ADDIS ABABA UNIVERSITY
DEPARTMENT OF MATHEMATICS

A MULTIPARAMETRIC PROGRAMMING APPROACH FOR MULTILEVEL OPTIMIZATION

By

Abay Molla

A THESIS

SUBMITTED TO THE DEPARTMENT OF MATHEMATICS OF ADDIS ABABA UNIVERSITY IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR MASTER OF SCIENCE DEGREE IN MATHEMATICS

Advisor

Dr. Semu Mitiku

Addis Ababa, Ethiopia

July 7, 2011

Addis Ababa University

School of Graduate Studies

THESIS APPROVAL SHEET

Thsis Title: *A multiparametric programming approach for multilevel optimization.*

Student's Name: Abay Molla

We verify that this thesis satisfies the requirements of the Graduate School

Chairman of ~~department~~ Exam. board **Signature**

Semu Mitiku

Advisor **Signature**

Berhanu Guta

~~External~~ Examiner 1 **Signature**

Mengistu Goa

~~Internal~~ Examiner 1 **Signature**

(Date)

Declaration

"I declare that this thesis has been composed by me and that no part of the thesis has formed the basis for the award of any Degree, Diploma, Associate ship, fellowship or any other similar title to me.

Abay Molla

Author's Signature"

Permission

"This is to certify that this thesis is compiled/a record of the research work done/by Mr. Abay Molla in the Department of Mathematics, Addis Ababa University, under my supervision. I hereby also confirm that the thesis can be submitted for evaluation by examiners and eventual defense.

Advisor's Signature"

Acknowledgement

During my study, I was fortunate to interact with a number of people from whom I have benefited greatly. I am grateful to those who have offered their advice, assistance, encouragement and friendship. Most of all, my thanks goes to Dr. Semu Mitiku. He always shares willingly of his knowledge and many ideas. I also recall and appreciate his invaluable unreserved material and moral support to start my study earlier and to complete it on time. His encouragement and appreciation throughout my study at the university has been very indispensable. Throughout my study, he has taken active interest in my work, and I consider myself very fortunate to have been under his direction. Our interaction has for me been an invaluable learning experience.

I am highly indebted and grateful to my lecturers and friends especially my best friend Ato Temesgen B. for his encouragement and material support in every aspects for the completion of this thesis.

I have to thank the Mathematics department of Addis Ababa University for giving me the opportunity to do my graduate work and teaching me far more about life than Mathematics. Special thanks should go to Dr. Semu Mitiku head department of Mathematics for all his help and encouragement.

Finally, I would like to forward my grateful thanks to my families. Specially, special thanks goto my sisters Eneye Molla and Asres Molla. They encourage, help and advise me from at the beginning of my life. without them none of this would have mattered.

Dedication

This thesis is dedicated to
my mother Mebirate Wolle and my father Molla Kassa.

Abstract

Multilevel optimization problems are mathematical programs which have a subset of their variables constrained to be an optimal solution of other programs parameterized by their remaining variables. It is implicitly determined by a series of optimization problems which must be solved in a predetermined sequence. To be more precise, the inner level problem is a multiparametric programming problem parameterized by upper level optimization variables. The solution approach for multilevel programming problems is to represent the inner level problem with sufficient conditions and augment it in the upper level constraints. As a result, without convexity assumption for multilevel optimization at inner levels, stationary points may not be sufficient or optimal for the inner level problems and the set of all stationary points may not be connected. This implies it is impossible to use the approach of augmenting the conditions of the lower level problems into the constraints of the upper level problem. Recently, researchers have proposed a solution strategy for multilevel optimization via multiparametric programming approach by considering the follower's problem as a multiparametric optimization problem. But, their proposed algorithms work only for problems with convex, quadratic or linear problems.

In this work, we present the foundations of a general global optimization algorithm for the solution of general multilevel problems based on the recent developments in multiparametric programming theory. Specifically, we outline the general global optimization strategy for the solution of bilevel and trilevel programming problems with nonconvexity formulation at the inner levels and we have proved $\varepsilon-$convergence of the algorithm.
Keywords: Multiparametric programming, Multilevel optimization

Contents

List of Tables

1 Introduction

1.1 Background of the study

In many real-world problems decisions have been made in a hierarchical order where the individual decision makers have no direct control or influence upon the decisions of the others, but their actions affect all other decision makers. Specially, the higher level (or leader) of the hierarchy have the power to strongly influence the performance and strategies of the decision makers at followers [34].

Group decision making process is an energetic force at the moment in humankind. One of the most widespread decision making structure throughout history has been the hierarchy [41]. Hierarchies exist virtually in every part of practical life. Forexample, many industrial situations involve several groups which are interconnected in hierarchical structure [23]. They also arise in the governments, business centers, firms, religious organizations, and in our families as well [30].

The design of practical systems involving a large number of subsystems with multiple load situations, involve excessive number of design variables and constraints [39]. As the result, the optimization problem becomes unmanageably large and the solution process becomes too expensive. In such cases the optimization problem can be broken into a series of smaller problems using different strategies. Among these strategies multilevel optimization is one of a decomposition technique in which the problem is reformulated as several smaller subproblems and a coordination problem to preserve the coupling among the problems [8], [34], [39].

Multilevel optimization problems are mathematical programs which have a subset of their variables constrained to be an optimal solution of other programs parameterized by their remaining variables [43]. It's implicitly determined by a series of optimization problems which must be solved in a predetermined sequence. In terms of modeling the constraint domain associated with multilevel programming problem consist of decision makers in the hierarchy who make decisions in a structured, "leader-follower" ordering.

The mathematical foundation and the internal uniformity of multilevel optimization theory make it a major tool for modeling and designing, a programmed decision making process in interactive environment [30]. The process of modeling a situation as a multilevel optimization requires the decision maker to enumerate implicitly the decision and their considered options as well as to consider their preference and reaction.

An important feature of the so called "Multilevel" programming is that a planner at one level of a hierarchy of planners may have his/her objective function determined in part, by variables controlled at other levels [24], [30]. However, each planner's control instruments may allow him/her to influence the policies at other levels and thereby improve his/her own objective function. Such problems have the following characteristics in common:

- The system has interacting decision making units within a hierarchical structure. Each subordinate level performs its policies after knowing completely the decisions of superior levels.
- Each unit maximizes net benefits independently of other units but may be influenced by actions and reactions of those units.
- The external effect on a decision-maker's problem can be reflected in both the objective function as well as on the set of feasible decisions.

Generally, in the hierarchy of decision makers each level of the hierarchy wishes to maximize its individual benefits in view of partial exogenous control exercise at other level.

Many resource planning problems require compromises among the objectives of several interacting individuals or agencies often, these groups are arranged within an administrative or hierarchical structure with independent and perhaps conflicting objectives.

The aim of multilevel optimization theory is to investigate the way in which rational peoples should interact when they have conflicting interests [30]. For instance, consider a programming problem in which the government is at the first level. During the planning period the government proposes a certain goal. In order to optimize the achievement of such goals, it formulates certain policy measures such as taxes and subsidies. The industries at the second level design their course of action keeping such policy measures in mind so that their objectives are fulfilled. Consequently, the industries supply their products to the customers in a certain area. The customers at the third level are at liberty to make their purchase from any industry. In doing so the customers will consider economic criteria such as cost optimization. This is a trilevel programming problem in which the government objectives are at least in partial conflict with the two sectors industry and customers, the policy makers face an optimization problem subject to the optimization problems for industries as well as for the customers. Very often they can benefit through cooperation. Cooperation behavior often emerges at a group rather than at individual level; in many instances we observe the formation of groups, teams, committees, and clubs cooperativeness each of them persecuting the same goal (in turn provision of commodities, maximization of profits, and so on). It is common to focus on what each level ought in some sense to agree. The natural restriction on such agreement is each level gets at least as large as payoffs as they can guarantee them self. Therefore, multilevel optimization can be applied in Government sectors, Market economy, Design of transportation networks and Process systems engineering problems.

Despite their significance, general solution strategies for solving such complex problems are limited, especially due to the multi-layer nature, non-linearities and nonconvexities occur [43]. In addition, the potential presence of logical decisions (which requires the inclusion of binary variables), fluctuations in resource, market requirements and prices increases further the complexity of the problem. Moreover, Ben-Ayed and Blair [9] in 1990 described the difficulty of bilevel programming problem.From the literature one can know that bilevel linear programming problem is NP-hard [9].

According to Pistikopoulos [37] in 1995, the model of many process engineering problems officially (or formally) involve varying parameters. As the literature said varying parameters can affect the feasibility and economics of the task. Eventhough Migdalas, Paradalos and Värbrand [34] in 1992 set out the difficulties of multilevel optimization during numerical approximation, for about half a century mathematical programming has been successfully applied to decision making problems that arise from various application fields and have covered many area of human activities

In the late nineties Bahatia and Bieger [6] have been explained an approach (Multi-period, to name) to handle such problems. Also, Acevedo and Pistikopoulos [2] in 1998, described another approach (stochastic programming, to name). Recently, Pistikopoulos *et al.*, [15], [16], [38] have proposed novel solution algorithms, based on parametric programming theory, which open the possibility to address general classes of multilevel programming problems. The advantage of using the proposed approach to address these problems is that one can obtain a complete map of all the optimal solutions. Moreover, Jezowski and Thullie [28] in 2009, have reported in the literature, parametric programming is a chief rule for analyzing the effect of varying parameters in any mathematical programming. However, each approach depends on the description of the varying parameters.

In 2002 Dua *et al.*, [16] highlighted the aim of multiparametric programming approach, and computational complexity of some class of multiparametric programming. Further, in 1996, Acevedo and Pistikopoulos [1] provided a detailed theory and algorithm for the solution of a wide range of parametric programs. In 2007 Pistikopoulos, Dua and Georgiads [38], Pardalos and Coleman [36] in 2009, have been expanded the general theory, algorithm, characteristics and applications of multiparametric programming.

1.2 Historical Development

Since the advent of linear programming and game theory in the 1940's and 1950's, a substantial effort has been directed toward analyzing the behaviors of interacting decision makers, each attempting to optimize individual objectives in view of decisions made by others. Many scientific disciplines have contributed to ward analyzing problems associated with hierarchical organizations including engineering, operations research, businesses, game theory, sociology, economics and political science [36].

The execution of many decisions in businesses is sequential, from a higher level (leader) to a lower levels (followers); each unit independently optimizes its own objective, but is affected by other unit's actions through externalities. This is called multilevel programming (MLP) problem (also called multilevel decision or multilevel optimization problems). Multilevel programming was motivated early in 1952 by the game theory of Von Stackelberg [42] in framework of unbalanced economic market. On wards in the beginning of seventies Bracken and McGill [10] open the novel formulation of bilevel programming. Furthermore, in 1980, Bialas and Karwan [8] plan the general multilevel linear programming problems. Moreover, the basic principles of multilevel optimization were encountered by Goertzel in 1989, on his PhD thesis [25]. However the basic idea of the multilevel philosophy had been proposed by the sociologist Etzoine (1968), in his adaptive society as a method for optimizing the social structure.

The multi-layer nature in multilevel problems results in non-linearities and nonconvexities hence, it is not surprising that general solution strategies for solving such complex problems are rather limited. But, it is widely accepted that a global optimization approach is needed for the solution of such multilevel problems [18]

As reported in the literature [30], many solution approaches have been developed for multilevel programming problems. The approach proposed by Candler and Townsly in 1982, focuses on generating and enumerating bases from the lower level activities. The limitation of the algorithm is that it may not stop as soon as the goal optimal is attained. In the same literature the approach commonly known as simplex method also described to address multilevel optimization problems. It is applicable for bilevel programming with bounded variables. The method finds the extreme points in the set of rational reactions of the most lower level problem, never allowing the upper level objective function to optimize. The main drawback of this method is that only the local optimal solution is obtained. Moreover, the K^{th}-best algorithm, which was proposed by Wen, Bialas and Karwan in 1983, moves sequentially through ordered extreme points of the overall solution space until the K^{th}-best one is found in the rational reaction set of the lower level problem and then terminates with global optimal solution. However, the algorithm consumes long time before the solution is found.

Faísca *et al.*, in 2007 have been developed multiparametric programming approach to multilevel problems [18], makes possible the development of a unified strategy for their solution to global optimality.

1.3 Scientific Motivation

The hierarchical optimization structure appears naturally in many applications when the followers' actions depend on upper level decisions. The applications of bilevel and multilevel programming include transportation(taxation, network design, trip demand estimation), management (coordination of multidivisional firms, credit allocation) and planning (agricultural policies, electric utility). Recently, multilevel optimization problems have been an area of active research in the world that focuses on the whole hierarchy structure [5], [30]. Such problems have been long recognized as an important decision making problem.

In 1981 Candler *et al.*, [11] described the role of multilevel programming in agricultural economics in their published paper. Furthermore, the availability of multilevel programming problem arrives at engineering design as described by Kocvara and Outrata [29] in 1993. According to Vicente and Calamai [43] in 1994, the structure of bilevel and multilevel programming lead to many practical problems that are useful for many applications. In such applications Leblanc and Boyce [31] and Marcotte [33], together described the claim/or application of multilevel optimization in transportation network design problem. Further, Multilevel optimization problems have attracted considerable attention from the scientific and economic community in recent years. Due to its many applications, multilevel and in particular bilevel programming have evolved significantly [17], [38].

Mathematical programming methods to solve such problems trace back early in the development of linear programming [30]. As a result, the majority of research on bilevel programming has centered on the linear version of the problem in which only one follower is involved. There have been few dozen algorithms, such as; the K^{th} best approach, Kuhn-Tucker approach, complementarity pivot approach, penalty function approach and Simplex-Cutting Plane algorithm approach proposed for solving linear bilevel programming (BLP) problems since the field being caught the attention of researchers in the last three decades [32]. Likely, in the beginning of 21^{th} century the algorithms [13] such as; A descent algorithm, A bundle algorithm, Penalty methods, A trust region method and Smoothing methods are also proposed for solving bilevel programming problems.

Using these approaches many researchers have produced useful solutions of multilevel optimization problems. In practice however, the proposed approaches have been faced by more complex problems. For example the problem like, the inclusion of nonconvexity formulation in the most inner-level optimization problem. This type of problem is not adequately addressed by the approaches discussed above, to my best knowledge.

Thereby a few untrustworthy approaches yet known in solving three level programming problems [30]. Among these, Lakie E., in 2007, adapted the " K^{th}-best" algorithm for solving trilevel linear programming problems and Zhang G. *et al.*, [45] in 2009, also used the same approach for solving the same type of problem. Likely, Hybrid method is one of the approaches offered by Wen in 1981, which combines the " K^{th}-best" vertex enumeration algorithm and complimentary pivot algorithm. Eventhough this method works satisfactorily for most problems, its computational complexity grows geometrically with the number of constraints (i.e. the number of levels in the hierarchy). Moreover, Bard, in 1983 extended the idea of the grid search algorithm, which is designed to solve two level of hierarchy to a model of three level hierarchies [30]. The algorithm that Bard proposed for solving three level programming problems includes a cutting plane approach for solving a bi-linear programming problem and a vertex search procedure for the third level at each iteration. One of the main advantages of this algorithm is that, it can be extended beyond three level hierarchies and can be used for general multilevel linear programming problems. Its principal limitation seems to be the bookkeeping burden imposed by the prospect of multiple optimal solutions.

In the papers [1], [17], [19], [38], the authors investigated the solutions of bilevel optimization problems using parametric approach and other approaches. In their work they were limited on problems with convex, quadratic and linear objective functions defined on convex sets. Even they did not indicate solution methods when the numbers of optimization variables are greater than two. Further, in [36], [18] the researchers investigated the solutions of bilevel programming problems with multi-follower and trilevel programming problems. Eventhough they overcome some limitations of the above all discussed approaches, their proposed algorithms work only for problems with convex, quadratic or linear problems.

Until now several algorithms have been developed that can find an optimal solution for the linear bilevel and trilevel programming problem and some algorithms that solves nonlinear programming problem. However the computational efficiency of these algorithms does not consistently perform well, due to the complexity of the problem. And the algorithms proposed for solving three level programming problems have some limitations as we have seen above.

It is evident that multilevel optimization problems are important in design of practical systems involving a large number of subsystems with multiple load situations; involve excessive number of design variables and constraints [39]. In most cases such problems may be nonconvex and nonlinear classes of multilevel optimization problems. So, one can ask, is it possible to address such classes of multilevel optimization problems using multiparametric programming approach? But, the existing studies do not give a full answer for such questions. So, it will be helpful to investigate more efficient algorithm for solving bilevel programming problem and extend it as well to the general k-level programming problems; In addition, there exist particular problems involving multilevel programming, further study of which will be interesting.

Multiparametric programming [16], [38] is a natural tool for analyzing the effects of parameters in the solution of parametric problems. Overall, it is a technique for solving multilevel optimization problem, where the objective is to minimize or maximize a performance criterion subject to a given set of constraints and where some of the parameters vary between specified lower and upper bounds [38]. It comprises a number of strategies and algorithms to deal with different problem formulations and where a key feature is identified: parametric uncertainty [38].

Many people have used multiparametric programming as a tool to answer questions related to multilevel optimization problem. Faísca *et al.*, [17] proposed a global optimization approach for the solution of various classes of bilevel programming problems (BLPP) based on recently developed parametric programming algorithms. Further, Faísca *et al.*, [19] in 2006 outlined the foundations of a general global optimization strategy for the solution of multilevel hierarchical and general decentralized multilevel problems based on parametric programming approach. Moreover, Faísca, Saraiva and Rustem [18] in 2007, brought into play the multiparametric programming approach for solving multilevel hierarchical and decentralized optimization problems.

Recently, researchers proposed the solution method for multilevel optimization problems using multiparametric programming approach, by considering the followers problem as a multiparametric optimization problem, where the decision variables at upper levels could be seen as parameters for the lower level problems. The core idea of the proposed approach is to recast each optimization subproblem in the multilevel hierarchy as a multiparametric programming problem and then transform the multilevel problem into a single level optimization problem [19].

The purpose of this study is to present a general and flexible framework for multilevel optimization problems over connected region through multiparametric programming approach. This new framework is a slight extension of the existing framework for multilevel hierarchical and decentralized optimization problems.

1.4 Overview

In each level there is decision maker who make decision heterogeneously and finally, decision makers decide in a coopratíve mode. The main focus of this thesis is to address the solution of general multilevel optimization up to three levels through multiparametric programming approach with a single decision maker in each level

In section 2, the general multilevel mathematical definition and formulation is formally introduced, which is used throughout the paper, and respective definitions of feasible and rational reaction set. For a decision maker to make a rational decision, he/she needs to determine the rational reaction of the decision maker who is lower in the hierarchy. This rational set is mathematically derived. It also briefly introduces the relevant multiparametric programming theory and algorithms in section 3. The proposed underestimators of some classes of nonconvex programming problems are then described in detail in chapter 4. The proposed methodology for the solution of bilevel programming problem with nonconvex formulation in the inner problems and trilevel programming problem with the inclusion of nonconvexity in the third-level is then described in detail in Section 5, and illustrated with examples.

2 Preliminaries

This section is devoted mainly to definitions and formulations of optimization problems. In the first subsection some mathematical concepts which contain some basic definitions of sets and functions are introduced. The concept on general definitions of multilevel optimization and problem formulations are presented in the second and third subsection respectively.

2.1 Some mathematical concepts and notations

This subsection contains some basic notations and basic definitions of mathematics which are available for this work.

2.1.1 Basic Notations

If S and Q are two sets and f is a mapping (function) defined on S with its values lying in Q, we can write as; $f : S \to Q$.
$\mathbb{R}-$Stand for the set of real numbers.
$\square-$ Denote end of the proof.
$\mathbb{R}^n-$ Euclidean space with axes, $x_1, x_2, \ldots, x_n$.
$X, Y \in \mathbb{R}^n$ will be denoted as $X = (x_1, x_2, \ldots, x_n)$ and $Y = (y_1, y_2, \ldots, y_n)$.
For $M \subseteq \mathbb{R}^n$, $C(M)$ denotes the space of all continuous function on M. If $\Omega \subseteq \mathbb{R}^n$ is an open set and $k \geq 1$, is an integer, then $C^k(\Omega)$ denotes the space of all functions which have continuous partial derivatives up to order k in Ω.

2.1.2 Definitions of Sets and Functions

As convexity of sets as well as functions plays an important role in the theory of optimization problems, we need to have the concept of convex sets and convex functions. Further, some other basic definitions of mathematical terms which are used throughout this work are introduced.

Definition 2.1 *A set S is said to be convex if for any $x, y \in S, \lambda x + (1-\lambda)y \in S, \forall \lambda \in [0, 1] \subseteq \mathbb{R}$.*

Definition 2.2 *Let $S \subseteq \mathbb{R}^n$ be a convex set. A functional $f : S \to \mathbb{R}$ is said to be convex if for any $x, y \in S$ and $\forall \lambda \in [0, 1], f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$.*

Definition 2.3 *A topological space is said to be compact if every open cover of the entire space has a finite subcover.*

Definition 2.4 *We say a function $L : \mathbb{R}^m \to \mathbb{R}^m$ is linear, if for any vector x and y in $\mathbb{R}^m$ and α, β in $\mathbb{R}$,*
$L(\alpha x + \beta y) = \alpha L(x) + \beta L(y)$

Definition 2.5 *We say a function $A : \mathbb{R}^m \to \mathbb{R}^n$ is affine if there is a linear function $L : \mathbb{R}^m \to \mathbb{R}^n$ and a vector b in $\mathbb{R}^n$ such that $A(x) = L(x) + b$, for all $x \in \mathbb{R}^m$.*

Definition 2.6 *Let $\{a_n\}$ be a sequence of real numbers and let d be a real number. The sequence $\{a_n\}$ is said to converge to d if for every positive number ϵ, there exists a natural number $\mathbb{N}$ such that for any $n \geq \mathbb{N}$, $| a_n - d | < \epsilon$*

Definition 2.7 *Let $(X, \| \cdot \|)$ be a normed space. A sequence $\{x_n\}_{n=1}^{\infty}$ of elements of X is called weakly convergent to some $x_0 \in X$ if for all continuous linear functional L on X*

$$\lim_{n \to \infty} L(x_n) = L(x_0)$$

Definition 2.8 *Let $(X, \| \cdot \|)$ be a normed space. A nonempty subset S of X is called weakly sequentially compact if every sequence in S contains a weakly convergent subsequence whose weak limit belongs to S.*

Definition 2.9 *A set P is a polyhedron if there is a system of finitely many inequalities $Ax \leq b$ such that, $P = \{x \in \mathbb{R}^n : Ax \leq b\}$*

Definition 2.10 *A point $x_0 \in S$ is said to be a minimal point of a real valued function f defined on S if for any point $x_1 \in S$ we have $f(x_1) \leq f(x_0)$. Where, S is the set of constraints and note that the optimal value of the objective function is a value of f evaluated at x_0.*

In an optimization problem were the objective function is to be maximized the optimal value is the least upper bound of the objective function values over the entire feasible region. If there is no upper bound, then we say that the optimal value is $+\infty$, while if the feasible region is the empty set, we define the optimal value of a maximization problem to be $-\infty$.

Conversely, in an optimization problem were the objective function is to be minimized the optimal value is the greatest lower bound of the objective function values over the entire feasible region. If there is no lower bound, then we say that the optimal value is $-\infty$, while if the feasible region is the empty set, we define the optimal value of a minimization problem to be $+\infty$

2.2 Multilevel Optimization

Hierarchical optimization deals with mathematical programming problems whose feasible set is implicitly determined by a sequence of nested optimization problems. In recent years, multilevel optimization approach has been found to be the most sensible way to handle large and complex design problems. This approach is based on different coordination methods, which have been found very efficient to solve such problems by horizontal or vertical decomposition of the original problem. Coordination and behavior variables are introduced to maintain interactions among individually optimized levels and subsystems at a level. This coordination process by the behavior variables leads the design towards an overall system optimum.

Multilevel programming techniques are developed to solve decentralized planning problems with multiple decision makers in a hierarchical organization. These become more important for modern decentralized organizations where each unit seeks its own interests. In the decentralized systems, there is one higher-level decision maker (who is referred to as the center/leader) and many lower level decision makers (who are referred to as divisions/followers) while in the multilevel systems, there are many levels with one or more decision maker(s) at each level. In both systems, the leader makes decision first and the follower reacts by optimizing the objective function conditioned on the leader's decision. But our work mainly is concerned in one decision maker at each level instead of many decision makers at the same level horizontally.

2.2.1 General Definition of Multilevel Optimization

To define a k-level optimization problem formally, consider an organization composed of k-levels, each characterized by individual objective functions f_i for $i = 1, 2, \ldots, k$, which are to be minimized/maximized by the respective decision makers. Let the decision variable space $\mathbb{R}^n$ be partitioned among k levels, such that $(x_1, x_2, \ldots, x_k) \in X_1 \times X_2 \times \cdots \times X_k$. Where, $X_i \subseteq \mathbb{R}^{n_i}$ for $i = 1, 2, \ldots, k$ and $\sum_{i=1}^{k} n_i = n$. Without loss of generality, assume that decisions are made sequentially beginning with decision maker 1 who has control over a vector $x_1 \in X_1$, followed by decision maker 2 who has control over a vector $x_2 \in X_2$ down through decision maker k who has control over a vector $x_k \in X_k$ where X_i for $i = 1, 2, \ldots, k$ is nonempty and compact subsets of $\mathbb{R}^{n_i}$.

The general multilevel programming problem can then be defined as:

$$
\begin{aligned}
&\min_{x_1} f_1(x_1, x_2, \ldots, x_k)\\
&\quad \text{s.t.}\\
&\quad g_1(x_1, x_2, \ldots, x_k) \le 0\\
&\qquad\qquad\qquad \text{where } [x_2, x_3, \ldots, x_k] \text{ solves}\\
&\quad \min_{x_2} f_2(x_1, x_2, \ldots, x_k)\\
&\qquad \text{s.t.}\\
&\qquad g_2(x_1, x_2, \ldots, x_k) \le 0\\
&\qquad\qquad\qquad \text{where } [x_3, x_4, \ldots, x_k] \text{ solves}\\
&\qquad \min_{x_3} f_3(x_1, x_2, \ldots, x_k)\\
&\qquad\quad \text{s.t.}\\
&\qquad\quad g_3(x_1, x_2, \ldots, x_k) \le 0\\
&\qquad\qquad\qquad \text{where } [x_4, \ldots, x_k] \text{ solves}\\
&\qquad\qquad\qquad \ddots\\
&\qquad\qquad\qquad \min_{x_k} f_k(x_1, x_2, \ldots, x_k)\\
&\qquad\qquad\qquad\quad \text{s.t.}\\
&\qquad\qquad\qquad g_k(x_1, x_2, \ldots, x_k) \le 0
\end{aligned} \tag{2.1}
$$

where, $f_i : X_1 \times X_2 \times \cdots \times X_k \to \mathbb{R}$ and g_k's are vector valued functions.

Let us denote the minimization of a function $f(\cdot)$ over a compact region $X \subseteq \mathbb{R}^n$ by varying only $x_k \subseteq X_k$ for given, $(x_1, x_2, \ldots, x_{k-1})$ by:
$\min\{f(x_1, x_2, \ldots, x_k) : (x_k | x_1, x_2, \ldots, x_{k-1}), (x_1, x_2, \ldots, x_k) \in X : g(x_1, x_2, \ldots, x_k) \le 0\}$,
where, $X = X_1 \times X_2 \times \cdots \times X_k$ be compact set in $\mathbb{R}^n$.

Definition 2.11 *The set $\Psi_f(X)$ defined by $\Psi_f(X) = \{(\bar{x}_1, \bar{x}_2 \ldots, \bar{x}_k) \in X | f(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k)$ $= \min\{f(x_1, x_2, \ldots, x_k) : (x_k | x_1, x_2, \ldots, x_{k-1}), g_k(x_1, x_2, \ldots, x_k) \le 0\}\}$ is known as the set of rational reactions of f over X*

Let $f_1(x_1, x_2, \ldots, x_k), f_2(x_1, x_2, \ldots, x_k), \ldots, f_k(x_1, x_2, \ldots, x_k)$ be bounded functions defined over a compact set X. After the decisions of level 1 to $k-1$ are made, the values of variables $x_1, x_2, \ldots, x_{k-1}$ are determined. For given $(x_1, x_2, \ldots, x_{k-1}) \in X_1 \times X_2 \times \cdots \times X_{k-1}$, the decision making problem of the k^{th}-level is equivalent to the following minimization problem.

$$\min_{x_k}\{f_k(x_1, x_2, \ldots, x_k) : (x_k | x_1, x_2, \ldots, x_{k-1}), \Omega_{k-1} = \{x_k \in X_k : g_k(x_1, x_2, \ldots, x_k) \le 0\}\}$$

where, g_k is a vector valued function. Based on this optimization problem, we give the following definitions: The set, Ω_{k-1}, is defined to be the k^{th}–level *feasible region*. Note that the feasible region for each follower is affected by the leaders' choice. The rational reaction set for the k^{th}–level, $\Psi_{k-1} = \Psi_{f_k}(X_k)$ may contain the $(k-1)^{th}$–level feasible region. Where, $\Psi_{f_k}(X_k) = \{x_k \in X_k : x_k \in \arg\min\{f_k(x_1, x_2, \ldots, x_k) : (x_k|x_1, x_2, \ldots, x_{k-1}), g_k(x_1, x_2, \ldots, x_k) \leq 0\}$. Thus, we have a relationship between rational reaction sets and feasible sets of the multilevel optimization problems, i.e. the feasible set for the $(k-1)^{th}$–level is a subset of the rational reaction set for the k^{th}–level in the hierarchy.

Note that, if $k = 1$ and f_k is nonlinear, problem (2.1) reduces to standard nonlinear programming problem. When $k = 2$ then problem (2.1) is known as *bilevel programming problem*. When $k = 3$ problem (2.1) is known as *trilevel programming problem*.

Definition 2.12 *A point $X_0 = \{x_1, x_2, \ldots x_k\}$ is an optimal point to a multilevel programming problem if X_0 is an optimal point to a Leader's problem, satisfying lower level problems as constraints.*

Note that at certain outer parameter values, the inner problem may have multiple optima, while the outer problem will be optimum only at specific inner variable values. When this situation arises, the optimum of the multilevel programming problem is achieved only if the inner optimizer cooperates with the outer optimizer.

2.3 Mathematical Formulation of Multilevel Optimization

Let us consider a multilevel programming problem in which one leader and one or multiple (if any) follower(s) are involved. In a multilevel programming problem model, a control for the decision variables is partitioned amongst the decision-makers who seek to minimize their individual payoff objective functions. Perfect information is assumed so that all planers know the objective and feasible choices available for others. The leader goes first and attempts to optimize her/his objective function. In order that, the leader must look forward to all possible responses of her/his challengers. Each follower executes her/his policies after and in view of the decisions of the leader [23].

Let the hierarchical system be comprised of k-levels of decision makers, where the higher level decision maker also called the leader controls decision variables x_1, the second level divisions control decision variables x_2, and continuing in this manner we can reach at k^{th} level divisions control decision variables x_k.

Let the k^{th}-level problem is introduced as follows:

$$\begin{aligned}\min_{x_k}\{f_k(x_1,x_2,\dots,x_k) : g_k(x_1,x_2,\dots,x_k) \le 0\}&\\ \text{where, } f_k : \mathbb{R}^{n_1}\times\mathbb{R}^{n_2}\times\cdots\times\mathbb{R}^{n_k}\to\mathbb{R}&\\ g_k : \mathbb{R}^{n_1}\times\mathbb{R}^{n_2}\times\cdots\times\mathbb{R}^{n_k}\to\mathbb{R}^{q}.&\end{aligned} \tag{2.2}$$

Let $\Omega_{k-1}(x_1,x_2,\dots,x_{k-1})$ stand for the feasible set of problem (2.2) for fixed $(x_1,x_2,\dots,x_{k-1}) \in \mathbb{R}^{n_1}\times\mathbb{R}^{n_2}\times\cdots\times\mathbb{R}^{n_{k-1}}$. Let for the moment $\Psi_{k-1}(x_1,x_2,\dots,x_{k-1})$ denote the set of optimal solutions of problem (2.2) and it's a point-to-set mapping from $(x_1,x_2,\dots,x_{k-1})$ into a power set of $\mathbb{R}^n$ defined by, $\Psi_{k-1} : \mathbb{R}^{n_1}\times\mathbb{R}^{n_2}\cdots\times\mathbb{R}^{n_{k-1}}\to 2^{\mathbb{R}^n}$.

And then the $(k-1)^{th}$-level problem can be formulated as follows:

$$\begin{aligned}\text{“}\min_{x_{k-1}}\text{”}\{f_{k-1}(x_1,x_2\dots x_k) : g_{k-1}(x_1,x_2\dots x_k) \le 0,&\\ x_k\in\Psi_{k-1}(x_1,x_2,\dots,x_{k-1})\}&\\ \text{where, } f_{k-1} : \mathbb{R}^{n_1}\times\mathbb{R}^{n_2}\times\cdots\times\mathbb{R}^{n_{k-1}}\to\mathbb{R}&\\ g_{k-1} : \mathbb{R}^{n_1}\times\mathbb{R}^{k_2}\times\cdots\times\mathbb{R}^{n_{k-1}}\to\mathbb{R}^{s}&\end{aligned} \tag{2.3}$$

Note that the formulation with the quotation marks is used to express the hesitation in the definition of the multilevel optimization problem in case of non-uniquely determined $k^{th}-$level optimal solutions and this results in finitely many problems in this level corresponding to the solutions obtained in the k^{th}-level problem.

Let $\Omega_{k-2}(x_1,x_2,\dots,x_{k-2})$ symbolize the feasible set for problem (2.3) for fixed $\{x_1,x_2,\dots,x_{k-2}\}$. Ψ_{k-2} is also a point-to-set mapping solutions from $\mathbb{R}^{n_1}\times\mathbb{R}^{n_2}\times\cdots\times\mathbb{R}^{n_{k-2}}$ into a power set $2^{\mathbb{R}^n}$.

Continuing in this manner one can arrive at the second level. Then the second level problem can be formulated as follows:

$$\begin{aligned}\text{“}\min_{x_2}\text{”}\{f_2(x_1,x_2,\dots,x_k) : g_2(x_1,x_2,\dots,x_k)\le 0, x_k\in\Psi_{k-1}(x_1,x_2,\dots,x_{k-1}),&\\ x_{k-1}\in\Psi_{k-2}(x_1,x_2,\dots,x_{k-2}),\dots,x_3\in\Psi_2(x_1,x_2)\}&\\ \text{where, } f_2 : \mathbb{R}^{n_1}\times\mathbb{R}^{n_2}\times\cdots\times\mathbb{R}^{n_k}\to\mathbb{R}&\\ g_2 : \mathbb{R}^{n_1}\times\mathbb{R}^{n_i}\times\cdots\times\mathbb{R}^{n_k}\to\mathbb{R}^{q}&\end{aligned} \tag{2.4}$$

Let $\Omega_1(x_1)$ represent the feasible set for problem (2.4) for fixed x_1. Let $\Psi_1(x_1)$ be a point-to-set mapping (solution set) from $\mathbb{R}^{n_1}$ in to a power set of $\mathbb{R}^n$. Then the final condition of k-level programming problem is to select the parameter vector x_1 for the followers problem such that this selection of x_1 is conducted so that a certain inequality constraint $g_1(x_1,x_2,\dots,x_k)\le 0$ are satisfied and an objective function $f_1(x_1,x_2,\dots,x_k)$ is minimized as desired. Therefore, the general multilevel optimization problem can be described as follows:

$$
\begin{aligned}
&\min_{x_1} f_1(x_1, x_2, \ldots, x_k) \\
&\qquad\qquad \text{s.t.} \qquad\qquad (1^{st} \text{ level}) \\
&g_1(x_1, x_2, \ldots, x_k) \leq 0 \\
&\qquad\qquad \text{Where } [x_2, \ldots, x_k] \text{ solve,} \\
&\min_{x_2} f_2(x_1, x_2, \ldots, x_k) \\
&\qquad\qquad \text{s.t.} \qquad\qquad (2^{nd} \text{ level}) \\
&g_2(x_1, x_2, \ldots, x_k) \leq 0 \\
&\qquad\qquad \text{Where } [x_3, \ldots, x_k] \text{ solves,} \\
&\qquad\qquad \ddots \\
&\min_{x_k} f_k(x_1, x_2, \ldots, x_k) \\
&\qquad \text{s.t.} \qquad (k^{th} \text{ level}) \\
&g_k(x_1, x_2, \ldots, x_k) \leq 0
\end{aligned} \tag{2.5}
$$

Here, f_i's are real functions, g_i's are vector-valued functions for each $i = 1, 2, \ldots, k$ defined on convex sets which belongs to the group of real numbers. Note that the objective function at each level p, $f_p(x_1, x_2, \ldots, x_k)$, is defined over the decision space of all levels. Thus, the p^{th}-level decision maker may have his/her objective function determined, in part by variables controlled at another levels. However, by controlling optimization variables, after decisions at level 1 to $p-1$, have been made, level p may influence the decisions made at level $p+1$ and all lower levels to improve his/her own objective function as well.

The main concern leftover to solve the leader's objective function to global optimality. Thus, we try to compute the set $\{x_1, x_2 \ldots, x_k\}$ which minimizes globally the leader's objective:

$$
\begin{aligned}
\min_{x_1} \{ f_1(x_1, x_2, \ldots, x_k) : g_1(x_1, x_2, \ldots, x_k) \leq 0, \\
x_k \in \Psi_{k-1}(x_1, x_2, \ldots, x_{k-1}), \\
x_{k-1} \in \Psi_{k-2}(x_1, x_2, \ldots, x_{k-2}), \ldots, x_3 \in \Psi_2(x_1, x_2), x_2 \in \Psi_1(x_1) \}
\end{aligned}
$$

which is a multilevel optimization problem in terms of the rational reaction sets of the followers problem.

For the sake of simplicity and without loss of generality, we analyze the relations in Problem (2.5) using two particular classes of multilevel programming problems: the bilevel programming problem which organizes vertically in two levels and the trilevel programming problem, which organizes vertically in three levels, with the inclusion of nonconvexity in the second and third level respectively.

2.3.1 Bilevel Programming with Nonconvexity Formulation in the Inner Problem

Bilevel programming problems with nonconvexity in the inner problem involve an optimization hierarchy of two levels where the set of all variables is partition between two vectors x_1 and x_2. Where x_2 is chosen as an optimal solution of the inner mathematical programming problem parameterized in x_1. Thus, the bilevel programming problem is hierarchical in the sense that leader's problem constraints are defined in part by the inner optimization problem.

Let the inner problem be introduced with the inclusion of nonconvexity in the objective function as well as in the constraint functions as follows and consider g_2 to be a vector valued function:

$$\min_{x_2}\{f_2(x_1,x_2) : g_2(x_1,x_2) \leq 0\}$$

Where,

$$\begin{aligned} f_2 &: \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R} \\ g_2 &: \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^q \end{aligned} \tag{2.6}$$

Thus, problem (2.6) is a nonconvex optimization problem with respect to x_2 for a given x_1.

Let $\Omega(x_1)$ denote a compact and connected feasible set of problem (2.6) for fixed $x_1 \in \mathbb{R}^{n_1}$ and $\Psi(x_1)$ be the set of solutions. Then Ψ is a so-called a point-to-set mapping from $\mathbb{R}^{n_1}$ into a power set $2^{\mathbb{R}^n}$.

The final state of the bilevel programming problem is to select a parameter vector x_{1_0} for the lower level problem which is optimal in a certain sense. To be more precise, this selection of x_{1_0} is conducted so that certain (nonlinear) inequality constraints $g_1(x_1,x_2) \leq 0$ are satisfied and an objective function $f_1(x_1,x_2)$ is minimized. That means, x_{1_0} minimizes the following optimization problem:

$$\min_{x_1} f_1(x_1,x_2)$$

S.t

$$g_1(x_1,x_2) \leq 0,\ x_2 \in \Psi(x_1)$$

$$\begin{aligned} \text{Where, } f_1 &: \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R} \\ g_1 &: \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^k \end{aligned} \tag{2.7}$$

The problem of determining a best solution x_1^* for the leader's problem can thus be described as that of finding a vector x_1^* of parameters for the parametric optimization problem (2.6) which together with the response $x_2(x_1) \in \Psi(x_1)$ proves to satisfy the constraint, $g_1(x_1,x_2) \leq 0$ and to give the best possible function value for $f_1(x_1,x_2(x_1))$.

Consequently, bilevel programming problem with the inclusion of nonconvexity in the inner problem can be reformulated as a problem that involve two optimization levels:

$$
\begin{aligned}
&\min_{x_1} f_1(x_1, x_2) \\
&\quad \text{s.t.} \qquad\qquad (1^{st} \text{ level}) \\
&g_1(x_1, x_2) \leq 0 \\
&\min_{x_2} f_2(x_1, x_2) \\
&\quad\quad \text{s.t.} \qquad\qquad (2^{nd} \text{ level}) \\
&\quad\quad g_2(x_1, x_2) \leq 0
\end{aligned} \tag{2.8}
$$

Where $x_1 \in X_1 \subseteq \mathbb{R}^{n_1}, x_2 \in X_2 \subseteq \mathbb{R}^{n_2}$ and both are compact convex sets. One can also define the feasible set and rational reaction set as follows:

Definition 2.13 *1. The set*

$$\Omega(x_1) = \{x_2 \in X_2 : g_2(x_1, x_2) \leq 0\} \tag{2.9}$$

is called a feasible set for the inner problem.

2. The set of solutions

$$\Psi(x_1) = \{x_2 \in X_2 : x_2 \in \arg\min\{f_2(x_1, x_2) : x_2 \in \Omega(x_1)\}\} \tag{2.10}$$

is called the rational reaction set for the inner level.

Likewise, when more than one problems are present at the same hierarchical optimization level the problem is called a decentralized bilevel optimization problem [25]. Such problems are assumed to behave according to the Nash equilibrium [21]. The Nash equilibrium is often a preferred strategy to coordinate such decentralized systems [10], [25]. Consequently, the optimization subproblems posed in the second level are assumed to reach a Nash equilibrium point [10], [17].

One can observe the parametric nature of the rational reaction set, Equ. (2.10) easily. However, each rational reaction set is a function of both the upper level decision variables and the decision variables of the other problems located in the same hierarchical level (if any). But, in this paper we consider only the follower has single decision maker in each level. The main concern remains to solve the leader's objective function to global optimality as well. Thus, we try to compute the set $\{x_1, x_2\}$ which minimizes globally the leader objective:

$$\min_{x_1}\{f_1(x_1, x_2) : g_1(x_1, x_2) \leq 0, x_2 \in \Psi\} \tag{2.11}$$

Note that problem (2.11) is a bilevel optimization problem in terms of the rational reaction set calculated in the 2^{nd}-level problem.

2.3.2 Trilevel Programming with Nonconvexity Formulation in the most Inner Problem

Three level (or trilevel) programming problems with the inclusion of nonconvexity formulation in the third level are mathematical optimization problems where the set of all variables are partitioned among three vectors x_1, x_2 and x_3 ; x_3 to be chosen as an optimal solution of the third level programming problem which is parameterized in x_1 and x_2. After finding the optimal solution of the third level problem x_2 is to be chosen as an optimal solution of the second level programming problem parameterized in x_1 using an optimal solution x_3 of the third level problem. Finally, x_1 is chosen as an optimal solution of the first (leader) level programming problem using the optimal solutions x_2 and x_3 of the second and the third level problems respectively as given values.

For a moment let the most inner level problem be introduced as follows:

$$\min_{x_3}\{f_3(x_1,x_2,x_3) : g_3(x_1,x_2,x_3) \le 0\}$$

Where,

$$\begin{aligned} f_3 &: \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{n_3} \to \mathbb{R} \\ g_3 &: \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{n_3} \to \mathbb{R}^{q} \end{aligned} \tag{2.12}$$

is a nonconvex programming problem with respect to x_3 for a given x_1 and x_2.

Let the connected and compact set $\Omega_2(x_1,x_2)$ stand for the feasible set of problem (2.12) for fixed $(x_1,x_2) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2}$ and for a moment let $\Psi_2(x_1,x_2)$ be a solution set for problem (2.12). Then $\Psi_2(x_1,x_2)$ is a point-to-set mapping from (x_1,x_2) into a power set of $\mathbb{R}^n$ defined by $\Psi_2 : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to 2^{\mathbb{R}^n}$. After we determine $\Psi_2(x_1,x_2)$ the second level optimization problem can be formulated as follows:

$$\text{``}\min_{x_2}\text{''}\{f_2(x_1,x_2,x_3) : g_2(x_1,x_2,x_3) \le 0, x_3 \in \Psi_2(x_1,x_2)\}$$

Where,

$$\begin{aligned} f_2 &: \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{n_3} \to \mathbb{R} \\ g_2 &: \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{n_3} \to \mathbb{R}^{k} \end{aligned} \tag{2.13}$$

Note that the formulation within the quotation marks is used to express the hesitation in the definition of the trilevel programming problem followed by non-uniquely determined lower level optimal solutions and this results in finitely many problems in this level corresponding to the solutions obtained in the 3^{rd}-level problem.

Let $\Omega_1(x_1)$ symbolize the feasible set for problem (2.13) may in general be a disconnected set and $\Psi_1(x_1,x_2)$ stands for the solution set of problem (2.13) for fixed x_1 and x_3. Here again Ψ_1 is also a point-to-set mapping from $\mathbb{R}^{n_1}$ into a power set $2^{\mathbb{R}^n}$. Then the final condition of the three level programming problem is to select the parameter vector x_1 such that this selection of x_1 is conducted so that a certain inequality constraint

$g_1(x_1, x_2, x_3) \leq 0$ is satisfied and an objective function $f_1(x_1, x_2, x_3)$ is minimized as well. i.e. find x_1 such that

$$\min_{x_1}\{f_1(x_1, x_2, x_3), x_2 \in \Psi_1(x_1), x_3 \in \Psi(x_1, x_2), g_1(x_1, x_2, x_3) \leq 0\}$$
$$\text{Where, } f_1 : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{n_3} \to \mathbb{R}$$
$$g_1 : \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \mathbb{R}^{n_3} \to \mathbb{R}^{s} \tag{2.14}$$

The problem of determining a best solution x_1^* for problem (2.14) can thus be described as that of finding best vectors of parameters x_1^*, x_2^* and x_1^* for the parametric optimization problems (2.12) and (2.13) respectively. This problem is the ***three level programming problem***.

As a result, we get the following nested optimization problem, known as the trilevel programming problem which represents the above problem:

$$\begin{aligned}
&\min_{x_1} f_1(x_1, x_2, x_3) \\
&\qquad \text{s.t.} \qquad (1^{st} \text{ level}) \\
&g_1(x_1, x_2, x_3) \leq 0, \\
&\qquad \text{where, } x_2, x_3 \text{ solves} \\
&\min_{x_2} f_2(x_1, x_2, x_3) \\
&\qquad \text{s.t.} \qquad (2^{nd} \text{ level}) \\
&g_2(x_1, x_2, x_3) \leq 0, \\
&\qquad \text{where } x_3 \text{ solves} \\
&\min_{x_3} f_3(x_1, x_2, x_3) \\
&\qquad \text{s.t.} \qquad (3^{rd} \text{level}) \\
&g_3(x_1, x_2, x_3) \leq 0
\end{aligned} \tag{2.15}$$

Where $x_1 \in X_1 \subseteq \mathbb{R}^{n_1}, x_2 \in X_2 \subseteq \mathbb{R}^{n_2}$ and $x_3 \in X_3 \subseteq \mathbb{R}^{n_3}$. Problem (2.15) comprises three subproblems, one at each optimization level with the following basic definitions of sets:

Definition 2.14 *1. The set*

$$\Omega_2(x_1, x_2) = \{x_3 \in X_3 : g_3(x_1, x_2, x_3) \leq 0\} \tag{2.16}$$

is called a feasible set for the third level.

2. The set of solutions defined as,

$$\Psi_2(x_1, x_2) = \{x_3 \in X_3 : x_3 \in \arg\min\{f_3(x_1, x_2, x_3)\} : x_3 \in \Omega_2\} \tag{2.17}$$

is called the rational reaction set for the third level.

3. *The set*

$$\Omega_1(x_1) = \{(x_2, x_3) \in X_2 \times X_3 : g_2(x_1, x_2, x_3) \leq 0,$$
$$g_3(x_1, x_2, x_3) \leq 0, x_3 \in \Psi_2(x_1, x_2)\} \tag{2.18}$$

is called a feasible set for the second level problem.

4. *The set of solutions*

$$\Psi_1(x_1) = \{(x_2, x_3) \in X_2 \times X_3 : x_2 \in \arg\min\{f_2(x_1, x_2, x_3) :$$
$$x_2 \in \Psi_1(x_1), x_3 \in \Psi_2(x_1, x_2)\}\} \tag{2.19}$$

is called the rational reaction set for the second level.

One can easily see the parametric nature of the rational reaction sets, Equ. (2.17) and (2.19), which returns the dependence of the decisions taken at the upper levels on the decisions taken at lower levels. This fact shows that in multilevel programming problems the relations between the levels differ from non-cooperative game theory, where each player must choose a strategy simultaneously [16].

2.4 Computational Difficulties of Multilevel Optimization with Nonconvexity Formulation in the Inner Levels

Consider a bilevel programming problem given in (2.8). The traditional solution approach for problem (2.8) is to represent the inner level problem and its constraints with its sufficient conditions augmented as constraints of the upper level problem. This procedure works well if the inner level problem is strictly convex. However, when only convexity (without strict convexity) assumption at inner level is considered, the optimality region may contain non-singleton set.

In the nonconvex case the KKT conditions are not sufficient and this approach will, hence, result in a much larger feasible set [14]. As a result, if the inner level problem is nonconvex, it is clear that a bilevel programming problem cannot be equivalent to the corresponding mathematical program with complementarity constraints. That is why many researchers considered strict convexity at the lower level in proposing the solution strategy for such problem.

In general, without convexity assumption at inner level, stationary points may not be sufficient or optimal for the inner level problem and the set of all stationary points may not be connected. This implies that it is impossible to use the approach of augmenting the conditions of the lower level problem into the constraints of the upper level problem.

Moreover, consider an approach for investigating bilevel programming problems in which direct substitution of the rational reaction set $\Psi(x_1)$ of the inner level problem into the upper level problem. This produces a mathematical program without equilibrium constraints depending on the upper level problem. But, if the lower level problem is nonconvex, the rational reaction set may be a disconnected set. It turns out that the upper level optimization problem may not have a solution in the rational reaction set, even if the upper level is a linear programming problem. Moreover, the global optimal solution cannot be efficiently computed and the behavior of a local solution is hard to analyze for the inner level optimization problem. In addition, the mapping $\Psi(\cdot)$ of global optimal solution can contain more than one optimal solution locally within the corresponding critical region (i.e. the region where a particular functional relationship between $x_2(x_1)$ and x_1 holds) and may not be lower semicontinuous. As a result, the function $x_2(x_1) \in \arg\min_{x_2}\{f_2(x_1, x_2) : x_2 \in \Psi(x_1)\}$ is in general not uniquely determined locally within the corresponding critical region and it can be discontinuous. Furthermore, the rational reaction set may not be closed. Hence a bilevel programming problem may not have an optimal solution even if the feasible set of the inner problem is compact.

These all discussed above makes attempts to formulate optimality conditions or solution algorithms very difficult and there is no general method of solving it yet.

In general, the computational complexity of the solution algorithms grows geometrically with the number of constraints and/or the number of levels in the hierarchy.

Eventhough, it is tricky and most challenging in practice, the conventional way of tackling this kind of problem theoretically is via convex relaxation and applying the Branch-and-Bound algorithm as will be discussed in Section 5.1 through a multiparametric programming approach.

3 Multiparametric Programming: Theory and Algorithm

Sensitivity analysis is a technique used to describe how the solution to a mathematical program changes with small changes in the problem parameters. Parametric programming is a closely related, but more advanced technique in which the solution is explicitly found for a full range of parameter values. Mathematical programs which contain a vector of parameters rather than a single one, are commonly referred to as multiparametric programs.

A historical survey of the main developments within parametric programming can be found in [7], [36], [38]. The first works on parametric programming were done in the early fifties, with researchers stating questions like "what happens, when some of the initial data change?. The early interest in this regard was within the field of linear programming, where a single parameter on the right hand side of the constraints was considered.

The purpose of this section is not to give a complete overview of the subject, but to introduce some central concepts and ideas used throughout this thesis.

3.1 Semicontinuity of the Solution Set Mapping

By the structure of multilevel optimization problems as stated in Section 2.2, the investigation of the properties of parametric optimization problems is crucial for our considerations. Since the optimal value function of parametric optimization problems is often not continuous and since the continuity properties of point-to-set mappings are of a different nature than those for functions, relaxed continuity properties of functions and point-to-set mappings will now be introduced:

Consider the following multiparametric programming problem:

$$\begin{aligned} Z(\theta) = \min_{x} f(x,\theta) & \\ \text{s.t.} & \\ g_i(x,\theta) \leq 0,\ \forall i = 1,2,\ldots,p, & \\ h_j(x,\theta) = 0,\ \forall j = 1,2,\ldots,q, & \\ x \in X \subseteq \mathbb{R}^n, \theta \in \Theta \subseteq \mathbb{R}^m, & \end{aligned} \tag{3.1}$$

Here $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, $g_i : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, and $h_j : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$, are sufficiently smooth functions.

Definition 3.1 *A function $f : \mathbb{R}^n \to \mathbb{R}$ is lower semicontinuous at a point $x \in \mathbb{R}^n$ if for each sequence $\{x^t\}_{t=1}^{\infty} \subset \mathbb{R}^n$ with $\lim_{t\to\infty} x^t = x$ we have $\liminf_{t\to\infty} f(x^t) \geq f(x)$ and upper semicontinuous at $x \in \mathbb{R}^n$ provided that $\limsup_{t\to\infty} f(x^t) \leq f(x)$ for each sequence*

$\{x^t\}_{t=1}^{\infty}$ *that converges to* x. *The function* f *is continuous at* $x \in \mathbb{R}^n$ *if it is lower as well as upper semicontinuous at the point* x.

Definition 3.2 *A point-to-set mapping* $\Psi : \mathbb{R}^m \to 2^{\mathbb{R}^n}$ *is called upper semicontinuous at a point* $\theta \in \mathbb{R}^m$ *if for each open set* O *with* $\Psi(\theta) \subset O$, *there exists an open neighborhood* $U_\delta(\theta)$ *of* θ *with* $\Psi(\theta_1) \subset O$ *for each* $\theta_1 \in U_\delta(\theta)$. Ψ *is lower semicontinuous at* $\theta \in \mathbb{R}^m$ *if for each open set* O *with* $\Psi(\theta) \bigcap O \neq \emptyset$, *there is an open neighborhood* $U_\delta(\theta)$ *such that* $\Psi(\theta_1) \bigcap O \neq \emptyset$ *for all* $\theta_1 \in U_\delta(\theta)$.

Let $\theta_0 \in \mathbb{R}^m$, $x_0 \in \mathbb{R}^n$. We say (MFCQ) is satisfied at (x_0, θ_0) if,

$$\begin{aligned} \nabla_x g_i(x_0, \theta_0) &< 0, \text{for each } i \in I(x_0, \theta_0) = \{k : g_k(x_0, \theta_0) = 0\} \\ \nabla_x h_j(x_0, \theta_0) &= 0, \text{for each } j = 1, 2, \ldots, q \end{aligned} \tag{3.2}$$

and the gradients $\{\nabla_x h_j(x_0, \theta_0) : j = 1, 2, \ldots, q\}$ are linearly independent and the set $SP(\theta) = \{x \in \Omega(\theta) : \Lambda(x, \theta) \neq \emptyset\}$ is called the set of stationary solutions of problem (3.1), where $\Lambda(x, \theta)$ is the set of Lagrange multipliers.

Assumption **(C)**: The set $\{(x, \theta) : g(x, \theta) \leq 0,\ i = 1, 2, \ldots, p,$ $h_j(x, \theta) = 0,\ j = 1, 2, \ldots, q\}$ is nonempty and compact. This assumption guaranties that the feasible set of problem (3.1) is nonempty and compact for each $\theta \in \{z : \Omega(z) \neq \emptyset\}$ as well. This implies that $\Psi(\theta)$ (The set of solutions of problem (3.1)) as well as $\Psi_{local}(\theta)$ (The set of local solutions of problem (3.1)) are nonempty and compact sets for each $\theta \in \{z : \Omega(z) \neq \emptyset\}$.

Theorem 3.3 *[13] Consider problem (3.1) at* $\theta_0 \in \mathbb{R}^m$ *and let the* ***Assumption (C)*** *and (MFCQ) be satisfied for each* $x_0 \in SP(\theta_0)$. *Then, the point-to-set mapping* $SP(\cdot)$ *is upper semicontinuous at* θ_0.

Note that the local solution set mapping $\Psi_{local}(\cdot)$ is in general not upper semicontinuous for nonconvex parametric optimization problems. However, the global solution mapping could be upper semicontinous.

Theorem 3.4 *[13] If for the problem (3.1) the* ***assumption (C)*** *as well as (MFCQ) are satisfied then the global solution set mapping* $\Psi(\cdot)$ *is upper semicontinues.*

We now turn to our attention to the lower semicontinuity of Ω and Ψ. Consider the mapping, $\Psi_\varepsilon(\theta) = \{x \in \Omega(\theta) | f(x, \theta) < Z(\theta) + \varepsilon\}$.

Theorem 3.5 *[7]* Z *is lower semicontinuous at* θ_0 *if* Ω *is upper semicontinuous at* θ_0, $\Omega(\theta_0)$ *is compact and* f *is lower semicontinuous on* $\Omega(\theta_0) \times \theta_0$.

Theorem 3.6 *[7] Let* f *be defined on* $\mathbb{R}^n \times \mathbb{R}^m$. *Then*

1. *Ψ_ε is lower semicontinuous at θ_0 for each $\varepsilon > 0$ if Ω is lower semicontinous and Z is lower semicontinuous at θ_0*

2. *Z is continuous at θ_0 if Ψ is lower semicontinuous at θ_0 and $\Psi(\theta_0)$ is nonempty.*

3. *Z is continuous at θ_0 if there exists an $\varepsilon_0 > 0$ such that Ψ_ε is lower semicontinuous at θ_0 for all $\varepsilon \in (0, \varepsilon_0)$.*

Corollary 3.7 *If the mapping Ω is lower semicontinous at θ_0 then the following statements are equivalent:*

1. *Z is continuous at θ_0*

2. *Ψ_ε is lower semicontinuous at θ_0 for each $\varepsilon > 0$.*

3.2 Existence Theorem

A known existence theorem for a given parameter is Weierstrass Theorem [27] which says that every continuous function attains its minimum on a compact set. This statement is modified for general parametric optimization problem (3.1) for a given parameter.

Definition 3.8 *Let $(X, \|\cdot\|)$ be a normed space, S be nonempty subsets of X and $f : S \to \mathbb{R}$ be a given function. The function f is said to be weakly lower semicontinuous at (x_0, θ_0) if for every sequence $\{(x_n, \theta_n)\}_{n=1}^{\infty} \subseteq \mathbb{R}^n \times \mathbb{R}^m$ converging weakly to $(\bar{x}, \bar{\theta}) \in \mathbb{R}^n \times \mathbb{R}^m$, we have,* $\liminf_{n\to\infty} f(x_n, \theta_n) \geq f(\bar{x}, \bar{\theta})$.

Theorem 3.9 *Let $(X, \|\cdot\|)$ be a normed space, S be nonempty subsets of X and $f : S \to \mathbb{R}$ be a given function. If the set S is weakly sequentially compact and the function f is weakly lower semicontinuous, then there is at least one $(x_0, \theta_0) \in \mathbb{R}^n \times \mathbb{R}^m$ with $f(x_0, \theta_0) \leq f(x, \theta)$ for all $(x, \theta) \in \mathbb{R}^n \times \mathbb{R}^m$.*

Proof. Let $\{(x_n, \theta_n)\}_{n=1}^{\infty}$ be an infimal sequence in $\mathbb{R}^n \times \mathbb{R}^m$, i.e. a sequence with the property $\lim_{n\to\infty} f(x_n, \theta_n) = \inf f(x, \theta)$. Since the set $\mathbb{R}^n \times \mathbb{R}^m$ is weakly sequentially compact, there is a subsequence $\{(x_{n_i}, \theta_{n_i})\}_{i=1}^{\infty}$ converging weakly to some $(x_0, \theta_0) \in \mathbb{R}^n \times \mathbb{R}^m$. Because of the weak lower semicontinuity of f at $(x_0, \theta_0) \in \mathbb{R}^n \times \mathbb{R}^m$ it follows that

$$f(x_0, \theta_0) \leq \liminf_{i\to\infty} f(x_{n_i}, \theta_{n_i}) = \inf_{(x,\theta)\in\mathbb{R}^n\times\mathbb{R}^m} f(x, \theta) \Box$$

3.3 Optimality Conditions

In a general programming, a point is called a local optimal solution if there is no better feasible point with respect to the objective function in a certain sufficiently small neighborhood of this point [13]. Qualitative and quantitative properties of the optimal solution and of the optimal objective function value of problem (3.1) with respect to parameter changes are of the main interest in this section.

Problem (3.1) is a convex parametric optimization problem if all functions $f(\cdot,\theta)$, $g_i(\cdot,\theta)$, $i = 1, 2, \dots, p$ are convex functions and $h_j(\cdot,\theta), j = 1, 2, \dots, q$ are affine-linear functions on $\mathbb{R}^n$ for each fixed θ in $\mathbb{R}^m$.

Definition 3.10 *The region where a particular functional relationship between $x(\theta)$and θ holds is known as a critical region and hereafter we shall denote it by CR.*

Let $Z : \mathbb{R}^m \to \mathbb{R}$ be the so-called optimal value function defined by the optimal objective function values of problem (3.1) as:

$$Z(\theta) = \{\min_x f(x,\theta)|g_i(x,\theta) \leq 0, \text{ for } i = 1, 2, \dots, p, \ h_j(x,\theta) = 0, \text{ for } j = 1, 2, \dots, q\}$$

and define the solution set mapping $\Psi : \mathbb{R}^m \to 2^{\mathbb{R}^n}$ by,

$$\Psi(\theta) = \arg\min_x\{f(x,\theta) : g_i(x,\theta) \leq 0, h_j(x,\theta) = 0\} \tag{3.3}$$

Here, $2^{\mathbb{R}^n}$ is the power set of $\mathbb{R}^n$ (i.e. the family of subsets of $\mathbb{R}^n$). By definition Ψ is a point-to-set mapping which maps $\theta \in \mathbb{R}^m$ to the set of global optimal solution of problem (3.1). For convex parametric optimization problems, $\Psi(\theta)$ is closed and convex but possibly empty.

Let us denote another point-to-set mapping, $\Omega(\theta) : \mathbb{R}^m \to 2^{\mathbb{R}^n}$ with $\Omega = \{x \in \mathbb{R}^n : g_i(x,\theta) \leq 0, h_j(x,\theta) = 0\}$, the feasible set mapping. The set of local minimizers of problem (3.1) is denoted by $\Psi_{loc} = \{x \in \Omega(\theta) : \exists\varepsilon > 0 : f(x,\theta) \leq f(z,\theta), \forall z \in \Omega(\theta) \bigcap V_\varepsilon(x)$, where $V_\varepsilon(x) = \{z \in \mathbb{R}^n : \|x - z\| < \varepsilon\}$ is any neighborhood of x.

Let us denote the linear combination of the objective function and the constraint functions by,

$$L(x,\theta,\lambda,\mu) = f(x,\theta) + \sum_{i=1}^{p} \lambda_i g_i(x,\theta) + \sum_{j=1}^{q} h_j(x,\theta),$$

the so-called Lagrangian function of problem (3.1) and consider the set $\Lambda = \{(\lambda,\mu) \in \mathbb{R}^n \times \mathbb{R}^m : \lambda_i > 0, \lambda_i g_i(x,\theta) = 0, \nabla_x L(x,\theta,\lambda,\mu) = 0\}$ of Lagrange multipliers corresponding to the point $(x,\theta) \in \mathbb{R}^n \times \mathbb{R}^m$. Then we have the following result,

Theorem 3.11 *[13] Consider problem (3.1) at $(x_0,\theta_0) \in \mathbb{R}^n \times \mathbb{R}^m$ with $x_0 \in \Psi(\theta_0)$. Then (MFCQ) is satisfied at (x_0,θ_0) if and only if $\Lambda(x_0,\theta_0)$ is nonempty, convex and polyhedron.*

Let us recall that the set $SP(\theta) = \{x \in \Omega(\theta)) : \Lambda(x,\theta) \neq \emptyset\}$ is called the set of stationary solutions of problem (3.1). It is a direct consequence of the definition that $x \in SP(\theta)$

if and only if there are vectors $\lambda \in \mathbb{R}^p$, $\mu \in \mathbb{R}^q$ such that the triple (x, λ, μ) solves the system of KKT conditions:

$$\begin{aligned} \nabla_x L(x, \theta, \lambda, \mu) \\ g_i(x, \theta) < 0 \\ h_j(x, \theta) = 0, \\ \lambda_i g_i(x, \theta) = 0 \\ \text{for each } i = 1, 2, \ldots, p \ \text{ and } j = 1, 2, \ldots, q \end{aligned}$$

3.4 Multiparametric Linear and Quadratic Programming

In this review we present an algorithm for the solution of multiparametric linear and quadratic programming problems, with linear constraints and linear or convex quadratic objective functions, the optimal solution of these optimization problems is given by a conditional piecewise linear function of the varying parameters. This function results from first-order estimations of the analytical nonlinear optimal function. The core idea of the algorithm is to approximate the analytical nonlinear function by affine functions, whose validity is confined to regions of feasibility and optimality. Therefore, the space of parameters is systematically characterized into different regions where the optimal solution is an affine function of the parameters.

3.4.1 Methodology

Consider the general parametric quadratic-linear programming problem:

$$\begin{aligned} Z(\theta) = \min_x c^T x + \frac{1}{2} x^T Q x \\ \text{s.t} \\ Ax \leq b + F\theta \\ x \in X \subseteq \mathbb{R}^n, \theta \in \Theta \subseteq \mathbb{R}^m \end{aligned} \tag{3.4}$$

where c is a constant vector of dimension n, Q is an $(n \times n)$ symmetric positive definite constant matrix, A is a $(p \times n)$ constant matrix, F is a $(p \times m)$ constant matrix, b is a constant vector of dimension p, and X and Θ are compact polyhedral convex sets of dimensions n and m, respectively. The first-order KKT optimality conditions for (3.4) are given as follows:

$$L = c^T x + \frac{1}{2} x^T Q x + \sum_{i=1}^{p} \lambda_i (A_i x - b_i - F_i \theta)$$

$$\nabla_x L = 0$$
$$\lambda_i A_i x - b_i - F_i\theta = 0$$
$$\lambda_i \geq 0,\ i = 1, 2, \dots p \tag{3.5}$$

An application of Basic Sensitivity Theorem [22] to (3.4) at $[x_0, \theta_0]$ gives the following result:

$$\begin{pmatrix} \frac{dx(\theta_0)}{d\theta} \\ \frac{d\lambda(\theta_0)}{d\theta} \end{pmatrix} = -M_0^{-1} \cdot N_0 \tag{3.6}$$

where,

$$M_0 = \begin{bmatrix} Q & A_1^T \dots A_p^T & \\ -\lambda_1 A_1 & -V_1 & \\ \vdots & & \ddots \\ -\lambda_p A_p & & -V_p \end{bmatrix}$$

$N_0 = (U, -\lambda_1 F_1, \dots, -\lambda_p F_p)^T$,

$V_i = A_i x_0 - b_i - F_i\theta_0$ and U is a null matrix of dimension $(n \times m)$. Thus, in the linear-quadratic optimization problem, the Jacobians reduce to a mere algebraic manipulation of the matrices declared in problem (3.4). In the neighborhood of the KKT point, $[x_0, \theta_0]$, and under the consideration of Basic Sensitivity Theorem [22], a first order approximation of $[x(\theta), \lambda(\theta)]$ in the neighborhood of θ_0 is rewritten as follows:

$$\begin{bmatrix} x(\theta) \\ \lambda(\theta) \end{bmatrix} = \begin{bmatrix} x_0 \\ \lambda_0 \end{bmatrix} - M_0^{-1} \cdot N_0 \cdot (\theta - \theta_0) \tag{3.7}$$

Note that when the assumptions in Basic Sensitivity Theorem [22] are satisfied, M_0 is always invertible. This is where parametric programming detaches from the sensitivity analysis theory. Whilst sensitivity analysis stops here, where we know what happens if the process conditions deviate from the nominal values to some value in its neighborhood, parametric programming is concerned with the whole range of the parametric variability.

The space of θ where this solution (3.7) remains optimal is defined as the critical region, CR and can be obtained by using feasibility and optimality conditions. Feasibility is ensured by substituting $x(\theta)$ into the inactive inequalities given in (3.4), whereas the optimality condition is given by $\tilde{\lambda}(\theta) \geq 0$, where $\tilde{\lambda}(\theta)$ corresponds to the vector of active inequalities, resulting in a set of parametric constraints. Let this set be represented by

$$CR = \{\tilde{A}x(\theta) \leq \tilde{b} + \tilde{F}\theta, \tilde{\lambda}(\theta) \geq 0, CR^{IG}\} \tag{3.8}$$

where $\tilde{A}$, $\tilde{b}$, and $\tilde{F}$ correspond to the inactive inequalities and CR^{IG} represents a set of linear inequalities defining an initial given region. From the parametric inequalities

Table 1: mp-QP algorithm

Step 1	In a given region solve (3.4) by treating θ as a free variable to obtain a feasible point θ_0
Step 2	Fix $\theta = \theta_0$ and solve (3.4) to obtain $[x(\theta), \lambda(\theta)]$
Step 3	Compute $[(M_0)^1 N_0]$ from (3.6)
Step 4	Obtain $[x(\theta), \lambda(\theta)]$ from (3.7)
Step 5	Form a set of inequalities, CR, as described in Equ. (3.8)
Step 6	Remove redundant inequalities from this set of inequalities and define the corresponding CR^Q as given in Equ. (3.9)
Step 7	Define the rest of the region, CR^{Rest} as given in Equ. (3.10)
Step 8	If no more regions to explore, go to next step, otherwise go to Step 1
Step 9	Collect all the solutions and unify the regions having the same solution to obtain a compact representation

thus obtained, the redundant inequalities are removed and a compact representation of CR is obtained as follows:

$$CR^Q = \Delta\{CR\} \tag{3.9}$$

where Δ is an operator which removes redundant constraints for a procedure to identify redundant constraints (see Appendix A of [20] for more detail). Note that CR^Q is a polyhedral region. Once CR^Q has been defined for a solution, $[x(\theta), \lambda(\theta)]$, the next step is to define the rest of the region, CR^{Rest}, (can see Appendix B of [20] for more detail).

$$CR^{Rest} = \{CR^{IG} - CR^Q\} \tag{3.10}$$

Another set of parametric solutions in each of these regions is then obtained and corresponding CRs are then calculated. The algorithm terminates when there are no more regions to be explored. Note that while defining the rest of the regions, some of the regions are split and hence the same optimal solution may be obtained in more than one regions. Therefore, the regions with the same optimal solution are united and a compact representation of the final solution is obtained. The main steps of the algorithm are outlined in Table 1.

3.5 Multiparametric Mixed Integer Linear Programming

A mixed integer LP (MILP) is a LP in which some of the optimization variables are restricted attaining integer values only, while a multiparametric mixed integer LP (mp-MILP) is a parameter-dependent MILP which should be solved for a range of parameter values. Several solution strategies exist for parametric MILP with a multiple parameters, see [15], [38]. Where the authors develop a branch and bound based method to solve the problem. The approach is based upon solving one multiparametric linear programming (mpLP) at each node of the BB (Branch and Bound) tree, and as in standard BB methods, complete enumeration of the integer variables is avoided by maintaining upper bounds on the value function.

In this review, we will focus on the case when the mathematical formulation is described by a mixed integer linear objective program (MILP) and the varying parameters are present on the right hand side (RHS) of the constraints. The presence of integer variables in parametric linear programming problems poses an extra challenge for parametric programming due to the discontinuous nature of the solution space.

3.5.1 An Algorithm for the Solution of Multiparametric MILP Problems

Consider an mp-MILP problem of the following form:

$$\begin{aligned} z(\theta) = \min_{x,y} c^T x + d^T y \\ \text{S.t} \\ Ax + Ey \leq b + F\theta \\ \theta^L \leq \theta \leq \theta^U, \\ x \in R^n, y \in \{0,1\} \end{aligned} \tag{3.11}$$

where x is a vector of continuous variables, y is a vector of binary variables, θ is a vector of parameters, A is an $(m \times n)$ constant matrix, E is an $(m \times l)$ constant matrix, F is an $(m \times s)$ constant matrix, b, c and d are constant vectors of dimension m, n and l, respectively.

The algorithm described in this section is based upon decomposing problem (3.11) into a mpLP and a MILP subproblems [15]. The solution of the mpLP, which is obtained by fixing the vector of binary variables, provides a parametric upper bound, whereas the solution of the MILP, which is obtained by treating θ as a free variable, provides a new integer vector. The parametric solutions corresponding to two different integer solutions are then compared, using a procedure proposed in appendix A of [15], in order to keep as tight upper bounds as possible. The steps of the algorithm are described in detail in the literature [15].

Table 2: mp-MILP programming algorithm

Step 1	Define an initial region of θ, CR, with best upper bound $\hat{Z}^*(\theta) = \infty$, and an initial integer solution $\bar{y}$.
Step 2	Form a multiparametric LP problem, for each region with a new integer solution, $\bar{y}$:
2.1	Solve multiparametric LP subproblem to obtain a set of parametric upper bounds $\hat{Z}(\theta)$ and corresponding critical regions CR as well.
2.2	If $\hat{Z}(\theta) \leq \hat{Z}^*(\theta)$ for some region of θ, update the best upper bound function, $\hat{Z}^*(\theta)$ and the corresponding integer solutions, y^*.
2.3	If an infeasibility is found in some region CR, go to step 2.
Step 3	For each region CR, formulate and solve the MILP master problem by (i) treating θ as a variable bounded in the region CR, (ii) introducing an integer cut, and (iii) introducing a parametric cut, $c^T x + d^T y \leq \mathbf{Z}^*(\theta)$. Return to Step 1 with new integer solutions and corresponding CRs.
Step 4	The algorithm terminates in a region where the solution of the MILP subproblem is infeasible. The final solution is given by the current upper bounds $\hat{Z}^*(\theta)$ in the corresponding CRs.

3.6 Multiparametric Nonlinear Programming

The multiparametric nonlinear programming problem, in its most general form, can be expressed as:

$$
\begin{aligned}
Z(\theta) = \min_x f(x, \theta) \\
\text{s.t.} \\
g_i(x,\theta) \leq 0, \text{ for all } i = 1, 2, \ldots, p, \\
h_j(x,\theta) = 0, \text{ for all } j = 1, 2, \ldots, q \\
x \in X \subseteq \mathbb{R}^n, \theta \in \Theta \subseteq \mathbb{R}^m,
\end{aligned} \tag{3.12}
$$

where f, g_i's and h_j's are twice continuously differentiable in x and θ. Assume also that f is a convex function and g_i's, h_j's are define a convex set. Note that f is the performance criterion to be minimized, $h_j = 0$ and $g_i(x, \theta) \leq 0$ are the constraints, x is the vector of optimization variables and θ is the vector of parameters. Driving the optimizer $x^*(\theta)$ as a function of the parameter θ is referred to as multiparametric programming. The objective is to obtain the optimal solution $x(\theta)$ which when substituted into $f(x, \theta)$ provides the optimal objective function value, $Z(\theta)$ as a function of θ.

Therefore, the first-order KKT optimality conditions for (3.12) are given as follows:

$$L = f(x,\theta) + \sum_{i=1}^{p} \lambda_i g_i(x,\theta) + \sum_{j=1}^{q} \mu_j\ h_j(x,\theta),$$
$$\nabla_x L = 0,$$
$$g_i(x,\theta) \leq 0,\ \lambda_i g_i(x,\theta) = 0,\ \lambda_i \geq 0, \text{ for all, } i = 1,2,\ldots,p$$
$$h_j(x,\theta) = 0, \text{ for all, } j = 1,2,\ldots,q \tag{3.13}$$

where, λ_i's and μ's are the vector of Lagrange multipliers.

Definition 3.12 : *Let x be a feasible solution to problem (3.12) for a given θ. We define active constraints as the set of constraints with $g_i(x,\theta) = 0$ and inactive constraints with $g_i(x,\theta) < 0$. The active set $A(x,\theta)$ is the set of constraints indices of the active constraints, that is,*

$$A(x,\theta) = \{i \in \{1,2,\ldots,p\} | g_i(x,\theta) = 0\}$$

Definition 3.13 *For an active set A, we say that the linear independence constraint qualification holds if the set of active constraint gradients are linearly independent.*

Definition 3.14 *Strict complementary slackness is said to hold at KKT point (x_0, λ_0, μ_0) if and only if for $i = 1,2,\ldots,p$ $\lambda_i > 0$ if $g_i(x_0,\theta_0) = 0$ and $\lambda_i = 0$ if $g_i(x_0,\theta_0) < 0$.*

The main sensitivity result for (3.12) drives directly from system (3.13) is given in the next theorem.

Theorem 3.15 *[22] Let θ_0 be a vector of parameter values and (x_0, λ_0, μ_0) be a KKT triple corresponding to (3.13), where, λ_0 is nonnegative and x_0 is feasible in (3.12). Also assume that,*

1. *strict complementary slackness (SCS) holds,*
2. *the binding constraint gradients are linearly independent (LICQ: Linear Independence Constraint Qualification), and*
3. *the second-order sufficiency conditions (SOSC) hold.*

Then, in the neighborhood of θ_0, there exists a unique, once continuously differentiable function, $Z(\theta) = [x(\theta), \lambda(\theta), \mu(\theta)]$, satisfying (3.13) with $Z(\theta_0) = [x(\theta_0), \lambda(\theta_0), \mu(\theta_0)]$, where $x(\theta)$ is a unique isolated minimizer for (3.12), and

$$\begin{pmatrix} \frac{dx(\theta_0)}{d\theta} \\ \frac{d\lambda(\theta_0)}{d\theta} \\ \frac{d\mu(\theta_0)}{d\theta} \end{pmatrix} = -M_0^{-1} \cdot N_0 \tag{3.14}$$

where, M_0 and N_0 are the Jacobian of system (3.13) with respect to x and θ:

$$M_0 = \begin{bmatrix} \nabla_{xx}L \; \nabla_x g_1 \ldots \nabla_x g_p \\ \nabla_x^T g_1 \quad -g_1 \quad\quad 0 \\ \vdots \quad\quad \ddots \\ \nabla_x^T g_p \quad -g_p \quad 0 \\ \nabla_x^T h_1 \\ \vdots \\ \nabla_x^T h_q \quad\quad\quad 0 \end{bmatrix}$$

$$N_0 = \left(\nabla_{\theta x}^{2T}L, -\lambda_1\nabla_\theta g_1, \ldots, -\lambda_1\nabla_\theta g_p, -\lambda_1\nabla_\theta h_1, \ldots, -\lambda_1\nabla_\theta h_q\right)^T$$

Proof:(see [22])

Note that the assumptions in the Theorem (3.15) ensure that the inverse of the Jacobian of Equ. (3.14) exists [15], [18], [38]. In other word, when M_0 is not invertible any violation of assumptions in Theorem 3.15 is easly detected.

In [18] Dua *et al.*, have proposed an algorithm to solve Equ. (3.14) in the entire range of the varying parameters for general convex problems. This algorithm is based on approximations of the nonlinear optimal expression, $x = \gamma^*(\theta)$ by a set of first order approximations.

Corollary 3.16 *First order estimation of $x(\theta)$, $\lambda(\theta)$, $\mu(\theta)$, near $\theta = \theta_0$ [21]: under the consideration of Theorem 3.15, a first order approximation of $[x(\theta), \lambda(\theta), \mu(\theta)]$ in the neighborhood of θ_0 is,*

$$\begin{bmatrix} x(\theta) \\ \lambda(\theta) \\ \mu(\theta) \end{bmatrix} = \begin{bmatrix} x_0 \\ \lambda_0 \\ \mu_0 \end{bmatrix} - M_0^{-1} \cdot N_0 \cdot (\theta - \theta_0) \tag{3.15}$$

where (x_0, λ_0, μ_0)=$(x(\theta_0), \lambda(\theta_0), \mu(\theta_0))$, $M_0 = M(\theta_0)$, $N_0 = N(\theta_0)$

The space of θ where this solution (3.15) remains optimal is defined as the *critical region*, CR, and can be obtained by using feasibility and optimality conditions. Feasibility is ensured by substituting $x(\theta)$ into the inactive inequalities given in problem (3.12), whereas the optimality condition is given by $\tilde{\lambda}(\theta) \geq 0$, where $\tilde{\lambda}(\theta)$ corresponds to the vector of active inequalities, resulting in a set of parametric constraints. Each piecewise linear approximation is confined to regions defined by feasibility and optimality conditions [18]. If $\breve{g}$ corresponds to the non-active constraints, and $\tilde{\lambda}$ to the lagrangian multipliers of the active constraints:

$$\begin{cases} \breve{g}(x(\theta), \theta) \leq 0 & \text{Feasibility conditions} \\ \tilde{\lambda}(\theta) \geq 0 & \text{Optimality conditions.} \end{cases}$$

Table 3: MP-NLP algorithm

Step 1	Solve problem (3.12) for a given θ and obtain x^*.
Step 2	Linearize (3.12) i.e. converting the original problem into an MP-LP problem.
Step 3	Solve the MP-LP problem and set the upper bound with in the newly identified region. Compare the value of the obtained linear profile with the real value of the parametric profile of in every corner of the critical region.
Step 4	If $\hat{Z} - \dot{Z} \leq \varepsilon$ in all the corners of the critical region under consideration, exclude this region from further consideration, update CR^{Rest} ensuring a convex formulation, and move to a different critical region with in CR^{Rest}.
Step 5	Use problem (3.12) restricting the domain of the parametric space to include only the new critical region under consideration, to find a suitable point and proceed from Step (2) if all the parameter space is explored, $CR^{Rest} = \emptyset$ terminate the algorithm, otherwise select the corner for which $\hat{Z}(\theta) - \dot{Z}(\theta)$ has the greatest value.
Step 6	Solve problem (3.12) within the parameter space that yield the greatest Z^{diff} in step (5)
Step 7	Solve the corresponding MP-LP problem and divide the space of parameters in to critical regions, according to the comparison procedure, so that the upper bound is updated go to Step (4)
Step 8	when all the parameter space is searched, $CR^{IG} = \emptyset$, check if there are repeated solutions in different critical regions and unify them in order to have a compact and convex solution.

Consequently, the explicit expressions are given by a conditional piecewise linear function [18]:

$$\begin{cases} x = C_1 + K_1 \cdot \theta, \theta \in CR^1 \\ x = C_2 + K_2 \cdot \theta, \theta \in CR^2 \\ \vdots \\ x = C_p + K_p \cdot \theta, \theta \in CR^p \end{cases}$$

Where C_i are column vectors and K_i are real matrices, whereas $CR^i \in R^m$ are critical regions and note that CR^i denotes the i^{th} critical region. The main steps of the algorithm are outlined in the Table 3.6.

Note that the solution of multi-parametric nonconvex programming problems can be approached by employing the principles of multiparametric nonlinear programming.

4 Convex Relaxation

A significant effort has been spent in the last five decades studying theoretical and algorithmic aspects of local optimization algorithms. Comparatively, there has been traditionally much less attention devoted to global optimization methods. However, in the last decade the area of global optimization has attracted, a lot of interest from the operations research, engineering and applied mathematics communities [3], [4]. This recent flow of interest can be attributed to the realization that there exists an abundance of optimization problems for which existing local optimization approaches can't consistently locate the global minimum solution.

Some section of all optimization problems arising in an industrial or scientific context are characterized by the presence of nonconvexities in the participating functions. The nonconvexities represent major barriers in attaining the global optimal solution [3]. To guarantee that KKT optimality conditions are both necessary and sufficient for obtaining the global optimum of the i^{th}-level problem via multiparametric programming approach, the functions f_i and g_i, for $i = \{1, 2, \ldots, k\}$, have to be convex.

If for fixed parameter value x, the assumptions of Basic sensitivity theorem hold, f_k and g_k are convex, then the KKT optimality conditions are necessary and sufficient for obtaining the global optimum of the most inner problem [26]. Whereas KKT conditions are only necessary conditions.

Many of the deterministic methods proposed to date rely on the generation of valid convex underestimatores for the nonconvex functions involved [3], [38]. Successive improvements of these estimates, together with the identification of the global solution of the resulting convex problems, eventually lead to the determination of the global solution of the original nonconvex problem.

In this section, we describe a method to formulate a class of nonconvex optimization problems as equivalent convex optimization problems. Our method closely follows the work of Adjimin *et al.*, [3] in 1997 and Gümü *et al.*, [26]. A convex relaxation of nonconvex terms can be constructed by replacing the occurrence of each nonconvex term, with a convex lower bounding functions.

4.1 Underestimation of Multiparametric Nonconvex Programming Problem

A major difficulty with multiparametric nonconvex formulation is that the global optimal solution cannot be efficiently computed, and the behavior of a local solution is hard to analyze. In practice, convex relaxation has been adopted to remedy the problem. The choice of convex formulation makes the solution unique and efficient to compute.

In this sub-section we discuss the way to get valid convex underestimators of some classes of multiparametric nonconvex programming problems which may be occurred on the most inner problem. In constructing a convex underestimator for the overall function, first we noted that the linear and convex terms do not require any transformation [3]. The convex envelope of the bilinear, trilinear, fractional, fractional trilinear, univariate concave and general nonconvex terms can be constructed by the following simple rule.

In the case of bilinear term (i.e. linear as a function of either of its argument when the other is fixed.) x_1x_2, Pistikopoulos *et al.*, [38] in 2007, showed that the tightest convex lower bound over the domain $[x_1^L, x_1^U] \times [x_2^L, x_2^U]$ is obtained by introducing a new variable w_2 which replaces every occurrences of x_1x_2 in the problem and satisfies the following relationship:

$$w_2 = \max\{x_1^L x_2 + x_2^L x_1 - x_1^L x_2^L, x_1^U x_2 + x_2^U x_1 - x_1^U x_2^U\} \tag{4.1}$$

This lower bound can be relaxed and included in the minimization problem by adding two linear inequality constraints,

$$\begin{cases} w_2 \geq x_1^L x_2 + x_2^L x_1 - x_1^L x_2^L \\ w_2 \geq x_1^U x_2 + x_2^U x_1 - x_1^U x_2^U \end{cases} \tag{4.2}$$

Trilinear terms of the form $x_1x_2x_3$ can be underestimated through replacing it by the new variable subject to sets of constraints in a way similar to the underestimation of bilinear terms [3]. i.e. a new variable w_3 is introduced and bounded by the following inequality constraints:

$$\begin{cases} w_3 \geq x_1 x_2^L x_3^L + x_1^L x_2 x_3^L + x_1^L x_2^L x_3 - 2x_1^L x_2^L x_3^L \\ w_3 \geq x_1 x_2^U x_3^U + x_1^U x_2 x_3^L + x_1^U x_2^L x_3 - x_1^U x_2^L x_3^L - x_1^U x_2^U x_3^U \\ w_3 \geq x_1 x_2^L x_3^L + x_1^L x_2 x_3^U + x_1^L x_2^U x_3 - x_1^L x_2^U x_3^U - x_1^L x_2^L x_3^L \\ w_3 \geq x_1 x_2^U x_3^L + x_1^U x_2 x_3^U + x_1^L x_2^U x_3 - x_1^L x_2^U x_3^L - x_1^U x_2^U x_3^U \\ w_3 \geq x_1 x_2^L x_3^U + x_1^L x_2 x_3^L + x_1^U x_2^L x_3 - x_1^U x_2^L x_3^U - x_1^L x_2^L x_3^L \\ w_3 \geq x_1 x_2^L x_3^U + x_1^L x_2 x_3^U + x_1^U x_2^U x_3 - x_1^L x_2^L x_3^U - x_1^U x_2^U x_3^U \\ w_3 \geq x_1 x_2^U x_3^L + x_1^U x_2 x_3^L + x_1^L x_2^L x_3 - x_1^U x_2^U x_3^L - x_1^L x_2^L x_3^L \\ w_3 \geq x_1 x_2^U x_3^U + x_1^U x_2 x_3^U + x_1^U x_2^U x_3 - 2x_1^U x_2^U x_3^U \end{cases} \tag{4.3}$$

Fractional terms of the form $\frac{x_1}{x_2}$ where $x_2 > 0$ are underestimated by introducing a new variable w_f and two new constraints which depends on the sign of the bounds on x_1 and x_2 [3].

$$\begin{aligned} w_f &\geq \begin{cases} \frac{x_1^L}{x_2} + \frac{x_1}{x_2^U} - \frac{x_1^L}{x_2^U} & , if\ x_1^L \geq 0 \\ \frac{x_1}{x_2^U} + \frac{x_1^L}{x_2^L} - \frac{x_1^L x_2}{x_2^L x_2^U} & , if\ x_1^L < 0 \end{cases} \\ w_f &\geq \begin{cases} \frac{x_1^U}{x_2} + \frac{x_1}{x_2^L} - \frac{x_1^U}{x_2^L} & , if\ x_1^U \geq 0 \\ \frac{x_1}{x_2^L} + \frac{x_1^U}{x_2^U} - \frac{x_1^U x_2}{x_2^L x_2^U} & , if\ x_1^U < 0 \end{cases} \end{aligned} \tag{4.4}$$

The fractional trilinear terms of the form $\frac{x_1 x_2}{x_3}$ are replaced by a new variable w_F as usual [3] and the constraints for $x_1^L, x_2^L \geq 0$ and $x_3^L > 0$ are given by:

$$\left\{\begin{array}{l}
w_F \geq \frac{x_1 x_2^L}{x_3^U} + \frac{x_1^L x_2}{x_3^U} + \frac{x_1^L x_2^L}{x_3} - \frac{2x_1^L x_2^L}{x_3^U} \\
w_F \geq \frac{x_1 y_2^L}{x_3^U} + \frac{x_1^L x_2}{x_3^L} + \frac{x_1^L x_2^U}{x_3^L} - \frac{x_1^L x_2^L}{x_3^U} \\
w_F \geq \frac{x_1 x_2^U}{x_3^L} + \frac{x_1^U x_2}{x_3^U} + \frac{x_1^U x_2^L}{x_3} - \frac{x_1^U x_2^L}{x_3^U} - \frac{x_1^U x_2^U}{x_3^L} \\
w_F \geq \frac{x_1 x_2^U}{x_3^U} + \frac{x_1^U x_2}{x_3^L} + \frac{x_1^L x_2^U}{x_3} - \frac{x_1^L x_2^U}{x_3^U} - \frac{x_1^U x_2^U}{x_3^L} \\
w_F \geq \frac{x_1 x_2^L}{x_3^U} + \frac{x_1^L x_2}{x_3^L} + \frac{x_1^U x_2^L}{x_3} - \frac{x_1^U x_2^L}{x_3^L} - \frac{x_1^L x_2^L}{x_3^U} \\
w_F \geq \frac{x_1 x_2^U}{x_3^U} + \frac{x_1^U x_2}{x_3^L} + \frac{x-1^L x_2^U}{x_3} - \frac{x_1^L x_2^U}{x_3^U} - \frac{x_1^U x_2^U}{x_3^L} \\
w_F \geq \frac{x_1 x_2^L}{x_3^U} + \frac{x_1^L x_2}{x_3^L} + \frac{x_1^U x_2^L}{x_3} - \frac{x_1^U x_2^L}{x_3^L} - \frac{x_1^L x_2^L}{x_3^U} \\
w_F \geq \frac{x_1 x_2^U}{x_3^L} + \frac{x_1^U x_2}{x_3^L} + \frac{x_1^U x_2^U}{x_3} - \frac{2x_1^U x_2^U}{x_3^L}
\end{array}\right. \tag{4.5}$$

Univariate concave functions can be trivially underestimated by their linearization at the lower bound of the variable range [3]. Thus the convex envelope of the concave function $f(x)$ over the interval $[x^L, x^U]$ is the linear function of x:

$$f(x^L) + \frac{f(x^U) - f(x^L)}{x^U - x^L}.(x - x^L) \tag{4.6}$$

Note that the generation of the best convex underestimator for a univariate concave function does not require the introduction of additional variables or constraints.

All other general multiparametric nonconvex terms for which customized underestimators do not exist are underestimated as proposed in the literature [3]. A function $f(x, y) \in C^2(\mathbb{R}^m \times \mathbb{R}^n)$ is underestimated over the entire domain $[y^L, y^U]$ by the function $F(x, y)$ defined as:

$$F(x, y) = f(x, y) + \sum_{i=1}^{n} \alpha_i (y_i^L - y_i)(y_i^U - y_i) \tag{4.7}$$

Where the α_i's are positive scalars.

Theorem 4.1 *(Properties of $F(x, y)$)*
(i) $F(x, y)$ is a valid underestimator of $f(x, y)$, for a given x
(ii) $F(x, y)$ matches $f(x, y)$ at all corner points.
(iii) $F(x, y)$ is convex in $y_i \in [y^L, y^U]$, $i = 1, 2, \ldots, n$.
(iv) The maximum separation between the nonconvex term of generic structure $f(x, y)$ and its convex relaxation $F(x, y)$ is bounded and proportional to the positive parameters α_i and to the square of the diagonal of the current box constraints

Proof: (i) For every $i = 1, 2, \ldots, n$, we have $(y_i^L - y_i)(y_i^U - y_i) \leq 0$ and also by definition we have $\alpha_i \geq 0$. Therefore, $y \in [y^L, y^U]$, implies $F(x, y) \leq f(x, y)$.

(ii) Let y^c be a corner point of $[y^L, y^U]$, then for every $i = 1, 2, \ldots, n$, $(y_i^L - y_i^c) = 0$ or $(y_i^U - y_i^c) = 0$. Therefore, $F(x, y^c) = f(x, y^c)$ in either case, for fixed x
(iii)direct result from the definition of α_i's
(iv) see ([4]) □

Since $F(x, y)$ is convex if and only if its Hessian matrix $H_F(x, y)$ is positive semi-definite, a useful convexity condition is derived by noting that $H_F(x, y)$ is related to the Hessian matrix $H_f(x, y)$ of $f(x, y)$ by [4]

$$H_F(x, y) = H_f(x, y) + 2\Delta \tag{4.8}$$

Where Δ is a diagonal matrix whose diagonal elements are the α_i's. Δ is referred to as the diagonal shift matrix, since the addition of the quadratic term to the function $f(x, y)$ as shown in equation (4.7), corresponds to the introduction of a shift in the diagonal elements of its Hessian matrix $H_f(x, y)$. The following theorem can then be used to ensure that $F(x, y)$ is indeed a convex underestimator:

Theorem 4.2 *[3] $F(x, y)$, as defined in equ. (4.7), is convex if and only if $H_f(x, y) + 2\Delta = H_f(x, y) + 2diag(\alpha_i)$ is positive semi-definite for all $y \in [y^L, y^U]$. however, x is given in its parameter space.*

A number of deterministic methods have been devised in order to identify an appropriate diagonal shift matrix. They are discussed in detail in the literature (eg. see [3])

4.2 Parametric Overestimator

An upper bound can be calculated either by solving the original problem locally over each subdomain of the solution space [3], (this type of calculating an upper bound is powerful in classical nonconvex optimization problems, but since parametric nonconvex programming problem may violate the second order sufficiency condition this method may fail to give appropriate solution) or by performing a problem evaluation at the solution of the lower bounding problem [38]. Substituting the solutions of the underestimating subproblem into the original nonconvex problem results in an overestimator which in general may be nonlinear and nonconvex (for the details see details in [38]).

For classes of bilinear terms, there are different overestimators for a given parametric optimization problems [38]. Convex parametric overestimators can also be similarly obtained by creating overestimators of the objective and constraint functions and solving the resulting multiparametric programming problem as described in [36], [38].

Parametric overestimator of type 1 denoted by $\hat{Z}_{0_1}(x)$ is perhaps the easiest one to obtain because it requires only the substitution of the solution of the underestimating subproblem into the original (nonconvex) one.

Another overestimator denoted by $\hat{Z}_{0_2}(x)$ can be created based up on the following lemmas, from [38].

Lemma 4.3 *[38] The maximum separation between bilinear terms denoted by y_1y_2 and the underestimating subproblem may be denoted by w in side the rectangle $[y_1^L, y_1^U] \times [y_2^L, y_2^U]$ is equal to $\delta_{12} = \frac{[y_1^U - y_1^L][y_2^U - y_2^L]}{4}$.*

Lemma 4.4 *[38] If $f(x,y) = f^c(x,y) + \sum_{i=1}^n \sum_{j=1}^n \alpha_{i,j} y_i y_j$ where $i \neq j$, $f^c(x,y)$ is convex function of y for a given x, $\alpha_{i,j}$ is a positive constant for all i and j and g is convex vector valued constraint set. Then, $\hat{Z}_{0_2}(x) = \check{Z}(x) + \sum_{i=1}^n \sum_{j=1}^n \alpha_{i,j}\delta_{ij}$, where $\delta_{ij} = \frac{[y_i^U - y_i^L][y_j^U - y_j^L]}{4}$*

Unlike the case of $\hat{Z}_{0_1}(x)$ and many cases of $\hat{Z}_{0_2}(x)$, an overestimator of the third type $\hat{Z}_{0_3}(x)$ requires solving a parametric optimization problem. The key factor is that this type of formulation is completely general. Overestimating subproblems are given by replacing the occurrence of each bilinear terms y_iy_j, for $i \neq j$ with w as follows:

$$\begin{cases} w \leq y_1^L \cdot y_2 + y_2^U \cdot y_2 - y_1^L \cdot y_2^U \\ w \leq y_1^U y_2 + y_2^L y_2 - y_1^U y_2^L \end{cases} \tag{4.9}$$

Note that Equ. (4.9) are used [38] except that the sign of the less than inequalities has been changed to the greater than inequalities as follows:

$$\begin{cases} w \geq y_1^L y_2 + y_2^U y_2 - y_1^L y_2^U \\ w \geq y_1^U y_2 + y_2^L y_2 - y_1^U y_2^L \end{cases} \tag{4.10}$$

The fourth type of overestimator $\hat{Z}_{0_4}(x)$ also requires solving a parametric optimization problem. The overestimating subproblem is formulated as maximization problem subject to additional linear constraints (4.9). However it is limited to the case when only bilinear terms are present in the objective function because it relies on maximization of the auxiliary variables w, which replaces the bilinear term [38].

5 Algorithm for Multilevel Optimization with Nonconvexity Formulation at Inner Levels

In this section, we try to show the way in which we can possibly solve multilevel optimization problems where the nonconvexity formulation occurs at inner levels. Eventhough it's inflexible, for the sake of clarity and without loss of generality, we will discuss and analyze the algorithm using two particular classes of multilevel programming problems: bilevel and trilevel programming problem with nonconvex formulation in the most inner problem and solve them to global optimality through the application of multiparametric programming. For this assume throughout this study that the objective as well as the constraint set functions are partially differentiable with respect to the optimization variables and the parameter variables in each level.

5.1 Bilevel Programming Problem with Nonconvexity Formulation at Inner Problem

Consider the following bilevel programming problem with nonconvex formulation in the inner problem:

$$\begin{aligned}
&\min_{x} f_1(x, y) \\
&\quad \text{s.t.} \\
&\qquad g_1(x, y) \leq 0 \\
&\qquad \min_{y} f_2(x, y) \\
&\qquad\quad \text{s.t} \\
&\qquad\qquad g_2(x, y) \leq 0 \\
&\qquad\qquad y^L \leq y \leq y^U \\
&\qquad\qquad x^L \leq x \leq x^U
\end{aligned} \tag{5.1}$$

where $x \in X$, $y \in Y$ and X and Y are compact convex sets.

Problem (5.1) comprises two subproblems, one at each optimization level. The inner problem can be recast as a multiparametric nonconvex programming problem, where the optimization variables corresponding to the upper optimization problem are classified as parameters.

Assume that the inner problem contains all nonconvex terms which have been discussed in Section 4. For the sake of clarity and with out losses of generality we analyze the discussions in the algorithm only with the occurrence of generic nonconvex and bilinear terms, throughout this study.

Recall, that problem (5.1) can be rewritten equivalently as,

$$\min_{x}\{f_1(x,y(x)) : g_1(x,y(x)) \leq 0, y(x) \in \Psi(x)\}$$

where, $\Psi(x) = \{y \in Y : y \in \arg\min\{f_2(x,y) : g_2(x,x) \leq 0\}\}$. Note that arg min stands for the set of points of the given argument for which the value of the given function attains its minimum.

Theorem 5.1 *Consider the inner level of problem (5.1) be general nonlinear parametric programming problem. Assume* ***Assumption (C)*** *is satisfied and the set of stationary points be nonempty. Let* $O_s = \{y \in Y : y^i = y_0^i - M_0^{-1} \cdot N_0(x - x_0),\ x \in CR^i\}$ *(where,* $i = 1, 2, \ldots, I_2$ *with* I_2 *being the number of critical regions) be an optimal solution set. Then* $O_s \equiv \Psi(x)$

Proof:($\subseteq$) Let $y^\star \in O_s$, we need to show that $y^\star \in \Psi(x)$.
Since $y^\star \in O_s$ we know that $y^\star$ is an optimal solution, for $x \in CR^i$ for some i. This implies that $f_2(x, y^\star(x)) = \min_y\{f_2(x,y) : g_2(x,y) \leq 0, x \in CR^i\}$. Now by definition of arg min we have $y^\star \in \Psi(x)$. Hence $O_s \subseteq \Psi(x)$.
($\supseteq$): Direct from the definition.
Therefore $O_s \equiv \Psi(x)$ as desired.□

As a result, the approach begins with rewriting the inner problem of (5.1) as:

$$\min_{y}\{cf(x,y) + \sum_{k=1}^{N} Nf_k(x,y) + \sum_{i=1}^{N-1}\sum_{j=i+1}^{N} b_{ij}y_iy_j\}$$

s.t.

$$Cf^p(x,y) + \sum_{k=1}^{n} Nf_k^p(x,y) + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} b_{ij}^p y_iy_j \leq 0,\ p = 1, 2, \ldots, m$$

$$x^L \leq x \leq x^U$$

$$y^L \leq y \leq y^U \tag{5.2}$$

Where, $Cf(x,y)$ is the convex part of the objective function, $Nf_k(x,y)$ is the set of generic nonconvex terms appearing in the objective function and $b_{ij}y_iy_j$'s are the bilinear terms appearing in the objective function. Similarly, for each constraint p, we have a convex part $Cf^p(x,y)$, generic nonconvex terms $Nf_k^p(x,y)$ and the bilinear terms $b_{ij}^p y_iy_j$.

Consequently, derivation of the convex relaxation of problem (5.2) is performed. It can be constructed by replacing each of the generic nonconvex terms and each bilinear terms with its tighter convex lower bounding functions. Thus, the convex lower bounding of the bilinear term can be obtained by replacing it by additional variable w_{ij} and introducing the following linear constraints:

$$\begin{cases} w_{ij} \geq y_i^L y_j + y_j^L y_i - y_i^L y_j^L \\ w_{ij} \geq y_i^U y_j + y_j^U y_i - y_i^U y_j^U \end{cases} \tag{5.3}$$

and for each one of the generic nonconvex function, one can establish a convex underestimating function $Nf_k^{p,\text{conv}}$ which can be defined by augmenting the original nonconvex expression with the addition of a separable convex quadratic function of y_i.
i.e.

$$Nf_k^{p,\text{conv}}(x,y) = \sum_{k=1}^{n} Nf_k^p(x,y) +$$

$$\sum_{k=1}^{n}\sum_{i=1}^{k} \alpha_{ki}^p (y_i^L - y_i)(y_i^U - y_i), p = 1, 2, \ldots, m$$

where, $\alpha_{ki}^p(y^L, y^U) \geq \max\{0, \frac{-1}{2} \min_{y^L \leq y \leq y^U} \lambda(y)\}$ and,
$\alpha_{kj}^p(x^L, x^U) \geq \max\{0, \frac{-1}{2} \min_{x^L \leq x \leq x^U} \lambda(x)\}$, with $\lambda(x)$ and $\lambda(y)$ are the set of eigenvalues of the Hessian matrix of $Nf_k^{p,\text{conv}}$ over $x^L \leq x \leq x^U$ and $y^L \leq y \leq y^U$, respectively. Based on these relaxation of different nonconvex terms appearing in the objective function and constraint functions formulation (5.2) can be rewritten equivalently as:

$$\check{Z}^{11}(x) = \min_y \{cf(x,y) + \sum_{k=1}^{n} Nf_k(x,y) + \sum_{k=1}^{n}\sum_{i=1}^{k} \alpha_{ki}(y_i^L - y_i)(y_i^U - y_i) + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} w_{ij}\}$$

S.t.

$$\sum_{k=1}^{n} Nf_k^p(x,y) + \sum_{k=1}^{n}\sum_{i=1}^{k} \alpha_{ki}^p(y_i^L - y_i)(y_i^U - y_i) + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} w_{ij} \leq 0$$

$$p = 1, 2, \ldots, m$$

$$\begin{aligned} w_{ij} &\geq y_i^L y_j + y_j^L y_i - y_i^L y_j^L \\ w_{ij} &\geq y_i^U y_j + y_j^U y_i - y_i^U y_j^U \\ x^L &\leq x \leq x^U \\ y^L &\leq y \leq y^U \end{aligned} \tag{5.4}$$

Problem (5.4) can be solved with suitable multiparametric programming algorithm given in [36], [38] which is also discussed in Section 3, resulting in:

$$\begin{aligned} y^i &= m_2^i + n_2^i \cdot x,\ H_2^i \cdot x \leq h_2^i \\ w_{ij}^i &= O_2^i + L_2^i \cdot x\ D_2^i \cdot x \leq d_2^i \end{aligned} \tag{5.5}$$

where, $i = 1, 2, \ldots, I_2$, with I_2 being the number of critical regions and consequently, the number of linear approximation on the optimal rational reaction set $\Psi(x)$ (see Corollary 3.16).

Substituting the expressions in (5.5) into the objective function of the underestimated problem (5.4) one can obtain a multiparametric lower bound $\check{Z}_i^{11}(x)$ for the solution of problem (5.2) with the corresponding critical regions, CR^i. Substitution of the expressions into the objective function of problem (5.2) provides a parametric upper bound $\hat{Z}_i^{11}(x)$ with the corresponding critical region CR^i. Since, the constraints in problem (5.2) are nonconvex, the solutions have to also be substituted into the constraints to check feasibility.

Each of the lower bounds, $\check{Z}_i^{11}(x)$ is then compared to $\hat{Z}_i^{11}(x)$, and the region of x where,

$$\hat{Z}_i^{11}(x) - \check{Z}_i^{11}(x) \leq \varepsilon$$

are fathomed in a small positive tolerance ε. If each $\check{Z}_i^1(x)$ are within ε of $\hat{Z}_i^1(x)$ in the space of x, then the expressions in (5.5) can be incorporated in the leader problem of (5.1). Note that since the expressions in (5.5) are piecewise linear functions of x, the complexity of the original problem does not increase.

Otherwise, the initial region of y is partitioned into smaller regions in the following way: Tighter lower bounds to the solution of problem (5.2) can be obtained by dividing the initial feasible region into two subregions by using one of the branching rule that are developed within the deterministic global optimization algorithm the so-called branch and bound [26]. Forexample, the initial feasible region, as determined by the variable bounds, can be subdivided into two subregions by halving along the longest side. After branching the corresponding subproblem in each subregion can be reformulated as:

$$\begin{aligned}
\check{Z}^{21}(x) = \min_y \{ cf(x,y) + \sum_{k=1}^{n}\sum_{i=1}^{k} \alpha_{ki}(y_i^L - y_i)(y_{inew}^U - y_i) \\
+ \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} w_{ij} \}
\end{aligned}$$

$$\text{S.t.}$$

$$\begin{aligned}
\sum_{k=1}^{n} Nf_k^p(x,y) + \sum_{k=1}^{n}\sum_{i=1}^{k} \alpha_{ki}^p(y_i^L - y_i)(y_{inew}^U - y_i) \\
+ \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} w_{ij} \leq 0 \\
p = 1,2,\ldots,m \\
w_{ij} \geq y_i^L y_j + y_j^L y_i - y_i^L y_j^L \\
w_{ij} \geq y_{inew}^U y_j + y_{jnew}^U y_i - y_{inew}^U y_{jnew}^U \\
x^L \leq x \leq x^U \\
y^L \leq y \leq y_{new}^U
\end{aligned} \tag{5.6}$$

and,

$$
\begin{aligned}
\check{Z}^{22}(x) = \min_{y} cf(x,y) + \sum_{k=1}^{n} Nf_k(x,y) + \sum_{k=1}^{n}\sum_{i=1}^{k} \alpha_{ki}(y_{inew}^L - y_i)(y_i^U - y_i) \\
+ \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} w_{ij} \\
\text{S.t.} \\
\sum_{k=1}^{n} Nf_k^p(x,y) + \sum_{k=1}^{n}\sum_{i=1}^{k} \alpha_{ki}^p(y_{inew}^L - y_i)(y_i^U - y_i) \\
+ \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} w_{ij} \leq 0 \\
p = 1, 2, \ldots, m \\
w_{ij} \geq y_{inew}^L y_j + y_{jnew}^L y_i - y_{inew}^L y_{jnew}^L \\
w_{ij} \geq y_i^U y_j + y_j^U y_i - y_i^U y_j^U \\
x^L \leq x \leq x^U \\
y_{new}^L \leq y \leq y^U
\end{aligned}
\tag{5.7}
$$

Solve each of the problems (5.6) and (5.7) using suitable multiparametric programming approach in the region of x and substitute each parametric solutions of problem (5.6) and (5.7) within the corresponding CR, into the objective function of problem (5.6) and (5.7) respectively, one can obtain a parametric lower bound for the solution of problem (5.2). Similarly, one can get a parametric upper bound for the solution of problem (5.2) by substituting each parametric solutions of the problems (5.6) and (5.7) into the original problem (5.2).

Compare each parametric upper bounds with the former upper bound for each subproblem and choose the least upper bound within the coresponding critical regions and we should denote them by $\hat{Z}_{lub}(x)$. Similarly, compare the lower bounds and take the maximum of them and update the lower bound as well and denote it by $\check{Z}_{glb}(x)$. Consequently, perform convergence test by comparing the updated parametric upper bound with the updated lower bounds. If the convergence test is satisfied the Branching rule stop there.

Otherwise, continue in the same manner as discussed above until the lower bound and the upper bound come together inside sufficiently small positive tolerance ε. Following this procedure, the bilevel optimization problems in (5.1) will be converted to single-level optimization problem:

$$\begin{aligned}
&\min_{x} f_1^*(x, y(x)) \\
&\text{s.t} \\
&g_1(x, y(x)) \leq 0 \\
&x \in S, \\
&S = \{x \in X : y \in Y, g_2(x, y) \leq 0\}
\end{aligned} \tag{5.8}$$

The solution of the first-level optimization problem of (5.1) depends on the number of critical regions obtained in the inner problem. Based upon the above developments an algorithm for the solution of (5.1) is presented in Table 4, and is illustrated with one example.

Remark 5.2 *If the optimization variables are unbounded the algorithm doesn't work.*

5.1.1 Convergence Justification for the Algorithm

The concept of convergence has been an important issue corresponding to the algorithm of optimization problems and other related disciplines. In this section, we summarize the convergence of the algorithms developed in the above section.

The algorithm for the special structure bilevel optimization problem with the inclusion of nonconvexity in the inner problem is a deterministic global optimization method based on parametric programming algorithm and branch and bound framework.

As it has been discussed in section 4, the maximum separation between the generic and bilinear nonconvex terms and their respective convex underestimator functions is bounded [4]. Further, as the size of the rectangular domains approaches zero, these maximum separation tends to zero as well. This implies that as the current box constraint $[y^L, y^U]$ collapse into a point;

- The maximum separation between the original nonconvex objective function and its convex underestimator function becomes zero and
- By the same argument the maximum separation between the original nonconvex constraint set and the convex underestimated constraint set goes to zero as well.

In order to prove the convergence of the algorithm stated in Table 4, we need the following Theorems.

Theorem 5.3 *Consider problem (5.1) and assume that the optimization variables are bounded. If the reformulated inner problem has a solution $y(x) \in 2^{R^m}$ at each iteration, then the sequence $\chi_k = \hat{Z}_k(x) - \check{Z}_k(x)$, is a decreasing sequence. Where $\hat{Z}_k(x)$ is the k^{th} least upper bound and $\check{Z}_k(x)$ is the k^{th} greatest lower bound.*

Table 4: A Parametric programming approach for bilevel programming problems with nonconvexities in the inner problem

Step	Description
1	Recast the inner problem as a multiparametric programming problem, with parameters being the upper level optimization problem variables
2	Initialize the current upper bound as $\hat{Z}(x) = \infty$, a region of x, CR, a space of continuous variable y determined by the lower and upper bounds y^L and y^U respectively, and tolerance, ε
3	For a given region of x, CR, and the corresponding space of y, formulate and solve the relaxed convex optimization problem using suitable multiparametric programming algorithm as described in Section 3 and obtain the parametric lower bound, $\check{Z}(x)$ and the parametric upper bound, $\hat{Z}(x)$, by using one of the methods described in Section 4
4	Check the difference between $\hat{Z}(x)$ and $\check{Z}(x)$ by solving the following maximization problem, $\max_x\{\hat{Z}(x) - \check{Z}(x) : x \in CR^R\}$. If, $\hat{Z}(x) - \check{Z}(x) \leq \max_x\{\hat{Z}(x) - \check{Z}(x) : x \in CR^R\} \leq \varepsilon$ and the relaxed convex optimization problems are infeasible for some rectangle, then in such cases fathom those regions accordingly in the space of parameters.
5	If no more space of y to explore, terminate, and go to the next Step. Otherwise, go to Step 9.
6	Substitute each of the solutions in the leader's problem and formulate one-level optimization problem
7	Solve each single-level problems
8	Compare the leader's optimal solutions and select the best as one need.
9	Branch and bound on y and formulate a convex underestimator in each subrectangles of y
10	Solve the convex underestimator problems in each subrectangles using multiparametric programming approach and obtain the lower and upper bounds for the solution in each subrectangle.
11	Compare the former lower bound with the lower bounds of each subrectangle within the corresponding critical regions and take, $\check{Z}_{glb}(x)$ to be the maximum within the corresponding critical region.
12	Compare the former upper bound with the upper bounds of each subrectangle within the corresponding critical regions and take $\hat{Z}_{lub}(x)$ to be the minimum within the corresponding critical region and go to Step (4)

Proof: To prove the theorem we need to show the following two things.

1. First we need to show that the sequence $\{\hat{Z}_k(x)\}_{k=1}^{n}$ is a decreasing sequence. Let $\hat{Z}^p(x)$ be the upper bound at p^{th} iteration. Hence, from the algorithm for the p^{th} iteration we have, $\hat{Z}_p(x) = \min\{\min\{\hat{Z}_i(x)\}_{i=1}^{p-1}, \hat{Z}^p(x)\}$, and similarly, at $p+1$, we've $\hat{Z}_{p+1}(x) = \min\{\min\{\hat{Z}_i(x)\}_{i=1}^{p}, \hat{Z}^{p+1}(x)\} = \min\{\min\{\hat{Z}_i(x)\}_{i=1}^{p-1}, \hat{Z}_p(x), \hat{Z}^{p+1}(x)\}$, but $\hat{Z}_p(x) = \min\{\min\{\hat{Z}_i(x)\}_{i=1}^{p-1}, \hat{Z}^p(x)\}$. This implies that $\hat{Z}_{p+1}(x) \leq \hat{Z}_p(x)$, as desired.
2. Secondly, we need to show that $\check{Z}_k(x)$ is an increasing sequence. Let $\check{Z}^p$ be the lower bound for p^{th} iteration. As a result of the algorithm, we have, $\check{Z}_p(x) = \max\{\max\{\check{Z}_i(x)\}_{i=1}^{p-1}, \check{Z}^p\}$ and similarly at $p+1$ we have,
$\check{Z}_{p+1}(x) = \max\{\max\{\check{Z}_i(x)\}_{i=1}^{p}, \check{Z}^{p+1}\} = \max\{\max\{\check{Z}_i(x)\}_{i=1}^{p-1}, \check{Z}_p, \check{Z}^{p+1}\}$, but $\check{Z}_p(x) = \max\{\max\{\check{Z}_i(x)\}_{i=1}^{p-1}, \check{Z}^p\}$ and as the size of the rectangular domain decreases, the maximum separation between the original nonconvex function and their respective convex underestimator function decreases. This implies that $\check{Z}_p(x) \leq \check{Z}_{p+1}(x)$.

Therefore, since $\hat{Z}_k(x)$ is a decreasing sequence and $\check{Z}_k(x)$ is an increasing sequence at each iteration, we can conclude that the difference χ_k is a decreasing and hence a convergent sequence.□

Theorem 5.4 *Let $X \subseteq R^m$ be a polyhedron and $CR^Q = \{x \in X : \tilde{g}_2(x) - \tilde{b} \leq 0\} \subseteq X$, be a critical region. Assume $CR^Q \neq \emptyset$. Also let $CR^i = \{x \in X : \tilde{g}_2^i(x) - \tilde{b}^i > 0, \tilde{g}_2^j(x) - \tilde{b}^j \leq 0, \forall j < i,\ i = 1, 2, \ldots, K\}$ where $K = dim(b)$, and let $CR^{Rest} = \bigcup_{i=1}^{K} CR^i$. Then*

1. *$CR^{Rest} \bigcup CR^Q = X$,*
2. *$CR^Q \bigcap CR^i = \emptyset$,*
3. *$CR^i \bigcap CR^j = \emptyset$, $\forall i \neq j$, i.e. $\{CR^Q, CR^1, \ldots, CR^K\}$ is a partition of X.*

Proof

1.
   - ($\subseteq$) Let $x \in CR^{Rest} \bigcup CR^Q$. If $x \in CR^Q$, we are done. Otherwise, there exists an index i such that $x \in CR^i$, this implies that $x \in X$ such that $\tilde{g}_2^i(x) > \tilde{b}^i$. Hence $CR^{Rest} \bigcup CR^Q \subseteq X$ as desired.
   - ($\supseteq$) Let $x \in X$. Assume that $x \notin CR^Q$. Then, there exists an index i such that $\tilde{g}_2^i(x) - \tilde{b}^i > 0$. Let $i^* = \min_{i \leq K}\{i : \tilde{g}_2^i(x) > \tilde{b}^i\}$. Then, $x \in CR^{i^*}$, because $\tilde{g}_2^{i^*}(x) > \tilde{b}^{i^*}$. This implies that $x \in CR^Q \bigcup CR^{Rest}$. Hence the result.
2. If $x \in CR^Q$ then by definition, there doesn't exist an index i that satisfy $\tilde{g}_2^i(x) - \tilde{b}^i > 0$. which implies that $x \notin CR^i$.
3. Let $x \in CR^i$ and take $i > j$. Since $x \in CR^i$, by definition of $CR^i(i > j)$ $\tilde{g}_2^j(x) - \tilde{b}^j \leq 0$, which implies that $x \notin CR^j$. □

As a direct consequence of the above two Theorems we have,

Corollary 5.5 *Let* $CR^{Rest}\bigcup CR^{Q} = X, CR^{Q}\bigcap CR^{i} = \emptyset$ *and* $CR^{i}\bigcap CR^{j} = \emptyset, \forall i \neq j$ *be a partition of* X. *If* $\{\chi_k\}_{k=1}^{n}$ *is a decreasing sequence, then the algorithm converges.*

Note that one can extend the idea discussed here for the trilevel programming algorithm.

5.1.2 Illustrative Example

While the ideas discussed in the previous section are quiet general, for the sake of simplicity in presentation the illustration will be centered on the case when the only nonconvexities in the inner problem are due to the presence of bilinear terms.

Consider the following bilevel programming problem, with the inner problem taken from the work of Pistikopoulos *et al.*, [18].

$$\begin{aligned}
&\min_{x} f_1 = -2x + y \\
&\quad \text{s.t} \\
&\frac{1}{2}x + y \leq 0 \\
&\quad\quad \text{where } [y, z]^T \text{ solves} \\
&\quad\quad\quad \min_{y,z} f_2 = yz \\
&\quad\quad \text{s.t} \\
&\quad\quad -2y - z + x \leq 0 \\
&\quad\quad -y - 3z + \frac{1}{2}x \leq 0 \\
&\quad\quad 0 \leq y \leq 1 \\
&\quad\quad -\frac{1}{6} \leq z \leq \frac{7}{12} \\
&\quad\quad 0 \leq x \leq 1
\end{aligned} \tag{5.9}$$

Reformulate the second-level optimization problem as follows:

$$\begin{aligned}
&\min_{y,z} f_2 = yz \\
&\text{s.t} \\
&-2y - z + x \leq 0 \\
&-y - 3z + \frac{1}{2}x \leq 0 \\
&0 \leq y \leq 1 \\
&-\frac{1}{6} \leq z \leq \frac{7}{12} \\
&0 \leq x \leq 1
\end{aligned} \tag{5.10}$$

Following the steps discussed in Table 5.1. Let the tolerance $\varepsilon = 0.01$; and perform a convex relaxation for the nonconvex problem (5.10) by replacing the bilinear term with

w and introduce the following linear inequalities :
$w \geq -\frac{1}{6}y$,
$w \geq z + \frac{7}{12}y - 1$, reformulate problem (5.10) by fixing $x = 1$ as follows,

$$\begin{aligned} &\min_{w,z,y} w \\ &\text{s.t} \\ &-w \leq \frac{1}{6}y \\ &-w \leq -z - \frac{7}{12}y + \frac{7}{12} \\ &-z - 2y + 1 \leq 0 \\ &-3z - y + \frac{1}{2} \leq 0 \\ &0 \leq x \leq 1 \\ &-\frac{1}{6} \leq z \leq \frac{7}{12} \\ &0 \leq y \leq 1 \end{aligned} \tag{5.11}$$

Solving problem (5.11) by using Lagrange method we can get, $(w_0, y_0, z_0) = (-\frac{1}{36}, 1, -\frac{1}{6})$; $(\lambda_{1_0}, \lambda_{2_0}, \lambda_{3_0}, \lambda_{4_0}) = (\frac{3}{5}, \frac{2}{5}, 0, \frac{2}{15})$. Computing $-M_0^{-1} \cdot N_0$ we get $-M_0^{-1} \cdot N_0 = (0.25, 0, 0.17, 0, 0, 0, 0)^T$, and then compute
$(w(x), z(x), y(x), \lambda_1(x), \lambda_2(x), \lambda_3(x), \lambda_4(x))^T$, from corollary 3.16 as:

$$\begin{bmatrix} w(x) \\ z(x) \\ y(x) \\ \lambda_1(x) \\ \lambda_2(x) \\ \lambda_3(x) \\ \lambda_4(x) \end{bmatrix} = \begin{bmatrix} -\frac{1}{36} \\ -\frac{1}{6} \\ 1 \\ \frac{3}{5} \\ \frac{2}{5} \\ 0 \\ \frac{2}{15} \end{bmatrix} - \begin{bmatrix} 0 \\ .17 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \cdot (x-1)$$

Form a set of inequalities corresponding to CR^R,
$CR^R = \begin{cases} x \leq 1 \\ 0 \leq x \leq 1 \end{cases}$ and the rest of the region is empty.

Substituting $w(x)$ and $y(x), z(x)$ into the objective of problem (5.11) and problem (5.10) respectively, we can get a lower bound $\check{Z}^{11}(x) = -0.17x$, and an upper bound $\hat{Z}^{11}(x) = -0.03$ for the global solution of problem (5.10). Compare $\check{Z}^{11}(x)$ and $\hat{Z}^{11}(x)$ in the region of x. For $x \in [0, 1]$, $\hat{Z}^{11}(x) - \check{Z}^{11}(x) \leq 0.2$, thus no fathoming in this region is performed.

Branch the feasible region along y as $\begin{cases} 0 \leq y \leq \frac{1}{2} \\ -\frac{1}{6} \leq z \leq \frac{7}{12} \end{cases}$, $\begin{cases} \frac{1}{2} \leq y \leq 1 \\ -\frac{1}{6} \leq z \leq \frac{7}{12} \end{cases}$ and underestimate the nonconvex problem (5.10) in the region, $\begin{cases} 0 \le y \le \frac{1}{2} \\ -\frac{1}{6} \le z \le \frac{7}{12} \end{cases}$ by fixing $x = 1$ as,

$$\begin{aligned} &\min_{w,z,y} w \\ &\text{s.t} \\ &-w - \frac{1}{6}y \le 0 \\ &-w + \frac{1}{2}z + \frac{7}{12}y \le \frac{7}{24} \\ &-2y - z \le -1 \\ &-y - 3z \le -\frac{1}{2} \\ &0 \le y \le \frac{1}{2} \\ &-\frac{1}{6} \le z \le \frac{7}{12} \end{aligned} \tag{5.12}$$

Solving problem (5.12) globally using KKT condition approach we get $(w_0, z_0, y_0) = (0, 0, \frac{1}{2})$; $(\lambda_{1_0}, \lambda_{2_0}, \lambda_{3_0}, \lambda_{4_0}) = (0, 1, 0.25, \frac{1}{12})$ consequently, compute $-M_0^{-1} \cdot N_0$ as $-M_0^{-1} \cdot N_0 = (0.292, 0, 0.5, 0, 0, 0, 0)^T$.

Compute $(w(x), z(x), y(x), \lambda_1(x), \lambda_2(x), \lambda_3(x), \lambda_4(x))^T$ from 3.16 as usual:

$$\begin{bmatrix} w(x) \\ z(x) \\ y(x) \\ \lambda_1(x) \\ \lambda_2(x) \\ \lambda_3(x) \\ \lambda_4(x) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{2} \\ 0 \\ 1 \\ 0.25 \\ \frac{1}{12} \end{bmatrix} + \begin{bmatrix} 0.292 \\ 0 \\ 0.5 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \cdot (x - 1)$$

Form a set of inequalities corresponding to CR^R,

$CR^R = \begin{cases} x \ge -1 \\ 0 \le x \le 1 \end{cases}$. By removing redundant constraints we can rewrite it as

$CR^R = \begin{cases} x \ge 0 \\ x \le 1 \end{cases}$. This implies that no rest of critical region to be explored. Similarly, we can obtain a parametric upper bound $\hat{Z}^{21} = 0$ and a parametric lower bound $\check{Z}^{21}(x) = 0.292x - 0.292$. Compare the lower bound with the former lower bound and update as $\check{Z}^{21}_{glb}(x) = 0.29 \cdot x - 0.29$. In the same manner compare the parametric upper bound with the former upper bound and update as, $\hat{Z}^{21}_{lub} = 0.06x - 0.02x^2 - 0.04, \forall x \in [0, 1]$. Consequently, test ε-convergence as follows, $\hat{Z}^{21}_{lub} - \check{Z}^{21}_{glb}(x) \le 0.25$ for $x \in [0, 1]$. Thus no fathoming is performed.

Here again underestimate the nonconvex problem (5.10) in the region, $\begin{cases} \frac{1}{2} \leq y \leq 1 \\ -\frac{1}{6} \leq z \leq \frac{7}{12} \end{cases}$ as follows:

$$\begin{aligned} &\min_{w,z,y} w \\ &\text{s.t} \\ w &\geq \frac{1}{2}z - \frac{1}{6}y + \frac{1}{12} \\ w &\geq z + \frac{7}{12}y - \frac{7}{12} \\ -2y - z &\leq -x \\ -y - 3z &\leq frac12x \\ \frac{1}{2} &\leq y \leq 1 \\ -\frac{1}{6} &\leq z \leq \frac{7}{12} \end{aligned} \tag{5.13}$$

And reformulate problem (5.13) as follows by fixing $x = 1$ as well,

$$\begin{aligned} &\min_{w,z,y} w \\ &\text{s.t} \\ w &\geq \frac{1}{2}z - \frac{1}{6}y + \frac{1}{12} \\ -2y - z &\leq -1 \\ -y - 3z &\leq -\frac{1}{2} \\ \frac{1}{2} &\leq y \leq 1 \\ -\frac{1}{6} &\leq z \leq \frac{7}{12} \end{aligned}$$

After solving the above problem we get the same optimal solution as problem (5.11) that is $w(x) = 0.25x - 0.54$, $z(x) = 0.05x - 0.05$ and $y(x) = -0.33x + 0.17$. This implies that no change at lower and upper bounding of the solution.

Branch and bound along z as; $\begin{cases} 0 \leq y \leq \frac{1}{2} \\ -\frac{1}{6} \leq z \leq \frac{5}{24} \end{cases}$, $\begin{cases} 0 \leq y \leq \frac{1}{2} \\ \frac{5}{24} \leq z \leq \frac{7}{12} \end{cases}$ and $\begin{cases} \frac{1}{2} \leq y \leq 1 \\ -\frac{1}{6} \leq z \leq \frac{5}{24} \end{cases}$, $\begin{cases} \frac{1}{2} \leq y \leq 1 \\ \frac{5}{24} \leq z \leq \frac{7}{12} \end{cases}$

Underestimate the original problem in the subrectangle, $\begin{cases} 0 \le y \le \frac{1}{2} \\ -\frac{1}{6} \le z \le \frac{5}{24} \end{cases}$ as:

$$
\begin{aligned}
&\min_{w,y,z} w \\
&\text{s.t} \\
&-w - \frac{1}{6}y \le 0 \\
&-w + \frac{1}{2}z + \frac{5}{24}y - \frac{5}{48} \le 0 \\
&-z - 2y \le -x \\
&-3z - y \le -\frac{1}{2}x \\
&-\frac{1}{6} \le z \le \frac{5}{24} \\
&0 \le y \le \frac{1}{2} \\
&0 \le x \le 1
\end{aligned}
\tag{5.14}
$$

Solving the above problem for a given $x = 1$ using KKT condition approach we can get, $(w_0, y_0, z_0) = (0, \frac{1}{2}, 0)$; $(\lambda_{1_0}, \lambda_{2_0}, \lambda_{3_0}, \lambda_{4_0}) = (0, 1, \frac{1}{40}, \frac{19}{120})$. Computing $-M_0^{-1} \cdot N_0$ for this problem gives us,

$$M_0^{-1} \cdot N_0 = (-\frac{17}{160}, -\frac{1}{2}, 0.002, 0, 0, 0, 0)^T,$$

Consequently, compute $(w(x), z(x), y(x), \lambda_1(x), \lambda_2(x), \lambda_3(x), \lambda_4(x))$ as:

$$
\begin{bmatrix} w(x) \\ y(x) \\ z(x) \\ \lambda_1(x) \\ \lambda_2(x) \\ \lambda_3(x) \\ \lambda_4(x) \end{bmatrix}
=
\begin{bmatrix} 0 \\ \frac{1}{2} \\ 0 \\ 0 \\ 1 \\ \frac{1}{40} \\ \frac{19}{120} \end{bmatrix}
-
\begin{bmatrix} -\frac{17}{160} \\ -0.5 \\ 0.002 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}
\cdot (x - 1)
$$

Form a set of inequalities corresponding to CR^R,

$CR^R = \begin{cases} x \ge 0 \\ x \le 1 \end{cases}$ and the rest of the region, CR^{Rest} is empty. Thus, $w(x) = \frac{17}{160} \cdot x - \frac{17}{160}$, $z(x) = -0.002 \cdot x + 0.002$, $y = \frac{1}{2} \cdot x$, $\forall x \in [0, 1]$ are solutions of the underestimated problem. Substituting w and y, z into the objective of problem (5.14) and problem (5.10) respectively, we can get a lower bound and an upper bound $\check{Z}^{31}(x) = \frac{17}{160} \cdot x - \frac{17}{160}$, and $\hat{Z}^{31}(x) = -0.001 \cdot x^2 + 0.001 \cdot x$, for the global solution of problem (5.10) respectively. Update the lower bound as, $\check{Z}^{31}_{glb} = 0.1 \cdot x - 0.1$. Similarly, Update the parametric upper bound $\hat{Z}^{31}_{lub} = 0.06 \cdot x - 0.02 \cdot x^2 - 0.04$. Test the convergence by selecting a small positive tolerance. But $\hat{Z}^{31}_{lub} - \check{Z}^{31}_{glb} \le 0.06$, Thus no fathom.

Similarly, underestimate the original problem in the region $\begin{cases} 0 \le y \le \frac{1}{2} \\ \frac{5}{24} \le z \le \frac{7}{12} \end{cases}$ and formulate as follows:

$$
\begin{aligned}
\min_{w,z,y} w& \\
\text{s.t}& \\
-w + \frac{5}{24}y \le 0& \\
-w + \frac{7}{12}y + \frac{1}{2}z - \frac{7}{24} \le 0& \\
-2y - z + x \le 0& \\
-y - 3z + \frac{1}{2}x \le 0& \\
0 \le y \le \frac{1}{2}& \\
\frac{5}{24} \le z \le \frac{7}{12}& \\
0 \le x \le 1&
\end{aligned} \tag{5.15}
$$

Solving the above problem by fixing $x = 1$ using lagrange method, we can get the following result; $(w_0, y_0, z_0) = (\frac{5}{72}, \frac{1}{3}, \frac{1}{3})$ and $(\lambda_{1_0}, \lambda_{2_0}, \lambda_{3_0}, \lambda_{4_0}) = (\frac{2}{3}, \frac{1}{3}, \frac{1}{6}, 0)$. Compute ${M_0}^{-1} \cdot N_0$ as ${M_0}^{-1} \cdot N_0 = (0.2, 0.8, 0.6, 0, 0, 0, 0)^T$. Calculate $(w(x), y(x), z(x), \lambda_1, \lambda_2, \lambda_3, \lambda_4)^T$, from Corollary 3.16 as

$$
\begin{bmatrix} w(x) \\ y(x) \\ z(x) \\ \lambda_1(x) \\ \lambda_2(x) \\ \lambda_3(x) \\ \lambda_4(x) \end{bmatrix} = \begin{bmatrix} \frac{5}{72} \\ \frac{1}{3} \\ \frac{1}{3} \\ \frac{2}{3} \\ \frac{1}{3} \\ \frac{1}{6} \\ 0 \end{bmatrix} - \begin{bmatrix} -\frac{1}{6} \\ -0.8 \\ 0.6 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \cdot (x-1)
$$

Form a set of inequalities by removing redundant inequalities as $CR^R = \begin{cases} x \ge 0 \\ x \le 0 \end{cases}$, this implies that the solution is optimal $\forall x \in [0, 1]$. Now again update the current parametric lower bound as $\check{Z}^{32}_{glb} = 0.2 \cdot x - 0.1$ and the current parametric upper bound as $\hat{Z}^{32}_{lub} = 0.6 \cdot x - 0.5 \cdot x^2 - 0.15$ for $x \in [0, 0.2] \bigcup [0.8, 1]$ and $\hat{Z}^{32}_{lub} = 0.06 \cdot x - 0.02 \cdot x^2 - 0.04$ for $x \in [0.2, 0.8]$. Test ε-convergence as $\hat{Z}^{32}_{lub} - \check{Z}^{32}_{glb} \le 0.01, \forall x \in (0.1, 0.2]$ and fathom $\forall x \in [0, 0.1] \bigcup [0.8, 1]$ Thus, ε-convergence is satisfied for $x \in [0, 0.2] \bigcup [0.8, 1]$. But, $\hat{Z}^{32}_{lub} - \check{Z}^{32}_{glb} \le 0.03\ \forall x \in [0.2, 0.8]$, thus no fathoming in this region of x.

Underestimate the original problem in the subrectangle $\begin{cases} \frac{1}{2} \leq y \leq 1 \\ -\frac{1}{6} \leq z \leq \frac{5}{24} \end{cases}$, as follows:

$$\begin{aligned} &\min_{w,z,y} w \\ &\text{s.t} \\ &-w + \frac{1}{2}z - \frac{1}{6}y + \frac{1}{12} \leq 0 \\ &-w + \frac{5}{24}y + z - \frac{7}{24} \leq 0 \\ &-2y - z + x \leq 0 \\ &-y - 3z + \frac{1}{2}x \leq 0 \\ &\frac{1}{2} \leq y \leq 1 \\ &-\frac{1}{6} \leq z \leq \frac{5}{24} \\ &0 \leq x \leq 1 \end{aligned} \tag{5.16}$$

But, there doesn't exist Lagrange multipliers that satisfy the KKT conditions.

Underestimate the original problem in the region $\begin{cases} \frac{1}{2} \leq y \leq 1 \\ \frac{5}{24} \leq z \leq \frac{7}{12} \end{cases}$, but the problem is infeasible in this region.

Branch and bound the previous subrectangles along y as:
$\begin{cases} 0 \leq y \leq \frac{1}{4} \\ -\frac{1}{6} \leq z \leq \frac{5}{24} \end{cases}$, $\begin{cases} \frac{1}{4} \leq y \leq \frac{1}{2} \\ -\frac{1}{6} \leq z \leq \frac{5}{24} \end{cases}$ and $\begin{cases} 0 \leq y \leq \frac{1}{4} \\ \frac{5}{24} \leq z \leq \frac{7}{12} \end{cases}$, $\begin{cases} \frac{1}{4} \leq y \leq \frac{1}{2} \\ \frac{5}{24} \leq z \leq \frac{7}{12} \end{cases}$

Relax the original problem in the rectangle $\begin{cases} 0 \leq y \leq \frac{1}{4} \\ -\frac{1}{6} \leq z \leq \frac{5}{24} \end{cases}$ and reformulate as follows:

$$\begin{aligned} &\min_{w,z,y} w \\ &\text{S.t} \\ &-w - \frac{1}{6}y \leq 0 \\ &-w + \frac{1}{4}z + \frac{5}{24}y - \frac{5}{96} \leq 0 \\ &-2y - z + x \leq 0 \\ &-y - 3z + \frac{1}{2}x \leq 0 \\ &0 \leq y \leq \frac{1}{4} \\ &-\frac{1}{6} \leq z \leq \frac{5}{24} \\ &0 \leq x \leq 1 \end{aligned} \tag{5.17}$$

But, this problem is infeasible for $x \in (\frac{1}{2}, 1]$, as a result solving the above problem for $x \in [0, \frac{1}{2}]$ by fixing $x = \frac{1}{4}$, we get $(w_0, y_0, z_0) = (-\frac{5}{192}, \frac{1}{8}, 0)$ and $(\lambda_{1_0}, \lambda_{2_0}, \lambda_{3_0}, \lambda_4) = (0, 1, \frac{3}{40}, \frac{7}{120})$. Compute $M_0^{-1} \cdot N_0$ from Basic Sensitivity Theorem as $M_0^{-1} \cdot N_0 = (0.2, -0.1, 0, 0, 0, 0, 0)^T$ and calculate $(w(x), y(x), z(x), \lambda_1(x), \lambda_2(x), \lambda_3(x), \lambda_4(x))^T$, from Corollary 3.16 as

$$\begin{bmatrix} w(x) \\ y(x) \\ z(x) \\ \lambda_1(x) \\ \lambda_2(x) \\ \lambda_3(x) \\ \lambda_4(x) \end{bmatrix} = \begin{bmatrix} -\frac{5}{192} \\ \frac{1}{8} \\ 0 \\ 0 \\ 1 \\ \frac{3}{40} \\ \frac{7}{120} \end{bmatrix} - \begin{bmatrix} -0.03 \\ -0.1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \cdot (x - \frac{1}{4})$$

Form a set of inequalities by removing redundant inequalities as $CR^R = \begin{cases} x \leq 1 \\ x \geq 0 \end{cases}$, this implies that the solution is optimal throughout the varying parameter space. Obtain the parametric upper and lower bounds for the above problem and update the upper bound as $\hat{Z}_{lub}^{41} = 0.06 \cdot x - 0.02 \cdot x^2 - 0.04$ for $x \in [0.2, 0.8]$ and a lower bound as $\check{Z}_{glb}^{41} = 0.03 \cdot x - 0.04$. As a result we have that $\hat{Z}_{lub}^{41} - \check{Z}_{glb}^{41} \leq 0.01$ for $x \in [0.2, 0.8]$. Thus, ε-convergence is satisfied for $\varepsilon = 0.01$.

Incorporate the solutions $z(x) = 0$ and $y(x) = 0.1x - 0.1$ for $x \in [0.2, 0.8]$ into the first level and reformulate as follows:

$$\begin{gathered} \min_x f_1 = -2x + y \\ \text{s.t} \\ \frac{1}{2}x + y \leq 0 \\ 0.2 \leq x \leq 0.8 \end{gathered}$$

Solving the above problem one can get $\lambda_0 = 5$, $x = \frac{1}{4}$, $y = 0.14$ and $z = 0$. Similarly, incorporate $y = \frac{5}{4}x - \frac{1}{2}$ and $z(x) = 0.6x - 0.3$ into the first optimization problem for $x \in (0.1, 0.2]$, and can be reformulated as follows:

$$\begin{gathered} \min_x f_1 = -\frac{3}{4}x - \frac{1}{2} \\ \text{s.t} \\ \frac{7}{4}x - \frac{1}{2} \leq 0 \\ 0.1 \leq x \leq 0.2 \end{gathered}$$

Solving the above problem we can get $x = \frac{2}{5}$, but it is not feasible. Since $\frac{1}{5} \notin [0.1, 0.2]$

Therefore we have an optimal solution; $x = \frac{1}{4}$, $y = 0.14$ and $z = 0$

5.2 Trilevel Programming with Nonconvexity Formulation in the Inner most Problem

Consider the following trilevel programming problem, with nonconvexity formulation in the most inner problem:

$$
\begin{aligned}
&\min_x f_1(x,y,z) \\
&\text{s.t.} \qquad\qquad (1^{st}\text{ level}) \\
&\quad g_1(x,y,z) \leq 0 \\
&\quad \min_y f_2(x,y,z) \\
&\quad \text{s.t.} \qquad\qquad (2^{nd}\text{ level}) \\
&\qquad g_2(x,y,z) \leq 0 \\
&\qquad \min_z f_3(x,y,z) \\
&\qquad \text{s.t.} \qquad\quad (3^{rd}\text{ level}) \\
&\qquad g_3(x,y,z) \leq 0 \\
&\qquad\qquad x^L \leq x \leq x^U \\
&\qquad\qquad y^L \leq y \leq y^U \\
&\qquad\qquad z^L \leq z \leq z^U
\end{aligned} \tag{5.18}
$$

Assume that the third level of problem (5.18)is a nonconvex programming problem. Reformulate this problem as a multiparametric nonconvex programming problem by taking x and y as parameters:

$$
\begin{aligned}
Z(x,y) = &\min_z f_3(x,y,z) \\
&\text{S.t.} \\
&\quad g_3(x,y,z) \leq 0 \\
&\quad x^L \leq x \leq x^U \\
&\quad y^L \leq y \leq y^U \\
&\quad z^L \leq z \leq z^U
\end{aligned} \tag{5.19}
$$

In order to solve problem (5.19) one can categorize each nonconvex terms of special structure (e.g. bilinear, trilinear, fractional univariate concave and etc.), and nonconvex of generic structure. But for a time being we analyze the problem with generic nonconvex and bilinear formulation. Based on this partitioning of different terms appearing in the objective function and constraints, formulation (5.19) is rewritten equivalently as follows:

$$Z(x,y) = \min_z \{cf(x,y,z) + Nf(x,y,z) + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} b_{ij} z_i z_j\}$$

$$\text{S.t}$$

$$cf^p(x,y,z) + Nf^p(x,y,z) + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} b^p_{ij} z_i z_j \leq 0$$

$$x^L \leq x \leq x^U$$

$$y^L \leq y \leq y^U$$

$$z^L \leq z \leq z^U \tag{5.20}$$

where, $cf(x,y,z)$ is the convex part of the objective function;
$Nf(x,y,z)$ is the generic nonconvex terms appearing in the objective function and $b_{ij}z_iz_j$'s are bilinear terms appearing in the objective function. Similarly, for each constraint g, there is a convex part $cf^p(x,y,z)$, generic nonconvex terms $Nf^p(x,y,z)$ and the bilinear terms $b^p_{ij}z_iz_j$.

Underestimate each nonconvex terms in problem (5.20) with a tight convex underestimator and can be constructed by replacing each of the generic nonconvex terms and each bilinear terms with its tighter convex lower bounding functions as discussed in Section 4. Thus, problem (5.20) can be relaxed as follows:

$$Z^1(x,y) = \min_z \{cf(x,y,z) + Nf(x,y,z) +$$

$$\sum_{k=1}^{K} \alpha_k(z^L, z^U)(z_i^L - z_i)(z_i^U - z_i) + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} b_{ij} w_{ij}\}$$

$$\text{S.t.}$$

$$cf^p(x,y,z) + Nf^p(x,y,z) + \sum_{k=1}^{K} \alpha^p_k(z^L, z^U)(z_i^L - z_i)(z_i^U - z_i)$$

$$+ \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} b^p_{ij} w_{ij} \leq 0, \; p = 1, 2, \ldots m$$

$$w_{ij} \geq z_i^L z_j + z_j^L z_i - z_i^L z_j^L$$

$$w_{ij} \geq z_i^U z_j + z_j^U z_i - z_i^U z_j^U$$

$$x^L \leq x \leq x^U$$

$$y^L \leq y \leq y^U$$

$$z^L \leq z \leq z^U \tag{5.21}$$

Solving the underestimated problem (5.21) above using suitable multiparametric programming algorithm [36], [38], resulting in:

$$z^k = m_3^k + P_1^k \cdot x + P_2^k \cdot y,\ \ H_{31}^k \cdot x + H_{32} \cdot y \leq h_3^k$$
$$w_{ij}^k = O_3^k + L_1^k \cdot x + L_2^k \cdot y,\ D_{31}^k \cdot x + D_{32}^k \cdot y \leq d_3^k \tag{5.22}$$

where, $k = 1, 2, \ldots, K_2$, with K_2 being the number of critical regions and consequently, the number of linear approximations done on the optimal rational reaction set $\Psi_2(x, y)$ (see Corollary 3.16). And substituting each expressions in (5.22) into the objective function of the underestimated problem (5.21) one can obtain a multiparametric lower bound $\check{Z}_k^{11}(x, y)$ for the solution of problem (5.20) within the corresponding critical regions CR^k.

On the other hand the upper bound $\hat{Z}_k^{11}(x, y)$ for the solution of problem (5.20) can be obtained by substituting each expressions in (5.22) into the original nonconvex problem (5.20).

Consequently, ε-convergence test is performed in the region of (x, y), as $\hat{Z}_k^{11}(x, y) - \check{Z}_k^{11}(x, y) \leq \max_{x,y}\{\hat{Z}_k^{11}(x, y) - \check{Z}_k^{11}(x, y)\} \leq \varepsilon$ where, ε is a small positive tolerance are fathomed. If each $\check{Z}_k^{11}(x, y)$ are within ε of $\hat{Z}_k^{11}(x, y)$ the algorithm terminates and the expressions in (5.22) can then be incorporated into the second optimization level of (5.18). Note that since the expressions in (5.22) are piecewise linear functions of x and y , the complexity of the original problem does not increase. Hence, the second level can be reformulated as the following K_2 optimization problems:

$$\begin{aligned} Z(x) = \min_y & f_2(x, y, z(x, y)) \\ & \text{S.t.} \\ & g_2(x, y, z(x, y)) \leq 0 \\ & y \in D_f \\ D_f = \{y \in Y : & \exists z \in Z : g_3(x, y, z) \leq 0\} \\ & x^L \leq x \leq x^U \\ & y^L \leq y \leq y^U \\ & z^L \leq z \leq z^U \end{aligned} \tag{5.23}$$

One can thus proceed with optimization levels 1 and 2. Following this procedure, trilevel optimization problems in (5.18) result in K_1 single level convex optimization problems:

$$
\begin{aligned}
&\min_x f_1(x, y(x), z(x,y)) \\
&\text{S.t.} \\
&g_1(x, y(x), z(x,y)) \leq 0 \\
&x \in C_f \\
C_f = \{x \in X, \exists (y,z) \in Y \times Z : g_2(x,y,z) \leq 0, g_3(x,y,z) \leq 0\} \\
&x^L \leq x \leq x^U \\
&y^L \leq y \leq y^U \\
&z^L \leq z \leq z^U
\end{aligned} \tag{5.24}
$$

Solve the K_1 final optimization problems (5.24) to the global optimal solutions. Note that the number of K_1 single level optimization problems directly depend on the solutions obtained in each optimization levels within the corresponding critical regions [20].

Otherwise, divide the space of z along the longest side and underestimate each nonconvex terms in problem (5.20) in each subregions of z, and reformulate the original problem as described below:

$$
\begin{aligned}
Z^{21}(x,y) = \min\{cf(x,y,z) + Nf(x,y,z) + \\
\sum_{k=1}^{K} \alpha_k(z^L_{new}, z^U)(z^L_{inew} - z_i)(z_i^U - z_i) + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} b_{ij}w_{ij}\} \\
\text{S.t.} \\
cf^p(x,y,z) + Nf^p(x,y,z) + \sum_{k=1}^{K} \alpha_k^p(z^L_{new}, z^U)(z^L_{inew} - z_i)(z_i^U - z_i) \\
+ \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} b^p_{ij}w_{ij} \leq 0 \; p = 1, 2, \ldots m \\
w_{ij} \geq z^L_{inew} z_j + z^L_{jnew} z_i - z^L_{inew} z^L_{jnew} \\
w_{ij} \geq z_i^U z_j + z_j^U z_i - z_i^U z_j^U \\
x^L \leq x \leq x^U \\
y^L \leq y \leq y^U \\
z^L_{new} \leq z \leq z^U
\end{aligned} \tag{5.25}
$$

and,

$$Z^{22}(x,y) = \min\{cf(x,y,z) + Nf(x,y,z) + \sum_{k=1}^{K} \alpha_k(z^L, z^U_{new})$$
$$(z_i^L - z_i)(z^U_{inew} - z_i) + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} b_{ij}w_{ij}\}$$
s.t.
$$cf^p(x,y,z) + Nf^p(x,y,z) + \sum_{k=1}^{K} \alpha_k^p(z^L, z^U_{new})(z_i^L - z_i)(z^U_{inew} - z_i)$$
$$+\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} b^p_{ij}w_{ij} \leq 0 \; p = 1,2,\ldots m$$
$$w_{ij} \geq z_i^L z_j + z_j^L z_i - z_i^L z_j^L$$
$$w_{ij} \geq z^U_{inew} z_j + z^U_{jnew} z_i - z^U_{inew} z^U_{jnew}$$
$$x^L \leq x \leq x^U$$
$$y^L \leq y \leq y^U$$
$$z^L \leq z \leq z^U_{new} \tag{5.26}$$

Solve each of problems (5.25) and (5.26) using the application of multiparametric programming approach resulting in linear functions of x and y in the same way as expression (5.22) as well, and substitute each parametric solutions of problem (5.25) and (5.26) into problem (5.25) and (5.26) to get a parametric lower bound $\check{Z}_k^{21}(x,y)$ and $\check{Z}_k^{22}(x,y)$ respectively, for the solution of problem (5.20). Similarly, we can get a parametric upper bound $\hat{Z}_k^{21}(x,y)$ and $\hat{Z}_k^{22}(x,y)$ for the global solution of problem (5.20) by substituting each solutions of problems (5.25) and (5.26) into the original problem (5.20) respectively.

Compare each parametric upper bound with the former upper bound and choose the least upper bound as well. In the same manner compare each parametric lower bounds with the former lower bound and take the greatest lower bond. Consequently, ε-convergence test is performed as usual.

Continue in this manner until a lower and an upper bound of the solution come together inside a small positive tolerance ε. Following this procedure, trilevel optimization problems in (5.18) result in K_1-single level optimization problem as folows:

$$\begin{aligned}
&\min_y f_1(x, y(x), z(x, y)) \\
&\text{s.t.} \\
&g_1(x, y(x), z(x, y)) \leq 0 \\
&x \in C_f \\
&C_f = \{x \in X, \exists (y, z) \in Y \times Z : g_2(x, y, z) \leq 0, g_3(x, y, z) \leq 0\} \\
&x^L \leq x \leq x^U \\
&y^L \leq y \leq y^U \\
&z^L \leq z \leq z^U
\end{aligned} \tag{5.27}$$

The number of K_1-optimization problem (5.27) depends on the number of critical regions obtained in the inners problem. Based upon the above developments an algorithm for the solution of (5.18) is presented in Table 5.

5.2.1 Mathematical explanation for the algorithm

Consider the following trilevel programming problem:

$$\begin{aligned}
&\min_x f(x, y, z) \\
&\text{S.t} \\
&g_1(x, y, z) \leq 0 \\
&\quad \text{where } [y, z] \text{ solves} \\
&\min_y f_2(x, y, z) \\
&\text{S.t} \\
&g_2(x, y, z) \leq 0 \\
&\quad \text{where } z \text{ solves} \\
&\min_z f_3(x, y, z) \\
&\text{S.t} \\
&g_3(x, y, z) \leq 0
\end{aligned} \tag{5.28}$$

From Section 2 , we can define the following definitions of sets for the third-level of optimization problem (5.28):

- feasible set for the third level,

$$\Omega_2(x, y) = \{z \in Z : g_2(x, y, z) \leq 0\, k = 1, 2, \ldots, K\} \tag{5.29}$$

- rational reaction set for the third level,

Table 5: Global parametric programming approach algorithm for trilevel programming problem

Step 1	Recast the third level of the optimization problem as a multiparametric programming problem with parameters being the upper levels optimization variables x and y
Step 2	Initialize the current upper bound as $\hat{Z} = \infty$, a region of (x, y), CR, and a space of continues variable z determined by the lower and upper bounds z^L and z^U, respectively and tolerance ε
Step 3	For a given region of (x, y), CR and the corresponding space of z formulate and solve the relaxed convex underestimator optimization problem using suitable multiparametric programming algorithm and obtain the parametric lower bound $\check{Z}(x, y)$ and upper bound $\hat{Z}(x, y)$ for the solution of the nonconvex problem by using the methods described in Section 4
Step 4	Check the difference between $\hat{Z}(x)$ and $\check{Z}(x)$ by solving the following maximization problem, $\max_{x,y}\{\hat{Z}(x,y) - \check{Z}(x,y) : [x,y]^T \in CR^R\}$. If $\hat{Z}(x,y) - \check{Z}(x,y) \leq \max_{x,y}\{\hat{Z}(x,y) - \check{Z}(x,y) \leq \varepsilon$ and the relaxed convex optimization problems infeasible for some rectangle fathom those regions and the corresponding region of parameters.
Step 5	If no more space of z to explore, terminate, and go to the next step. Otherwise, go to Step 11
Step 6	Substitute each of the solutions into the second level optimization problem and formulate the corresponding multiparametric programming problems with the variables from the leader optimization variables being the parameter. If the second optimization problem is nonconvex refer the algorithm for bilevel optimization problem. Else,
Step 7	Solve the resulting problem using parametric programming algorithm
Step 8	Substitute each of the solutions into the leader's problem and formulate one-level optimization problem
Step 9	Solve each K_1 problems
Step 10	Compare the leader's optimal solutions and select the best one. Otherwise, go to the next step.
Step 11	Branch and bound on y and formulate a convex underestimator in each subrectangles of y
Step 12	Solve the convex underestimator problems in each subrectangles using multiparametric programming approach and obtain the lower and upper bounds for the solution in each subrectangle.
Step 13	Compare the former lower bound with the lower bounds of each subrectangle within the corresponding critical regions and take the maximum within the corresponding critical region.
Step 14	Compare the former upper bound with the upper bounds of each subrectangle within the corresponding critical regions and take the minimum within the corresponding critical region and go to Step (4)

$$\Psi_2(x,y) = \{z \in Z : z \in \arg\min\{f_3(x,y,z) : z \in \Omega_2\}\} \tag{5.30}$$

Consider the third level programming problem with x and y are taken as parameters:

$$\begin{aligned} &\min_z f_3(x,y,z) \\ &\quad \text{S.t} \\ &g_2(x,y,z) \leq 0 \end{aligned} \tag{5.31}$$

Let us recall, the result from Fiacco [30], under the assumption of Basic Sensitivity Theorem a first order approximation of $[z(x,y), \lambda_k(x,y)]$, for k=1,2,...,q in a neighborhood of (x_0, y_0) is,

$$\begin{bmatrix} z(x,y) \\ \lambda_k(x,y) \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ \lambda_{k0} \end{bmatrix} - (M_0)^{-1} \cdot N_0 \cdot \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \tag{5.32}$$

Consider equation (5.32) as the optimal solution of problem (5.31) for a given (x_0, y_0). The space of parameters where this solution (5.32) remains optimal is defined as the critical region and obtained by using feasibility and optimality conditions [19]. Feasibility is ensured by substituting $z(x,y)$ into the inactive inequalities within problem (5.31), whereas, the optimality condition is given by $\check{\lambda}_Q(x,y) \geq 0$, where $Q \in \{1,2,\ldots,q\}$ corresponds to the vector of active inequalities [19], [41]. In short, equation (5.32) is the optimal solution of problem (5.31) in the region:

$$CR = \left\{ \begin{array}{l} \breve{g}_Q(z(x,y),x,y) \leq 0 \\ \check{\lambda}_Q(x,y) \geq 0 \end{array} \right. \tag{5.33}$$

From the multiparametric inequalities thus obtained in (5.33), the redundant inequalities are removed and a compact representation of CR is obtained as follows:
$CR^Q = \Delta\{CR\}$
where Δ is an operator which removes redundant constraints, (for the detail of this procedure see appendix A of [20]). Once, CR^Q has been obtained for a solution $[z(x,y), \lambda_k(x,y)]$ the next step is to define the rest of the regions of the parameter space, CR^{Rest} (if any) as:

$$CR^{Rest} = CR^{IG} - CR^Q \tag{5.34}$$

where CR^{IG} is the initial given parameter space. In each of those regions, another set of parametric solutions of problem (5.31) are then obtained and corresponding critical regions are calculated as in [20], (for more detail see Appendex B of [20]). Consequently, the explicit solutions are given by piecewise linear functions corresponding to critical regions [18]:

$$z = \begin{cases} c^1 + k^1 \cdot \begin{bmatrix} x \\ y \end{bmatrix}, (x,y) \in CR^1 \\ c^2 + k^2 \cdot \begin{bmatrix} x \\ y \end{bmatrix}, (x,y) \in CR^2 \\ \vdots \\ c^L + k^L \cdot \begin{bmatrix} x \\ y \end{bmatrix}, (x,y) \in CR^L \end{cases} \tag{5.35}$$

where, c^i are column vectors, and $k^i, i = 1, 2, \ldots, L$, are either real matrices or row vectors. The algorithm stops when there are no more regions to be explored. In other words, the algorithm terminates when the solution of the differential equation of Basic Sensitivity Theorem [22] has been fully approximated by first order expansions [20]. Otherwise, define the rest of the parameter space and solve the problem in that region and determine the region where this solution remains optimal as well.

Once, all optimal solutions with the corresponding critical regions are obtained as the expression in (5.35), can then be incorporated in the second optimization level of problem (5.28). Note that, at the second level there may be L different problems, each problem can then be recast as parametric programming problem where x is considered as parameters. One can solve each problem for the given feasible x_0 in the same procedure as the third level problem solved provided that the solution is a function of x. Incorporate each parametric solutions of the second optimization level problem with the corresponding critical regions into the first optimization problem of (5.35). Thus, the trilevel programming problem in (5.35) results in K_1-single level optimization problem. After solving each K_1-single level optimization problem compare the value of the objective function and choose the most appropriate value as one need.

5.2.2 Illustrative example

Consider the following trilevel programming problem with the inclusion of nonconvex term in the third level:

$$
\begin{aligned}
&\min_x f_1(x,y,z) = -x + 4.y \\
&\quad \text{s.t} \\
&\quad x + y \le 1 \\
&\quad \min_y f_2(x,y,z) = 2y + z \\
&\quad\quad \text{s.t} \\
&\quad\quad -2x + y \le -z \\
&\quad\quad \min_z f_3(x,y,z) = -z^2 + y \\
&\quad\quad\quad \text{s.t} \\
&\quad\quad\quad z \le x \\
&\quad\quad\quad 0 \le y, z \le 1 \\
&\quad\quad\quad 0 \le x \le \frac{1}{2}
\end{aligned} \tag{5.36}
$$

Initialize $\varepsilon = 0.002$ and reformulate the third-level optimization problem as a multiparametric programming problem, with parameters being x and y

$$
\begin{aligned}
&\min_z f_3(x,y,z) = -z^2 + y \\
&\quad \text{s.t} \\
&\quad z \le x \\
&\quad 0 \le y, z \le 1 \\
&\quad 0 \le x \le \frac{1}{2}
\end{aligned} \tag{5.37}
$$

Underestimate problem (5.37) as:

$$
\begin{aligned}
&Z(x,y) = \min -z + y \\
&\quad \text{s.t} \\
&\quad z \le x \\
&\quad 0 \le y, z \le 1 \\
&\quad 0 \le x \le \frac{1}{2}
\end{aligned} \tag{5.38}
$$

Solve the resulting problem using a multiparametric optimization algorithm [38]:

$CR = \begin{cases} z = x \\ 0 \le y \le 1 \\ 0 \le x \le \frac{1}{2} \end{cases}$, and perform an upper bound $\hat{Z}(x,y) = -x^2 + y$ and a lower bound $\check{Z}(x,y) = -x + y$ for the global solution of problem (5.37).

Compare the upper bound and the lower bound with in a maximum tolerance but, $\hat{Z}(x,y) - \check{Z}(x,y) \leq 0.25$, implies that no fathoming in the space of z is performed.

Branch and Bound on z as: $0 \leq z \leq \frac{1}{2}$ and $\frac{1}{2} \leq z \leq 1$ and underestimate the original problem on $0 \leq z \leq \frac{1}{2}$,

$$\begin{aligned} Z(x,y) &= \min -\frac{1}{2}z + y \\ \text{s.t} & \\ z &\leq x \\ 0 \leq\ & y \leq 1 \\ 0 \leq\ & x, z \leq \frac{1}{2} \end{aligned} \tag{5.39}$$

Solve the resulting problem using a multiparametric optimization algorithm:

$CR = \begin{cases} z = x \\ 0 \leq y \leq 1 \\ 0 \leq x \leq \frac{1}{2} \end{cases}$. Compare the lower bound $\check{Z}(x,y) = -x + y$ and an upper bound $\hat{Z}(x,y) = x^2 + y$ with in positive tolerance as: $\hat{Z}(x,y) - \check{Z}(x,y) \leq 0.06$, this implies that no fathoming in the region $z \in [0, \frac{1}{2}]$.

Underestimate the original problem (5.37) on $\frac{1}{2} \leq z \leq 1$ as,

$$\begin{aligned} Z(x,y) &= \min -\frac{3}{2}z + y \\ \text{s.t} & \\ z &\leq x \\ 0 \leq\ & y \leq 1 \\ 0 \leq\ & x \leq \frac{1}{2} \\ \frac{1}{2} \leq\ & z \leq 1 \end{aligned} \tag{5.40}$$

But problem (5.40) is infeasible for the region of the parameter space $(x, y) \in [0, \frac{1}{2}] \times [1, 1]$ and fathom the region of z and the corresponding critical region.

Branch and Bound again on z as: $0 \leq z \leq \frac{1}{4}$ and $\frac{1}{4} \leq z \leq \frac{1}{2}$. And underestimate the original problem (5.37) on $z \in [0, \frac{1}{4}]$ as,

$$\begin{aligned} Z(x,y) &= \min -\frac{1}{4}z + y \\ \text{s.t} & \\ z &\leq x \\ 0 \leq\ & y \leq 1 \\ 0 \leq\ & x \leq \frac{1}{2} \\ 0 \leq\ & z \leq \frac{1}{4} \end{aligned} \tag{5.41}$$

Solve the resulting problem using a multiparametric optimization algorithm,

$CR = \begin{cases} z = x \\ 0 \le y \le 1 \\ 0 \le x \le \frac{1}{4} \end{cases}$, and problem (5.41) is infeasible in region of $(x, y) \in (\frac{1}{4}, \frac{1}{2}] \times [0, 1]$ and the problem fathom in this region as well.

Compare the upper bound and the lower bound in the parameter space of $(x, y) \in [0, \frac{1}{4}]$ as, $\hat{Z}(x, y) - \check{Z}(x, y) \le 0.015$ and again no fathoming in this region.

Underestimate the original problem (5.37) on $\frac{1}{4} \le z \le \frac{1}{2}$ as:

$$
\begin{aligned}
Z(x, y) = \min & -\frac{3}{4}z + y \\
\text{s.t} & \\
& z \le x \\
& 0 \le y \le 1 \\
& 0 \le x \le \frac{1}{2} \\
& \frac{1}{4} \le z \le \frac{1}{2}
\end{aligned} \tag{5.42}
$$

Solve the resulting problem using a multiparametric optimization algorithm,

$CR = \begin{cases} z = x \\ 0 \le y \le 1 \\ 0 \le x \le \frac{1}{4} \end{cases}$, and problem (5.42) is infeasible in the region of $(x, y) \in [0, \frac{1}{4}) \times [0, 1]$, so fathoming is performed in this region. And compare the upper bound with the lower bound for the global solution of problem (5.37) as $\hat{Z}(x, y) - \check{Z}(x, y) \le 0.015$

Branch and Bound on z as: $0 \le z \le \frac{1}{8}$, $\frac{1}{8} \le z \le 1$ and underestimate the nonconvex problem (5.37) on $0 \le z \le \frac{1}{8}$ as;

$$
\begin{aligned}
Z(x, y) = \min & -\frac{3}{4}z + y \\
\text{s.t} & \\
& z \le x \\
& 0 \le y \le 1 \\
& 0 \le x \le \frac{1}{2} \\
& 0 \le z \le \frac{1}{8}
\end{aligned} \tag{5.43}
$$

Solve the resulting problem using a multiparametric optimization algorithm,

$CR = \begin{cases} z = x \\ 0 \le y \le 1 \\ 0 \le x \le \frac{1}{8} \end{cases}$, and problem (5.43) is infeasible in region of $(x, y) \in (\frac{1}{8}, \frac{1}{2}] \times [0, 1]$, so fathoming is performed in this region. And compare the upper bound with the lower bound for the global solution of problem (5.37) as $\hat{Z}(x,y) - \check{Z}(x,y) \leq 0.002$. Thus, the upper bound and the lower bounds are with in a specified tolerance $\varepsilon = 0.002$ in the space of (x, y) and all the space of (x, y) are explored then the algorithm terminates here.

Incorporate the rational reaction set $CR = \begin{cases} z = x \\ 0 \leq y \leq 1 \\ 0 \leq x \leq \frac{1}{2} \end{cases}$, into the 2^{nd}-level optimization problem:

$$\min_{y} f_2(x,y,z) = 2y + x$$
$$\text{s.t}$$
$$-2x + y \leq -x$$
$$0 \leq y \leq 1$$
$$0 \leq x \leq \frac{1}{2} \tag{5.44}$$

Solve the resulting problem using multiparametric linear programming algorithm [20]

$$CR = \begin{cases} y(x) = x \\ \lambda = \frac{1}{2} > 0 \\ 0 \leq x \leq \frac{1}{2} \end{cases}.$$

Incorporate the above rational reaction set into leader's problem

$$\min_{x} f_1(x,y,z) = -\frac{1}{2}x$$
$$\text{s.t}$$
$$2x \leq 1$$
$$0 \leq x \leq \frac{1}{2} \tag{5.45}$$

Solve the resulting optimization problem using KKT condition and the solution is given as $x = \frac{1}{2}$, $y = 1$ and $z = 1$.

6 Conclusion and recommendation

In a hierarchical organization interactive decision makers exist within a predominantly hierarchical structure and execution of decisions is sequential from the top level to the bottom level. Each decision-maker independently minimizes its own objectives, but is affected by the action of other decision-makers at various levels through externalities.

Multilevel programming problems belong to global optimization problems. Even when the involved functions are linear, the resulted problem will be nonconvex because of its hierarchical nature [43], [34]. Especially, when the inner problem is nonconvex the nonconvexity of the resulting problem is too high and there is no general method of solving it yet.

Recently, a method through a multiparametric programming approach is developed to solve multilevel convex problems, which made it possible for the development of a unified strategy for their solution to global optimality. The main idea is to divide the followers parameter space into different rational reaction sets, and search for the global optimum of a convex programming problem in each area for the higher levels.

The difficulty and complexity of the solution approach for the multilevel optimization is easily confirmed by looking at what might be considered its simplest version, the linear multilevel optimization. For simplicity assume that the initial parameter space is unbounded. Thus, let the critical region be defined by p constraints, then the rest of the critical regions consist of p convex polyhedra CR^R defined by at most p inequalities. A major difficulty with nonconvex formulation in the inner most problem is that the global optimal solution cannot be efficiently computed and the behavior of a local solution is hard to analyze for the inner most problem. Further, it may not satisfy one of the assuptions in Basic Sensitivity Theorem and the critical regions may not be convex polyhedra in general. Thus, solving multilevel optimization is very difficult.

The key advantage of using multiparametric programming approach, however, is that the optimal solution for the inner problem is obtained as a function of varying upper level optimization variables without exhaustively enumerating the entire space of the varying parameters.

A Branch-and-Bound global optimization procedure, is proposed for locating the global minimum solution of the inner most problem based on the refinement of converging lower and upper bounds for problems with box-like domains. The partitioning strategy involves the successive subdivision of a rectangle into two subrectangles by halving on the middle point of the longest side of the rectangle (bisection). Therefore, at each iteration a lower bound of the objective function value of the inner nonconvex parametric problem is simply the minimum over all the minima of the relaxed problem in every subrectangle and an upper bound of the objective function value of the inner nonconvex parametric problem is obtained by substitution of the relaxed solution into the objective function of the inner nonconvex parametric problem.

We have described a multiparametric programming approach strategy for the solution of hierarchical multilevel with nonconvexity formulation in the inner problems, based on the combination of multiparametric programming approach and Branch-and-Bound algorithm. The main limitation of our algorithm is that it depends on the boundedness of the optimization variables when special terms (e.g bilinear, trilinear, fractional, fractional trilinear to name fews) as well as generic nonconvex term occur in the problem.

A further study is required to refine the proposed algorithm so that it can be used to solve general trilevel optimization problems. In addition, we believe that the algorithm can be extended to solving multilevel hierarchical control problems and multilevel decentralized optimization problems.

References

[1] Acevedo J., and Pistikopoulos E. N., A parametric mixed integer nonlinear programming algorithm for process synthesis under uncertainty, *Industrial and engineering chemistry research*, 35(1996)147-158.

[2] Acevedo J., and Pistikopoulos E.N., Stochastic optimization Based algorithms for process synthesis under uncertainty, *computers and chemical engineering*, 22(1998)647-671. In Vicente N.L. and Calamai H.P., *Bilevel and Multilevel programming: A Bibliography Review Journal of Global Optimization, 5(1994)1-9.*

[3] Adjiman S. C., Dallwing S., Floudas A. C., Neumaier A., A global optimization method, αBB, for general twice-defferentiable constrained NLPs I.Theoretical advances, *Elsevier science Ltd*, 22(1997)1137-1158.

[4] Androulakis P. I., Maranas D. C., and Floudas A. C., αBB: A Global Optimization Method for General Constrained Nonconvex Problems, *Kluwer Acadaminc publisher*, (1995)1-27.

[5] Arora R. S. and Gaur A., A fuzzy algorithm for multilevel programming problems, Operational Research Society of India, 2010.

[6] Bahatia T. K., and Biegler L. T., Multi-period design and planning with interior point method, 1999.

[7] Bank B., Guddat J., Klatte D., Kummer B. and Tammer K., Nonlinear parametric programming problems, Academie-Verlag Berlin, 1982.

[8] Bialas F. W., Karwan H. M., Multilevel Otimization: A Mathematical Programming Perspective, M.Sc. Thesis, State University of New York, 1980.

[9] Blair C., Computational difficulties of bilevel linear programming, *Operations Research*, 38(1990)556-560.

[10] Braken J., and McGill M. J., Mathematical programs with optimization problems in the constraints, *operations research*, 21(1973)37-44.

[11] Candler W., Fortuny-Amat J. and McCarl B., The potential role of multilevel programming in agricultural economics, *American Journal of Agricultural Economics*, 63(1981)521-531.

[12] Dejenee A., Three-Person cooperative game and its application in decision Making process of hierarchical organization, M.Sc. Thesis, Addis Ababa university, 2007.

[13] Dempe S., Foundations of bilevel Programming, Kluwer Academic Publishers, 2002.

[14] Dempe S. and Dutta J., Is bilevel programming a special case of a mathrmatical program with comlementarity constraints, *Mathematical programming*, 124(2010)1-12.

[15] Dua V., Pistikopoulos, E. N., An algorithm for the solution of multiparametric mixed integer linear programming problems, *Kluwer Acadamic Publisher*, 99(2001)124-128.

[16] Dua V., Bozinis A., Pistikopoulos E. N., A multiparametric programming approach for mixed-integer quadratic engineering problems, *Elsevier Science Ltd*, 26(2002)715-721.

[17] Faísca P. N., Dua V., Rustem B., Saraiva M. P. and Pistikopoulos N. E., Parametric global optimisation for bilevel programming, Springer *Science+Business Media B.V.*, 38(2006)610-617.

[18] Faísca P. N., Saraiva M. P., Rustem B., Pistikopoulos N. E. A multiparametric programming approach for multilevel hierarchical and decentralized optimization problems, *Springer-Verlag*, 6(2007)377-397.

[19] Faísca P. N, Dua V., Saraiva M. P., Rustem B., and Pistikopoulos E. N., A Global Parametric Programming Optimization Strategy for Bilevel programming Problems, Elsevier B.V., 2006.

[20] Faísca P. N., Dua V. and Pistikopoulos N. E., Multiparametric Linear and Quadratic Programming, WILEY-VCH Verlag GmbH and Co. KGaA, 2007.

[21] Fiacco V. A., Introduction to Sensitivity and Stability Analysis in Nonlinear Programming, Acadamic press, 1983.

[22] Fiacco V. A., Sensitivity analysis for nonlinear programming using penalty methods, *Mathematical Programming*, 10(1976)287-311.

[23] Floudas A. C., and Pardalos M. P., Frontiers in global Optimization, Kluamer Acadamic Publisher, 2004.

[24] Gaur A. and Arora R. S., Multilevel Multiobjective integer linear Programming problems, *Advanced Modeling and Optimization*, 10(2008)298-301.

[25] Goertzel B., The structure of intelligence: A new Mathematical Model of mind, Springer-Verlag, 1993.

[26] Gümüs H. Z., Floudas A. C., Global optimization of nonlinear bilevel programming problems, *Kluwer Acadamin publisher*, 20(2001)1-31.

[27] Jahn J., Introduction to the theory of nonlnear optimization, Third Edition, Springer, 2007.

[28] Jezowski J. and Thullie J., 19th European symposium on computer aided process engineering , Elsevier B.V, 2009.

[29] Kocvara M. and Outrata J., A numerical solutions of two selected shape optimization problems, Technical Report DFG, University of Bayreuth, 1993. In Vicente N. L. and Calamai H. P., *Bilevel and Multilevel programming: A Bibliography Review, Journal of Global Optimization, 5(1994)1-9.*

[30] Lakie E., Linear Three Level Programming Problem with the Application to Hierarchical Organizations, M.Sc Thesis, Addis Ababa University, 2007.

[31] Leblanc L., and Boyce D., A bilevel programming algorithm for exact solution of the network design problem with user optimal flow, *Transportation research*, 20(1986)259-265. In Vicente N. L. and Calamai H. P., *Bilevel and Multilevel programming: A Bibliography Review, Journal of Global Optimization, 5(1994)1-9.*

[32] Lu J., Shi C., Zhang G., Ruan D., Multi-follower linear bilevel programming: Model and Kuhn-Tucker approach, University of Technology, 2007.

[33] Marcotte P., Network design problem with congestion effects, A case of bilevel programming , *Mathematical programming*, 43(1986)142-162.

[34] Migdalas A., Pardalos M. P., and Värbrand P., Multilevel optimization: Algorithm, theory and applications, Klumer Acadamic Publlisher, 1992.

[35] Önal H., A modified simplex approach for solving bilevel linear programming problems, *European Journal of Operations Research*, 67(1993)126-135. In Vicente N. L. and Calamai H. P., *Bilevel and Multilevel programming: A Bibliography Review Journal of Global Optimization, 5(1994)1-9.*

[36] Pardalos M. P., Coleman F. T., Lectures On Global Optimization, American Mathematical Society, 2009.

[37] Pistikopoulos E. N., Uncertainty in process design and operations, *computers and chemical engineering*, 19(1995)553-563.

[38] Pistikopous, N. E., Georgiads C. M. and Dua V., Multiparametric programming: Theory, algorithm, and application, WILEY-VCH Verlag GmbH and Co. KGaA, 2007.

[39] Rao S. S., Engineering Optimization: theory and practice, 4th eds, John Wiley and sons.Inc., 2009.

[40] Ruan Z. G., Wang Y. S., Yamamoto Y., and Zhu S. S., Optimality conditions and geoametric properties of a linear multilevel programming problem with dominated objective functions, *Journal of optimization theory and applications*, 123(2004)409-429.

[41] Sakawa M., and Nishizak I., Cooperative and Non-cooperative Multilevel Programming, Springer Science+Business Media, 2009.

[42] Stackelberg V. H., The Theory of the Market Economy, Oxford University Press, 1952.

[43] Vicente N. L. and Calamai H. P., Bilevel and Multilevel programming: A Bibliography Review, *Journal of Global Optimization*, 5(1994)1-9.

[44] White J. D., Penality function approach to linear trilevel programming, *Journal of optimization theory and applications*, 93(1997)183-197.

[45] Zhang G., Lu J., Montero J., and Zeng Y., Model, solution concept, and K^{th}-best algorithm for linear trilevel programming, *Elsevier Inc.*, 180(2010)481-49