear decision rule based approximation method, the stochastic model was efficiently solved. To study the effect of uncertainty set size on the solution, experiments with different uncertainty set sizes were conducted. Using samples from the widest uncertainty set, the evaluation results demonstrated that the solution obtained from the widest uncertainty set results demonstrates the lowest number of violations subject to uncertainty. Finally, since in this study a linear decision rule was employed in order to solve the stochastic problem, the obtained solution was suboptimal. The suboptimality of the solution was investigated by fixing the binary component of the stochastic solution and solving the LP problem for sampled uncertainty values; it was found that LDR solution can further be improved.

Nomenclature

Indices and Sets

$i \in \mathcal{I}$	Set of drainage areas $\{1, \cdots, N_{DA}\}$
$k \in \mathcal{K}_i$	Set of time periods within the lifetime of drainage area i $\{1, \cdots, T_i^{life}\}$
$t \in \mathcal{T}$	Set of time periods within the lifetime of whole project $\{1, \cdots, T\}$
$l \in \mathcal{L}$	Set of uncertain parameters $\{2, 3, ..., N_{DA} + T + 1\}$

Continuous Variables

NPV	Net present value
$q_{i,t}^{steam}$	Steam injection rate for i-th DA at year t of project horizon (m^3/year)
$q_{i,t}^{oil}$	Oil production rate for i-th DA at year t of project horizon (m^3/year)
$q_{i,t}^{oil}(\xi)$	Adjustable oil production rate for i-th DA at year t of project horizon (m^3/year)
$Q_{i,t,l}$	Coefficient of linear decision rule
t_i^{start}	Production start time of drainage area i
TAR_t	Total annual revenue at year t
TCC_t	Total capital cost at year t
TFC_t	Total fuel cost at year t
$\Lambda_{i,t}$,λ_t,$\mu_{i,t}$	Dual variables

Binary Variables

$z_{i,t}$	$z_{i,t} = 1$ if drainage area i has started production at year t otherwise 0
$x_{l,t}$	$x_{l,t} = 1$ if uncertainty variable l is revealed at year t otherwise 0

Uncertain Parameters

ξ	Set of uncertain variables within uncertainty set Ξ
$\tilde{q}_{i,k}^{maxoil}(\xi)$	Uncertain oil production of i-th DA at year k of its life time
$OP_t(\xi)$	Uncertain oil (bitumen) prices at year t
ζ_i	Oil production uncertainty for i-th drainage area
ϵ_t	Oil price uncertainty at year t

Parameters

A	Planar area of steam chamber (m^2)
$C_{\nu o}$	Overburden volumetric heat capacity (MJ/m^3K)
$C_{\nu r}$	Reservoir volumetric heat capacity (MJ/m^3K)
Cap_{CPF}	Oil processing capacity ($m^3/year$)
Cap_{steam}	Steam generation capacity ($m^3/year$)
CC_{DA}	Capital cost for drainage area installation per year (M)
DA_{cost}	Installation cost for each drainage area (M$)
DR	Discount rate at each year
FC	Fuel cost (M$/$m^3$)
g	Gravitational acceleration (m/s^2)
h	Effective drainage height (m)
H_c	Heat inside the steam chamber (J)
H_o	Heat loss to overburden (J)
H_t	Total cumulative heat consumption (J)
$H_{l\nu}$	Steam latent heat (MJ/m^3)
h	Steam chamber height (m)
k	Effective reservoir permeability(m^2)
k_t	Overburden thermal conductivity($MJ/(m \cdot K \cdot year)$)
L	Well length
m	Viscosity-temperature correlation coefficient
N_{DA}	Number of drainage areas
N_w	Number of wells in each drainage area
NCB	Net continuous bitumen
OP_t	Oil price at year t (CAN $/bbl)
$q_{i,k}^{maxoil}$	Max oil production for each well in i-th DA at k-th year of its life time (m^3/year)
$q_{i,k}^{steam}$	Steam injection for each well in i-th DA at k-th year of its life time (m^3/year)
RY	Royalty
S_o	Reservoir initial oil saturation
S_{oil}	Initial oil saturation
S_{or}	Residual oil saturation
T_i^{life}	Total life time of i-th DA (years)
T_{steam}	Steam temperature (°C)
V_{Steam}	Steam volume (m^3)
w	Half-well spacing (m)
α	Reservoir thermal diffusivity (m^2/s)
η_s	Reservoir effective sweep efficiency
ν_s	Oil viscosity at steam temperature (m^2/s)
ϕ	Reservoir porosity

Chapter 7

Conclusion and Future work

7.1 Concluding remarks

This study is an endeavor to propose new methods that can resolve problems that traditional scenario-tree method faces in solving multistage stochastic problems. Scenario-tree method usually result in exponential growth of problem size in multistage settings that require long run times and very powerful computational resources. By employing linear decision rules for binary and continuous variables based on the lifting method and integrating these decision rules into a single framework, a flexible piece-wise linear decision rule for continuous variables is obtained that results in a significant computational efficiency compared to scenario-tree method. Also a novel framework that integrates both exogenous and endogenous uncertain parameters into a single framework is proposed while studies that address both types of uncertainty are rare in the literature. A new hybrid method that combines scenario-tree and linear decision rule methods is proposed that results in run time and computational efficiency in problems where both constraints variables and coefficients are uncertain.

7.2 Future work

This study can be extended by integrating other methods that result in improvement of solution quality or reduction of runtime and computational expense. For example, decision-dependent uncertainty sets can be integrated into the frameworks proposed in chapters 3, 4 and 5. Also methods that has been recently proposed in the literature such as muli-parametric programming can be integrated with the methods proposed in this study to improve the solution quality. Chapters 3 and 4 propose new methods that deal with exogenous uncertainty. These methods can be extended for endogenous uncertainty.

Bibliography

1. K. Miura, J. Wang, *et al.*, "An analytical model to predict cumulative steam/oil ratio (csor) in thermal-recovery sagd process," *Journal of Canadian Petroleum Technology*, vol. 51, no. 04, pp. 268–275, 2012.

2. J. Guo, C. Zan, D. Ma, and L. Shi, "Oil production rate predictions for steam assisted gravity drainage based on high-pressure experiments," *Science China Technological Sciences*, vol. 56, no. 2, pp. 324–334, 2013.

3. S. Rahim and Z. Li, "Reservoir geological uncertainty reduction: an optimization-based method using multiple static measures," *Mathematical Geosciences*, vol. 47, no. 4, pp. 373–396, 2015.

4. B.-T. Aharon, G. Boaz, and S. Shimrit, "Robust multi-echelon multi-period inventory control," *European Journal of Operational Research*, vol. 199, no. 3, pp. 922–935, 2009.

5. C.-T. See and M. Sim, "Robust approximation to multiperiod inventory management," *Operations research*, vol. 58, no. 3, pp. 583–594, 2010.

6. P. J. Goulart, E. C. Kerrigan, and J. M. Maciejowski, "Optimization over state feedback policies for robust control with constraints," *Automatica*, vol. 42, no. 4, pp. 523–533, 2006.

7. J. Skaf and S. P. Boyd, "Design of affine controllers via convex optimization," *IEEE Transactions on Automatic Control*, vol. 55, no. 11, pp. 2476–2487, 2010.

8. G. C. Calafiore, "Multi-period portfolio optimization with linear control policies," *Automatica*, vol. 44, no. 10, pp. 2463–2473, 2008.

9. G. C. Calafiore, "An affine control method for optimal dynamic asset allocation with transaction costs," *SIAM Journal on Control and Optimization*, vol. 48, no. 4, pp. 2254–2274, 2009.

10. V. Gupta and I. E. Grossmann, "An efficient multiperiod minlp model for optimal planning of offshore oil and gas field infrastructure," *Industrial & Engineering Chemistry Research*, vol. 51, no. 19, pp. 6823–6840, 2012.

11. M. Colvin and C. T. Maravelias, "Modeling methods and a branch and cut algorithm for pharmaceutical clinical trial planning using stochastic programming," *European Journal of Operational Research*, vol. 203, no. 1, pp. 205–215, 2010.

12. S. Ahmed, A. J. King, and G. Parija, "A multi-stage stochastic integer programming approach for capacity expansion under uncertainty," *Journal of Global Optimization*, vol. 26, no. 1, pp. 3–24, 2003.

13. J. R. Birge and C. H. Rosa, "Incorporating investment uncertainty into greenhouse policy models," *The Energy Journal*, vol. 17, no. 1, 1996.

14. L. F. Escudero, P. V. Kamesam, A. J. King, and R. J. Wets, "Production planning via scenario modelling," *Annals of Operations research*, vol. 43, no. 6, pp. 309–335, 1993.

15. S. Takriti, J. R. Birge, and E. Long, "A stochastic model for the unit commitment problem," *IEEE Transactions on Power Systems*, vol. 11, no. 3, pp. 1497–1508, 1996.

16. R. Schultz, "Stochastic programming with integer variables," *Mathematical Programming*, vol. 97, no. 1, pp. 285–309, 2003.

17. N. V. Sahinidis, "Optimization under uncertainty: state-of-the-art and opportunities," *Computers & Chemical Engineering*, vol. 28, no. 6, pp. 971–983, 2004.

18. I. E. Grossmann, R. M. Apap, B. A. Calfa, P. García-Herreros, and Q. Zhang, "Recent advances in mathematical programming techniques for the optimization of process systems under uncertainty," *Computers & Chemical Engineering*, vol. 91, pp. 3–14, 2016.

19. R. M. Apap and I. E. Grossmann, "Models and computational strategies for multistage stochastic programming under endogenous and exogenous uncertainties," *Computers & Chemical Engineering*, vol. 103, pp. 233–274, 2017.

20. V. Goel and I. E. Grossmann, "A class of stochastic programs with decision dependent uncertainty," *Mathematical programming*, vol. 108, no. 2, pp. 355–394, 2006.

21. V. Gupta and I. E. Grossmann, "A new decomposition algorithm for multistage stochastic programs with endogenous uncertainties," *Computers & chemical engineering*, vol. 62, pp. 62–79, 2014.

22. A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski, "Adjustable robust solutions of uncertain linear programs," *Mathematical Programming*, vol. 99, no. 2, pp. 351–376, 2004.

23. A. Atamtürk and M. Zhang, "Two-stage robust network flow and design under demand uncertainty," *Operations Research*, vol. 55, no. 4, pp. 662–673, 2007.

24. N. H. Lappas and C. E. Gounaris, "Robust optimization for decision-making under endogenous uncertainty," *Computers & Chemical Engineering*, vol. 111, pp. 252–266, 2018.

25. C. Ning and F. You, "A data-driven multistage adaptive robust optimization framework for planning and scheduling under uncertainty," *AIChE Journal*, vol. 63, no. 10, pp. 4343–4369, 2017.

26. Q. Zhang, M. F. Morari, I. E. Grossmann, A. Sundaramoorthy, and J. M. Pinto, "An adjustable robust optimization approach to scheduling of continuous industrial processes providing interruptible load," *Computers & Chemical Engineering*, vol. 86, pp. 106–119, 2016.

27. X. Chen, M. Sim, P. Sun, and J. Zhang, "A linear decision-based approximation approach to stochastic programming," *Operations Research*, vol. 56, no. 2, pp. 344–357, 2008.

28. X. Chen and Y. Zhang, "Uncertain linear programs: Extended affinely adjustable robust counterparts," *Operations Research*, vol. 57, no. 6, pp. 1469–1482, 2009.

29. A. Georghiou, W. Wiesemann, and D. Kuhn, "Generalized decision rule approximations for stochastic programming via liftings," *Mathematical Programming*, vol. 152, no. 1-2, pp. 301–338, 2015.

30. J. Goh and M. Sim, "Distributionally robust optimization and its tractable approximations," *Operations research*, vol. 58, no. 4-part-1, pp. 902–917, 2010.

31. D. Bampou and D. Kuhn, "Scenario-free stochastic programming with polynomial decision rules," in *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, pp. 7806–7812, IEEE, 2011.

32. D. Bertsimas, D. A. Iancu, and P. A. Parrilo, "A hierarchy of near-optimal policies for multistage adaptive optimization," *IEEE Transactions on Automatic Control*, vol. 56, no. 12, pp. 2809–2824, 2011.

33. S. Avraamidou and E. N. Pistikopoulos, "Adjustable robust optimization through multiparametric programming," *Optimization Letters*, pp. 1–15, 2019.

34. D. Bertsimas and C. Caramanis, "Adaptability via sampling," in *Decision and Control, 2007 46th IEEE Conference on*, pp. 4717–4722, IEEE, 2007.

35. D. Bertsimas and A. Georghiou, "Design of near optimal decision rules in multistage adaptive mixed-integer optimization," *Operations Research*, vol. 63, no. 3, pp. 610–627, 2015.

36. G. A. Hanasusanto, D. Kuhn, and W. Wiesemann, "K-adaptability in two-stage robust binary programming," *Operations Research*, vol. 63, no. 4, pp. 877–891, 2015.

37. D. Bertsimas and C. Caramanis, "Finite adaptability in multistage linear optimization," *IEEE Transactions on Automatic Control*, vol. 55, no. 12, pp. 2751–2766, 2010.

38. K. Postek and D. d. Hertog, "Multistage adjustable robust mixed-integer optimization via iterative splitting of the uncertainty set," *INFORMS Journal on Computing*, vol. 28, no. 3, pp. 553–574, 2016.

39. D. Bertsimas and A. Georghiou, "Binary decision rules for multistage adaptive mixed-integer optimization," *Mathematical Programming*, vol. 167, no. 2, pp. 395–433, 2018.

40. D. Bertsimas, D. B. Brown, and C. Caramanis, "Theory and applications of robust optimization," *SIAM review*, vol. 53, no. 3, pp. 464–501, 2011.

41. V. Gabrel, C. Murat, and A. Thiele, "Recent advances in robust optimization: An overview," *European journal of operational research*, vol. 235, no. 3, pp. 471–483, 2014.

42. J. Dupacová, "Optimization under exogenous and endogenous uncertainty," *University of West Bohemia in Pilsen*, 2006.

43. J. H. Lee and J. M. Lee, "Approximate dynamic programming based approach to process control and scheduling," *Computers & chemical engineering*, vol. 30, no. 10-12, pp. 1603–1618, 2006.

44. J. M. Lee and J. H. Lee, "Approximate dynamic programming-based approaches for input–output data-driven control of nonlinear processes," *Automatica*, vol. 41, no. 7, pp. 1281–1288, 2005.

45. W. B. Powell, *Approximate Dynamic Programming: Solving the curses of dimensionality*, vol. 703. John Wiley & Sons, 2007.

46. J. M. Lee and J. H. Lee, "Approximate dynamic programming strategies and their applicability for process control: A review and future directions," *International Journal of Control, Automation, and Systems*, vol. 2, no. 3, pp. 263–278, 2004.

47. J. R. Birge and S. W. Wallace, "A separable piecewise linear upper bound for stochastic linear programs," *SIAM Journal on Control and Optimization*, vol. 26, no. 3, pp. 725–739, 1988.

48. J. R. Birge and R. J.-B. Wets, "Sublinear upper bounds for stochastic programs with recourse," *Mathematical Programming*, vol. 43, no. 1-3, pp. 131–149, 1989.

49. S. W. Wallace, "A piecewise linear upper bound on the network recourse function," *Mathematical Programming*, vol. 38, no. 2, pp. 133–146, 1987.

50. S. W. Wallace and T. Yan, "Bounding multi-stage stochastic programs from above," *Mathematical programming*, vol. 61, no. 1-3, pp. 111–129, 1993.

51. J. Birge and F. Louveaux, "Introduction to stochastic programming. series in operations research and financial engineering," 1997.

52. A. J. Conejo, M. Carrión, J. M. Morales, *et al.*, *Decision making under uncertainty in electricity markets*, vol. 1. Springer, 2010.

53. T. W. Jonsbråten, "Optimization models for petroleum field exploitation," 1998.

54. V. Goel and I. E. Grossmann, "A stochastic programming approach to planning of offshore gas field developments under uncertainty in reserves," *Computers & chemical engineering*, vol. 28, no. 8, pp. 1409–1429, 2004.

55. A. Shapiro and A. Nemirovski, "On complexity of stochastic programming problems," in *Continuous optimization*, pp. 111–146, Springer, 2005.

56. D. Kuhn, W. Wiesemann, and A. Georghiou, "Primal and dual linear decision rules in stochastic and robust optimization," *Mathematical Programming*, vol. 130, no. 1, pp. 177–209, 2011.

57. P. Vayanos, D. Kuhn, and B. Rustem, "Decision rules for information discovery in multi-stage stochastic programming," in *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*, pp. 7368–7373, IEEE, 2011.

58. A. Ben-Tal, B. Golany, A. Nemirovski, and J.-P. Vial, "Retailer-supplier flexible commitments contracts: A robust optimization approach," *Manufacturing & Service Operations Management*, vol. 7, no. 3, pp. 248–271, 2005.

59. Á. Lorca, X. A. Sun, E. Litvinov, and T. Zheng, "Multistage adaptive robust optimization for the unit commitment problem," *Operations Research*, vol. 64, no. 1, pp. 32–51, 2016.

60. N. H. Lappas and C. E. Gounaris, "Multi-stage adjustable robust optimization for process scheduling under uncertainty," *AIChE Journal*, vol. 62, no. 5, pp. 1646–1667, 2016.

61. C. Ning and F. You, "Data-driven adaptive nested robust optimization: general modeling framework and efficient computational algorithm for decision making under uncertainty," *AIChE Journal*, vol. 63, no. 9, pp. 3790–3817, 2017.

62. B. Christian and S. Cremaschi, "A branch and bound algorithm to solve large-scale multistage stochastic programs with endogenous uncertainty," *AIChE Journal*, vol. 64, no. 4, pp. 1262–1271, 2018.

63. D. Bertsimas and I. Dunning, "Multistage robust mixed-integer optimization with adaptive partitions," *Operations Research*, vol. 64, no. 4, pp. 980–998, 2016.

64. J. A. Filar and A. Haurie, *Uncertainty and environmental decision making*. Springer, 2010.

65. Z. Li and M. G. Ierapetritou, "Capacity expansion planning through augmented lagrangian optimization and scenario decomposition," *AIChE Journal*, vol. 58, no. 3, pp. 871–883, 2012.

66. C. E. Gounaris, W. Wiesemann, and C. A. Floudas, "The robust capacitated vehicle routing problem under demand uncertainty," *Operations Research*, vol. 61, no. 3, pp. 677–693, 2013.

67. M. Dyer and L. Stougie, "Computational complexity of stochastic programming problems," *mathematical programming*, vol. 106, no. 3, pp. 423–432, 2006.

68. S. J. Garstka and R. J.-B. Wets, "On decision rules in stochastic programming," *Mathematical Programming*, vol. 7, no. 1, pp. 117–143, 1974.

69. A. Ben-Tal and A. Nemirovski, "Robust convex optimization," *Mathematics of operations research*, vol. 23, no. 4, pp. 769–805, 1998.

70. A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust optimization*, vol. 28. Princeton University Press, 2009.

71. D. Bertsimas, D. A. Iancu, and P. A. Parrilo, "Optimality of affine policies in multistage robust optimization," *Mathematics of Operations Research*, vol. 35, no. 2, pp. 363–394, 2010.

72. B. D. Anderson and J. B. Moore, *Optimal control: linear quadratic methods*. Courier Corporation, 2007.

73. A. Ben-Tal and D. Den Hertog, "Immunizing conic quadratic optimization problems against implementation errors," 2011.

74. C. C. M. Caramanis, *Adaptable optimization: Theory and algorithms*. PhD thesis, Massachusetts Institute of Technology, 2006.

75. R. Misener and C. A. Floudas, "Antigone: algorithms for continuous/integer global optimization of nonlinear equations," *Journal of Global Optimization*, vol. 59, no. 2-3, pp. 503–526, 2014.

76. F. Motamed Nasab and Z. Li, "Multistage adaptive optimization using hybrid scenario and decision rule formulation," *AIChE Journal*, vol. 65, no. 12, p. e16764, 2019.

77. J. Goh, N. G. Hall, and M. Sim, "Robust optimization strategies for total cost control in project management," *Submitted for publication*, 2010.

78. P. Rocha and D. Kuhn, "Multistage stochastic portfolio optimisation in deregulated electricity markets using linear decision rules," *European Journal of Operational Research*, vol. 216, no. 2, pp. 397–408, 2012.

79. İ. Yanıkoğlu, B. L. Gorissen, and D. den Hertog, "A survey of adjustable robust optimization," *European Journal of Operational Research*, vol. 277, no. 3, pp. 799–813, 2019.

80. T. Takamatsu, I. Hashimoto, and S. Shioya, "On design margin for process system with parameter uncertainty," *Journal of Chemical Engineering of Japan*, vol. 6, no. 5, pp. 453–457, 1974.

81. I. Grossmann and R. Sargent, "Optimum design of chemical plants with uncertain parameters," *AIChE Journal*, vol. 24, no. 6, pp. 1021–1028, 1978.

82. A. Gupta and C. D. Maranas, "A two-stage modeling and solution framework for multisite midterm planning under demand uncertainty," *Industrial & Engineering Chemistry Research*, vol. 39, no. 10, pp. 3799–3813, 2000.

83. G. Applequist, J. Pekny, and G. Reklaitis, "Risk and uncertainty in managing chemical manufacturing supply chains," *Computers & Chemical Engineering*, vol. 24, no. 9-10, pp. 2211–2222, 2000.

84. N. Boland, I. Dumitrescu, and G. Froyland, "A multistage stochastic programming approach to open pit mine production scheduling with uncertain geology," *Optimization online*, pp. 1–33, 2008.

85. M. Colvin and C. T. Maravelias, "A stochastic programming approach for clinical trial planning in new drug development," *Computers & Chemical Engineering*, vol. 32, no. 11, pp. 2626–2642, 2008.

86. S. Solak, J.-P. B. Clarke, E. L. Johnson, and E. R. Barnes, "Optimization of r&d project portfolios under endogenous uncertainty," *European Journal of Operational Research*, vol. 207, no. 1, pp. 420–433, 2010.

87. S. Peeta, F. S. Salman, D. Gunnec, and K. Viswanath, "Pre-disaster investment decisions for strengthening a highway network," *Computers & Operations Research*, vol. 37, no. 10, pp. 1708–1719, 2010.

88. L. F. Escudero, M. A. Garín, J. F. Monge, and A. Unzueta, "On preparedness resource allocation planning for natural disaster relief under endogenous uncertainty with time-consistent risk-averse management," *Computers & Operations Research*, vol. 98, pp. 84–102, 2018.

89. L. Hellemo, P. I. Barton, and A. Tomasgard, "Decision-dependent probabilities in stochastic programs with recourse," *Computational Management Science*, vol. 15, no. 3-4, pp. 369–395, 2018.

90. S. Mitra, I. E. Grossmann, J. M. Pinto, and N. Arora, "Robust scheduling under time-sensitive electricity prices for continuous power-intensive processes," 2012.

91. R. Vujanic, S. Mariéthoz, P. Goulart, and M. Morari, "Robust integer optimization and scheduling problems for large electricity consumers," in *2012 American Control Conference (ACC)*, pp. 3108–3113, IEEE, 2012.

92. S. Amaran, T. Zhang, N. V. Sahinidis, B. Sharda, and S. J. Bury, "Medium-term maintenance turnaround planning under uncertainty for integrated chemical sites," *Computers & Chemical Engineering*, vol. 84, pp. 422–433, 2016.

93. O. Nohadani and K. Sharma, "Optimization under decision-dependent uncertainty," *SIAM Journal on Optimization*, vol. 28, no. 2, pp. 1773–1795, 2018.

94. P. Vayanos, A. Georghiou, and H. Yu, "Robust optimization with decision-dependent information discovery," *Under review at Management Science*, 2019.

95. J. Dillinger, "The world's largest oil reserves by country," *Retrived from https://www.worldatlas.com/articles/the-world-s-largest-oil-reserves-by-country.html*, 2018.

96. Alberta Energy Regulator, "Alberta's energy reserves 2013 and supply/demand outlook 2014–2023," *Statistical Series (ST) 2015–98*, 2015.

97. R. Iyer, I. Grossmann, S. Vasantharajan, and A. Cullick, "Optimal planning and scheduling of offshore oil field infrastructure investment and operations," *Industrial & Engineering Chemistry Research*, vol. 37, no. 4, pp. 1380–1397, 1998.

98. S. A. Van Den Heever and I. E. Grossmann, "An iterative aggregation/disaggregation approach for the solution of a mixed-integer nonlinear oilfield infrastructure planning model," *Industrial & engineering chemistry research*, vol. 39, no. 6, pp. 1955–1971, 2000.

99. V. Kosmidis, J. Perkins, and E. Pistikopoulos, "A mixed integer optimization strategy for integrated gas/oil production," *Computer Aided Chemical Engineering*, vol. 10, pp. 697–702, 2002.

100. H. Shahandeh, S. Rahim, and Z. Li, "Strategic optimization of the oil sands development with sagd: Drainage area arrangement and development planning," *Journal of Petroleum Science and Engineering*, vol. 137, pp. 172–184, 2016.

101. J. Charry-Sanchez, A. Betancourt-Torcat, and L. Ricardez-Sandoval, "An optimization energy model for the upgrading processes of canadian unconventional oil," *Energy*, vol. 68, pp. 629–643, 2014.

102. A. Betancourt-Torcat, G. Gutierrez, A. Elkamel, and L. Ricardez-Sandoval, "Integrated energy optimization model for oil sands operations," *Industrial & Engineering Chemistry Research*, vol. 50, no. 22, pp. 12641–12663, 2011.

103. H. Shahandeh and Z. Li, "Optimal design of bitumen upgrading facility with co2 reduction," *Computers & Chemical Engineering*, vol. 106, pp. 106–121, 2017.

104. K. O. Jørnsten, "Sequencing offshore oil and gas fields under uncertainty," *European Journal of Operational Research*, vol. 58, no. 2, pp. 191–201, 1992.

105. J. R. Birge, "State-of-the-art-survey—stochastic programming: computation and applications," *INFORMS journal on computing*, vol. 9, no. 2, pp. 111–133, 1997.

106. W. K. Klein Haneveld and M. H. van der Vlerk, "Stochastic integer programming: General models and algorithms," *Annals of Operations Research*, vol. 85, pp. 39–57, 1999.

107. 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 & chemical engineering*, vol. 30, no. 6, pp. 1076–1092, 2006.

108. A. Betancourt-Torcat, A. Almansoori, A. Elkamel, and L. Ricardez-Sandoval, "Stochastic modeling of the oil sands operations under greenhouse gas emission restrictions and water management," *Energy & Fuels*, vol. 27, no. 9, pp. 5559–5578, 2013.

109. V. Gupta and I. E. Grossmann, "Multistage stochastic programming approach for offshore oilfield infrastructure planning under production sharing agreements and endogenous uncertainties," *Journal of Petroleum Science and Engineering*, vol. 124, pp. 180–197, 2014.

110. R. M. Butler, "Thermal recovery of oil and bitumen," 1991.

111. R. Chan, "Application of field performance data in developing simple analytical models to predict the performance of steam assisted gravity drainage.," 2013.

112. N. Edmunds, J. Peterson, *et al.*, "A unified model for prediction of csor in steam-based bitumen recovery," in *Canadian International Petroleum Conference*, Petroleum Society of Canada, 2007.

113. G. Giacchetta, M. Leporini, and B. Marchetti, "Economic and environmental analysis of a steam assisted gravity drainage (sagd) facility for oil recovery from canadian oil sands," *Applied Energy*, vol. 142, pp. 1–9, 2015.

114. K. A. Lawal, "Economics of steam-assisted gravity drainage for the nigerian bitumen deposit," *Journal of Petroleum Science and Engineering*, vol. 116, pp. 28–35, 2014.

115. Alberta Energy Regulator, "Oil prices," *Retrived from http://economicdashboard.alberta.ca/OilPrice*, 2018.

Appendix A

A.1 Chapter 1

1. Derivation procedure for illustrating example Hybrid 1 method

Deriving the dual counterpart for constraint 10d:

$$\sum_i \bar{\gamma}_i(1+O_i\xi)x_{i,1}+y_k \geq d_k \qquad \forall k \in \{2,3,4\}, \xi \in \Xi$$
$$\sum_i \bar{\gamma}_i x_{i,1}+\sum_i \bar{\gamma}_i x_{i,1}O_i\xi+y_k \geq d_k \qquad \forall k \in \{2,3,4\}, \xi \in \Xi$$

$$\begin{cases} \sum_i \bar{\gamma}_i\ x_{i,1}+\min_{\xi\in\Xi}\sum_i \bar{\gamma}_i x_{i,1}O_i\xi+y_k \geq d_k & \forall k \in \{2,3,4\} \\ -W\xi \geq -V \end{cases}$$

Applying duality:

$$\begin{cases} \sum_i \bar{\gamma}_i\ x_{i,1}+\max_\theta[-V^\top\theta_k]+y_k \geq d_k & \forall k \in \{2,3,4\} \\ -W^\top\theta_k=(\sum_i \bar{\gamma}_i x_{i,1}O_i)^\top & \forall k \in \{2,3,4\} \\ \theta_k \geq 0 & \forall k \in \{2,3,4\} \end{cases}$$

Dropping the max operator:

$$\begin{cases} \sum_i \bar{\gamma}_i\ x_{i,1}-V^\top\theta_k+y_k \geq d_k & \forall k \in \{2,3,4\} \\ -W^\top\theta_k=(\sum_i \bar{\gamma}_i x_{i,1}O_i)^\top & \forall k \in \{2,3,4\} \\ \theta_k \geq 0 & \forall k \in \{2,3,4\} \end{cases}$$

Deriving the dual counterpart for constraint 10e:

$$\begin{cases} \sum_i \bar{\gamma}_i(1+O_i\xi)(x_{i,1}+x_{i,a(k)})+y_k \geq d_k & \forall k \in \{5,...,13\} \\ -W\xi \geq -V \end{cases}$$

Applying duality:

$$\begin{cases} \sum_i \bar{\gamma}_i \ (x_{i,1} + x_{i,a(k)}) + \max_\phi[-V^\top \phi_k] + y_k \geq d_k & \forall k \in \{5, ..., 13\} \\ -W^\top \phi_k = (\sum_i \bar{\gamma}_i \ (x_{i,1} + x_{i,a(k)}) \ O_i)^\top & \forall k \in \{5, ..., 13\} \\ \phi_k \geq 0 & \forall k \in \{5, ..., 13\} \end{cases}$$

Dropping the max operator:

$$\begin{cases} \sum_i \bar{\gamma}_i \ (x_{i,1} + x_{i,a(k)}) - v^\top \phi_k + y_k \geq d_k & \forall k \in \{5, ..., 13\} \\ -w^\top \phi_k = (\sum_i \bar{\gamma}_i \ (x_{i,1} + x_{i,a(k)}) \ O_i)^\top & \forall k \in \{5, ..., 13\} \\ \phi_k \geq 0 & \forall k \in \{5, ..., 13\} \end{cases}$$

2. Derivation procedure for illustrating example Hybrid 2 method

Expectation of the objective function:

Note that $\xi_1$and ξ_2 are uniformly distributed around zero, therefore $\mathbb{E}(\xi_1) = 0$ and $\mathbb{E}(\xi_2) = 0$,

$$\begin{aligned} &\sum_i \alpha_i x_{i,1} + \mathbb{E}_\xi \left(\sum_i \alpha_i x_{i,2}(\xi) + \sum_{t=1}^{2} \beta y_t(\xi) \right) \\ &= \sum_i \alpha_i x_{i,1} + \mathbb{E}_\xi \left[\sum_i \alpha_i (\bar{x}_{i,2} + \tilde{x}_{i,2} \ \xi_1) + \beta (\bar{y}_1 + \tilde{y}_1 \ \xi_1 + \bar{y}_2 + \tilde{y}_2 \ \xi_1 + \hat{y}_2 \ \xi_2) \right] \\ &= \sum_i \alpha_i x_{i,1} + \left[\sum_i \alpha_i (\bar{x}_{i,2} + \tilde{x}_{i,2} \ \mathbb{E}(\xi_1)) + \beta \left[\bar{y}_1 + \tilde{y}_1 \ \mathbb{E}(\xi_1) + \bar{y}_2 + \tilde{y}_2 \ \mathbb{E}(\xi_1) + \hat{y}_2 \ \mathbb{E}(\xi_2) \right] \right] \\ &= \sum_i \alpha_i x_{i,1} + \left[\sum_i \alpha_i \bar{x}_{i,2} + \beta \left(\bar{y}_1 + \bar{y}_2 \right) \right] \end{aligned}$$

Expanding the variables and parameters:

Apply the decision rules in 20a-20d, and rearrange the constraints into vectorized form of uncertain parameters $\xi = [1, \xi_1, \xi_2]^\top$.

$$\begin{aligned} \min \ & \sum_i \alpha_i x_{i,1} + \left[\sum_i \alpha_i \bar{x}_{i,2} + \beta \left(\bar{y}_1 + \bar{y}_2 \right) \right] \\ s.t. \ & \sum_{i=1}^{2} x_{i,1} \leq 100 \\ & \sum_{i=1}^{2} [\bar{x}_{i,2} \quad \tilde{x}_{i,2} \quad 0] \ \xi \leq 100 && \forall \xi \in \Xi \\ & \sum_i \gamma_{i,l} \ x_{i,1} + [\bar{y}_1 \quad \tilde{y}_1 - \bar{d}_1 \quad 0] \ \xi \geq \bar{d}_1 && \forall l, \forall \xi \in \Xi \end{aligned}$$

$$\sum_i \gamma_{i,l}\ x_{i,1} + \left(\sum_i \gamma_{i,l}[\bar{x}_{i,2} \quad \tilde{x}_{i,2} \quad 0] + [\bar{y}_2 \quad \tilde{y}_2 \quad \hat{y} - \bar{d}_2]\right)\ \xi \geq \bar{d}_2 \qquad \forall l, \forall \xi \in \Xi$$

$$x_{i,1} \geq 0 \qquad \forall i$$

$$[\bar{x}_{i,2} \quad \tilde{x}_{i,2} \quad 0]\ \xi \geq 0 \qquad \forall i, \forall \xi \in \Xi$$

$$[\bar{y}_1 \quad \tilde{y}_1 \quad 0]\ \xi \geq 0 \qquad \forall \xi \in \Xi$$

$$[\bar{y}_2 \quad \tilde{y}_2 \quad \hat{y}_2]\xi \geq 0 \qquad \forall \xi \in \Xi$$

Duality theorem is applied to the six semi-infinite constraints to obtain robust counterpart constraints 21c-21e, 21g-21i. The derivation procedure is similar to 10d provided in section 1 of the appendix.

3. Parameters for Capacity Expansion Problem

Table A.1: Buying prices $\Gamma_{j,t}$ for j=1,2,3,4 (10^2\$/ton)

$\Gamma_{j,t}$	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	$t=6$	$t=7$
$j=1$	22.98	34.67	27.89	25.62	32.42	28.64	31.00
$j=2$	60.23	75.21	79.25	62.10	65.42	72.20	78.80
$j=3$	18.36	34.81	15.12	16.40	32.60	21.80	28.20
$j=4$	17.44	21.89	27.74	22.40	19.40	25.60	19.40

Table A.2: Selling price $\gamma_{j,t}$ for the fifth product j=5 (10^2\$/ton)

$\gamma_{j,t}$	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	$t=6$	$t=7$
$j=5$	58.03	55.14	48.58	49.10	53.20	52.50	57.80

Table A.3: Expansion investment cost (10^2\$/ton), $\alpha_{i,t}$

$\alpha_{i,t}$	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	$t=6$	$t=7$
$i=1$	11.52	6.82	9.24	7.80	8.40	10.20	9.80
$i=2$	10.08	9.35	5.34	7.60	6.50	8.20	9.80
$i=3$	11.91	7.40	9.25	8.20	7.60	9.60	11.00
$i=4$	4.89	10.10	8.61	6.80	7.40	9.84	8.40
$i=5$	7.78	7.13	11.81	8.30	9.70	7.90	11.20
$i=6$	6.69	5.82	6.74	5.92	6.10	6.80	6.50

Table A.4: Fixed investment cost (10^6\$), $\beta_{i,t}$

$\beta_{i,t}$	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	$t=6$	$t=7$
$i=1$	34	58.4	47.9	42	56	36	62
$i=2$	31.6	38.2	35.2	32.4	35.1	33.8	34.6
$i=3$	46.8	23.8	41.3	25.4	45.6	34.2	38.2
$i=4$	33.8	30.7	40.1	31.4	38.6	34.2	40.3
$i=5$	43.9	22.6	20.1	34.6	22.8	42.4	38.6
$i=6$	26.9	20.3	29.1	22.4	30.8	28.6	24.2

Table A.5: Unit operation cost (\$/ton), $\delta_{i,t}$

$\delta_{i,t}$	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	$t=6$	$t=7$
$i=1$	50	50	30	40	30	50	40
$i=2$	30	30	40	30	40	40	30
$i=3$	30	40	30	30	40	30	50
$i=4$	30	50	30	50	30	40	40
$i=5$	20	50	60	30	40	60	50
$i=6$	50	50	30	40	50	40	30

Table A.6: Availability of chemicals 1 to 4 (kton/year), $a_{j,t}$

$a_{j,t}$	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	$t=6$	$t=7$
$j=1$	370	330	390	340	350	40	360
$j=1$	680	360	320	350	440	500	660
$j=1$	450	580	280	300	500	540	290
$j=1$	680	390	460	400	480	660	540

Table A.7: Demand for chemical 5 (kton/year), $d_{5,t}$

$d_{j,t}$	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	$t=6$	$t=7$
$j=5$	91	40	94	44	64	93	52

Table A.8: Capital investment restriction (10^6\$), $CI(t)$

	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$	$t=6$	$t=7$
$CI(t)$	400	70.8	82.3	74.6	81.2	86.4	71.8

Table A.9: Yield rate for each process i and product j, $\eta_{i,j}$

$\bar{\eta}_{i,j}$	$j=1$	$i=2$	$i=3$	$t=4$	$t=5$
$i=1$	0	0	0	1	0.89
$i=2$	0	0	1.06	0	1
$i=3$	0	0	0	0	1
$i=4$	0	0	1	0	0.65
$i=5$	0	0	1	0	0
$i=6$	0	1	0	0	0

Table A.10: Consumption rate for each process i and product j, $\mu_{i,j}$

$\mu_{i,j}$	$j=1$	$i=2$	$i=3$	$t=4$	$t=5$
$i=1$	1.22	0	0.65	0	0
$i=2$	0	0.52	0	0	0
$i=3$	0	0	0	0.7	0
$i=4$	0.68	0	0	0.79	0
$i=5$	0	0.82	0	0	0
$i=6$	1.07	0	0	0	0

A.2 Chapter 2

Table A.11: Results for lifting method, first configuration

T,Br	Objective	Run time	Opt. Gap	Constraints	Cont. Vars	Discrete Vars
T=2, Br=1	118.0	0.04 s	0	405	277	14
T=2, Br=2	118.0	0.04 s	0	545	363	20
T=2, Br=3	103.0	0.06 s	0	685	449	26
T=2, Br=5	99.25	0.08 s	0	965	621	38
T=2, Br=7	99.25	0.27 s	0	1245	793	50
T=2, Br=15	94.09	17.42 s	0	2365	1481	98
T=5, Br=1	403.25	0.11 s	0	2358	1603	50
T=5, Br=2	399.50	6.30 s	0	3233	2133	80
T=5, Br=3	390.12	29.53 s	0	4108	2663	110
T=5, Br=5	378.87	1 hr 57 min	0	5858	3723	170
T=5, Br=7	371.37	10 hrs	6.05%	7608	4783	230
T=5, Br=15	369.03	10 hrs	22.98%	14608	9023	470
T=10, Br=1	1115.25	18.83 s	0	10893	5768	155
T=10, Br=2	1088.99	10 hrs	51.56%	12713	8363	260
T=10, Br=3	1012.12	10 hrs	6.13%	16213	10473	370
T=10, Br=5	1004.62	10 hrs	12.83%	23213	14693	590
T=10, Br=7	1004.62	10 hrs	20.23%	30213	18913	810
T=10, Br=15	1073.06	10 hrs	25.76%	58213	35557	1454
T=20, Br=1	3575.0	10 hrs	1.55%	36423	24813	610
T=20, Br=2	3579.99	10 hrs	9.14%	50423	33343	1140
T=20, Br=3	3421.25	10 hrs	11.4%	64423	41873	1670
T=20, Br=5	3441.87	10 hrs	19.68%	92423	58889	2686
T=20, Br=7	3892.81	10 hrs	29.35%	120423	75069	2866
T=20, Br=15	4549.99	10 hrs	21.66%	232423	139657	3454

Note: "T" denotes the number of stages, "Br" denotes the number of breakpoints.

Table A.12: Results for lifting method, second configuration

T,Br	Objective	Run time	Opt. Gap	Constraints	Cont. Vars	Discrete Vars
T=2, Br=1	124.5	0.11 s	0	565	388	21
T=2, Br=3	118.25	0.11 s	0	957	630	39
T=2, Br=5	107.62	0.23 s	0	1349	872	57
T=2, Br=7	107.62	1.38 s	0	1741	1114	75
T=2, Br=15	107.62	18 min 52 s	0	3309	2082	147
T=5, Br=1	498.0	0.13 s	0	3298	2248	75
T=5, Br=3	461.75	26 min 1s	0	5748	3738	165
T=5, Br=5	433.0	10 hrs	2.64%	8198	5228	255
T=5, Br=7	418.62	10 hrs	6.28%	10648	6718	345
T=5, Br=15	409.87	10 hrs	23.56%	20448	12678	705
T=10, Br=1	1306	3.7 s	0	12893	8768	225
T=10, Br=3	1202.25	10 hrs	5.64%	22693	14698	555
T=10, Br=5	1149.12	10 hrs	11.95%	32493	20628	885
T=10, Br=7	1196.62	10 hrs	19.27%	42293	26558	1215
T=10, Br=15	1309.12	10 hrs	26.77%	81493	49924	2181
T=20, Br=1	4456.52	10 hrs	3.27%	50983	34798	915
T=20, Br=3	4091.24	10 hrs	9.5%	90183	58788	2505
T=20, Br=5	4006.25	10 hrs	19.27%	129383	82712	4029
T=20, Br=7	4984.37	10 hrs	31.27%	168583	105382	4299
T=20, Br=15	6457.18	10 hrs	27.51%	325383	195864	5181

Table A.13: Results for partitioning method, first configuration

T,Br	Objective	Run time	Opt. Gap	Constraints	Cont. Vars	Discrete Vars
T=2, Br=1	118.0	8.42 s	0	89	71	20
T=2, Br=2	100.5	8.8 s	0	197	151	40
T=2, Br=3	96.43	8.64 s	0	349	263	68
T=2, Br=5	87.76	8.533 s	0	785	583	148
T=2, Br=7	85.30	8.48 s	0	1397	1031	260
T=2, Br=15	79.53	8.69 s	0	5605	4103	1028
T=5, Br=1	378.17	8.82 s	0	3244	2253	330
T=5, Br=2	320.30	9.52 s	0	24797	17023	2440
T=5, Br=3	291.70	11.65 s	0	104800	71693	10250
T=5, Br=5	272.34	1 min 22s	0	797828	544333	77770
T=5, Br=7	243.10	25 min 10 s	0	3365752	2293773	327690
T=10, Br=1	903.15	19.22 s	0	364561	245783	20500

Table A.14: Results for partitioning method, second configuration

T,Br	Objective	Run time	Opt. Gap	Constraints	Cont. Vars	Discrete Vars
T=2, Br=1	124.5	8.41 s	0	91	81	30
T=2, Br=2	129.22	8.66 s	0	203	271	60
T=2, Br=3	108.56	8.78 s	0	361	297	102
T=2, Br=5	101.37	8.71 s	0	815	657	222
T=2, Br=7	97.07	8.40 s	0	1453	1161	390
T=2, Br=15	93.21	8.49 s	0	5845	4617	1542
T=5, Br=1	476.75	8.9 s	0	3342	2418	495
T=5, Br=2	393.06	9.46 s	0	25649	18243	3660
T=5, Br=3	359.45	11.94 s	0	108556	76818	15375
T=5, Br=5	323.29	1 min 56 s	0	827378	583218	116655
T=5, Br=7	291.70	38 min 16 s	0	3492144	2457618	491535
T=10, Br=1	1184.07	21.4 s	0	372755	256033	30750

Table A.15: Results for lifting method with breakpoint optimization, first configuration

T,Br	Objective	Run time	Opt. Gap	Constraints	Cont. Vars	Discrete Vars
T=2, Br=1	94.07	5.98 s	0	756	630	14
T=2, Br=2	94.07	2 min 40 s	0	1212	1032	20
T=2, Br=3	90.95	48 min 48 s	0	1764	1530	26
T=2, Br=5	90.94	10 hrs	15.27%	3156	2814	38
T=2, Br=7	90.94	10 hrs	32.03%	4932	4482	50
T=2, Br=15	94.02	10 hrs	48.63%	15876	14994	98
T=5, Br=1	383.37	5 hrs	19.26%	4311	3561	50
T=5, Br=2	384.55	10 hrs	34.93%	7056	5961	80
T=5, Br=3	428.56	10 hrs	49.75%	10401	9861	110
T=5, Br=5	479.42	10 hrs	62.52%	18891	16761	170
T=10, Br=1	1089.23	10 hrs	30.6%	16716	13766	150
T=10, Br=2	1126.13	10 hrs	49.30%	27556	23216	260

Table A.16: Results for partitioning method with breakpoint optimization, first configuration

T,Br	Objective	Run time	Opt. Gap	Constraints	Cont. Vars	Discrete Vars
T=2, Br=1	94.07	1.92 s	0	137	117	20
T=2, Br=2	86.26	1 min 47 s	0	297	249	40
T=2, Br=3	80.01	10 hrs	1.28%	521	433	68
T=2, Br=5	78.06	10 hrs	2.72%	1161	957	148
T=2, Br=7	79.03	10 hrs	34.61%	2057	1689	260
T=2, Br=15	101.66	10 hrs	56.63%	8201	6697	1028
T=5, Br=1	371.86	10 hrs	17.22%	3872	2876	330

A.3 Chapter 3

Parameters for capacity planning example

Table A.17: Buying prices $\Gamma_{j,t}$ for $j = 1, 2, 3, 4$ (10^2\$/ton)

$\Gamma_{j,t}$	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$
$j=1$	22.98	34.67	27.89	25.62	32.42
$j=2$	60.23	75.21	79.25	62.10	65.42
$j=3$	18.36	34.81	15.12	16.40	32.60
$j=4$	17.44	21.89	27.74	22.40	19.40

Table A.18: Selling price $\bar{\gamma}_{j,t}$ for the fifth product $j = 5$ (10^2\$/ton)

$\bar{\gamma}_{j,t}$	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$
$j=5$	58.03	55.14	48.58	49.10	53.20

Table A.19: Expansion investment cost (10^2\$/ton), $\alpha_{i,t}$

$\alpha_{i,t}$	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$
$i=1$	11.52	6.82	9.24	7.80	8.40
$i=2$	10.08	9.35	5.34	7.60	6.50
$i=3$	11.91	7.40	9.25	8.20	7.60
$i=4$	4.89	10.10	8.61	6.80	7.40
$i=5$	7.78	7.13	11.81	8.30	9.70
$i=6$	6.69	5.82	6.74	5.92	6.10

Table A.20: Fixed investment cost (10^6\$), $\beta_{i,t}$

$\beta_{i,t}$	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$
$i=1$	34	58.4	47.9	42	56
$i=2$	31.6	38.2	35.2	32.4	35.1
$i=3$	46.8	23.8	41.3	25.4	45.6
$i=4$	33.8	30.7	40.1	31.4	38.6
$i=5$	43.9	22.6	20.1	34.6	22.8
$i=6$	26.9	20.3	29.1	22.4	30.8

Table A.21: Unit operation cost ($/ton), $\delta_{i,t}$

$\delta_{i,t}$	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$
$i=1$	50	50	30	40	30
$i=2$	30	30	40	30	40
$i=3$	30	40	30	30	40
$i=4$	30	50	30	50	30
$i=5$	20	50	60	30	40
$i=6$	50	50	30	40	50

Table A.22: Availability of chemicals 1 to 4 (kton/year), $a_{j,t}$

$a_{j,t}$	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$
$j=1$	370	330	390	340	350
$j=1$	680	360	320	350	440
$j=1$	450	580	280	300	500
$j=1$	680	390	460	400	480

Table A.23: Demand for chemical 5 (kton/year), $\bar{d}_{5,t}$

	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$
$\bar{d}_{5,t}$	91	40	94	44	64

Table A.24: Capital investment restriction ($10^6$$), $\bar{CI}(t)$

	$t=1$	$t=2$	$t=3$	$t=4$	$t=5$
$\bar{CI}(t)$	400	70.8	82.3	74.6	81.2

Table A.25: Yield rate for each process i and product j, $\nu_{i,j}$

$\nu_{i,j}$	$j=1$	$i=2$	$i=3$	$t=4$	$t=5$
$i=1$	0	0	0	1	0.89
$i=2$	0	0	1.06	0	1
$i=3$	0	0	0	0	1
$i=4$	0	0	1	0	0.65
$i=5$	0	0	1	0	0
$i=6$	0	1	0	0	0

Table A.26: Consumption rate for each process i and product j, $\mu_{i,j}$

$\mu_{i,j}$	$j = 1$	$i = 2$	$i = 3$	$t = 4$	$t = 5$
$i = 1$	1.22	0	0.65	0	0
$i = 2$	0	0.52	0	0	0
$i = 3$	0	0	0	0.7	0
$i = 4$	0.68	0	0	0.79	0
$i = 5$	0	0.82	0	0	0
$i = 6$	1.07	0	0	0	0

A.4 Chapter 4

1. Lifting of exogenous uncertainty

Continuous lifting

Figure A.1 illustrates the breakpoint locations along η_i axis and figure A.2 demonstrates the lifted uncertain parameters $G_{i,j}, j = 1, ..., r_i$ for general number of breakpoints $r_i - 1$.

Figure A.1: General number of breakpoints on η_i axis

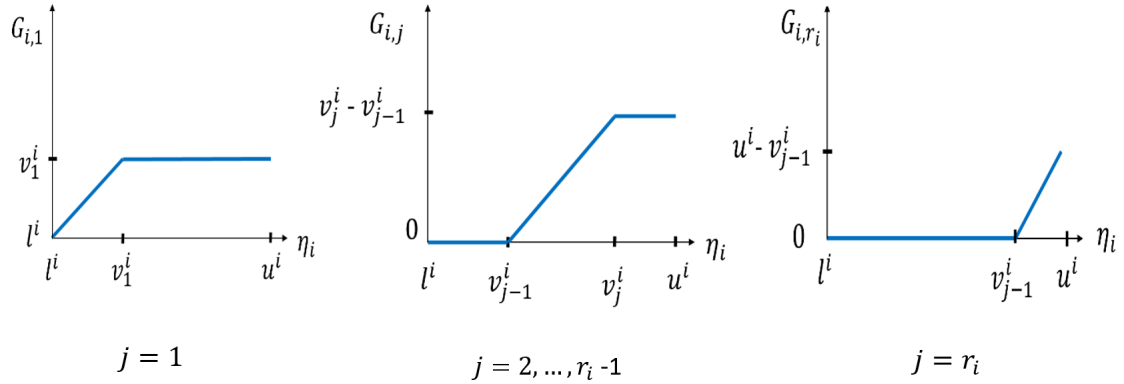


Figure A.2: Illustration of lifted uncertain parameters ($G_{i,j}$) for general number of breakpoints $r_i - 1$

In order to comprehend the fundamentals of the lifting method, this method is illustrated using 1 breakpoint. Figure A.3 shows the lifted uncertain parameters for one breakpoint. The sequence for lifting the uncertain parameter η_i is presented in figure A.4. The left figure shows that one breakpoint is selected for lifting the uncertain parameter. The middle figure plots the lifted non-convex set in the space of lifted uncertain parameters and the right figure convexifies the non-convex set using its extreme points and the definition of the convex hull.

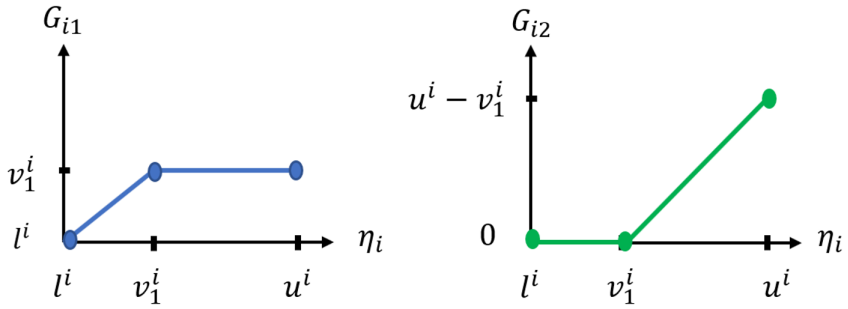


Figure A.3: Illustration of lifted uncertain parameter for one-breakpoint case

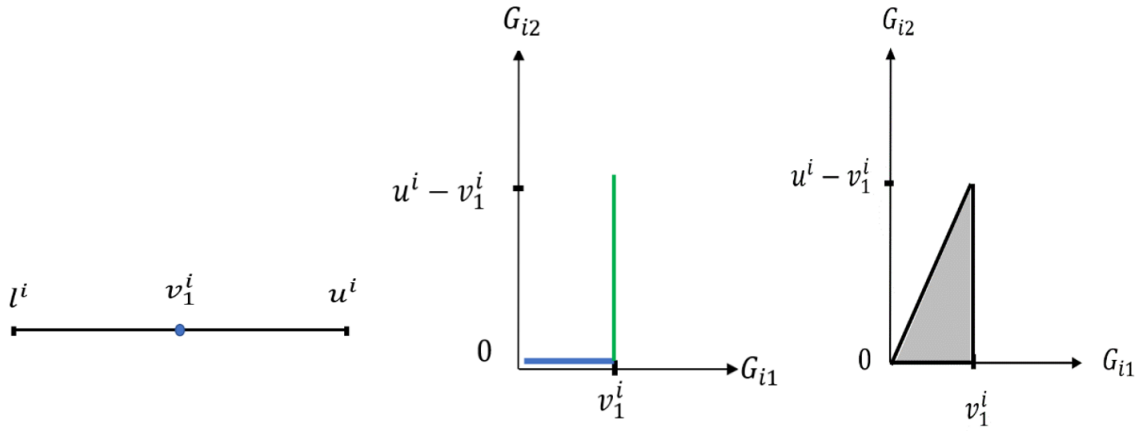


Figure A.4: Lifted uncertain parameter and convexified uncertainty set

The continuous variable is a linear combination of lifted parameters. For instance for one breakpoint setting, it is written as: $x(\eta) = X_0 + X_1 G_1 + X_2 G_2 = X[1, G_1, G_2] = X\mathbf{G}$. This formulation results in a piecewise linear continuous variable with respect to the original uncertain parameter. This formulation is more flexible compared to decision rules where the continuous variable is simply a linear function of the original uncertain parameter.

Binary lifting

The binary lifting is formualted using on 0-1 indicator functions. Figure A.5 illustrates the indicator function for the breakpoint located at v_1^i along the η_i axis. Figure A.6 shows the sequence for convexifying the lifted set for one breakpoint. The left figure shows that one breakpoint is selected. The middle figure shows the lifted set consists of disconnected pieces and therefore it is non-convex. The right figure shows that the non-convex set is convexified using its extreme points which results in a polyhedron.

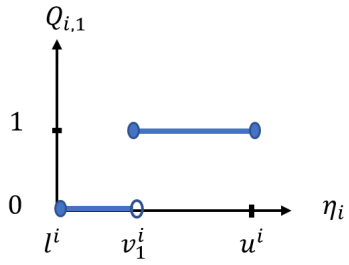


Figure A.5: Indicator functions for one breakpoint on uncertain parameter η_i

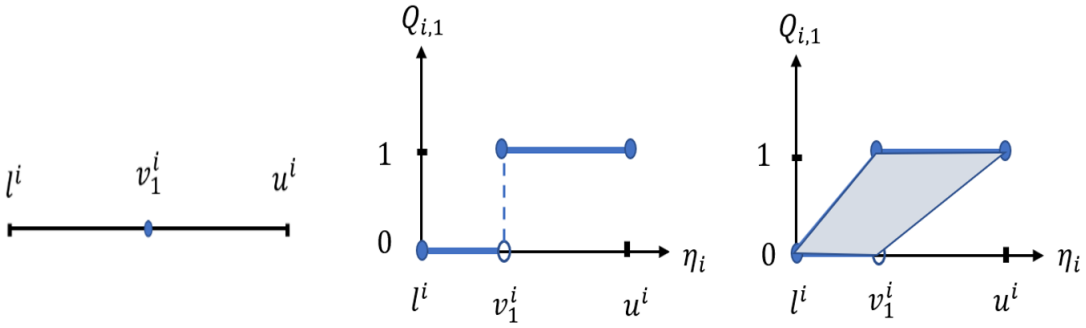


Figure A.6: Lifted uncertain parameter and convexified uncertainty set to be used for adaptive binary decision rule

The binary variable is a linear combination of 0-1 indicator functions. For instance for one breakpoint, the binary variable is formulated as: $y(\eta) = Y_0 + Y_1 Q_1 = \mathbf{Y}[1, Q_1] = \mathbf{YQ}$. The range of integer values for the $\mathbf{Y}$ variable is $\{-1, 0, +1\}$. Figure A.7 illustrates the 4 different possible solutions for the adaptive binary variable when 1 break point is applied.

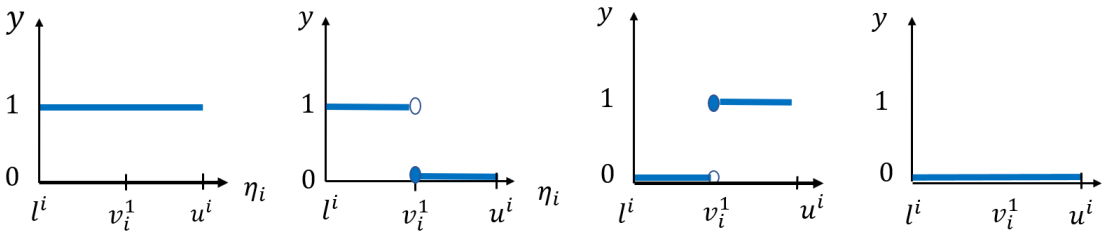


Figure A.7: Adaptive binary solution for 1 breakpoint

2. Shale gas problem

Nomenclature

Sets	
$t \in T$	Time horizon, in years, $T = \{1, ..., 10\}$
$j \in P$	Set of candidate production platforms, $P = \{1, 2, 3\}$
$i \in L$	Set of possible pipelines, $L = \{1, 2, 3, 4, 5\}$
$L^+(p)$	Set of ingoing pipelines into platform p
$L^-(p)$	Set of outgoing pipelines from platform p
O	Main pipeline
Parameters	
ξ_j	Gas field size, random variable, uniform distribution, $U(0, 10), U(0, 5), U(0, 5)$
η_t	gas price uncertainty at year t, uniform distribution, $\eta_t \in [0, 1]$
r_j	Maximum annual production rate for platform j
	$(3, 1.5, 1.5), (2, 1, 1), (1, 0.5, 0.5)$ for $T = 3, 5, 10$, respectively
d_t	Discount factor in year t , $\frac{1}{1.01^{t-1}}$
$c_t^g(\eta_t)$	Unit price for gas at time step t, $2 + \eta_t$
c_j^c	Cost for unit capacity expansion for production platform j, $(0.2, 0.2, 0.2)$
c_j^e	Cost for unit gas extraction for platform j, $(0.1, 0.1, 0.1)$
c_i^l	Cost for building pipeline i, $(2, 1, 3, 1, 5)$
c_j^p	Cost for building platform j, $(4, 2, 2)$
Variables	
$y_{j,t}^p(\xi)$	Binary, $x_{j,t}^p(\xi) = 1$ if platform j exist in year t, otherwise 0
$y_{i,t}^l(\xi, \eta)$	Binary, $x_{i,t}^l(\xi, \eta) = 1$ if pipeline i exist in year t, otherwise 0
$x_{i,t}^f(\xi, \eta)$	Amount of gas flow through pipeline i in year t
$x_{j,t}^e(\xi, \eta)$	Amount of gas extraction for platform j in year t
$x_{j,t}^c(\xi, \eta)$	Capacity of production platform j in year t

Formulation of the 3-platform case

Equations A.1a to A.1n present the problem formulation for the case with 3 platforms

$$\max \quad \mathbb{E}_{\xi,\eta} \sum_{t=1}^{10} d_t [c_t^g(\eta_t)(x_{2,t}^f(\xi, \eta) + x_{4,t}^f(\xi, \eta)) - \sum_{i=1}^{5} c_i^l (y_{i,t}^l(\xi, \eta) - y_{i,t-1}^l(\xi, \eta)) \tag{A.1a}$$
$$- \sum_{j=1}^{3} c_j^p (y_{j,t}^p(\xi) - y_{j,t-1}^p(\xi)) - \sum_{j=1}^{3} c_j^c x_{j,t}^c(\xi, \eta) - \sum_{j=1}^{3} c_j^e \, x_{j,t}^e(\xi, \eta)]$$

$$
\begin{aligned}
s.t. \quad & \sum_{t'=1}^{t} x^e_{j,t'}(\xi,\eta) \le \xi^j && \forall j,t,\xi\in\Xi,\eta\in\Psi && \text{(A.1b)}\\
& 0 \le x^e_{j,t}(\xi,\eta) \le r_j+\eta_t && \forall j,t,\xi\in\Xi,\eta\in\Psi && \text{(A.1c)}\\
& x^e_{1,t}(\xi,\eta) \ge x^f_{1,t}(\xi,\eta)+x^f_{3,t}(\xi,\eta) && \forall t,\xi,\eta && \text{(A.1d)}\\
& x^e_{2,t}(\xi,\eta)+x^f_{1,t}(\xi,\eta) \ge x^f_{2,t}(\xi,\eta)+x^f_{5,t}(\xi,\eta) && \forall t,\xi,\eta && \text{(A.1e)}\\
& x^e_{3,t}(\xi,\eta)+x^f_{3,t}(\xi,\eta)+x^f_{5,t}(\xi,\eta) \ge x^f_{4,t}(\xi,\eta) && \forall t,\xi,\eta && \text{(A.1f)}\\
& \sum_{t'=1}^{t} x^c_{1,t'}(\xi,\eta) \ge x^f_{1,t}(\xi,\eta)+x^f_{3,t}(\xi,\eta) && \forall t,\xi,\eta && \text{(A.1g)}\\
& \sum_{t'=1}^{t} x^c_{2,t'}(\xi,\eta) \ge x^f_{2,t}(\xi,\eta)+x^f_{5,t}(\xi,\eta) && \forall t,\xi,\eta && \text{(A.1h)}\\
& \sum_{t'=1}^{t} x^c_{3,t'}(\xi,\eta) \ge x^f_{4,t}(\xi,\eta) && \forall t,\xi,\eta && \text{(A.1i)}\\
& 0 \le x^f_{i,t}(\xi,\eta) \le M y^l_{i,t}(\xi,\eta) && \forall i,t,\xi\in\Xi,\eta\in\Psi && \text{(A.1j)}\\
& 0 \le x^c_{j,t}(\xi,\eta) \le M y^p_{j,t}(\xi) && \forall j,t,\xi\in\Xi && \text{(A.1k)}\\
& y^p_t(\xi) \ge y^p_{t-1}(\xi) && \forall t,\xi\in\Xi && \text{(A.1l)}\\
& y^l_t(\xi,\eta) \ge y^l_{t-1}(\xi,\eta) && \forall \xi\in\Xi,\eta\in\Psi && \text{(A.1m)}\\
& \sum_{j\in P} x^p_{j,t}(\xi) - \sum_{j\in P} x^p_{j,t-1}(\xi) \le 2 && \forall t,\xi\in\Xi && \text{(A.1n)}\\
& y^p_{t+1}(\xi) = y^p_{t+1}(y^p_t(\xi)\circ\xi) && \forall \in t,\xi && \text{(A.1o)}\\
& x^f_{i,t+1}(\xi,\eta) = x^f_{i,t+1}(y^p_t(\xi)\circ\xi,\eta_{[t+1]}) && \forall i,t,p,\xi,\eta && \text{(A.1p)}\\
& x^c_{j,t+1}(\xi,\eta) = x^c_{j,t+1}(y^p_t(\xi)\circ\xi,\eta_{[t+1]}) && \forall j,t,p,\xi,\eta && \text{(A.1q)}\\
& x^e_{j,t+1}(\xi,\eta) = x^e_{j,t+1}(y^p_t(\xi)\circ\xi,\eta_{[t+1]}) && \forall j,t,p,\xi,\eta && \text{(A.1r)}
\end{aligned}
$$

The objective of the deterministic counterpart formulation can be derived as following:

Part 1

$$
\begin{aligned}
&E_{\xi,\eta}\textstyle\sum_t d_t\left[c^g_t(\eta_t)(x^f_{2,t}(\xi,\eta)+x^f_{4,t}(\xi,\eta))\right]\\
&=\textstyle\sum_s p_s\sum_t d_t\left[(2+E(\eta_t))(X^f_{2,t,s}+X^f_{4,t,s})E_{\Xi_s}(\boldsymbol{\xi})\right]\\
&+\textstyle\sum_s p_s\sum_t d_t\left[2(X^{'f}_{2,t,s}+X^{'f}_{4,t,s})E(\eta'_{[t]})+(X^{'f}_{2,t,s}+X^{'f}_{4,t,s})E(\eta_t\boldsymbol{\eta}'_{[t]})\right]
\end{aligned}
$$

Part 2

$$
\begin{aligned}
&-E_{\xi,\eta}\textstyle\sum_t d_t\sum_i c^l_i(y^l_{i,t}(\xi,\eta)-y^l_{i,t-1}(\xi,\eta))\\
&=-\textstyle\sum_t d_t\sum_i c^l_i\left[\sum_s(p_s y^l_{i,t,s})Y^l_{i,t,s}E(Q_{[t]}(\eta))-\sum_s(p_s y^l_{i,t-1,s})Y^l_{i,t-1,s}E(Q_{[t-1]}(\eta))\right]
\end{aligned}
$$

Part 3

$$
\begin{aligned}
&E\textstyle\sum_t d_t\sum_j\left[-c^p_j(y^p_{j,t}(\xi)-y^p_{j,t-1}(\xi))-c^c_j x^c_{j,t}(\xi,\eta)-c^e_j x^e_{j,t}(\xi,\eta)\right]\\
&=\textstyle\sum_t d_t\sum_s p_s\sum_j[-c^p_j(y^p_{j,t,s}-y^p_{j,t-1,s})-c^c_j(X^c_{j,t,s}E_{\Xi_s}(\boldsymbol{\xi})+X^{'c}_{j,t,s}E(\boldsymbol{\eta}'))-c^e_j(X^e_{j,t,s}E_{\Xi_s}(\boldsymbol{\xi})+X^{'e}_{j,t,s}E(\boldsymbol{\eta}'))]
\end{aligned}
$$

The following constraints model that the built infrastructure dost not vanish

$$y^l_{i,t,s} Y^l_{i,t,s} Q_t(\eta) \geq y^l_{i,t-1,s} Y^l_{i,t-1,s} Q_{t-1}(\eta) \qquad \forall t, s, i, \eta \tag{A.2}$$

$$y^p_{j,t,s} \geq y^p_{j,t-1,s} \qquad \forall j, t, s \tag{A.3}$$

Non-anticipativity for endogenous uncertainty is modeled below

$$\left|y^p_{j',t,s} - y^p_{j',t,s'}\right| \leq y^p_{j,t-1,s} \qquad \forall j = 1,2,3, j' = 1,2,3, s, s' : s_{-j} = s'_{-j} \tag{A.4}$$

$$\left|y^l_{i,t,s} - y^l_{i,t,s'}\right| \leq y^p_{j,t-1,s} \qquad \forall j = 1,2,3, j' = 1,2,3, s, s' : s_{-j} = s'_{-j} \tag{A.5}$$

$$\left|X^e_{i,t,s,j'} - X^e_{i,t,s',j'}\right| \leq M y^p_{j,t-1,s} \qquad \forall i = 1,2,3, j = 1,2,3, j' = 1,2,3,4, s, s' : s_{-j} = s'_{-j} \tag{A.6}$$

$$\left|X^c_{i,t,s,j'} - X^c_{i,t,s',j'}\right| \leq M y^p_{j,t-1,s} \qquad \forall i = 1,2,3, j = 1,2,3, j' = 1,2,3,4, s, s' : s_{-j} = s'_{-j} \tag{A.7}$$

$$\left|X^f_{i,t,s,j'} - X^f_{i,t,s',j'}\right| \leq M y^p_{j,t-1,s} \qquad \forall i = 1,2,3,4,5, j = 1,2,3, j' = 1,2,3,4, s, s' : s_{-j} = s'_{-j} \tag{A.8}$$

$$\left|X^f_{i,t,s,j}\right| \leq M y^p_{j,t-1,s} \qquad \forall i = 1,2,3,4,5, j = 1,2,3, s, t \tag{A.9}$$

$$\left|X^e_{i,t,s,j}\right| \leq M y^p_{j,t-1,s} \qquad \forall i = 1,2,3, j = 1,2,3, s, t \tag{A.10}$$

$$\left|X^c_{i,t,s,j}\right| \leq M y^p_{j,t-1,s} \qquad \forall i = 1,2,3, j = 1,2,3, s, t \tag{A.11}$$

Sample derivation for robust constraints

Consider following constraint

$$\sum_{t'=1}^{t} x^c_{2,t'}(\xi, \eta) \geq x^f_{2,t}(\xi, \eta) + x^f_{5,t}(\xi, \eta) \qquad \forall t, \xi, \eta$$

Apply the decision rule

$$\sum_{t'=1}^{t} (X^c_{2,t',s}\xi + X'^c_{2,t',s}\eta'_{[t']}) \geq X^f_{2,t,s}\xi + X'^f_{2,t,s}\eta'_{[t]} + X^f_{5,t,s}\xi + X'^f_{5,t,s}\eta'_{[t]} \qquad \forall t, \xi, \eta' \in \hat{\Psi}'$$

Further use the truncate matrix $R_{[t]}\eta' = \eta'_{[t]}$

$$(\sum_{t'=1}^{t} X^c_{2,t',s} - X^f_{2,t,s} - X^f_{5,t,s})\xi + (\sum_{t'}^{t} X^c_{2,t',s} R_{[t']} - X'^f_{2,t,s} R_{[t]} - X'^f_{5,t,s} R_{[t]})\eta' \geq 0 \qquad \forall t, \xi, \eta' \in \hat{\Psi}'$$

Based on the uncertainty set $\{\xi : w\xi \geq v\}$ and $\{\eta' : w'\eta' \geq v'\}$, then duality theorem is applied to the constraint with respect to ξ and η' uncertain parameters:

$$v_s^\top \theta(t,s) + v'^\top \phi(t,s) \geq 0 \qquad \forall t, s$$

$$w_s^\top \theta(t,s) = \sum_{t'=1}^{t} X^c_{2,t',s} - X^f_{2,t,s} - X^f_{5,t,s} \qquad \forall t, s$$

$$\theta(t,s) \geq 0 \qquad \forall t, s$$

$$w'^\top \phi(t,s) = \sum_{t'=1}^{t} X'^c_{2,t',s} R_{[t']} - X'^f_{2,t,s} R_{[t]} - X'^f_{5,t,s} R_{[t]} \qquad \forall t, s$$

$$\phi(t,s) \geq 0 \qquad \forall t, s$$

A.5 Chapter 5

This section include table for geological properties of Athabasca oil reservoir, table for deterministic model parameters, derivation for deterministic dual counter parts of stochastic model constraints, calculation procedure for expectation of objective function with respect to the uncertainty set.

Table A.28 provides the physical and geological parameters of Athabasca reservoir and Table A.29 presents the parameters of the deterministic model.

Table A.28: Athabasca reservoir properties [1, 2]

Parameter	Value	Parameter	Value
$w(m)$	50	γ_1	1.224
α (m^2/s)	7.06×10^{-7}	γ_2	0.816
m	4	$T_{steam}(^\circ C)$	263.9, 242.5
$\nu_s(m^2/s)$	10	$\tau(\%)$	81
$h(m)$	26.5	$H_{l\nu}(MJ/m^3)$	1640, 1753
S_{or}	0.15	$C_{\nu r}(MJ/m^3K)$	2.35
DA_{cell}	510	$C_{\nu o}(MJ/m^3K)$	2.38
$d_p(m)$	90×10^{-6}	$k_t(MJ/m \cdot K \cdot year)$	85
$A_\nu(m^{-1})$	6.67×10^{-4}	η_s	0.5

Table A.29: Model parameters [2, 3]

Parameter	Value	Parameter	Value
N_w	6	DR (%)	6
$L(m)$	850	RY (%)	10
N_{DA}	43	$CC_{CPF}(M\$)$	2480
$SAGD_{lifetime}$ (yr)	25	$DA_{cost}(M\$)$	78
Cap_{CPF} (m^3/yr)	5.25×10^6	CC_{DA} (M\$)	300
Cap_{steam} (m^3/yr)	16×10^6	$FC(M\$/m^3)$	0.3023×10^{-6}

Dual counter part derivation

Derivation procedure for the deterministic counterpart of the multistage stochastic programming model using linear decision rule is presented below.

Constraint 2b: $DA_{cost} \sum_{i \in \mathcal{I}} z_{i,t}(\xi) \leq CC_{DA}, \forall t \in \mathcal{T}, \forall \xi \in \Xi$

Apply decision rule of binary variable

$$DA_{cost} \sum_{i \in \mathcal{I}} z_{i,t} \leq CC_{DA}, \forall t \in \mathcal{T}$$

Constraint 2c: $\sum_{t \in \mathcal{T}} z_{i,t}(\xi) \leq 1, \forall i \in \mathcal{I}, \forall \xi \in \Xi$

Apply decision rule of binary variable

$$\sum_{t \in \mathcal{T}} z_{i,t} \leq 1, \forall i \in \mathcal{I}$$

Constraint 2d: $q_{i,t}^{oil}(\xi) \leq \sum_{k \in \mathcal{K}_i} z_{i,t-k+1}(\xi)\, \tilde{q}_{i,k}^{maxoil}(\xi) N_w, \forall i \in \mathcal{I}, t \in \mathcal{T}, \xi \in \Xi$

Use the uncertainty model, and apply the linear decision rule of continuous variable

$$Q_{i,t}\, \xi \leq \sum_{k \in \mathcal{K}_i} z_{i,t-k+1}\, q_{i,k}^{maxoil}(1+\zeta_i) N_w, \forall i \in \mathcal{I}, t \in \mathcal{T}, \xi \in \Xi$$

$$\max_{\xi \in \Xi} Q_{i,t}\, \xi - (1+\zeta_i) \sum_{k \in \mathcal{K}_i} z_{i,t-k+1}\, q_{i,k}^{maxoil} N_w \leq 0, \forall i \in \mathcal{I}, t \in \mathcal{T}$$

Define $\alpha_{i,t} = \sum_{k \in \mathcal{K}_i} z_{i,t-k+1}\, q_{i,k}^{maxoil} N_w$, $e_i = [0, \cdots, 0, 1, 0, \cdots, 0]$ (all zero except the $i+1$-th element being 1, such that $e_i \xi = \zeta_i$), and $b_{i,t} = Q_{i,t} - \alpha_{i,t} e_i$

$$\left\{ \max_{\xi \in \Xi} b_{i,t} \xi \right\} - \alpha_{i,t} \leq 0, \forall i \in \mathcal{I}, t \in \mathcal{T}$$

$$\left\{ \begin{array}{c} \max\ b_{i,t} \xi \\ s.t. - W\xi \leq -\nu \end{array} \right\} - \alpha_{i,t} \leq 0, \forall i \in \mathcal{I}, t \in \mathcal{T}$$

Introduce dual variables $\Lambda_{i,t}$ and apply duality to the inner LP problem

$$\left\{ \begin{array}{c} \min \quad -\nu \Lambda_{i,t} \\ s.t. \quad -W\Lambda_{i,t} = b_{i,t} \\ \Lambda_{i,t} \geq 0 \end{array} \right\} - \alpha_{i,t} \leq 0, \forall i \in \mathcal{I}, t \in \mathcal{T}$$

Drop the minimization operator

$$\left\{ \begin{array}{l} -\nu \Lambda_{i,t} - \alpha_{i,t} \leq 0, \forall i \in \mathcal{I}, t \in \mathcal{T} \\ -W\Lambda_{i,t} = b_{i,t}, \forall i \in \mathcal{I}, t \in \mathcal{T} \\ \Lambda_{i,t} \geq 0, \forall i \in \mathcal{I}, t \in \mathcal{T} \end{array} \right.$$

Replace $\alpha_{i,t}$ and $b_{i,t}$

$$\begin{cases} -\nu\Lambda_{i,t} - \sum_{k\in\mathcal{K}_i} z_{i,t-k+1}\, q_{i,k}^{maxoil} N_w \le 0, \forall i \in \mathcal{I}, t \in \mathcal{T} \\ -W\Lambda_{i,t} = Q_{i,t} - \sum_{k\in\mathcal{K}_i} z_{i,t-k+1}\, q_{i,k}^{maxoil} N_w e_i, \forall i \in \mathcal{I}, t \in \mathcal{T} \\ \Lambda_{i,t} \ge 0, \forall i \in \mathcal{I}, t \in \mathcal{T} \end{cases}$$

Constraint 2e: $\sum_{i\in\mathcal{I}} q_{i,t}^{oil}(\xi) \le Cap_{CPF}, \qquad \forall t \in \mathcal{T}, \xi \in \Xi$

Apply the linear decision rule of continuous variable

$$\sum_{i\in\mathcal{I}} Q_{i,t}^{oil}\xi \le Cap_{CPF} \qquad \forall t \in \mathcal{T}, \xi \in \Xi$$

$$\max_{\xi\in\Xi} \sum_{i\in\mathcal{I}} Q_{i,t}^{oil}\xi \le Cap_{CPF} \qquad \forall t \in \mathcal{T}$$

Introduce dual variable λ_t

$$\left\{ \begin{array}{c} \min \quad -\nu\lambda_t \\ s.t. \quad -W\lambda_t = \sum_{i\in\mathcal{I}} Q_{i,t}^{oil} \\ \lambda_t \ge 0 \end{array} \right\} \le Cap_{CPF} \qquad \forall t \in \mathcal{T}$$

Constraint 2e minimum is removed:

$$\begin{cases} -\nu\lambda_t \le Cap_{CPF} \qquad \forall t \in \mathcal{T} \\ -W\lambda_t = \sum_{i\in\mathcal{I}} Q_{i,t}^{oil} \qquad \forall t \in \mathcal{T} \\ \lambda_t \ge 0 \qquad \forall t \in \mathcal{T} \end{cases}$$

Constraint 2f: $\sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{K}_i} z_{i,t-k+1}(\xi)\, q_{i,k}^{steam} N_w \le Cap_{steam}, \forall t \in \mathcal{T}$

Apply binary decision rule

$$\sum_{i\in\mathcal{I}} \sum_{k\in\mathcal{K}_i} z_{i,t-k+1}\, q_{i,k}^{steam} N_w \le Cap_{steam}, \forall t \in \mathcal{T}$$

Constraint 2g: $x_{i+1,t}(\xi) = \sum_{t'\le t} z_{i,t'}(\xi), \forall i \in \mathcal{I}, t \in \mathcal{T}, \xi \in \Xi$

Apply binary decision rule

$$x_{i+1,t} = \sum_{t'\le t} z_{i,t'}, \forall i \in \mathcal{I}, t \in \mathcal{T}$$

Constraint 2h: $q_{i,t}(\xi) \ge 0, \forall i \in \mathcal{I}, t \in \mathcal{T}, \xi \in \Xi$

Apply decision rule for continuous variable

$$\left\{ \begin{array}{c} \min_{\xi} Q_{i,t}\xi \\ s.t.\ \ W\xi \geq \nu \end{array} \right\} \geq 0, \forall i \in \mathcal{I}, t \in \mathcal{T}$$

Introduce dual variable $\mu_{i,t}$

$$\left\{ \begin{array}{c} \max \ \ \nu\mu_{i,t} \\ s.t.\ \ W\mu_{i,t} = Q_{i,t} \\ \mu_{i,t} \geq 0 \end{array} \right\} \geq 0, \forall i \in \mathcal{I}, t \in \mathcal{T}$$

Finally, after dropping maximizing operator

$$\left\{ \begin{array}{ll} \nu\mu_{i,t} \geq 0 & \forall i \in \mathcal{I}, t \in \mathcal{T} \\ W\mu_{i,t} = Q_{i,t} & \forall i \in \mathcal{I}, t \in \mathcal{T} \\ \mu_{i,t} \geq 0 & \forall i \in \mathcal{I}, t \in \mathcal{T} \end{array} \right.$$

Expectation of objective function with respect to uncertainty set

Calculation procedure for expectation of objective function with respect to the uncertainty set is provided below.

$$\sum_t \frac{1}{(1+DR)^t} \mathbb{E}\left[(1-RY) OP_t(\xi) \sum_{i \in \mathcal{I}} q_{i,t}^{oil}(\xi) - DA_{cost} \sum_i z_{i,t}(\xi) - FC \sum_i \sum_{k \in \mathcal{K}_i} z_{i,t-k+1}(\xi) q_{i,k}^{steam} N_w \right]$$

First term

$$\sum_t \frac{1}{(1+DR)^t} \mathbb{E}\left[(1-RY) OP_t(\xi) \sum_{i \in \mathcal{I}} q_{i,t}^{oil}(\xi) \right]$$

$$= \sum_t \frac{1}{(1+DR)^t} \mathbb{E}\left[(1-RY)(A_t P\xi + b_t) \sum_{i \in \mathcal{I}} Q_{i,t}^{oil}\xi \right]$$

$$= \sum_t \frac{1-RY}{(1+DR)^t} \mathbb{E}\left[A_t P\xi \sum_{i \in \mathcal{I}} Q_{i,t}\xi + b_t \sum_{i \in \mathcal{I}} Q_{i,t}\xi \right]$$

$$= \sum_t \frac{1-RY}{(1+DR)^t} \mathbb{E}\left[(A_t P\xi)^T \sum_{i \in \mathcal{I}} Q_{i,t}\xi + b_t \sum_{i \in \mathcal{I}} Q_{i,t}\xi \right]$$

$$= \sum_t \frac{1-RY}{(1+DR)^t} \mathbb{E}\left[\xi^T (P^T A_t^T \sum_{i \in \mathcal{I}} Q_{i,t})\xi + b_t \sum_{i \in \mathcal{I}} Q_{i,t}\xi \right]$$

$$= \sum_t \frac{1-RY}{(1+DR)^t} \left[tr(P^T A_t^T \sum_{i \in \mathcal{I}} Q_{i,t}\ \mathbb{E}(\xi\xi^T)) + b_t \sum_{i \in \mathcal{I}} Q_{i,t} \mathbb{E}(\xi) \right]$$

Second term

$$\sum_t \frac{1}{(1+DR)^t}\mathbb{E}\left[-DA_{cost}\sum_i z_{i,t}(\xi)\right] = \sum_t \frac{1}{(1+DR)^t}\left[-DA_{cost}\sum_i z_{i,t}\right]$$

Third term

$$\sum_t \frac{1}{(1+DR)^t}\mathbb{E}\left[-FC\sum_i\sum_{k\in\mathcal{K}_i} z_{i,t-k+1}(\xi)q_{i,k}^{steam}N_w\right]$$

$$= \sum_t \frac{1}{(1+DR)^t}\left[-FC\sum_i\sum_{k\in\mathcal{K}_i} z_{i,t-k+1,s}q_{i,k}^{steam}N_w\right]$$

Combined result:

$$\sum_{t\in\mathcal{T}} \frac{1-RY}{(1+DR)^t}\left[tr(P^T A_t^T \sum_{i\in\mathcal{I}} Q_{i,t}\,\mathbb{E}(\xi\xi^T)\,) + b_t\sum_{i\in\mathcal{I}} Q_{i,t}\mathbb{E}(\xi)\right]$$
$$-\sum_{t\in\mathcal{T}}\frac{DA_{cost}}{(1+DR)^t}\sum_i z_{i,t} - \sum_{t\in\mathcal{T}}\frac{N_w FC}{(1+DR)^t}\sum_i\sum_{k\in\mathcal{K}_i} z_{i,t-k+1}q_{i,k}^{steam}$$