Algorithms, analysis and software for the global optimization of two-stage stochastic programs

by

Rohit Kannan

B.Tech. Chemical Engineering, Indian Institute of Technology Madras (2012)
M.S. Chemical Engineering Practice, Massachusetts Institute of Technology (2014)

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

Doctor of Philosophy in Chemical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2018

Author . . . .
Department of Chemical Engineering
December 29, 2017

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

Accepted by . . . .
Patrick S. Doyle
Robert T. Haslam (1911) Professor of Chemical Engineering
Chairman, Committee for Graduate Students

Algorithms, analysis and software for the global optimization of two-stage stochastic programs

by

Rohit Kannan

Submitted to the Department of Chemical Engineering
on December 29, 2017, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Chemical Engineering

Abstract

Optimization models in the chemical process industries often include uncertain model parameters due to uncertainties in market forces and the environment, use of reduced-order and surrogate process models, and difficulty in measuring parameters accurately. Optimal solutions to formulations that simply ignore uncertainties in the model parameters can be economically worthless or even disastrous in safety-critical applications. Rigorously accounting for uncertainties in optimization models arising out of the process industries is usually computationally prohibitive because of their inherent nonconvex and combinatorial nature. This thesis develops branch-and-bound (B&B) algorithms and a software framework for the scalable solution of a rich class of optimization problems under parametric uncertainty called two-stage stochastic programs, which finds several applications within the petrochemical, pharmaceutical, and energy industries. Additionally, the convergence rates of broad classes of B&B algorithms for constrained optimization problems are analyzed to determine favorable characteristics of such algorithms that can help mitigate the cluster problem in constrained optimization.

Two-stage stochastic programming formulations assume that a finite number of scenarios of the uncertain parameters may be realized, and provide a suitable framework for modeling applications with economic objectives. General-purpose B&B algorithms for two-stage stochastic programs suffer from a worst-case exponential increase in solution times with the number of scenarios, which makes the solution of practical applications impractical. This thesis presents a decomposable B&B algorithm for the efficient solution of large-scenario instances of a broad class of two-stage stochastic programs. Furthermore, this thesis details a software framework, named `GOSSIP`, that was developed for solving such problems. `GOSSIP`, which implements state-of-the-art decomposition techniques for the global solution of two-stage stochastic programs, is shown to perform favorably on a diverse set of test problems from the process systems engineering literature, and is a step towards the efficient solution of two-stage stochastic programming applications from the chemical process industries.

Branch-and-bound algorithms that do not possess good convergence properties suffer from the so-called cluster problem wherein a large number of boxes are visited in the vicinity of global optimizers. While the convergence rates of B&B algorithms for unconstrained problems and the cluster problem in unconstrained optimization had been well-studied prior to this thesis, the analyses for constrained problems were lacking, and are the focus of the second part of this thesis. This section of the thesis begins by developing a notion of convergence order for bounding schemes for B&B algorithms, establishes conditions under which

first-order and second-order convergent bounding schemes may be sufficient to mitigate the cluster problem, and determines sufficient conditions for widely applicable bounding schemes to possess first-order and second-order convergence. In addition, this section analyzes the convergence orders of some reduced-space B&B algorithms in the literature and establishes that such algorithms may suffer from the cluster problem if domain reduction techniques are not employed. Determining sufficient conditions on the domain reduction techniques to be able to mitigate the above cluster problem can help identify efficient reduced-space B&B algorithms for solving two-stage stochastic programs.

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

Acknowledgments

Oftentimes, one reads a paper, a thesis, or even a work of literature solely by looking at its content, rather oblivious to the significant contributions of the many people in the lives of the authors to the work. In this section, I wish to convey my gratitude to the many people who have had a significant impact on me during the course of my PhD, and, in the process, provide the reader with some context about the environment in which the work in this thesis was carried out.

I would like to begin by thanking my thesis advisor, Paul Barton. Paul not only gave me the freedom to chart my own course, but was available to discuss my research for hours on end. Not one to shy away from technical discussions, Paul's thorough review of my drafts has helped me inculcate high standards for my work. The curiosity with which he tries to analyze underlying phenomena is something that I hope to emulate. Paul has also given me several opportunities to present my work at conferences, which have been vital to my professional development.

I also want to thank the members of my thesis committee, Professors Richard Braatz and Youssef Marzouk, for their helpful feedback and guidance during my committee meetings. Financial support from BP, as part of the BP-MIT conversion research program, is also gratefully acknowledged. I am especially thankful for the guidance of Randy Field at the MIT Energy Initiative, and feedback from Casey Hetrick, George Huff, Yung Lin, and Steven Mitchell at BP at various stages of the project.

I would like to thank my external collaborators and mentors, Professors Ruth Misener, Christodoulos Floudas, and Johannes Jäschke, and Adriaen Verheyleweghen, for their mentorship and assistance at various stages of my thesis. Ruth and Prof. Floudas gave me access to a C++-version of ANTIGONE that was crucial for generating the computational results in this thesis. Ruth also helped set up the link to ANTIGONE within my software and has been very supportive of my work, especially at conferences. Johannes, who made the (fortunate) mistake of sitting next to me while visiting our lab, has been a great friend and mentor at an important period. Despite his busy schedule, Johannes was always available to discuss my half-baked ideas, and listened to my (occasional?) grievances over lunch and coffee with empathy. He also put me in touch with his student Adriaen, who, after picking up the requisite skills at an impressive pace, helped test my software and provided valuable feedback. Many thanks are also due to Professor Arun Tangirala, my undergraduate advisor at IIT Madras, for his continued mentorship from my undergraduate days.

The departmental staff have always been very welcoming and quick to help with my administrative needs. I am particularly grateful to Angelique, Paul's administrative assistant, Eileen, Joel, and Suzanne in the student office, Adrienne and Sandra in the department's headquarters, and Jean and Jim, who provided much-needed computer support, for their assistance through the years. Practice school provided a unique and grueling industrial experience that helped improve my communication skills and develop the ability to dive into diverse, challenging problems quickly. I would like to thank the organizers of the practice school program, especially Bob Hanlon for his invaluable inputs on technical communication, and Claude and Marlyse Lupis for the many social outings during our stint in Fort Worth. I am also thankful for the company of my classmates who shared my practice school experience; some of our friendships have grown stronger as a result.

This section would be rather incomplete without acknowledging the support of the past and present members of the Process Systems Engineering Laboratory. My labmates have always been around to discuss ideas, participate in non-technical discussions and social events, and provide essential moral support in times of difficulty. They have made my experience in the lab truly enjoyable. Achim, Kamil, and Stuart helped me get started in the lab and were happy to answer naïve questions in my early years. Ali and Harry were always willing to answer any coding-related questions. Yu afforded me great flexibility in my research by bearing the bulk of the responsibility for the BP project, was always around to discuss my research, and provided useful feedback by testing my software. Garrett and Peter were always happy to discuss mathematical ideas, research-related or otherwise, and taught me the importance of working through the math in a methodical fashion. Amir, Garrett, José, Paul, Peter, Siah, Siva, and Suzane lent a sympathetic ear during tough times. I could not have made it through this difficult journey without all of their support.

I am extremely grateful to my friends, both within and outside the department, for their invaluable support and camaraderie over these past five years. I will forever remain indebted to Ankur, Chibueze, Karthick, Manish, and Mikhil for the many hours of companionship, especially over lunch and coffee. I could not have made it through my PhD without your encouragement and support during testing times, and I can only hope that I made your MIT experience a fraction-as-much better by being part of it. I hope that we stay in touch for years to come. I would also like to thank several other friends at MIT: Ananth, Gang, Krishna, Vishnu, and Yamini for the many hours of chit-chat within the confines of our department, Deepak, Divya, Rajesh, and Sai Gautam for the numerous memorable social outings, and Anil, Karthik, Rahul, Tapovan, and Vaibhav for the many joyful hours spent playing badminton.

Making the transition from IIT Madras to MIT meant moving away from several people I have been fortunate to call friends since middle/high school. I would like to thank Karthikram, Rakesh, Sreekanth, Vignesh, and the several other folks from IIT Madras who have made an effort to stay in touch with me over these past five years. I will cherish the meetups we have had in the US for many years to come, and hope that we stay in touch. Other than my close friends within the department, perhaps the people to whom I've vented much of my feelings of indignation over these past few years are my roommates. I've had the pleasure of sharing an apartment with Jaichander for more than four years during which we have shared snippets of our research, reveled in our successes, vented our frustrations, and simply been there for each other. Although I've only had the chance to spend time with Giridhari and Sudarshan more recently, it feels like we have been good friends for quite a while. I will cherish our frequent trips to Punjabi Dhaba and Dosa-n-Curry, our chai sessions, and the many hours spent sharing the happenings in our lives.

Finally, I would like to express my gratitude to my family for their unconditional love and support throughout my life. They have truly set the best possible initial conditions for the chaotic journey of life, and I could not have gotten to this point without their many sacrifices, endless love, constant encouragement, and steadfast support. Amma, Appa, Thathi, Thatha, Chitti, Chittu, Mami, Mama, Sneha, Rohit, and Sudharshini, thank you! I hope that I will continue to make you proud, wherever life leads me.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The broad objective of this thesis is to develop algorithms, analyses, and software to aid in optimization under uncertainty. Specifically, we consider the development of decomposition algorithms and a software framework for the global solution of a broad class of scenario-based two-stage stochastic programs, and analyze the convergence orders of widely applicable lower bounding techniques for (full-space and reduced-space) global optimization to determine favorable characteristics of associated global optimization algorithms from the viewpoint of the cluster effect [68]. We also present our computational experience with the developed algorithms and software on two-stage stochastic programming instances primarily from the process systems engineering literature. Finally, in Chapter 7, we list some potential directions for future work on the software implementation, and outline some open questions that have arisen out of our convergence-order analysis, answers to (some of) which could help identify some reduced-space branch-and-bound algorithms in the literature with favorable convergence properties that could be applied to solve stochastic programs (and other structured optimization problems) of interest (efficiently).

1.1 Optimization under uncertainty

In this section, we present some motivation for our work in Chapters 3 and 4 that develops a decomposition algorithm and a software framework for a class of optimization problems with parametric uncertainty (two-stage stochastic programs), formally state our research goals, provide a high-level overview of popular approaches for formulating and solving optimization problems under uncertainty, and list the contributions of this thesis in the context of the research goals stated in this section.

1.1.1 Motivation

Optimization formulations encountered in the process industries often involve uncertain process models. Even if any discrepancies in the structures of process models are ignored, there are usually significant uncertainties in model parameters due to our imprecise knowledge of market forces and the environment, and difficulties in estimating/measuring some model parameters accurately. Disregarding uncertainties in such models can either lead to solutions that are quite suboptimal, or even lead to infeasible decisions in practice, which then makes the computed solutions worthless. Over the past few decades, several approaches for formulating and solving optimization problems under uncertainty have been proposed, including stochastic programming [35, 189], chance-constrained programming [181], robust optimization [21], and dynamic programming [29, 189]. We briefly review existing approaches for formulating and solving optimization problems under parametric uncertainty (including the ones mentioned above) in Section 1.1.3.

Stochastic programming and robust optimization are the two most commonly used approaches for formulating and solving optimization problems under parametric uncertainty, especially for the optimization of chemical process systems under uncertainty. Scenario-based stochastic programming is typically employed either for problems with economic objectives, or in applications where recourse decisions can be made to take corrective action once the uncertain parameter values are realized[1], whereas robust optimization is usually employed in safety-critical applications in which certain constraints have to be necessarily satisfied for all possible realizations of the uncertain parameters. Both stochastic programming and robust optimization have found several applications in the process systems engineering literature, including the design and operation of process networks [111, 135, 137] and chemical plants [1, 53, 245, 249], planning and operation of offshore oil structures and refineries [220, 241–243], scheduling of chemical plants [133, 144, 186, 244], and turnaround planning for chemical sites [8].

[1] While robust optimization formulations can also be adapted to include recourse decisions, their inclusion typically complicates the solution of such models much more significantly and one often has to resort to simplified formulations, see [22, 248] for instance.

1.1.2 Research goals

A major aim of this thesis (see Chapters 3 and 4) is to develop decomposition algorithms and a software framework for the global optimization of the following class of (deterministic equivalent) two-stage stochastic mixed-integer nonlinear programs (MINLPs):

$$\begin{aligned}
\min_{\mathbf{x}_1,\cdots,\mathbf{x}_s,\mathbf{y},\mathbf{z}} \quad & \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}) \\
\text{s.t.} \quad & \mathbf{g}_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\}, \\
& \mathbf{r}_{y,z}(\mathbf{y}, \mathbf{z}) \leq \mathbf{0}, \\
& \mathbf{x}_h \in X_h, \quad \forall h \in \{1, \cdots, s\}, \\
& \mathbf{y} \in Y, \ \ \mathbf{z} \in Z,
\end{aligned} \tag{SP}$$

where $X_h = \{0,1\}^{n_{x_b}} \times \Pi_{x,h}$ with $\Pi_{x,h} \in \mathbb{IR}^{n_{x_c}}$, $\forall h \in \{1, \cdots, s\}$, $Y = \{0,1\}^{n_y}$, $Z \in \mathbb{IR}^{n_z}$, and functions $f_h : [0,1]^{n_{x_b}} \times \Pi_{x,h} \times [0,1]^{n_y} \times Z \to \mathbb{R}$, $\mathbf{g}_h : [0,1]^{n_{x_b}} \times \Pi_{x,h} \times [0,1]^{n_y} \times Z \to \mathbb{R}^m$, $\forall h \in \{1, \cdots, s\}$, and $\mathbf{r}_{y,z} : [0,1]^{n_y} \times Z \to \mathbb{R}^{m_{y,z}}$ are assumed to be continuous. The variables $\mathbf{y}$ and $\mathbf{z}$ in Problem (SP) denote the discrete and continuous first-stage/complicating decisions, respectively, that are made before the realization of the uncertainties (these are typically design-related decisions), while, for each $h \in \{1, \cdots, s\}$, the mixed-integer variables $\mathbf{x}_h$ denote the second-stage/recourse decisions made after the uncertain model parameters realize their 'scenario h' values (these are typically operational decisions). The quantity $p_h > 0$ represents the probability of occurrence of scenario h (with $\sum_{h=1}^{s} p_h = 1$)[2]. Equality constraints in the formulation are assumed to be modeled using pairs of inequalities and bounded general integer variables are assumed to be equivalently reformulated using binary variables in Problem (SP) purely for ease of exposition.

Although we can attempt to solve Problem (SP) directly using general-purpose off-the-shelf deterministic global optimization solvers, the techniques implemented therein typically face a worst-case exponential increase in solution times with a linear increase in the number of scenarios since they do not exploit Problem (SP)'s nearly-decomposable structure, which makes this option unattractive from a practical viewpoint. Consequently, although the modeler would like to consider a sufficiently-large number of scenarios to account for the ef-

[2] While the decomposition approaches considered in this thesis can possibly be adapted to risk-averse stochastic programs (see, for instance, [5]), we only consider the risk-neutral formulation presented in Problem (SP) in this thesis.

fects of uncertainty accurately, the solution of such large-scenario instances of Problem (SP) using general-purpose global optimization software is usually impractical for applications of interest. A major goal of this thesis is to develop efficient (parallelizable) decomposition techniques for (the global optimization of) Problem (SP) whose solution times scale linearly with an increase in the number of scenarios on a serial computer, and to develop a decomposition software toolkit for the numerical solution of instances of Problem (SP) and use it to solve applications of interest efficiently.

1.1.3 Existing approaches

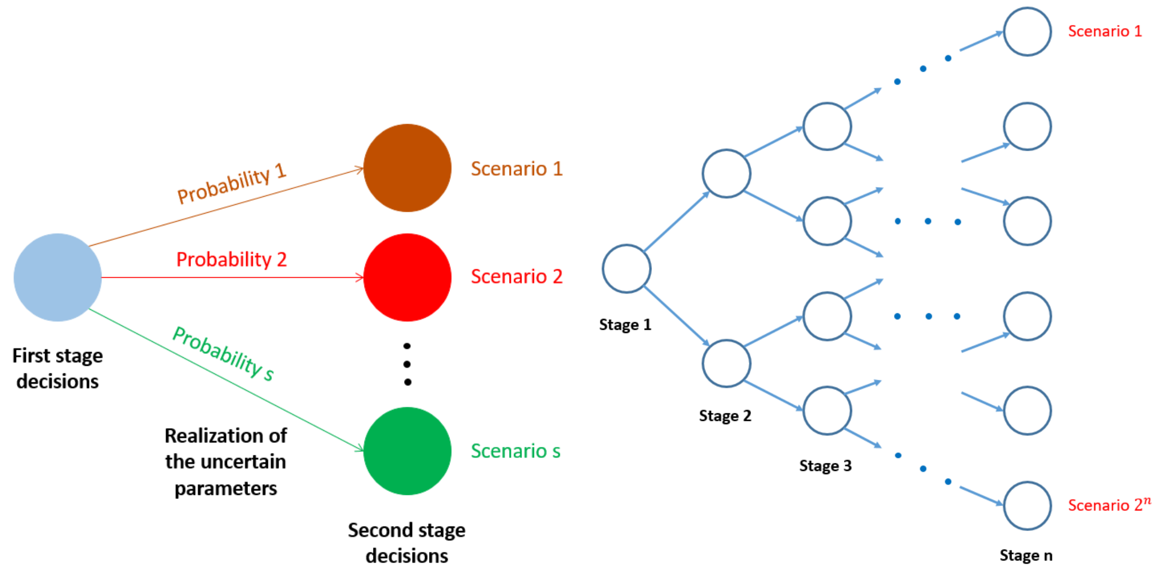
This section briefly introduces some existing approaches for formulating optimization problems under parametric uncertainty and lists some techniques in the literature for solving them.

1.1.3.1 Stochastic programming and chance-constrained optimization

Stochastic programming and chance-constrained optimization-based formulations take distributional information of the uncertain parameters into account, either by assuming that the uncertain parameters can be modeled as random variables with finite support (such as the formulation in Problem (SP)), or by using more general probabilistic descriptions of the uncertain parameters (see Problem (CCP)). These modeling frameworks are usually appropriate for applications in which a small, but non-negligible, sampling/distributional error is acceptable in practice[3], particularly when recourse decisions can be taken to mitigate the influence of uncertainty.

Dantzig [61] (originally in 1955) and Beale [14] are usually credited with being among the first to develop the stochastic programming framework we consider in this thesis. In a two-stage stochastic programming framework such as the one employed by Problem (SP), a set of first-stage decisions have to be made before the realization of the uncertain parameters after which it is assumed that the uncertain parameters take on one of a finite number of (known) values, each with a known probability. Subsequently, a set of second-stage/recourse decisions can be taken to react to the values assumed by the uncertain parameters. A schematic

[3]Although so-called ambiguous chance-constrained formulations can mitigate such errors (see, for instance, [77]), efficient solution approaches are known only for restrictive classes of problems that employ such formulations (a related framework is that of distributionally-robust optimization formulations, see [240] for instance). We also refer the reader to [146] for empirical studies on sampling-based methods for stochastic programs.



(a) Scenario tree for a two-stage stochastic program.

(b) Sample scenario tree for a multi-stage stochastic program.

Figure 1-1: Schematics of two-stage and multi-stage stochastic programming frameworks.

of the above two-stage stochastic programming framework is presented in Figure 1-1a.

Two-stage stochastic programming problems are usually solved using duality-based decomposition techniques such as Benders decomposition (BD) [25], generalized Benders decomposition (GBD) [85], and Lagrangian relaxation/decomposition [48] depending on the characteristics of the functions involved in their formulation. The reader is directed to references [35, 125, 196] for an overview of algorithms and software for stochastic mixed-integer linear programs. Few decomposition approaches exist for solving two-stage stochastic MINLPs in the form of Problem (SP) to global optimality. Li et al. [138, 139] generalized BD and GBD to a class of stochastic MINLPs that only contain bounded integer first-stage decisions, and called the resulting method 'nonconvex generalized Benders decomposition' (NGBD). Lagrangian relaxation-based approaches [112, 119] can be employed to solve the general class of Problem (SP), but are usually seen to be ineffective in solving large-scenario instances of Problem (SP) for applications of interest (see Chapters 3 and 4). In Chapter 3, we develop a modified Lagrangian relaxation algorithm that integrates Lagrangian relaxation with NGBD and scalable domain reduction techniques to solve Problem (SP).

Planning and operational problems in the process industries usually involve sequences of decisions that are taken over time with the aid of new (market) information and updated operational strategies as opposed to a simple two-stage decision framework such as the one outlined above. This leads to the concept of multi-stage stochastic programs, which are largely thought to be intractable even for the case of linear programs with parametric uncertainty [73][4] (this is illustrated by the sample scenario tree of a multi-stage stochastic program in Figure 1-1b, where even considering only two independent realizations of the uncertain parameters at each stage leads to an exponential explosion in the overall number of scenarios with the number of stages considered). Therefore, we restrict our attention to the two-stage stochastic programming formulation (SP) throughout this thesis.

One of the first introductions to chance-constrained programming was through the work of Charnes, Cooper, and Symonds [50, 51]. The general chance-constrained program of interest can be expressed as:

$$\begin{aligned} \min_{\mathbf{y} \in Y} \; & f(\mathbf{y}) \\ \text{s.t.} \; & \mathbb{P}(\mathbf{y} \in \mathbf{\Pi_y}(\boldsymbol{\omega})) \geq 1 - \varepsilon, \\ & \mathbf{g}(\mathbf{y}) \leq \mathbf{0}, \end{aligned} \tag{CCP}$$

where $Y = \{0,1\}^{n_{y_b}} \times \mathbb{R}^{n_{y_c}}$, $\mathbf{\Pi_y} : \Omega \to \mathcal{P}\left(\mathbb{R}^{n_{y_b}+n_{y_c}}\right)$, $\mathcal{P}\left(S\right)$ denotes the power set of S, $\boldsymbol{\omega}$ is a random variable from a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with 'support' $\Omega \subset \mathbb{R}^u$ (ignoring the slight abuse of the definition of support here), functions $f : [0,1]^{n_{y_b}} \times \mathbb{R}^{n_{y_c}} \to \mathbb{R}$ and $\mathbf{g} : [0,1]^{n_{y_b}} \times \mathbb{R}^{n_{y_c}} \to \mathbb{R}^m$ are assumed to be continuous, and $\varepsilon \in (0,1)$ is a user-defined parameter that specifies the maximum acceptable probability of constraints violation. The constraint $\mathbb{P}(\mathbf{y} \in \mathbf{\Pi_y}(\boldsymbol{\omega})) \geq 1 - \varepsilon$, called a (joint) chance constraint, imposes that certain constraints have to be satisfied (jointly) with a probability of at least $1 - \varepsilon$. A general setting for $\mathbf{\Pi_y}$ can be expressed as

$$\mathbf{\Pi_y}(\boldsymbol{\omega}) := \left\{ \mathbf{y} \in \mathbb{R}^{n_{y_b}+n_{y_c}} : \mathbf{h}(\mathbf{y}, \mathbf{x}(\boldsymbol{\omega}), \boldsymbol{\omega}) \leq \mathbf{0} \text{ for some } \mathbf{x}(\boldsymbol{\omega}) \in \{0,1\}^{n_{x_b}} \times \mathbb{R}^{n_{x_c}} \right\}, \;\; \forall \boldsymbol{\omega} \in \Omega,$$

where $\mathbf{x}(\bar{\boldsymbol{\omega}}) \in \{0,1\}^{n_{x_b}} \times \mathbb{R}^{n_{x_c}}$ denotes a vector of recourse decisions for the case when the uncertain parameters $\boldsymbol{\omega}$ assume the value $\bar{\boldsymbol{\omega}}$, and $\mathbf{h} : \mathbb{R}^{n_{y_b}+n_{y_c}} \times [0,1]^{n_{x_b}} \times \mathbb{R}^{n_{x_c}} \times \Omega \to \mathbb{R}^{m_u}$ is a continuous function.

Few algorithms are known for solving the general chance-constrained program Prob-

[4] The reader is directed to stochastic dual dynamic programming-based [179, 185, 252] and decision rule-based [128] techniques for two approaches that have been gaining increasing attention for (approximately) solving multi-stage stochastic (mixed-integer) linear programs.

lem (CCP) to optimality. This is partly due to the fact that: i. even linear programs with affine chance constraints can result in nonconvex feasible regions, and ii. checking the feasibility of joint chance-constrained problems involves multi-dimensional integration in general, which is nontrivial. Prékopa [181, Chapter 10] presents some strategies for reformulating a subclass of Problem (CCP) into equivalent convex optimization problems. Nemirovski and Shapiro [176] present convex optimization-based approaches that could help determine conservative feasible solutions for a subclass of Problem (CCP). Yang et al. [243] develop a 'near-global' optimization algorithm for solving a class of linear programs with joint chance constraints involving normally distributed random variables with applications to terminal blending in refineries. Luedtke and Ahmed [148] present a sample-average approximation-based approach for finding feasible solutions and statistical bounds for Problem (CCP). Finally, we mention that Calfa and coworkers [46] have proposed a data-driven approach to solving chance-constrained problems with right-hand side uncertainties based on the work of Jiang and Guan [105].

1.1.3.2 Robust optimization

Robust optimization-based formulations model the space of possible uncertain parameter realizations using so-called 'uncertainty sets' (rather than relying on probabilistic descriptions). These formulations are usually suitable for safety-critical applications in which considering the worst-case situation is vital since they require guaranteed satisfaction of the constraints for all possible realizations of the uncertain parameters (i.e., they optimize principally with the worst case in mind). Robust optimization formulations can be viewed as belonging to the class of semi-infinite programs (SIPs) [21, 23, 216].

Soyster [215] is usually credited with being one of the first researchers to develop a modern robust linear programming formulation along with a tractable solution approach. More than a couple of decades after Soyster's article, the works of Ben-Tal and Nemirovski (see [21, 23] and the references therein) and El-Ghaoui and coworkers [74, 75] (along with the subsequent works of other researchers, see [31, 33, 63] and related references) spurred a renewed interest in robust optimization-based approaches, particularly in the case of robust convex optimization and robust mixed-integer linear programming. The (single-stage/static) robust optimization analogue to Problem (SP) can be written as:

$$\begin{aligned} \min_{\mathbf{y}\in Y} \ & \max_{\boldsymbol{\omega}\in\Omega} f(\mathbf{y},\boldsymbol{\omega}) \\ \text{s.t.} \ & \mathbf{g}(\mathbf{y},\boldsymbol{\omega}) \leq \mathbf{0}, \quad \forall \boldsymbol{\omega}\in\Omega, \\ & \mathbf{h}(\mathbf{y}) \leq \mathbf{0}, \end{aligned} \tag{RP}$$

where $Y = \{0,1\}^{n_{y_b}} \times \Pi_y$ with $\Pi_y \in \mathbb{IR}^{n_{y_c}}$, $\Omega \subset R^u$ is a nonempty compact set, and functions $f : [0,1]^{n_{y_b}} \times \Pi_y \times \Omega \to \mathbb{R}$, $\mathbf{g} : [0,1]^{n_{y_b}} \times \Pi_y \times \Omega \to \mathbb{R}^{m_u}$, and $\mathbf{h} : [0,1]^{n_{y_b}} \times \Pi_y \to \mathbb{R}^m$ are assumed to be continuous. The constraint $\mathbf{g}(\mathbf{y},\boldsymbol{\omega}) \leq \mathbf{0}$, $\forall \boldsymbol{\omega} \in \Omega$, called a semi-infinite constraint when the cardinality of Ω is (uncountably) infinite, imposes that a feasible decision $\mathbf{y} \in Y$ should satisfy the constraint $\mathbf{g}(\mathbf{y},\boldsymbol{\omega}) \leq \mathbf{0}$ for every possible realization of the uncertain parameters $\boldsymbol{\omega} \in \Omega$.

Solution approaches for robust MINLPs essentially rely on general-purpose algorithms for solving SIPs, which are broadly classified into discretization and reduction-based methods [98]. Most global optimization approaches for semi-infinite programs [34, 65, 168] employ (adaptive) discretization-based techniques [37], which, at their core, iteratively replace each semi-infinite constraint with an increasing but finite number of constraints corresponding to well-chosen samples of the uncertain parameter values to guarantee convergence. We note that techniques for solving so-called 'bilevel programs' can also be used to solve SIPs, see [167, 170] for instance. When the functions in Problem (RP) possess certain special structures, there are a few techniques in the literature [64, 97] that can equivalently reformulate Problem (RP) into an 'ordinary' (mixed-integer) nonlinear program. We close this section by noting that some formulations and (conservative) solution approaches for solving two-stage robust optimization problems, which incorporate the ability to make recourse decisions, have been proposed in the literature in the past few decades [22, 32, 89, 218, 232, 248].

1.1.3.3 Other approaches

The previous two sections introduced some of the popular approaches for formulating and solving optimization problems under uncertainty. In this section, we briefly list some other approaches that have been explored in the literature with varying degrees of success for problems with specific structures. Before we proceed, we mention that the formulations considered in Sections 1.1.3.1 and 1.1.3.2 are not as mathematically dissimilar as one might imagine. For instance, if the uncertainty set Ω in Problem (RP) is replaced with the support of the probability distribution in Problem (CCP) (denoted therein by Ω as well), then the chance constraints and the robust constraints in those formulations essentially enforce similar sets of constraints.

Another popular approach for modeling optimization problems under parametric uncertainty that we do not adopt in this thesis is dynamic programming [15], which provides a framework for modeling multi-stage decision problems under uncertainty (see [180] for connections between dynamic programming and other modeling approaches). Dynamic programming models for optimization under uncertainty typically use multi-stage, finite-state and finite-policy formulations [180, 189] (usually with some sort of Markov decision process structure), and solutions of these models usually proceed using Bellman's 'principle of optimality' by breaking down the problem of interest into similar but simpler subproblems and storing these solutions for later (re)use. Since we do not know of techniques in the literature by which dynamic programming-based approaches can be readily employed to solve many of the applications of interest rigorously, we do not investigate the topic further.

Fuzzy programming provides another alternative to optimization under uncertainty using so-called fuzzy numbers and fuzzy sets that track the degree of violation of constraints involving random variables [16, 196]. Subcategories of fuzzy programming include: flexible programming, which deals with right-hand-side uncertainties in constraints, and possibilistic programming, which considers uncertainties in constraint and objective coefficients. We conclude this section by mentioning that several approaches, such as the framework of 'light robustness' [78], exist that combine the modeling frameworks discussed in the last few sections.

1.1.4 Contributions

Chapters 3 and 4 are devoted to the development of a decomposition algorithm and software framework for the global optimization of the two-stage stochastic program Problem (SP). Chapter 3 presents a modified Lagrangian relaxation-based (MLR-based) B&B algorithm for the decomposable solution of Problem (SP) by integrating Lagrangian relaxation (LR), NGBD, and decomposable bounds tightening techniques (this chapter develops, to the best of our knowledge, the first fully-decomposable algorithm for Problem (SP) that provably converges to an ε-optimal solution in finite time). We also establish a number of favorable theoretical results of our proposed algorithm in Chapter 3 such as the fact that the MLR lower bounding problem provides tighter lower bounds than the lower bounding problem of the conventional LR algorithm, and the result that it suffices to branch on fewer (first-stage) variables for the MLR B&B algorithm to converge compared to the LR B&B algorithm.

Chapter 4 presents the details of our software, `GOSSIP`, that can be used for the decomposable solution of Problem (SP) (under appropriate assumptions on the functions involved in its definition). `GOSSIP` includes implementations of NGBD, LR, and the MLR algorithms in conjunction with domain reduction methods and several advanced state-of-the-art techniques from the literature for preprocessing user input to solve Problem (SP). At the time of this writing, `GOSSIP` involves more than a hundred thousand lines of source code written primarily in C++. To the best of our knowledge, `GOSSIP` will be the first publicly available decomposition toolkit that can handle the general form of Problem (SP). We also compile and institute the first (soon-to-be publicly available) test library for two-stage stochastic MINLPs in Chapter 4.

1.2 Convergence-order analysis of branch-and-bound algorithms

We begin this section by motivating our work on the convergence-order analysis of B&B algorithms for constrained optimization problems (see Chapter 6) from the purview of the cluster problem in constrained optimization (see Chapter 5). Next, we state our research goals and briefly review related prior analyses for unconstrained problems in the literature. Finally, we close this section by listing the contributions of this thesis in the above areas.

1.2.1 Motivation

A key issue faced by deterministic branch-and-bound algorithms for continuous global optimization is the so-called cluster problem, where a large number of small boxes may be visited by the algorithm in neighborhoods of global optimizers (see [237, Figure 2.1] for a visual illustration of the cluster problem in unconstrained optimization). Du and Kearfott [68, 116] were the first to formally define and analyze this phenomenon in the context of interval arithmetic-based B&B algorithms for unconstrained global optimization. Subsequent analyses by Neumaier [178] and Wechsung et al. [237, 238] refined the cluster problem analysis of Du and Kearfott for the unconstrained case. The major conclusion of the above analyses is that the convergence order of the bounding scheme employed by the B&B algorithm plays a major role in determining the extent to which the cluster problem is mitigated; for unconstrained optimization problems with twice continuously differentiable objective functions, second-order (Hausdorff) convergent bounding schemes with small-enough prefactors can avoid ‘clustering’, whereas first-order convergent bounding schemes and second-order convergent bounding schemes with large-enough prefactors can result in an exponential number of boxes (as a function of the termination tolerance of the B&B algorithm) being visited in neighborhoods of global optimizers. Recent analyses on the convergence orders of bounding schemes for unconstrained problems [38, 39] (also see the related work [203]) can help determine sufficient conditions on bounding schemes for unconstrained optimization to possess the requisite convergence order to mitigate the cluster problem.

While the cluster problem for the unconstrained case had been quite well-studied prior to our work, much less was known (or at least explicitly stated in the literature) for the case of constrained optimization from a theoretical standpoint (and, to a lesser degree, from a computational perspective as well). Kearfott [114, 115] and Schichl et al. [201, 202] prescribe auxiliary techniques to exclude regions of the search space near global optimizers to mitigate the cluster problem (under suitable conditions). Goldsztejn et al. [87] also develop rejection tests for multi-objective constrained optimization, and present some academic case studies (some of which are particularly interesting to us because they empirically validate our analysis of the cluster problem in Chapter 5) to demonstrate the effectiveness of their proposed schemes. We present numerical results for two constrained optimization examples below (see Examples 5.2.8 and 5.2.9 in Chapter 5) to illustrate how the cluster effect for the constrained case can manifest differently compared to its unconstrained counterpart. For both examples, we implement a basic B&B framework (without any domain reduction, and only using basic branching and node selection rules) to compare the performance of three different lower bounding schemes (since these case studies are academic in nature, we assume that local optimization techniques locate a global optimal solution at the root node of the B&B tree).

Example 1.2.1. Consider the equality-constrained problem:

$$\begin{aligned} \min_{\substack{x\in[0,2],\\ y\in[0,3]}} \quad & y^2 - 7y - 12x \\ \text{s.t.} \quad & y = 2 - 2x^4. \end{aligned}$$

The optimal solution to the above problem is $(x^*, y^*) \approx (0.7175, 1.4698)$. Figure 1-2 compares the performance of three lower bounding schemes for the above problem.

Example 1.2.2. Consider the inequality-constrained problem:

$$\begin{aligned} \min_{\substack{x\in[0,3],\\ y\in[0,4]}} \quad & -x - y \\ \text{s.t.} \quad & y \leq 2x^4 - 8x^3 + 8x^2 + 2, \\ & y \leq 4x^4 - 32x^3 + 88x^2 - 96x + 36. \end{aligned}$$

The optimal solution to the above problem is $(x^*, y^*) \approx (2.3295, 3.1785)$. Figure 1-3 compares the performance of three lower bounding schemes for the above problem.

Figures 1-2 and 1-3 compare the performance of natural interval extensions-based [172, Section 5.4], centered form-based [172, Section 6.4], and McCormick relaxation-based [154, 213] lower bounding schemes for the two examples. The reader can check that while the natural interval extension and centered form-based lower bounding schemes suffer from the cluster problem for Example 1.2.1 (as seen from the fact that the number of iterations of the associated B&B methods grows significantly with a decrease in the termination tolerance; note that the McCormick relaxation-based method is able to mitigate the cluster problem for both the cases), their behavior is qualitatively different for the inequality constrained Example 1.2.2 where the number of iterations of the associated B&B algorithms is relatively insensitive to the termination tolerance[5]. We note that the cluster problem and convergence-

[5] At this stage, it is pertinent to note that although both the natural interval extension-based and centered form-based lower bounding schemes exhibit the same qualitative behavior with a decrease in the termination tolerance for Example 1.2.2, the centered form-based lower bounding scheme results in a much more efficient B&B algorithm for this case. This highlights a key limitation of relying on the cluster problem analysis as a primary measure of efficiency of B&B schemes in that schemes that do not suffer from the cluster problem are not necessarily 'efficient'/practical. We also add that the comparison plots of the number of iterations of the B&B algorithms with the different lower bounding schemes in Figures 1-2 and 1-3 use a natural logarithm scale.

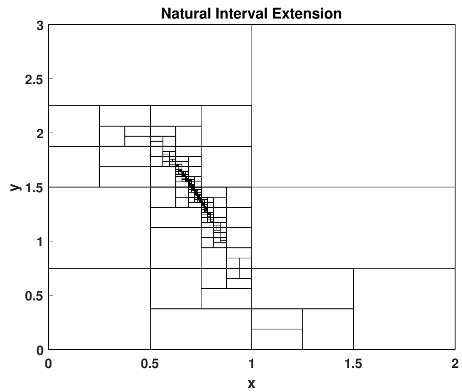
order analyses in Chapters 5 and 6 can qualitatively help explain the computational results for Examples 1.2.1 and 1.2.2.

1.2.2 Research goals

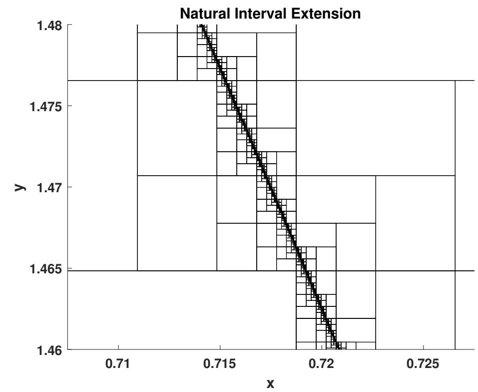
A second major goal of this thesis involves: i. formulating and analyzing the cluster problem in constrained global optimization, which involves providing conservative estimates of the number of B&B nodes visited in the vicinity of a constrained global minimizer as a function of the termination tolerance under appropriate assumptions, to determine necessary characteristics of B&B algorithms in order to mitigate clustering, and ii. developing a notion of convergence order of B&B algorithms for constrained problems, and analyzing the convergence orders of widely applicable full-space and reduced-space B&B schemes in the literature.

The first step in the convergence-order analysis of lower bounding schemes for constrained optimization involves defining an appropriate notion of convergence order for such schemes (see Definition 6.3.12). Thereafter, our aim is to determine conservative bounds on the convergence orders of widely used lower bounding schemes (such as McCormick relaxation-based schemes [76, 154, 213], interval arithmetic-based schemes [172], and Lagrangian duality-based schemes [69]), possibly by using bounds on the convergence orders of schemes of convex and concave relaxations of the functions involved in the optimization formulation that can be obtained from the previous analyses for unconstrained optimization [38, 39]. Roughly, we wish to determine the rate at which lower bounds converge to optimal objective function values on sequences of successively refined nodes converging to feasible points, and wish to determine the efficiency with which sequences of successively refined nodes all of which are contained within the infeasible region are detected to be infeasible by the lower bounding scheme.

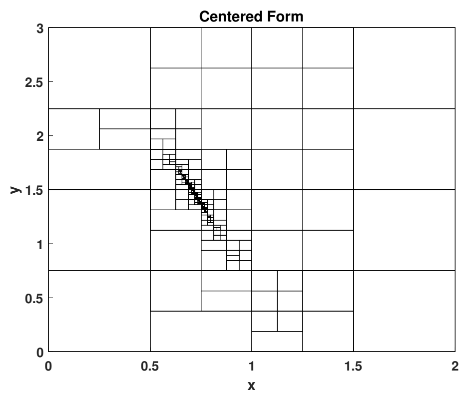
As part of our analysis of the cluster problem in constrained global optimization, we wish to determine when first-order and second-order convergent lower bounding schemes are sufficient to mitigate the cluster problem around constrained global optimizers. Additionally, uncovering the (worst-case) dependence of the extent of clustering on the convergence order prefactor (similar to the analysis in [237, 238]) is a desirable outcome of the analysis. By using the cluster problem and convergence order analyses, our aim is to be able to explain, at least in theory, disparities in the performances of various lower bounding schemes within



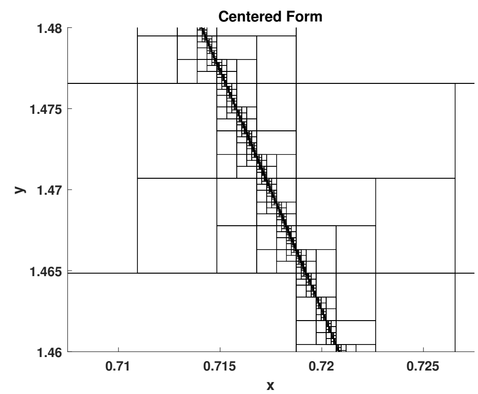
(a) Boxes visited by the natural interval extension-based B&B algorithm.



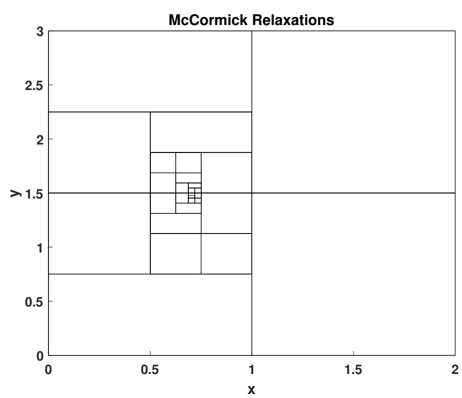
(b) Boxes visited by the natural interval extension-based B&B algorithm in a vicinity of the global minimizer.



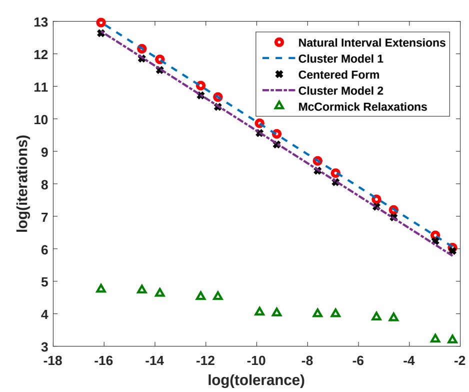
(c) Boxes visited by the centered form-based B&B algorithm.



(d) Boxes visited by the centered form-based B&B algorithm in a vicinity of the global minimizer.

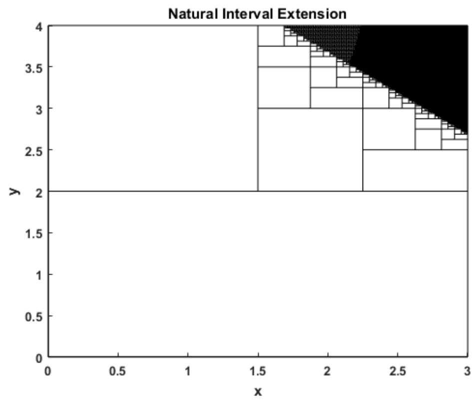


(e) Boxes visited by the McCormick relaxation-based B&B algorithm.

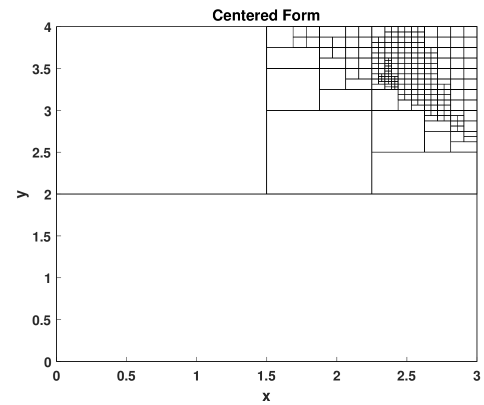


(f) Comparison of number of B&B iterations of the algorithms corresponding to three different bounding schemes for different termination tolerances (log denotes the natural logarithm).

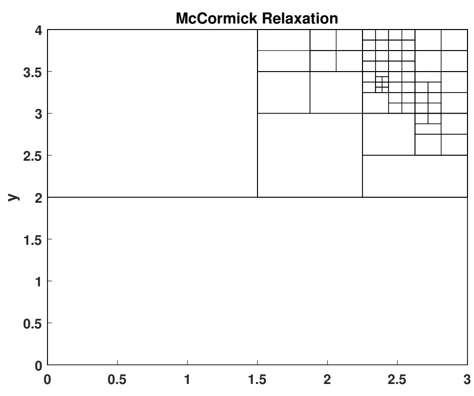
Figure 1-2: Summary of computational experiments for Example 1.2.1 using three different lower bounding schemes in a basic B&B framework.



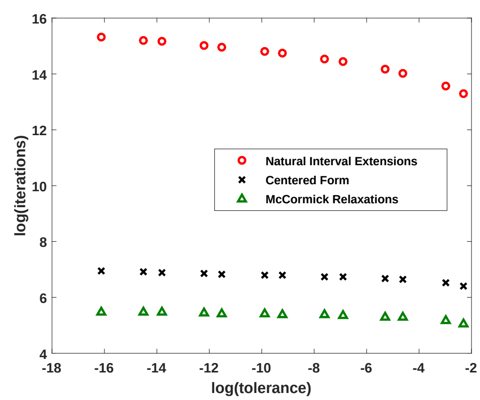
(a) Boxes visited by the natural interval extension-based B&B algorithm.



(b) Boxes visited by the centered form-based B&B algorithm.



(c) Boxes visited by the McCormick relaxation-based B&B algorithm.



(d) Comparison of number of B&B iterations of the algorithms corresponding to three different lower bounding schemes for different termination tolerances (log denotes the natural logarithm).

Figure 1-3: Summary of computational experiments for Example 1.2.2 using three different lower bounding schemes in a basic B&B framework.

B&B algorithms such as the relative performances of interval arithmetic-based and (non-constant) convex relaxation-based lower bounding schemes, and the relative performances of full-space and reduced-space lower bounding schemes.

1.2.3 Prior analyses

Previous work that deals with analyzing the convergence rates of lower bounding schemes for unconstrained problems has mainly been presented in the context of the Hausdorff convergence order, which determines the rate at which an estimate of the range of a function converges to its true range as the interval over which the estimate is constructed is refined. The works of Moore et al. [172], Ratschek and Rokne [184] and other researchers in the interval arithmetic community [93, 127, 205] established tight bounds on the Hausdorff convergence orders of interval arithmetic-based schemes of inclusion functions. Schöbel and Scholz [203] develop a notion of convergence order for bounding schemes for unconstrained optimization by considering the rate at which a scheme of lower bounds converges to the objective value of a scheme of feasible points on successively refined nodes. Their analysis also provides a very conservative estimate of the worst-case number of boxes explored by B&B algorithms for unconstrained optimization.

The convergence order framework that our analysis in Chapter 6 builds on was only recently introduced by Bompadre and Mitsos [38]. The authors therein introduce the notion of pointwise convergence order of schemes of convex and concave relaxations of a function[6], develop a theory of propagation of (Hausdorff and pointwise) convergence orders of schemes of McCormick relaxations [154], and establish tight bounds on the pointwise convergence orders of schemes of αBB relaxations [4] and convex and concave envelopes. Subsequently, Mitsos and coworkers [39, 174] developed the analysis in [38] further to analyze the propagation of convergence orders of schemes of Taylor models [184], McCormick-Taylor models [197], and multivariate McCormick relaxations [227]. Chapter 6 reviews other relevant analyses of the convergence rates of bounding schemes for unconstrained problems.

[6]The notion of pointwise convergence order will turn out to be critical in our analysis of the convergence orders of lower bounding schemes for constrained optimization in Chapter 6 (and, consequently, in our analysis of the cluster problem in Chapter 5). This is in contrast to the case of unconstrained optimization where the notion of Hausdorff convergence order assumes greater significance in the analysis of the cluster problem for this case.

1.2.4 Contributions

Chapter 5 presents our analysis of the cluster problem in constrained global optimization, and Chapter 6 presents our convergence-order analysis of B&B algorithms for constrained optimization. Our analysis of the cluster problem in Chapter 5 generalizes the previous analyses for unconstrained optimization [68, 237, 238], and establishes conditions under which first-order and second-order convergent lower bounding schemes may be sufficient to mitigate the cluster problem in neighborhoods of global optimizers[7]. Additionally, we develop refinements of the above analysis for problems with specific structures (such as problems with equality constraints), and establish conditions under which at least second-order convergence of the lower bounding scheme is required to mitigate clustering.

Our analysis of the convergence rate of lower bounding schemes for constrained optimization in Chapter 6 institutes a (fairly general) definition of convergence order for constrained optimization that also generalizes the corresponding definition for unconstrained optimization. Additionally, we analyze the convergence orders of convex relaxation-based and Lagrangian duality-based full-space lower bounding schemes, and show that our analysis reduces to the results for unconstrained optimization under appropriate assumptions. We also determine that the pointwise convergence order of schemes of relaxations of the functions involved in constrained optimization problems plays a crucial role rather than their Hausdorff convergence orders. Along with the analysis of the cluster problem summarized above, this establishes why basic implementations of interval arithmetic-based B&B algorithms may perform quite well for certain classes of inequality-constrained problems, but perform rather poorly for classes of equality-constrained problems. Finally, we also determine bounds on the convergence orders of two widely applicable reduced-space B&B algorithms in the literature [69, 76]. One significant finding of the above analysis is that reduced-space B&B algorithms may suffer from the cluster problem if appropriate domain reduction techniques are not employed (in fact, we show in Chapter 6 that certain reduced-space B&B algorithms can face severe clustering even for unconstrained problems). This insight could form the basis of future work that attempts to determine sufficient properties of domain reduction techniques for reduced-space B&B algorithms to mitigate clustering.

[7] Although our analysis of the cluster problem truly generalizes previous analyses for unconstrained optimization, we note that the proofs of our most important results utilize techniques from [237, 238].

Chapter 2

Background

This chapter presents relevant background concepts along with mathematical definitions and results that act as a common thread across this thesis. In particular, definitions and background results concerning local optimality conditions for unconstrained and constrained minimization, branch-and-bound algorithms for global optimization, convergence properties of schemes of convex and concave relaxations of functions, and decomposition algorithms for two-stage stochastic programming problems are provided. The ensuing chapters of this thesis provide additional chapter-specific mathematical background over the material presented in this chapter.

2.1 Notation

Throughout this thesis, we use $\mathbf{0}$ to denote a vector of zeros of appropriate dimension (it is sometimes also used to denote a matrix of zeros, which will be clear from the context), $\mathbb{N}$ to denote the set of natural numbers (starting from zero or one, depending on the context), $\mathbb{R}_+$ and $\mathbb{R}_-$ to denote the nonnegative and nonpositive reals, respectively, and $\mathbb{I}Z$ to denote the set of nonempty, closed and bounded interval subsets of $Z \subset \mathbb{R}^n$.

Sets are denoted using uppercase letters. The closure of a set $Z \subset \mathbb{R}^n$ is denoted by $\mathrm{cl}\,(Z)$, and $\mathrm{conv}(Z)$ and $\mathrm{int}(Z)$ are respectively used to denote its convex hull and interior. The difference between two sets X and Y is denoted by $X \backslash Y$.

Scalars and scalar-valued functions are denoted using lowercase letters, whereas vectors and vector-valued functions are denoted using lowercase bold letters. Matrices are denoted using uppercase bold letters. The transposes of a (column) vector $\mathbf{z} \in \mathbb{R}^n$ and a matrix

$\mathbf{M} \in \mathbb{R}^{m \times n}$ are denoted by $\mathbf{z}^{\mathrm{T}}$ and $\mathbf{M}^{\mathrm{T}}$, respectively. The expression $\|\mathbf{z}\|$ is used to denote the Euclidean norm of $\mathbf{z} \in \mathbb{R}^n$ (unless otherwise specified), and, for any $p \in \mathbb{N}$, $\|\mathbf{z}\|_p$ is used to denote the p-norm of $\mathbf{z}$. The expression z_j is used to denote the j^{th} component of a vector $\mathbf{z}$, $(z_1, z_2, \cdots, z_n)$ is used to denote a vector $\mathbf{z} \in \mathbb{R}^n$ with components $z_1, z_2, \cdots, z_n \in \mathbb{R}$ (note that (z_1, z_2) will be used to denote both an open interval in $\mathbb{R}$ and a vector in $\mathbb{R}^2$; the intended use will be clear from the context), and the notation $(\mathbf{u}, \mathbf{v}, \mathbf{w})$ is used to denote the column vector $[\mathbf{u}^{\mathrm{T}}\ \mathbf{v}^{\mathrm{T}}\ \mathbf{w}^{\mathrm{T}}]^{\mathrm{T}}$ corresponding to (column) vectors $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$.

The symbol $\lceil \cdot \rceil$ is used to denote the ceiling function, log is used to denote the natural logarithm, $\begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix}$ denotes a vector-valued function with domain Y and codomain $\mathbb{R}^{m+n}$ corresponding to vector-valued functions $\mathbf{g} : Y \to \mathbb{R}^m$ and $\mathbf{h} : Y \to \mathbb{R}^n$, and the quantity $\bar{\mathbf{f}}(Z)$ denotes the image of $Z \subset Y$ under the function $\mathbf{f} : Y \to \mathbb{R}^m$.

Consider a function $f : Z \to \mathbb{R}$ with Z denoting a nonempty open subset of $\mathbb{R}^n$. The symbol $\nabla f(\mathbf{z}) \in \mathbb{R}^n$ is used to denote the gradient (vector) of f at $\mathbf{z} \in Z$, $\nabla^2 f(\mathbf{z}) \in \mathbb{R}^{n \times n}$ denotes the Hessian (matrix) of f at $\mathbf{z} \in Z$, $f'(\mathbf{z}; \mathbf{d})$ is used to denote the (scalar) directional derivative of f at $\mathbf{z} \in Z$ in the direction $\mathbf{d}$, and the symbol $f^{(k)}(z)$ is used to denote the (scalar) k^{th} derivative of f at $z \in Z$ when $n = 1$ (i.e., $Z \subset \mathbb{R}$). The term 'differentiable' is used to refer to differentiability in the Fréchet sense.

Finally, consider the optimization problem

$$f^* = \inf_{\mathbf{z} \in Z} f(\mathbf{z}),$$

where $Z \subset \mathbb{R}^n$ is the *feasible set* and $f : Z \to \mathbb{R}$ is the *objective function*. We adopt the convention that $f^* = +\infty$ when $Z = \emptyset$. If the function f is continuous and the set Z is compact, we replace the infimum, 'inf', by a minimum, 'min' (even if Z could be empty).

2.2 Basic definitions and results

This section provides some basic background definitions and results, familiarity with which is assumed throughout this thesis. Background results that are more specific to the contributions of this thesis are reviewed in Section 2.3.

Definition 2.2.1. [Neighborhood of a Point] Let $\mathbf{x} \in X \subset \mathbb{R}^n$. For any $\alpha > 0$, $p \in \mathbb{N}$, the set

$$\mathcal{N}_\alpha^p(\mathbf{x}) := \Big\{ \mathbf{z} \in X : \|\mathbf{z} - \mathbf{x}\|_p < \alpha \Big\}$$

is called the α-neighborhood of $\mathbf{x}$ relative to X with respect to the p-norm.

Although the set X is not part of the notation for a neighborhood, its definition will either be clear from the context, or specified explicitly.

Lemma 2.2.2. [Equivalence of Norms on $\mathbb{R}^n$] All norms on $\mathbb{R}^n$ are equivalent. Specifically, if $\|\cdot\|_p$ and $\|\cdot\|_q$ are two norms in $\mathbb{R}^n$ for any $p, q \in \mathbb{N} \cup \{+\infty\}$ with $p \neq q$, then there exist constants $c_1, c_2 \in \mathbb{R}_+$ such that $c_1\|\mathbf{z}\|_p \leq \|\mathbf{z}\|_q \leq c_2\|\mathbf{z}\|_p$, $\forall \mathbf{z} \in \mathbb{R}^n$. Furthermore, for $(p, q) = (1, 2)$, $c_2 = 1$ provides a valid upper bound and for $(p, q) = (+\infty, 2)$, $c_2 = \sqrt{n}$ provides a valid upper bound.

Proof. For the first part of the lemma, see, for instance, [200, Theorem 4.2]. The second part of the lemma follows from the inequalities

$$\|\mathbf{z}\|_2^2 = \sum_{i=1}^{n} z_i^2 \leq \sum_{i=1}^{n} z_i^2 + \sum_{i=1}^{n} \sum_{j=i+1}^{n} 2|z_i||z_j| = \|\mathbf{z}\|_1^2$$

and

$$\|\mathbf{z}\|_2^2 = \sum_{i=1}^{n} z_i^2 \leq n \max_{i=1,\cdots,n} z_i^2 = n\|\mathbf{z}\|_\infty^2$$

for any $\mathbf{z} \in \mathbb{R}^n$. □

Definition 2.2.3. [Lower and Upper Semicontinuity] Let $Z \subset \mathbb{R}^n$. A function $f : Z \to \mathbb{R}$ is said to be lower semicontinuous at $\bar{\mathbf{z}} \in Z$ if for every $\varepsilon > 0$, there exists a neighborhood $\mathcal{N}_\alpha^p(\bar{\mathbf{z}})$ of $\bar{\mathbf{z}}$ (where the exact values of $\alpha > 0$ and $p \in \mathbb{N}$ are immaterial) such that

$$f(\mathbf{z}) \geq f(\bar{\mathbf{z}}) - \varepsilon, \quad \forall \mathbf{z} \in \mathcal{N}_\alpha^p(\bar{\mathbf{z}}).$$

The condition for lower semicontinuity can equivalently be expressed as $\liminf\limits_{\mathbf{z} \to \bar{\mathbf{z}}} f(\mathbf{z}) \geq f(\bar{\mathbf{z}})$.

Analogously, f is said to be upper semicontinuous at $\bar{\mathbf{z}}$ if for every $\varepsilon > 0$, there exists a neighborhood $\mathcal{N}_\alpha^p(\bar{\mathbf{z}})$ of $\bar{\mathbf{z}}$ such that

$$f(\mathbf{z}) \leq f(\bar{\mathbf{z}}) + \varepsilon, \quad \forall \mathbf{z} \in \mathcal{N}_\alpha^p(\bar{\mathbf{z}}).$$

The condition for upper semicontinuity can equivalently be expressed as $\limsup\limits_{\mathbf{z} \to \bar{\mathbf{z}}} f(\mathbf{z}) \leq f(\bar{\mathbf{z}})$.

Note that f is continuous at $\bar{\mathbf{z}}$ if and only if it is both lower and upper semicontinuous at $\bar{\mathbf{z}}$.

Definition 2.2.4. [Lipschitz and Locally Lipschitz Continuous Functions] Let $Z \subset \mathbb{R}^n$. A function $f : Z \to \mathbb{R}$ is Lipschitz continuous on Z with Lipschitz constant $M \geq 0$ if

$$|f(\mathbf{z}_1) - f(\mathbf{z}_2)| \leq M\|\mathbf{z}_1 - \mathbf{z}_2\|, \quad \forall \mathbf{z}_1, \mathbf{z}_2 \in Z.$$

The function f is locally Lipschitz continuous on Z if $\forall \bar{\mathbf{z}} \in Z$, there exist $\alpha, M > 0$ such that

$$|f(\mathbf{z}_1) - f(\mathbf{z}_2)| \leq M\|\mathbf{z}_1 - \mathbf{z}_2\|, \quad \forall \mathbf{z}_1, \mathbf{z}_2 \in \mathcal{N}^2_\alpha(\bar{\mathbf{z}}).$$

Note that locally Lipschitz continuous functions are Lipschitz continuous on compact subsets of their domains.

Since the cluster problem and convergence order analyses in this thesis (see Chapters 5 and 6) are asymptotic in nature (see Remark 6.3.6 and Lemma 6.3.8, for instance), we will need the following asymptotic notations.

Definition 2.2.5. [Big O and Little o Notations] Let $Y \subset \mathbb{R}$, $f : Y \to \mathbb{R}$, and $g : Y \to \mathbb{R}$. We say that $f(y) = O(g(y))$ as $y \to \bar{y} \in Y$ if and only if there exist $\delta, M > 0$ such that

$$|f(y)| \leq M|g(y)|, \quad \forall y \in Y \text{ with } |y - \bar{y}| < \delta.$$

Similarly, we say that $f(y) = o(g(y))$ as $y \to \bar{y} \in Y$ if and only if for all $M' > 0$ there exists $\delta' > 0$ such that

$$|f(y)| \leq M'|g(y)|, \quad \forall y \in Y \text{ with } |y - \bar{y}| < \delta'.$$

Note that unless otherwise specified, we consider $\bar{y} = 0$ in this thesis.

We next state a basic, single-variable version of Taylor's theorem for differentiable functions, which provides local polynomial approximations of such functions around points of interest. Multivariable generalizations of Taylor's theorem also exist, and will be relied on at various points in this thesis. Theorem 2.2.7 provides a second-order version of Taylor's theorem for multivariable functions.

Theorem 2.2.6. [Taylor's Theorem] Let $n \in \mathbb{N}$, and suppose the function $f : \mathbb{R} \to \mathbb{R}$ is

n-times differentiable at the point $\bar{z} \in \mathbb{R}$. Then, there exists a function $h : \mathbb{R} \to \mathbb{R}$ such that

$$f(z) = f(\bar{z}) + \sum_{k=1}^{n} \frac{f^{(k)}(\bar{z})}{k!}(z - \bar{z})^k + h(z)(z - \bar{z})^n, \quad \forall z \in \mathbb{R}, \text{ and}$$
$$\lim_{z \to \bar{z}} h(z) = 0.$$

A stronger version of the above result, which provides a semi-explicit form for the remainder, can be derived under additional assumptions (see [192, Theorem 5.15], for instance).

Theorem 2.2.7. [Multivariable Taylor's Theorem] Let $Z \subset \mathbb{R}^n$ be a nonempty open set, and suppose the scalar-valued function $f : Z \to \mathbb{R}$ is twice differentiable at $\bar{\mathbf{z}} \in Z$. Then, there exists a function $h : Z \to \mathbb{R}$, which depends on $\bar{\mathbf{z}}$, such that

$$f(\mathbf{z}) = f(\bar{\mathbf{z}}) + \nabla f(\bar{\mathbf{z}})^{\mathrm{T}}(\mathbf{z} - \bar{\mathbf{z}}) + \frac{1}{2}(\mathbf{z} - \bar{\mathbf{z}})^{\mathrm{T}}\nabla^2 f(\bar{\mathbf{z}})(\mathbf{z} - \bar{\mathbf{z}}) + h(\mathbf{z})\|\mathbf{z} - \bar{\mathbf{z}}\|^2, \quad \forall \mathbf{z} \in \mathrm{int}(Z),$$
$$\lim_{\mathbf{z} \to \bar{\mathbf{z}}} h(\mathbf{z}) = 0.$$

Definition 2.2.8. [Directional Derivative] Let $Z \subset \mathbb{R}^n$ be an open set, and consider a function $f : Z \to \mathbb{R}$. The function f is said to be directionally differentiable at a point $\mathbf{z} \in Z$ in the direction $\mathbf{d} \in \mathbb{R}^n$ if the following limit exists:

$$f'(\mathbf{z}; \mathbf{d}) := \lim_{t \to 0^+} \frac{f(\mathbf{z} + t\mathbf{d}) - f(\mathbf{z})}{t},$$

in which case the vector $f'(\mathbf{z}; \mathbf{d})$ is called the directional derivative of f at $\mathbf{z}$ in the direction $\mathbf{d}$. If the above limit exists in $\mathbb{R}$ for each $\mathbf{d} \in \mathbb{R}^n$, then f is said to be directionally differentiable at $\mathbf{z}$.

If f is differentiable at $\mathbf{z} \in Z$, then it is also directionally differentiable at $\mathbf{z}$ with the directional derivative for any $\mathbf{d} \in \mathbb{R}^n$ given by $f'(\mathbf{z}; \mathbf{d}) = \nabla f(\mathbf{z})^{\mathrm{T}}\mathbf{d}$. We close this section by presenting a couple of definitions [57, 183, 204] from nonsmooth analysis that will find some use in Chapter 5. These definitions rely on Rademacher's theorem, which essentially states that a locally Lipschitz continuous function defined on an open subset of $\mathbb{R}^n$ is differentiable almost everywhere (i.e., outside a subset of the domain of Lebesgue measure zero) on that subset (see, for instance, [57]).

Definition 2.2.9. [B-subdifferential] Let $Z \subset \mathbb{R}^n$ be an open set, and $f : Z \to \mathbb{R}$ be a locally Lipschitz continuous function. Let $\Omega_f \subset Z$ be the set of measure zero on which f

is not differentiable. The B-subdifferential of f at a point $\mathbf{z} \in Z$, denoted by $\partial_{\mathrm{B}} f(\mathbf{z})$, is defined as the set

$$\partial_{\mathrm{B}} f(\mathbf{z}) := \left\{ \mathbf{v} \in \mathbb{R}^{1 \times n} : \mathbf{v} = \lim_{m \to \infty} \left(\nabla f(\mathbf{z}^m)\right)^{\mathrm{T}},\ \mathbf{z}^m \to \mathbf{z},\ \mathbf{z}^m \in Z \backslash \Omega_f,\ \forall m \in \mathbb{N} \right\}.$$

Definition 2.2.10. [Clarke's Generalized Gradient] Let $Z \subset \mathbb{R}^n$ be an open set, and $f : Z \to \mathbb{R}$ be a locally Lipschitz continuous function. Clarke's generalized gradient of f at a point $\mathbf{z} \in Z$, denoted by $\partial f(\mathbf{z})$, is defined as the set

$$\partial f(\mathbf{z}) := \mathrm{conv}\left(\left\{ \mathbf{v} \in \mathbb{R}^n : \mathbf{v}^{\mathrm{T}} \in \partial_{\mathrm{B}} f(\mathbf{z}) \right\}\right).$$

2.3 Mathematical optimization

Throughout this section, we consider the following optimization formulation unless otherwise stated:

$$\begin{aligned} \inf_{\mathbf{x}}\ & f(\mathbf{x}) \\ \text{s.t. } & \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \\ & \mathbf{h}(\mathbf{x}) = \mathbf{0}, \\ & \mathbf{x} \in X \subset \{0,1\}^{n_{x_b}} \times \mathbb{R}^{n_{x_c}}. \end{aligned} \tag{P}$$

We assume throughout this section that the functions $f : X \to \mathbb{R}$, $\mathbf{g} : X \to \mathbb{R}^{m_I}$, and $\mathbf{h} : X \to \mathbb{R}^{m_E}$ are continuous on X. We will also tacitly assume that the functions f, $\mathbf{g}$, and $\mathbf{h}$ and the set X in Problem (P) can be finitely expressed on a computer (for instance, these functions and sets can be expressed using 'factorable functions', see [206, Section 2.2] or [237, Section 3.1]). When the number of discrete variables in the formulation equals zero (i.e., $n_{x_b} = 0$), the formulation contains no inequality and equality constraints (i.e., $m_I = m_E = 0$), and $X = \mathbb{R}^{n_{x_c}}$ (with a slight abuse of notation; more generally, X can be any open set), then Problem (P) is called an unconstrained optimization problem; otherwise, we call Problem (P) a constrained optimization problem. Note that problems with bounded general integer variables can be equivalently reformulated to the form of Problem (P). Some special cases of the above formulation include:

- linear programming (LP), when $n_{x_b} = 0$, $X = \mathbb{R}^{n_{x_c}}$, and $f, \mathbf{g}$, and $\mathbf{h}$ are affine on X,

- mixed-integer linear programming (MILP), when $X = \{0,1\}^{n_{x_b}} \times \mathbb{R}^{n_{x_c}}$ and $f, \mathbf{g}$, and $\mathbf{h}$ are affine on X,

- convex programming, when $n_{x_b} = 0$, X is a convex subset of $\mathbb{R}^{n_{x_c}}$, f and $\mathbf{g}$ are convex on X, and $\mathbf{h}$ is affine on X,

- (nonconvex) nonlinear programming (NLP), when $n_{x_b} = 0$,

- convex mixed-integer nonlinear programming (convex MINLP), when $X = \{0,1\}^{n_{x_b}} \times \mathbb{R}^{n_{x_c}}$, f and $\mathbf{g}$ are convex on $\text{conv}(X)$, and $\mathbf{h}$ is affine on X, and

- (nonconvex) mixed-integer nonlinear programming (MINLP), when no particular restrictions on the sets and functions in Problem (P) are imposed.

Heuristics and local optimization algorithms for Problem (P)[1] aim to determine good or locally-optimal feasible points with little computational effort, whereas the classes of global optimization algorithms for Problem (P) that we are interested in (called 'complete' and 'rigorous' algorithms, see [178, Section 1.2]) aim to determine a provably (near-) optimal solution(s). Some popular 'local optimization' approaches for solving nonconvex NLPs and MINLPs include: multi-start methods [228], variable neighborhood search methods and their variants [95, 141], and feasibility pump methods [41, 59]. Some widely applicable global optimization approaches for solving such problems include: branch-and-bound/branch-and-reduce approaches [101, 195, 225], generalized Benders decomposition and its variants [85, 139], and outer-approximation and its variants [40, 72, 79, 118].

In the next section, we review some local optimality conditions for unconstrained and constrained nonlinear programming problems.

[1]The term 'local optimization' is usually used to describe algorithms for nonconvex NLPs, since the introduction of discrete variables makes the definition of a local minimum (see Definition 2.3.2) lose some meaning.

2.3.1 Local optimality conditions

The definitions and results in this section will primarily be used in Chapters 5 and 6. We consider the nonlinear programming formulation:

$$\begin{aligned} \inf_{\mathbf{x}} \ & f(\mathbf{x}) \\ \text{s.t. } & \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \\ & \mathbf{h}(\mathbf{x}) = \mathbf{0}, \\ & \mathbf{x} \in X \subset \mathbb{R}^{n_{x_c}}. \end{aligned} \tag{NLP}$$

We assume that the functions $f : X \to \mathbb{R}$, $\mathbf{g} : X \to \mathbb{R}^{m_I}$, and $\mathbf{h} : X \to \mathbb{R}^{m_E}$ are continuous on X. Additional assumptions on the objective function, the constraint functions, and the set X will be imposed as necessary.

Definition 2.3.1. [Feasible Region] The feasible region for Problem (NLP) on the 'domain' X, denoted by $\mathcal{F}(X)$, is defined as $\mathcal{F}(X) := \{\mathbf{x} \in X : \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\}$.

Any point $\mathbf{x} \in X$ that is not feasible (or even any point that is not an element of X) is called an infeasible point.

Definition 2.3.2. [Local Minimum] A feasible point $\bar{\mathbf{x}} \in \mathcal{F}(X)$ is called a local minimum for Problem (NLP) if $\exists \alpha > 0$ such that $f(\mathbf{x}) \geq f(\bar{\mathbf{x}})$, $\forall \mathbf{x} \in \mathcal{N}^2_\alpha(\bar{\mathbf{x}}) \cap \mathcal{F}(X)$.

Definition 2.3.3. [Strict Local Minimum] A feasible point $\bar{\mathbf{x}} \in \mathcal{F}(X)$ is called a strict local minimum for Problem (NLP) if $\bar{\mathbf{x}}$ is a local minimum, and $\exists \alpha > 0$ such that $f(\mathbf{x}) > f(\bar{\mathbf{x}})$, $\forall \mathbf{x} \in \mathcal{N}^2_\alpha(\bar{\mathbf{x}}) \cap \mathcal{F}(X)$ such that $\mathbf{x} \neq \bar{\mathbf{x}}$.

The reader can contrast the above *local* conditions with the following *global* condition for a minimum.

Definition 2.3.4. [Global Minimum] A feasible point $\mathbf{x}^* \in \mathcal{F}(X)$ is called a global minimum for Problem (NLP) if $f(\mathbf{x}) \geq f(\mathbf{x}^*)$, $\forall \mathbf{x} \in \mathcal{F}(X)$.

The following definition of distance does not define a metric; however, it will prove useful in defining a measure of infeasibility for points in X for Problem (NLP).

Definition 2.3.5. [Distance Between Two Sets] Let $Y, Z \subset \mathbb{R}^n$. The distance between Y and Z, denoted by $d(Y, Z)$, is defined as

$$d(Y, Z) := \inf_{\substack{\mathbf{y} \in Y, \\ \mathbf{z} \in Z}} \|\mathbf{y} - \mathbf{z}\|.$$

Lemma 2.3.6. Let $\mathbf{z}, \mathbf{v} \in \mathbb{R}^n$, and let $K \subset \mathbb{R}^n$ be a (nonempty) convex cone. Then

$$d(\{\mathbf{z}\}, K) - d(\{\mathbf{v}\}, K) \leq d(\{\mathbf{z} - \mathbf{v}\}, K).$$

Proof. See [214]. □

Corollary 2.3.7. Let $\mathbf{z}, \mathbf{v} \in \mathbb{R}^{m+n}$. Then

$$d(\{\mathbf{z}\}, \mathbb{R}^m_- \times \{\mathbf{0}\}) - d(\{\mathbf{v}\}, \mathbb{R}^m_- \times \{\mathbf{0}\}) \leq d(\{\mathbf{z} - \mathbf{v}\}, \mathbb{R}^m_- \times \{\mathbf{0}\}).$$

Proof. This result is a direct consequence of Lemma 2.3.6. □

The above results ensure that the quantity $d\left(\{(\mathbf{g}(\mathbf{x}), \mathbf{h}(\mathbf{x}))\}, \mathbb{R}^{m_I}_- \times \{\mathbf{0}\}\right)$ provides a (useful) measure of infeasibility (measure of constraint violation) for any point $\mathbf{x} \in X$ for Problem (NLP) (see Chapters 5 and 6); we leave the task of verifying that the above quantity is strictly positive if and only if the corresponding point $\mathbf{x}$ is infeasible to the reader.

The next definition of a nonisolated feasible point will prove useful for the cluster problem analysis, see Section 5.3.1 of Chapter 5.

Definition 2.3.8. [Nonisolated Feasible Point] A feasible point $\mathbf{x} \in \mathcal{F}(X)$ is said to be nonisolated if $\forall \alpha > 0$, $\exists \mathbf{z} \in \mathcal{N}^2_\alpha(\mathbf{x}) \cap \mathcal{F}(X)$ such that $\mathbf{z} \neq \mathbf{x}$.

While the definitions thus far in this section can be adapted to the more general formulation (P), we will, for the most part, restrict their usage to formulation (NLP) (see Chapters 5 and 6).

2.3.1.1 Unconstrained optimization

In this section, we assume that Problem (NLP) is unconstrained (i.e., $m_I = m_E = 0$ and X is an open set). Differentiability assumptions on the objective function f are assumed as necessary.

Theorem 2.3.9. [First-order necessary optimality condition for unconstrained minimization] Suppose f is differentiable at a point $\bar{\mathbf{x}} \in X$. If $\bar{\mathbf{x}}$ is a local minimum for Problem (NLP), then $\nabla f(\bar{\mathbf{x}}) = \mathbf{0}$.

Proof. See the corollary to Theorem 4.1.2 in [13]. □

Theorem 2.3.10. [First-order necessary optimality condition for nonsmooth unconstrained minimization] Suppose f is locally Lipschitz continuous on X. If $\bar{\mathbf{x}} \in X$ is a local minimum for Problem (NLP), then $\mathbf{0} \in \partial f(\bar{\mathbf{x}})$, i.e., $\mathbf{0}$ is an element of Clarke's generalized gradient of f at $\bar{\mathbf{x}}$.

Proof. See Proposition 2.3.2 in [57]. □

We close this section with a second-order necessary optimality condition and a second-order sufficient optimality condition for local minima for Problem (NLP). The reader is directed to Section 4.1 in Chapter 4 of [13] for higher-order optimality conditions for unconstrained minimization, which may be used to develop the cluster problem analysis in unconstrained optimization [68, 238] further.

Theorem 2.3.11. [Second-order necessary optimality condition for unconstrained minimization] Suppose f is twice differentiable at a point $\bar{\mathbf{x}} \in X$. If $\bar{\mathbf{x}}$ is a local minimum for Problem (NLP), then $\nabla^2 f(\bar{\mathbf{x}})$ is positive semidefinite.

Proof. See Theorem 4.1.3 in [13]. □

Theorem 2.3.12. [Second-order sufficient optimality condition for unconstrained minimization] Suppose f is twice differentiable at a point $\bar{\mathbf{x}} \in X$. If $\nabla f(\bar{\mathbf{x}}) = \mathbf{0}$ and $\nabla^2 f(\bar{\mathbf{x}})$ is positive definite, then $\bar{\mathbf{x}}$ is a strict local minimum for Problem (NLP).

Proof. See [13, Theorem 4.1.3]. □

Note that $\bar{\mathbf{x}}$ need not be a local minimum if the 'positive definiteness' condition on the Hessian in Theorem 2.3.12 is relaxed to 'positive semidefiniteness' (a typical counterexample would be to consider the behavior at the inflection point $\bar{x} = 0$ in the 'minimization' of $f(x) = x^3$ on $X = (-1, 1)$).

2.3.1.2 Constrained optimization

In this section, we assume that X is a nonempty open set. Differentiability assumptions on the objective and constraint functions are assumed as necessary.

The following definition of the set of active inequality constraints will play a part in setting up local optimality conditions because the inequality constraints that are not active at a candidate local optimal solution will not be active on some neighborhood of it (this is a consequence of the assumption of continuity of the functions in Problem (NLP)). Therefore, conditions for local optimality will not involve the set of inactive constraints (in an essential way). We remark that a consequence of the above observation is that the optimality conditions detailed in this section for a local minimum of Problem (NLP) at which none of its constraints are active essentially reduce to the conditions presented in Section 2.3.1.1. We also note that the inactive inequality constraints do not feature (in an essential way) in sufficient conditions for global optimality of convex programs, see Theorem 2.3.19.

Definition 2.3.13. [Set of Active Inequality Constraints] Let $\mathbf{x} \in \mathcal{F}(X)$ be a feasible point for Problem (NLP). The set of active inequality constraints at $\mathbf{x}$, denoted by $\mathcal{A}(\mathbf{x})$, is given by

$$\mathcal{A}(\mathbf{x}) := \{j \in \{1, \cdots, m_I\} : g_j(\mathbf{x}) = 0\}.$$

The next definition of the cone of tangents at a reference feasible point can be thought of as an estimate of the set of directions from the reference point that *locally* lead to feasible points. Naturally, this definition will be used to formulate a first-order necessary local optimality condition in Theorem 2.3.15.

Definition 2.3.14. [Tangent and Cone of Tangents] Let $\mathbf{x} \in \mathcal{F}(X) \subset \mathbb{R}^{n_{x_c}}$ be a feasible point for Problem (NLP). A vector $\mathbf{d} \in \mathbb{R}^{n_{x_c}}$ is said to be a tangent of $\mathcal{F}(X)$ at $\mathbf{x}$ if there exists a sequence $\{\lambda_k\} \to 0$ with $\lambda_k > 0$, and a sequence $\{\mathbf{x}_k\} \to \mathbf{x}$ with $\mathbf{x}_k \in \mathcal{F}(X)$ such that

$$\mathbf{d} = \lim_{k\to\infty} \frac{\mathbf{x}_k - \mathbf{x}}{\lambda_k}.$$

The set of all tangents of $\mathcal{F}(X)$ at $\mathbf{x}$, denoted by $T(\mathbf{x})$, is called the tangent cone of $\mathcal{F}(X)$ at $\mathbf{x}$.

Theorem 2.3.15. [First-Order Necessary Optimality Condition] Consider Problem (NLP), and suppose f is differentiable at a local minimum $\bar{\mathbf{x}}$. Then

$$\left\{\mathbf{d} : \nabla f(\bar{\mathbf{x}})^{\mathrm{T}}\mathbf{d} < 0\right\} \cap T(\bar{\mathbf{x}}) = \emptyset.$$

Proof. See Theorem 5.1.2 in [13]. □

A limitation of the above first-order necessary optimality condition is that it is difficult to verify in practice; this is because the tangent cone is a geometrical object that is hard to compute numerically/algebraically. Fritz John [106] developed (originally in circa 1948) the following first-order optimality condition that is more amenable to numerical verification (see [87]).

Theorem 2.3.16. [Fritz John necessary optimality conditions] Consider Problem (NLP), and suppose $\bar{\mathbf{x}} \in \mathcal{F}(X)$. Furthermore, suppose f and g_j, $\forall j \in \mathcal{A}(\bar{\mathbf{x}})$, are differentiable at $\bar{\mathbf{x}}$, and h_k, $\forall k \in \{1, \cdots, m_E\}$, are continuously differentiable at $\bar{\mathbf{x}}$. If $\bar{\mathbf{x}}$ is a local minimum of Problem (NLP), then there exist scalars μ_0, μ_j, $\forall j \in \mathcal{A}(\bar{\mathbf{x}})$, and λ_k, $\forall k \in \{1, \cdots, m_E\}$, such that

$$\begin{aligned}
&\mu_0 \nabla f(\bar{\mathbf{x}}) + \sum_{j \in \mathcal{A}(\bar{\mathbf{x}})} \mu_j \nabla g_j(\bar{\mathbf{x}}) + \sum_{k=1}^{m_E} \lambda_k \nabla h_k(\bar{\mathbf{x}}) = \mathbf{0},\\
&\mu_0, \ \ \mu_j \geq 0, \quad \forall j \in \mathcal{A}(\bar{\mathbf{x}}),\\
&(\mu_0, \boldsymbol{\mu}_{\bar{\mathbf{x}}}, \boldsymbol{\lambda}) \neq \mathbf{0},
\end{aligned}$$

where $\boldsymbol{\mu}_{\bar{\mathbf{x}}}$ denotes the vector with components μ_j, $\forall j \in \mathcal{A}(\bar{\mathbf{x}})$, and $\boldsymbol{\lambda}$ denotes the vector $(\lambda_1, \cdots, \lambda_{m_E})$. If g_j, $\forall j \in \{1, \cdots, m_I\}$, are differentiable at $\bar{\mathbf{x}}$, then the above conditions can be equivalently written as:

$$\begin{aligned}
&\mu_0 \nabla f(\bar{\mathbf{x}}) + \sum_{j=1}^{m_I} \mu_j \nabla g_j(\bar{\mathbf{x}}) + \sum_{k=1}^{m_E} \lambda_k \nabla h_k(\bar{\mathbf{x}}) = \mathbf{0},\\
&\mu_j g_j(\bar{\mathbf{x}}) = 0, \quad \forall j \in \{1, \cdots, m_I\},\\
&\mu_0, \ \ \mu_j \geq 0, \quad \forall j \in \{1, \cdots, m_I\},\\
&(\mu_0, \boldsymbol{\mu}, \boldsymbol{\lambda}) \neq \mathbf{0},
\end{aligned}$$

where $\boldsymbol{\mu}$ denotes the vector $(\mu_1, \cdots, \mu_{m_I})$.

Proof. See Theorem 4.3.2 in [13]. $\square$

A key limitation of the above Fritz John necessary conditions is that they can be trivially satisfied by points that are nowhere close to being locally optimal for Problem (NLP). For instance, Fritz John's conditions are trivially satisfied if any one of the objective or constraint function gradients vanishes at $\bar{\mathbf{x}}$. A less-than-ideal situation also occurs if equality constraints are replaced by pairs of inequality constraints, in which case every feasible point turns out to be a Fritz John point (see [13, p. 186]). The following necessary condition for local optimality, credited to Karush [113], Kuhn and Tucker [129] provides a stronger necessary condition for local optimality under an additional 'constraint qualification' assumption. Under appropriate assumptions, the Karush-Kuhn-Tucker conditions are both necessary and sufficient for global optimality, see Theorem 2.3.19.

Theorem 2.3.17. [KKT necessary optimality conditions] Consider Problem (NLP), and suppose $\bar{\mathbf{x}} \in \mathcal{F}(X)$. Furthermore, suppose f and g_j, $\forall j \in \mathcal{A}(\bar{\mathbf{x}})$, are differentiable at $\bar{\mathbf{x}}$, and h_k, $\forall k \in \{1, \cdots, m_E\}$, are continuously differentiable at $\bar{\mathbf{x}}$. Additionally, suppose that $\nabla g_j(\bar{\mathbf{x}})$, $\forall j \in \mathcal{A}(\bar{\mathbf{x}})$, and $\nabla h_k(\bar{\mathbf{x}})$, $\forall k \in \{1, \cdots, m_E\}$, are linearly independent (or, alternatively, some other constraint qualification holds, see [13, Chapter 5]). If $\bar{\mathbf{x}}$ is a local minimum of Problem (NLP), then there exist scalars μ_j, $\forall j \in \mathcal{A}(\bar{\mathbf{x}})$, and λ_k, $\forall k \in \{1, \cdots, m_E\}$, (these scalars are unique if linear independence of the constraint gradients is assumed) such that

$$\begin{aligned}
&\nabla f(\bar{\mathbf{x}}) + \sum_{j \in \mathcal{A}(\bar{\mathbf{x}})} \mu_j \nabla g_j(\bar{\mathbf{x}}) + \sum_{k=1}^{m_E} \lambda_k \nabla h_k(\bar{\mathbf{x}}) = \mathbf{0}, \\
&\mu_j \geq 0, \quad \forall j \in \mathcal{A}(\bar{\mathbf{x}}).
\end{aligned}$$

If g_j, $\forall j \in \{1, \cdots, m_I\}$, are differentiable at $\bar{\mathbf{x}}$, then the above conditions can be equivalently written as:

$$\begin{aligned}
&\nabla f(\bar{\mathbf{x}}) + \sum_{j=1}^{m_I} \mu_j \nabla g_j(\bar{\mathbf{x}}) + \sum_{k=1}^{m_E} \lambda_k \nabla h_k(\bar{\mathbf{x}}) = \mathbf{0}, \\
&\mu_j g_j(\bar{\mathbf{x}}) = 0, \quad \forall j \in \{1, \cdots, m_I\}, \\
&\mu_j \geq 0, \quad \forall j \in \{1, \cdots, m_I\}.
\end{aligned}$$

Proof. See Theorem 4.3.7 in [13]. □

Next, we define KKT points based on the KKT necessary conditions for local optimality. It is worth noting that local optimization solvers [67, 86, 235] for Problem (NLP) typically aim to determine a KKT point that satisfies the (first-order) KKT necessary optimality conditions (while possibly imposing some second-order necessary optimality conditions).

Definition 2.3.18. [KKT Point] Consider Problem (NLP), and suppose $\bar{\mathbf{x}} \in \mathcal{F}(X)$. Furthermore, suppose f and g_j, $\forall j \in \mathcal{A}(\bar{\mathbf{x}})$, and h_k, $\forall k \in \{1, \cdots, m_E\}$, are differentiable at $\bar{\mathbf{x}}$. A point $(\bar{\mathbf{x}}, \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\lambda}}) \in \mathbb{R}^{n_{x_c}+m_I+m_E}$ is said to be a KKT point[2] if it satisfies the following conditions:

$$
\begin{aligned}
&\nabla f(\bar{\mathbf{x}}) + \sum_{j \in \mathcal{A}(\bar{\mathbf{x}})} \bar{\mu}_j \nabla g_j(\bar{\mathbf{x}}) + \sum_{k=1}^{m_E} \bar{\lambda}_k \nabla h_k(\bar{\mathbf{x}}) = \mathbf{0}, \\
&\mathbf{g}(\bar{\mathbf{x}}) \leq \mathbf{0}, \;\; \mathbf{h}(\bar{\mathbf{x}}) = \mathbf{0}, \;\; \bar{\mathbf{x}} \in X, \\
&\bar{\boldsymbol{\mu}} \geq \mathbf{0}, \;\; \bar{\mu}_j g_j(\bar{\mathbf{x}}) = 0, \;\; \forall j \in \{1, \cdots, m_I\}.
\end{aligned}
$$

Theorem 2.3.19. [KKT sufficient optimality conditions] Consider Problem (NLP), and suppose X is a convex set with $\bar{\mathbf{x}} \in \mathcal{F}(X)$. Furthermore, suppose f and g_j, $\forall j \in \mathcal{A}(\bar{\mathbf{x}})$, are convex on X and differentiable at $\bar{\mathbf{x}}$, and h_k, $\forall k \in \{1, \cdots, m_E\}$, are affine on X. Additionally, suppose there exist $\bar{\boldsymbol{\mu}} \in \mathbb{R}^{m_I}$ and $\bar{\boldsymbol{\lambda}} \in \mathbb{R}^{m_E}$ such that $(\bar{\mathbf{x}}, \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\lambda}})$ is a KKT point for Problem (NLP). Then $\bar{\mathbf{x}}$ is a global minimum for Problem (NLP). If the convexity and affinity ('linearity') assumptions only hold on a neighborhood of $\bar{\mathbf{x}}$, then $\bar{\mathbf{x}}$ is a local minimum for Problem (NLP).

Proof. See Theorem 4.3.8 in [13]. □

Definition 2.3.20. [Slater Point] Consider Problem (NLP), and suppose the equality constraint functions h_k, $\forall k \in \{1, \cdots, m_E\}$, are affine on X. A feasible point $\mathbf{x}^S \in X$ is called a Slater point if it satisfies:

$$
\begin{aligned}
&g_j(\mathbf{x}^S) < 0, \quad \forall j \in \{1, \cdots, m_I\} \text{ s.t. } g_j \text{ is not affine}, \\
&g_j(\mathbf{x}^S) \leq 0, \quad \forall j \in \{1, \cdots, m_I\} \text{ s.t. } g_j \text{ is affine}.
\end{aligned}
$$

[2] Occasionally, we abuse the definition of a KKT point by simply assuming that a point $\bar{\mathbf{x}} \in X$ is a KKT point, instead of assuming the existence of 'KKT multipliers' $\bar{\boldsymbol{\mu}}$ and $\bar{\boldsymbol{\lambda}}$ such that $(\bar{\mathbf{x}}, \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\lambda}})$ is a KKT point.

If X is convex, f and g_j, $\forall j \in \{1, \cdots, m_I\}$, are convex on X and satisfy appropriate differentiability conditions, h_k, $\forall k \in \{1, \cdots, m_E\}$, are affine on X, and Problem (NLP) has a Slater point and a minimum, then there exists a (global) minimum for Problem (NLP) that is also a KKT point. The following result is from [13, Section 4.4].

Theorem 2.3.21. [Second-order sufficient optimality conditions for constrained minimization] Consider Problem (NLP) with $\bar{\mathbf{x}} \in X$, and suppose f and g_j, $\forall j \in \mathcal{A}(\bar{\mathbf{x}})$, and h_k, $\forall k \in \{1, \cdots, m_E\}$, are twice differentiable at $\bar{\mathbf{x}}$. Furthermore, suppose there exist $\bar{\boldsymbol{\mu}} \in \mathbb{R}^{m_I}$, $\bar{\boldsymbol{\lambda}} \in \mathbb{R}^{m_E}$ such that $(\bar{\mathbf{x}}, \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\lambda}})$ is a KKT point for Problem (NLP). Define the restricted Lagrangian function on X as $L(\cdot; \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\lambda}}) := f(\cdot) + \bar{\boldsymbol{\mu}}^{\mathrm{T}}\mathbf{g}(\cdot) + \bar{\boldsymbol{\lambda}}^{\mathrm{T}}\mathbf{h}(\cdot)$.

1. If $\nabla^2 L(\mathbf{x}; \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\lambda}})$ is positive semidefinite for all $\mathbf{x} \in \mathcal{F}(X)$, then $\bar{\mathbf{x}}$ is a global minimum for Problem (NLP).

2. If there exists $\alpha > 0$ such that $\nabla^2 L(\mathbf{x}; \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\lambda}})$ is positive semidefinite for all $\mathbf{x} \in \mathcal{F}(X) \cap \mathcal{N}^2_{\alpha}(\bar{\mathbf{x}})$, then $\bar{\mathbf{x}}$ is a local minimum for Problem (NLP).

3. If $\nabla^2 L(\bar{\mathbf{x}}; \bar{\boldsymbol{\mu}}, \bar{\boldsymbol{\lambda}})$ is positive definite, then $\bar{\mathbf{x}}$ is a strict local minimum for Problem (NLP).

Proof. See Lemma 4.4.1 in [13]. □

The reader is directed to Section 4.4 in [13] (specifically, Theorems 4.4.2 and 4.4.3 in [13]) for KKT second-order necessary and sufficient conditions for constrained minimization.

2.3.2 The branch-and-bound approach to global optimization

The first few paragraphs of Section 2.3 listed some widely applicable global optimization approaches for solving nonconvex NLPs and MINLPs. In this section, we briefly outline a basic version of one of those approaches, branch-and-bound (B&B), that is implemented in state-of-the-art MINLP solvers such as ANTIGONE [162], BARON [225], Couenne [19], LINDOGlobal [145], and SCIP [233]. Throughout this section, we consider the general formulation (P) unless otherwise specified and assume that the set X is a compact interval intersected with some integrality constraints (while it is not necessary to assume that $\text{conv}(X)$ is an interval in general, see [101, Chapter VII.2] for instance, the applications in this thesis will only use interval partition elements within a B&B framework). Along with the continuing assumptions of continuity of the functions in Problem (P), the above

Algorithm 2.1 Outline of a generic branch-and-bound algorithm for Problem (P)

Initialize:

a. Iteration counter $k = 0$, bounds X^0 on the variables $\mathbf{x}$ after the optional application of preprocessing techniques to the input data (we assume that X^0 is a compact interval intersected with the appropriate integrality restrictions), and tolerances $\varepsilon > 0, \varepsilon^l > 0$, and $\varepsilon^u > 0$ such that $\varepsilon^l + \varepsilon^u \leq \varepsilon$.

b. Domain of the root node $M^0 := X^0$, and the initial partition $\mathcal{P}^0 = \{M^0\}$.

c. Objective function value of the best found feasible point, $UBD = +\infty$; lower bound on the optimal objective value on the root node, $LBD^0 = -\infty$; and the best found feasible point, $\{\mathbf{x}^*\} = \emptyset$.

repeat

1. (Node Selection) Pick an active node $n \in \left\{n \in \mathbb{N} \cup \{0\} : M^n \in \mathcal{P}^k\right\}$ using some node selection heuristic and set $\mathcal{P}^{k+1} = \mathcal{P}^k \backslash \{M^n\}$.
2. (Optional Upper Bounding Step) Solve an upper bounding problem on M^n with a termination tolerance of ε^u to try and determine a feasible point. Update UBD, $\mathbf{x}^*$ if a feasible solution better than the current best solution is obtained.
3. (Optional Bounds Tightening Step) Apply finite bounds tightening techniques to obtain $\bar{M}^n \subset M^n$, and set $M^n = \bar{M}^n$. If M^n is empty, **goto** Step 6.
4. (Lower Bounding Step) Solve a lower bounding problem on M^n to ε^l-optimality to obtain the lower bound LBD^n (if the lower bounding problem on M^n is infeasible, set $LBD^n = +\infty$). If node n can be fathomed, **goto** Step 6.
5. (Branching Step) Partition M^n into M^{n_1} and M^{n_2} by branching one of the variables (once again, we assume that M^{n_1} and M^{n_2} are intervals intersected with the appropriate integrality restrictions). Set $\mathcal{P}^{k+1} = \mathcal{P}^{k+1} \cup \{M^{n_1}\} \cup \{M^{n_2}\}$, $LBD^{n_1} = LBD^{n_2} = LBD^n$.
6. Set $\mathcal{P}^{k+1} = \mathcal{P}^{k+1} \backslash \left\{M^p \in \mathcal{P}^{k+1} : LBD^p \geq UBD - \varepsilon\right\}$, $k = k + 1$.

until $\mathcal{P}^k = \emptyset$

If $UBD < +\infty$, then $\mathbf{x}^*$ provides an ε-optimal solution to Problem (P).

assumption ensures that Problem (P) is either infeasible, or has an optimal solution. Algorithm 2.1 outlines the basic steps involved in a B&B algorithm for Problem (P). We note that Algorithm 2.1 merely provides the backbone of a generic B&B algorithm. In practice, the order in which the subproblems are solved may vary and additional subproblems may be solved to speed up the convergence of the algorithm. In the following paragraph, we briefly describe the various steps involved in Algorithm 2.1.

At any particular iteration of the B&B algorithm, a node that has not yet been 'fathomed' (discarded) is selected for processing by the algorithm using some heuristic that tries to minimize the overall number of nodes visited. Commonly used node selection heuristics are 'best-bound' selection, which selects a node with the lowest lower bound, and 'depthfirst' selection, which selects the 'deepest node' in the B&B tree (see [18, Section 3.1.2]). Once an unfathomed node is selected, an upper bounding problem is solved (typically using 'local optimization' approaches, see [18, Section 6]) with the variables restricted to the domain of the node to try and generate a better feasible solution than the current best solution. The reader is directed to the excellent survey [18] and the thesis [28] for popular upper bounding approaches for Problem (P). Next, domain reduction techniques are used to try and eliminate parts of the domain of the node that do not contain feasible points that are better than the current best solution [19, 182]. If the domain reduction techniques determine, for instance, that Problem (P) does not contain any feasible points on the domain of the node (that have better objective value than the current upper bound UBD), then the node is 'fathomed by infeasibility' ('fathomed by optimality') and a new node is selected for consideration. While the bounds tightening steps are not necessary to guarantee convergence, they typically play an important role in boosting the performance of global optimization solvers since these steps not only help fathom infeasible/suboptimal regions of the search space, but also help in the construction of tighter relaxations (and consequently, tighter lower bounding problems). The reader is directed to the articles [19, 182] for a survey of bounds tightening techniques for Problem (P) that are employed by state-of-the-art global optimization software. Section 2.3.2.2 presents some bounds tightening techniques that are of particular interest to this thesis. The fourth step in the B&B algorithm involves the generation of guaranteed lower bounds on the optimal objective value of Problem (P) when its variables are restricted to the domain of the node under consideration. If the computed lower bound on a node exceeds the objective function value of the best known feasible point, then the node can be 'fathomed by value dominance' and a new node is selected for consideration. In addition to the computational cost of generating lower bounds, both the strength of the lower bounds generated by the lower bounding problem and the rate at which the lower bounds converge to the optimal objective value as the node is partitioned influence the effectiveness of the B&B scheme (see Chapters 5 and 6). Section 2.3.2.1 will briefly review popular lower bounding techniques for Problem (P). If the node under consideration has not been fathomed, the last key step of Algorithm 2.1 subdivides (the domain of) the node into two (in general, multiple) subnodes (subdomains) on each of which all of the previously described steps may be carried out at a later point in the course of the algorithm (this step makes the B&B algorithm a divide-and-conquer algorithm). Once again, the choice of branching heuristic can greatly impact the number of nodes visited by the B&B algorithm; therefore, several advanced heuristics have been proposed in the literature to try and make effective subdivision choices [18, 19]. The B&B procedure selects a new node using the selection rule and repeats the above steps so long as the termination criteria of the algorithm have not been satisfied.

The next few subsections present overviews of some common approaches for the lower bounding steps in Algorithm 2.1. These sections also introduce associated definitions and background results that will be important for the rest of this thesis.

2.3.2.1 Lower bounding techniques

This section mentions some popular, automatable techniques for generating guaranteed lower bounds on the optimal objective value of Problem (P) on any given node of the B&B tree, and provides some background definitions and results that will be used in Chapters 5 and 6 to analyze their effectiveness (from a particular viewpoint). First, we discuss interval arithmetic-based lower bounding procedures that can generate rigorous lower bounds. Next, we outline convex relaxation-based lower bounding techniques that form the basis of complete commercial global optimization solvers. Finally, we present a Lagrangian duality-based lower bounding procedure that is of particular interest in stochastic programming and structured optimization applications.

While some of the techniques that we mention in this section can be adapted to construct lower bounding problems for more general classes of functions, we restrict our attention here to the subclass of Problem (P) in which the functions f, $\mathbf{g}$, and $\mathbf{h}$ are 'factorable' (see [206, Section 2.2] or [237, Section 3.1], for instance, for formal definitions of factorable functions). We also note that factorable functions can be represented as a directed acyclic graph (DAG), see Section 2.3.2.2 (also see references [237] and [239], for instance).

2.3.2.1.1 Interval-based techniques

Given two real numbers a and b with $a \leq b$, a one-dimensional interval $[a, b]$ is defined as the set of all real numbers that lie between a and b. More generally, given two vectors $\mathbf{z}^{\mathrm{L}}$ and $\mathbf{z}^{\mathrm{U}}$ in $\mathbb{R}^n$ with $\mathbf{z}^{\mathrm{L}} \leq \mathbf{z}^{\mathrm{U}}$, the set $\left\{\mathbf{z} \in \mathbb{R}^n : \mathbf{z}^{\mathrm{L}} \leq \mathbf{z} \leq \mathbf{z}^{\mathrm{U}}\right\}$ defines an n-dimensional interval (also sometimes called a 'box') in $\mathbb{R}^n$. One of the earliest introductions to interval analysis was presented by Moore as part of his thesis[3] [171]. Although interval analysis provides a computationally efficient and relatively easily implementable method for generating rigorous bounds on the ranges of functions, key limitations of many applications of the technique include: significant overestimation of the ranges of functions in practical problems, and slow 'rates of convergence of the overestimation gap' in a way that affects the performance of B&B algorithms (see Definition 2.3.33). In this thesis, our primary use of interval analysis [172, 177] will be to bound the ranges of functions on intervals (see Definition 2.3.25). Popular interval-based techniques for bounding the ranges of functions on intervals include [172, 177, 184]: natural interval extensions, centered forms, slope forms, and Taylor forms.

While interval analysis techniques find uses within other noninterval techniques for constructing lower bounding problems (for example, see [4, 154]), interval analysis-based techniques can also be used as standalone methods for constructing lower bounding problems for Problem (P) (that can be used in Algorithm 2.1) as detailed below. The ranges of each of the inequality constraint functions in Problem (P) can be bounded on the domain, M^n, of node n in the B&B tree by bounding the ranges of these functions on an interval that contains M^n using interval arithmetic (the interval hull of M^n, see Definition 2.3.22, is usually easy to construct for the domains of B&B nodes and is used in practice to enclose M^n). Once valid bounds on the ranges of the inequality constraint functions are obtained, we can check if those bounds do not intersect with the interval $(-\infty, 0]$ to determine whether any of those constraints are necessarily (individually) violated on M^n. Similarly, the nonsatisfaction of any of the equality constraints on M^n can potentially be detected by checking for empty intersections between valid bounds on the ranges of the equality constraint functions and the degenerate interval $[0, 0]$. If any of the equality or inequality constraints have been proven to be violated at each point in the domain of node n using the above techniques, the node can be fathomed by infeasibility; otherwise, a guaranteed lower bound, LBD^n, on the optimal objective value of Problem (P) with its variables restricted to M^n can be obtained by computing a lower bound on the range of the objective function on M^n via interval analysis.

The following definitions are adapted from [38], and will be used, in particular, by the analyses in Chapters 5 and 6.

[3] A major aim of Moore's thesis was to develop an efficient and automatable technique for constructing rigorous bounds on the ranges of functions on intervals, with applications to automated rigorous error analysis.

Definition 2.3.22. [Interval Hull] Let $Z \subset \mathbb{R}^n$ be a nonempty bounded set. The interval hull of Z, denoted by $\square Z$, is the smallest element of $\mathbb{IR}^n$ that encloses Z, i.e., for every interval $\bar{Z} \in \mathbb{IR}^n$ that encloses Z (i.e., $Z \subset \bar{Z}$), we have $\square Z \subset \bar{Z}$.

Definition 2.3.23. [Width of an Interval] Let $Z = [z_1^{\mathrm{L}}, z_1^{\mathrm{U}}] \times \cdots \times [z_n^{\mathrm{L}}, z_n^{\mathrm{U}}]$ be an element of $\mathbb{IR}^n$. The width of Z, denoted by $w(Z)$, is given by

$$w(Z) := \max_{i \in \{1,\cdots,n\}} (z_i^{\mathrm{U}} - z_i^{\mathrm{L}}).$$

The choice of the ∞-norm in the above definition of the interval width is not restrictive because all norms on $\mathbb{R}^n$ are equivalent, see Lemma 2.2.2. The next definition specializes the notion of Hausdorff metric to one-dimensional intervals.

Definition 2.3.24. [Hausdorff Metric] Let $X = [x^{\mathrm{L}}, x^{\mathrm{U}}]$ and $Y = [y^{\mathrm{L}}, y^{\mathrm{U}}]$ be two intervals in $\mathbb{IR}$. The Hausdorff metric between X and Y, denoted by $d_H(X, Y)$, is given by

$$d_H(X,Y) = \max\{\left|x^{\mathrm{L}} - y^{\mathrm{L}}\right|, \left|x^{\mathrm{U}} - y^{\mathrm{U}}\right|\} = \max\left\{\max_{x \in X} \min_{y \in Y} |x-y|, \max_{y \in Y} \min_{x \in X} |x-y|\right\}.$$

The following definition introduces the notion of inclusion functions. Note that interval arithmetic techniques can be readily used to derive inclusion functions for factorable functions [49, 62, 193].

Definition 2.3.25. [Inclusion Function] Let $V \subset \mathbb{R}^n$, and suppose $\mathbf{f} : V \to \mathbb{R}^m$ is continuous. For any $Z \subset V$, let $\overline{\mathbf{f}}(Z)$ denote the image of Z under $\mathbf{f}$. A mapping $F : \mathbb{I}V \to \mathbb{IR}^m$ is called an inclusion function for $\mathbf{f}$ on $\mathbb{I}V$ if, for every $Z \in \mathbb{I}V$, we have $\overline{\mathbf{f}}(Z) \subset F(Z)$.

The next definitions introduce the notion of Hausdorff convergence order of an inclusion function (cf. Definition 6.3.2 in Chapter 6) and the range order of a scalar-valued function.

Definition 2.3.26. [Hausdorff Convergence Order of an Inclusion Function] Let $V \subset \mathbb{R}^n$ be a nonempty set, $h : V \to \mathbb{R}$ be a continuous function, and H be an inclusion function of h on $\mathbb{I}V$.

The inclusion function H is said to have Hausdorff convergence of order $\beta > 0$ at a point $\mathbf{v} \in V$ if for each bounded $Q \subset V$ with $\mathbf{v} \in Q$, there exists $\tau \geq 0$ such that

$$d_H(\overline{h}(Z), H(Z)) \leq \tau w(Z)^{\beta}, \quad \forall Z \in \mathbb{I}Q \text{ with } \mathbf{v} \in Z.$$

Moreover, H is said to have Hausdorff convergence of order $\beta > 0$ on V if it has Hausdorff convergence of order (at least) β at each $\mathbf{v} \in V$, with the constant τ independent of $\mathbf{v}$.

Definition 2.3.27. [Range Order] Let $V \subset \mathbb{R}^n$ be a bounded set. Let $f : V \to \mathbb{R}$ be continuous, and let F be an inclusion function for f on $\mathbb{I}V$. The inclusion function F is said to have range of order $\alpha > 0$ at a point $\mathbf{v} \in V$ if there exists $\tau \geq 0$ such that for every $Z \in \mathbb{I}V$ with $\mathbf{v} \in Z$,

$$w(F(Z)) \leq \tau w(Z)^{\alpha}.$$

The function f itself is said to have a range of order $\alpha > 0$ at $\mathbf{v} \in V$ if its image $\overline{f}$ has range of order α at $\mathbf{v}$. The functions F and f are said to have ranges of order $\alpha > 0$ on V if they have ranges of order (at least) α at each $\mathbf{v} \in V$, with the constant τ independent of $\mathbf{v}$.

2.3.2.1.2 Convex relaxation-based techniques

A second popular approach for constructing lower bounding problems for Problem (P) proceeds via the construction of convex underestimators/relaxations (and concave overestimators/relaxations) of the functions in Problem (P). The notions of convex and concave relaxations of a function are central to the developments in this thesis, and are formally defined below. Many of the definitions and results in this section are adapted from [38], and will especially lay the groundwork for the analyses in Chapters 5 and 6.

Definition 2.3.28. [Convex and Concave Relaxations] Given a convex set $Z \subset \mathbb{R}^n$ and a function $f : Z \to \mathbb{R}$, a convex function $f^{\text{cv}}_Z : Z \to \mathbb{R}$ is called a convex relaxation of f on Z if $f^{\text{cv}}_Z(\mathbf{z}) \leq f(\mathbf{z})$, $\forall \mathbf{z} \in Z$. Similarly, a concave function $f^{\text{cc}}_Z : Z \to \mathbb{R}$ is called a concave relaxation of f on Z if $f^{\text{cc}}_Z(\mathbf{z}) \geq f(\mathbf{z})$, $\forall \mathbf{z} \in Z$.

Given a (continuous) convex relaxation f^{cv} of the objective function f on $\text{conv}(X)$, (continuous) convex relaxations g^{cv}_j, for each $j \in \{1, \cdots, m_I\}$, of the corresponding inequality constraint functions g_j on $\text{conv}(X)$, and (continuous) convex and concave relaxations h^{cv}_k and h^{cc}_k, for each $k \in \{1, \cdots, m_E\}$, respectively, of the corresponding equality constraint functions h_k on $\text{conv}(X)$, the following convex optimization problem provides a valid lower bound on the optimal objective value of Problem (P):

$$\begin{aligned}
\min_{\mathbf{x}} \ & f^{\text{cv}}(\mathbf{x}) \\
\text{s.t. } & g_j^{\text{cv}}(\mathbf{x}) \leq 0, \quad \forall j \in \{1, \cdots, m_I\}, \\
& h_k^{\text{cv}}(\mathbf{x}) \leq 0, \ h_k^{\text{cc}}(\mathbf{x}) \geq 0, \quad \forall k \in \{1, \cdots, m_E\}, \\
& \mathbf{x} \in \text{conv}(X).
\end{aligned} \tag{P$^{\text{cv}}$}$$

Since the extent of the gap between convex and concave relaxations and the corresponding functions determines, in part, the efficiency of B&B algorithms, we are typically interested in constructing the tightest possible convex and concave relaxations of the functions in Problem (P). The next definition introduces the corresponding notion of convex and concave envelopes.

Definition 2.3.29. [Convex and Concave Envelopes] Given a convex set $Z \subset \mathbb{R}^n$ and a function $f : Z \to \mathbb{R}$, a convex function $f_Z^{\text{cv,env}} : Z \to \mathbb{R}$ is called the convex envelope of f on Z if $f_Z^{\text{cv,env}}$ is a convex relaxation of f on Z and for every convex relaxation $f_Z^{\text{cv}} : Z \to \mathbb{R}$, we have $f_Z^{\text{cv,env}}(\mathbf{z}) \geq f_Z^{\text{cv}}(\mathbf{z})$, $\forall \mathbf{z} \in Z$. Similarly, a concave function $f_Z^{\text{cc,env}} : Z \to \mathbb{R}$ is called the concave envelope of f on Z if $f_Z^{\text{cc,env}}$ is a concave relaxation of f on Z and for every concave relaxation $f_Z^{\text{cc}} : Z \to \mathbb{R}$, we have $f_Z^{\text{cc,env}}(\mathbf{z}) \leq f_Z^{\text{cc}}(\mathbf{z})$, $\forall \mathbf{z} \in Z$.

While the convex and concave envelopes of so-called 'elementary functions' are known on interval domains [142, 154] and the library of known convex envelopes is ever-increasing [121, 122, 147, 155, 157, 187, 221, 223], computing the envelopes of most functions is usually a nontrivial task since the envelopes of functions that are defined as compositions of other functions are not easily calculable in general. Consequently, global optimization algorithms and software work with popular and easily automatable techniques for the construction of convex and concave relaxations such as the auxiliary variable method [213], McCormick's relaxation technique and its generalizations [124, 154, 207, 227], αBB relaxations and its variants [3, 4, 6], and advanced relaxation strategies for problems with special structures [12, 45, 120, 161–163, 208, 209, 250].

The remainder of the results in this section are best viewed, in the context of this thesis, as building up the framework for the cluster problem and convergence order analyses of B&B algorithms for Problem (NLP). The following result establishes sufficient conditions for lower semicontinuity of the convex envelope. Note that a weaker version of this result is presented in [188, Corollary 17.2.1], and stronger versions of this result are stated without proof in [69, p. 349] (where the assumption that the function f is bounded above is relaxed) and in [222, p. 253] (where the assumptions that the function f is bounded above and the set W is bounded are relaxed).

Lemma 2.3.30. Let $Z \subset \mathbb{R}^n$ be a nonempty compact convex set and $f : Z \to \mathbb{R}$ be a lower semicontinuous function on Z bounded above by $M \in \mathbb{R}$. Let $f_Z^{\mathrm{cv,env}}$ denote the convex envelope of f on Z. Then $f_Z^{\mathrm{cv,env}}$ is lower semicontinuous on Z.

Proof. The function f is lower semicontinuous on the compact set Z if and only if its epigraph $\{(\mathbf{x}, r) : \mathbf{x} \in Z, r \geq f(\mathbf{x})\}$ is closed. Consequently, we have that the set $S := \{(\mathbf{x}, r) : \mathbf{x} \in Z, r \geq f(\mathbf{x}), r \leq M\}$ is compact. Theorem 17.2 in [188] implies that $\mathrm{conv}(S)$ is a compact convex set. Therefore, the set $\mathrm{conv}(S) \cup \{(\mathbf{x}, r) : \mathbf{x} \in Z, r \geq f(\mathbf{x})\}$ is closed, which implies that the epigraph of the convex envelope $\left\{(\mathbf{x}, r) : \mathbf{x} \in Z, r \geq f_Z^{\mathrm{cv,env}}(\mathbf{x})\right\}$ is closed, which in turn implies that $f_Z^{\mathrm{cv,env}}$ is lower semicontinuous on Z. $\square$

Because B&B algorithms involve the construction of convex and concave relaxations on successively refined partition elements, the following notion of schemes of relaxations [38] will prove useful in the analysis of their convergence rates.

Definition 2.3.31. [Schemes of Convex and Concave Relaxations] Let $V \subset \mathbb{R}^n$ be a nonempty convex set, and let $f : V \to \mathbb{R}$. Assume that for every $Z \in \mathbb{I}V$, we can construct functions $f_Z^{\mathrm{cv}} : Z \to \mathbb{R}$ and $f_Z^{\mathrm{cc}} : Z \to \mathbb{R}$ that are convex and concave relaxations, respectively, of f on Z. The sets of functions $(f_Z^{\mathrm{cv}})_{Z \in \mathbb{I}V}$ and $(f_Z^{\mathrm{cc}})_{Z \in \mathbb{I}V}$ define schemes of convex and concave relaxations, respectively, of f in V, and the set of pairs of functions $(f_Z^{\mathrm{cv}}, f_Z^{\mathrm{cc}})_{Z \in \mathbb{I}V}$ defines a scheme of relaxations of f in V. The schemes of relaxations are called continuous when f_Z^{cv} and f_Z^{cc} are continuous on Z for each $Z \in \mathbb{I}V$.

The next definition specializes the definition of Hausdorff convergence order of an inclusion function (see Definition 2.3.26) to schemes of relaxations.

Definition 2.3.32. [Hausdorff Convergence Order of Schemes of Relaxations] Let $V \subset \mathbb{R}^n$ be a nonempty convex set, and $f : V \to \mathbb{R}$ be a continuous function. Let $(f_Z^{\mathrm{cv}})_{Z \in \mathbb{I}V}$ and $(f_Z^{\mathrm{cc}})_{Z \in \mathbb{I}V}$ respectively denote schemes of convex and concave relaxations of f in V.

The scheme of relaxations $(f_Z^{\mathrm{cv}}, f_Z^{\mathrm{cc}})_{Z\in\mathbb{I}V}$ is said to have Hausdorff convergence of order $\beta > 0$ at $\mathbf{v} \in V$ if for each bounded $Q \subset V$ with $\mathbf{v} \in Q$, there exists $\tau \geq 0$ such that

$$d_H(\overline{f}(Z), H_f(Z)) \leq \tau w(Z)^{\beta}, \quad \forall Z \in \mathbb{I}Q \text{ with } \mathbf{v} \in Z,$$

where

$$H_f(Z) := \left[\inf_{\mathbf{z}\in Z} f_Z^{\mathrm{cv}}(\mathbf{z}), \sup_{\mathbf{z}\in Z} f_Z^{\mathrm{cc}}(\mathbf{z})\right], \quad \forall Z \in \mathbb{I}Q.$$

The scheme $(f_Z^{\mathrm{cv}}, f_Z^{\mathrm{cc}})_{Z\in\mathbb{I}V}$ is said to have Hausdorff convergence of order β on V if it has Hausdorff convergence of order (at least) β at each $\mathbf{v} \in V$, with a constant τ independent of $\mathbf{v}$.

The following definition of pointwise convergence order [38] (cf. Definition 6.3.4) provides a stronger notion of convergence order than Hausdorff convergence order (as demonstrated by Lemma 2.3.36; also see [38, Theorem 1]), and will be particularly useful in the analysis of the convergence rates of B&B algorithms for constrained problems in Chapter 6.

Definition 2.3.33. [Pointwise Convergence Order of Schemes of Relaxations] Let $V \subset \mathbb{R}^n$ be a nonempty convex set, and $f : V \to \mathbb{R}$ be a continuous function. Let $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}V}$ and $(f_Z^{\mathrm{cc}})_{Z\in\mathbb{I}V}$ respectively denote schemes of convex and concave relaxations of f in V. The scheme of convex relaxations $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}V}$ is said to have pointwise convergence of order $\gamma > 0$ at $\mathbf{v} \in V$ if for each bounded $Q \subset V$ with $\mathbf{v} \in Q$, there exists $\tau^{\mathrm{cv}} \geq 0$ such that[4]

$$\sup_{\mathbf{z}\in Z} |f(\mathbf{z}) - f_Z^{\mathrm{cv}}(\mathbf{z})| \leq \tau^{\mathrm{cv}} w(Z)^{\gamma}, \quad \forall Z \in \mathbb{I}Q \text{ with } \mathbf{v} \in Z.$$

Similarly, the scheme of concave relaxations $(f_Z^{\mathrm{cc}})_{Z\in\mathbb{I}V}$ is said to have pointwise convergence of order $\gamma > 0$ at $\mathbf{v} \in V$ if for each bounded $Q \subset V$ with $\mathbf{v} \in Q$, there exists $\tau^{\mathrm{cc}} \geq 0$ such that

$$\sup_{\mathbf{z}\in Z} |f_Z^{\mathrm{cc}}(\mathbf{z}) - f(\mathbf{z})| \leq \tau^{\mathrm{cc}} w(Z)^{\gamma}, \quad \forall Z \in \mathbb{I}Q \text{ with } \mathbf{v} \in Z.$$

The scheme of relaxations $(f_Z^{\mathrm{cv}}, f_Z^{\mathrm{cc}})_{Z\in\mathbb{I}V}$ is said to have pointwise convergence of order $\gamma > 0$ at $\mathbf{v} \in V$ if the corresponding schemes of convex and concave relaxations have pointwise convergence of orders (at least) γ at $\mathbf{v}$. Furthermore, the schemes of relaxations are said to have pointwise convergence of order $\gamma > 0$ on V if they have pointwise convergence of order

[4]While the use of the absolute value function in the following expressions is redundant, we use it instead of parentheses anyway because it looks better...

at least γ at each $\mathbf{v} \in V$, with constants τ^{cv} and τ^{cc} independent of $\mathbf{v}$.

The next definition presents a notion of convergence order of individual schemes of convex and concave relaxations [238] (cf. Definition 6.3.3) based on the notion of Hausdorff convergence order of a scheme of relaxations (see Definition 2.3.32), which will prove to be more relevant in the context of B&B algorithms for global optimization.

Definition 2.3.34. [Convergence Orders of Schemes of Convex and Concave Relaxations] Let $V \subset \mathbb{R}^n$ be a nonempty bounded convex set, and $f : V \to \mathbb{R}$ be a continuous function. Let $(f_Z^{\text{cv}})_{Z \in \mathbb{I}V}$ and $(f_Z^{\text{cc}})_{Z \in \mathbb{I}V}$ respectively denote continuous schemes of convex and concave relaxations of f in V.

The scheme of convex relaxations $(f_Z^{\text{cv}})_{Z \in \mathbb{I}V}$ is said to have convergence of order $\beta > 0$ at $\mathbf{v} \in V$ if there exists $\tau^{\text{cv}} \geq 0$ such that

$$\min_{\mathbf{z} \in Z} f(\mathbf{z}) - \min_{\mathbf{z} \in Z} f_Z^{\text{cv}}(\mathbf{z}) \leq \tau^{\text{cv}} w(Z)^{\beta}, \quad \forall Z \in \mathbb{I}V \text{ with } \mathbf{v} \in Z.$$

Similarly, the scheme of concave relaxations $(f_Z^{\text{cc}})_{Z \in \mathbb{I}V}$ is said to have convergence of order $\beta > 0$ at $\mathbf{v} \in V$ if there exists $\tau^{\text{cc}} \geq 0$ such that

$$\max_{\mathbf{z} \in Z} f_Z^{\text{cc}}(\mathbf{z}) - \max_{\mathbf{z} \in Z} f(\mathbf{z}) \leq \tau^{\text{cc}} w(Z)^{\beta}, \quad \forall Z \in \mathbb{I}V \text{ with } \mathbf{v} \in Z.$$

The schemes $(f_Z^{\text{cv}})_{Z \in \mathbb{I}V}$ and $(f_Z^{\text{cc}})_{Z \in \mathbb{I}V}$ are said to have convergence of order $\beta > 0$ on V if they have convergence of order (at least) β at each $\mathbf{v} \in V$, with the constants τ^{cv} and τ^{cc} independent of $\mathbf{v}$.

The following result from [102] will be useful in the convergence order analysis of lower bounding schemes for B&B algorithms.

Lemma 2.3.35. Let $Z \subset \mathbb{R}^n$ be nonempty, and functions $f : Z \to \mathbb{R}$ and $g : Z \to \mathbb{R}$ be bounded on Z. Then

$$\left|\sup_{\mathbf{z} \in Z} f(\mathbf{z}) - \sup_{\mathbf{z} \in Z} g(\mathbf{z})\right| \leq \sup_{\mathbf{z} \in Z} |f(\mathbf{z}) - g(\mathbf{z})|,$$
$$\left|\inf_{\mathbf{z} \in Z} f(\mathbf{z}) - \inf_{\mathbf{z} \in Z} g(\mathbf{z})\right| \leq \sup_{\mathbf{z} \in Z} |f(\mathbf{z}) - g(\mathbf{z})|.$$

Proof. See Proposition 11.7 in [102]. $\square$

Lemma 2.3.36. Let $V \subset \mathbb{R}^n$ be a nonempty compact convex set, and $(f_Z^{\text{cv}})_{Z \in \mathbb{I}V}$ and $(f_Z^{\text{cc}})_{Z \in \mathbb{I}V}$ respectively denote schemes of convex and concave relaxations of a bounded function $f : V \to \mathbb{R}$ in V. If either scheme has pointwise convergence of order $\gamma > 0$, it has convergence of order $\beta \geq \gamma$.

Proof. By noting from Definition 2.3.33 that

$$\sup_{\mathbf{z} \in Z} |f(\mathbf{z}) - f_Z^{\text{cv}}(\mathbf{z})| \leq \tau^{\text{cv}} w(V)^{\gamma}, \quad \forall Z \in \mathbb{I}V,$$
$$\sup_{\mathbf{z} \in Z} |f_Z^{\text{cc}}(\mathbf{z}) - f(\mathbf{z})| \leq \tau^{\text{cc}} w(V)^{\gamma}, \quad \forall Z \in \mathbb{I}V,$$

which implies that the schemes of convex and concave relaxations are bounded, the result follows from Lemma 2.3.35 via

$$\inf_{\mathbf{z} \in Z} f(\mathbf{z}) - \inf_{\mathbf{z} \in Z} f_Z^{\text{cv}}(\mathbf{z}) \leq \sup_{\mathbf{z} \in Z} |f(\mathbf{z}) - f_Z^{\text{cv}}(\mathbf{z})|, \quad \forall Z \in \mathbb{I}V,$$

and

$$\sup_{\mathbf{z} \in Z} f_Z^{\text{cc}}(\mathbf{z}) - \sup_{\mathbf{z} \in Z} f(\mathbf{z}) \leq \sup_{\mathbf{z} \in Z} |f_Z^{\text{cc}}(\mathbf{z}) - f(\mathbf{z})|, \quad \forall Z \in \mathbb{I}V. \qquad \square$$

Finally, we review a couple of key results from Bompadre and Mitsos [38] that will be relevant to our work on convergence order analysis in Chapter 6. The following two results together imply that the schemes of envelopes of a nonlinear twice continuously differentiable function have exactly second-order pointwise convergence (on nondegenerate intervals).

Theorem 2.3.37. Let $V \subset \mathbb{R}^n$ be a nonempty open bounded convex set, and $f : V \to \mathbb{R}$ be a nonlinear, twice continuously differentiable function. Let $(f_Z^{\text{cv}}, f_Z^{\text{cc}})_{Z \in \mathbb{I}V}$ denote a scheme of relaxations of f in V. Then the scheme $(f_Z^{\text{cv}}, f_Z^{\text{cc}})_{Z \in \mathbb{I}V}$ has pointwise convergence of order at most two on V.

Proof. See Theorem 2 in [38]. □

Theorem 2.3.38. Let $V \subset \mathbb{R}^n$ be a nonempty open bounded convex set, and $f : V \to \mathbb{R}$ be a twice continuously differentiable function. Let $(f_Z^{\text{cv,env}}, f_Z^{\text{cc,env}})_{Z \in \mathbb{I}V}$ denote the scheme of envelopes of f in V. Then the scheme $(f_Z^{\text{cv,env}}, f_Z^{\text{cc,env}})_{Z \in \mathbb{I}V}$ has pointwise convergence of order at least two on V.

Proof. See Theorem 10 in [38]. □

2.3.2.1.3 Lagrangian duality-based techniques

The last lower bounding technique that we consider in this chapter is based on Lagrangian duality (see [13, Chapter 6] and [30, Chapter 5] for additional information, especially for geometric interpretations of the Lagrangian dual problem). Given a particular instance of the constrained optimization problem, Problem (P), we can construct the following associated Lagrangian dual problem[5] that provides a guaranteed lower bound (see Theorem 2.3.39) on the optimal objective value of Problem (P):

$$\sup_{\boldsymbol{\mu} \geq \mathbf{0}, \boldsymbol{\lambda}} \min_{\mathbf{x} \in X} f(\mathbf{x}) + \boldsymbol{\mu}^{\mathrm{T}} \mathbf{g}(\mathbf{x}) + \boldsymbol{\lambda}^{\mathrm{T}} \mathbf{h}(\mathbf{x}). \tag{D}$$

Under appropriate convexity assumptions on the functions and sets in Problem (P) (which requires, in part, that Problem (P) contain no discrete variables in its formulation; recall the standing assumptions of the continuity of the functions f, $\mathbf{g}$, and $\mathbf{h}$ and the compactness of X) and additional constraint qualification assumptions, the 'primal problem', Problem (P), and the Lagrangian dual problem (hereafter simply referred to as the dual problem), Problem (D), have the same optimal objective function value (when these two values are equal, we say that 'strong duality' holds). The dual function, defined as $d : \mathbb{R}^{m_I}_{+} \times \mathbb{R}^{m_E} \to \mathbb{R}$ with $d : \mathbb{R}^{m_I}_{+} \times \mathbb{R}^{m_E} \ni (\boldsymbol{\mu}, \boldsymbol{\lambda}) \longmapsto \min_{\mathbf{x} \in X} f(\mathbf{x}) + \boldsymbol{\mu}^{\mathrm{T}} \mathbf{g}(\mathbf{x}) + \boldsymbol{\lambda}^{\mathrm{T}} \mathbf{h}(\mathbf{x})$ when X is nonempty, is a convex function that is typically nondifferentiable[6]; consequently, the outer maximization in Problem (D) is usually solved by applying a nonsmooth optimization algorithm to d [100]. Even when conditions for strong duality are not satisfied, solving the dual problem provides a valid lower bound as implied by the following theorem.

Theorem 2.3.39. [Weak Duality] The optimal objective value of Problem (D) provides a lower bound on the optimal objective value of Problem (P).

Proof. See Theorem 6.2.1 in [13] or Proposition 5.1.3 in [30], for instance. □

An immediate consequence of the above result is that not solving the outer maximization in Problem (D) to optimality still yields a valid lower bound for Problem (P) (this is particularly useful from a practical viewpoint because numerical nonsmooth optimization

[5] In fact, depending on which constraints are 'dualized', we can construct several Lagrangian dual problems (cf. Section 2.3.3.1.2); however, we only consider one particular way of constructing the Lagrangian dual problem in this section.

[6] Although the dual function is typically nondifferentiable only on a set of measure zero, it is usually nondifferentiable at optimal solutions to Problem (D).

algorithms are not yet as computationally efficient as their smooth counterparts). While it is not apparent how solving Problem (D) is computationally relevant for nonconvex NLPs and MINLPs (since the inner minimization of Problem (D) is itself a nonconvex NLP/MINLP in these cases), Section 2.3.3.1.2 presents one application of Lagrangian duality-based lower bounds for stochastic programming problems, where the dual lower bounding problem can exploit the nearly-decomposable structure of stochastic programs to generate lower bounds in a 'computationally scalable' manner. Another advantage of solving the Lagrangian dual problem is that it usually provides tighter lower bounds than the convex relaxation-based lower bounding approaches described in Section 2.3.2.1.2, as seen from the following result (also see [71]). A formal proof of this result is deferred to Lemma 6.4.20 in Chapter 6 where it fits in more naturally.

Lemma 2.3.40. Let f^{cv} and $\mathbf{g}^{\mathrm{cv}}$ denote (any) convex relaxations of f and $\mathbf{g}$, respectively, on $\mathrm{conv}(X)$, and let $\mathbf{h}^{\mathrm{cv}}$ and $\mathbf{h}^{\mathrm{cc}}$ denote (any) convex and concave relaxations, respectively, of $\mathbf{h}$ on $\mathrm{conv}(X)$. Assume that strong duality holds for the convex relaxation-based lower bounding problem that uses the relaxations f^{cv}, $\mathbf{g}^{\mathrm{cv}}$, $\mathbf{h}^{\mathrm{cv}}$, and $\mathbf{h}^{\mathrm{cc}}$ in its construction (see Problem (P^{cv})). Then the lower bound obtained by solving Problem (D) is at least as strong as that obtained by solving the above convex relaxation-based lower bounding problem.

2.3.2.2 Bounds tightening techniques

Bounds tightening/domain reduction techniques typically play a crucial role in accelerating the convergence of B&B algorithms for Problem (P) by discarding regions of the search space that are guaranteed to exclude optimal solutions, and can enable the solution of otherwise challenging instances of Problem (P) in reasonable times [19, 182]. Bounds tightening techniques not only help shrink the search space for optimization algorithms, but also play an important role in generating tight bounds by means of generating tighter relaxations of the nonconvex functions involved, for instance. Several bounds tightening techniques have been proposed in the literature for nonconvex MINLPs [19, 182, 224] that are usually either feasibility-based [17, 94], or optimality-based [194, 246]. This section reviews some bounds tightening techniques that are relevant to this thesis. Section 2.3.3.1.1 and Chapter 3 detail additional tailored domain reduction techniques for two-stage stochastic programming problems.

The least computationally expensive bounds tightening technique considered in this thesis is a feasibility-based bounds tightening (FBBT) technique called forward-backward interval propagation [234, 239]. In this approach, interval bounds on the variables are first propagated forward through a computational graph representation of Problem (P) to deduce interval bounds on intermediate node and constraint expressions using interval arithmetic [172]. Next, constraint information (such as the types of constraints and their 'right-hand side values') is enforced to potentially shrink the estimated ranges of the constraint expressions. Finally, the updated interval bounds on the constraint expressions are propagated backward through the graph, by applying interval-based inverses of operations, to potentially determine tighter bounds on the problem variables. Since propagating interval bounds is a relatively cheap operation, the above procedure is usually repeated until the improvement in variables' bounds falls below a predefined threshold. We consider the following example to illustrate an application of the forward-backward interval propagation technique (also see Section 2.2.1 and Figure 3 in [233]).

Example 2.3.41. Consider the problem:

$$\begin{aligned} \min_{x,y} \quad & x + y \\ \text{s.t.} \quad & xy = 1, \\ & x^2 + y^2 \leq 100, \\ & x \in [0, 10], \;\; y \in [0, 10]. \end{aligned}$$

Figure 2-1 depicts a computational graph representation of the above problem (with `EQL` corresponding to the equality constraint, `OBJ` corresponding to the objective, and `LEQ` corresponding to the inequality constraint in the above example). Figure 2-2 notes the intervals obtained after a single round of application of the above-described forward-backward interval propagation technique. The reader can verify that the resulting interval bounds on the variables provide a tight description of the feasible set in this case.

Although tight interval bounds were obtained using the forward-backward interval propagation technique for Example 2.3.41 in a computationally inexpensive manner, a key limitation of the technique is that its application usually does not yield such a tight description of the feasible region in practical applications.

A second feasibility-based bounds tightening technique that we consider in this thesis

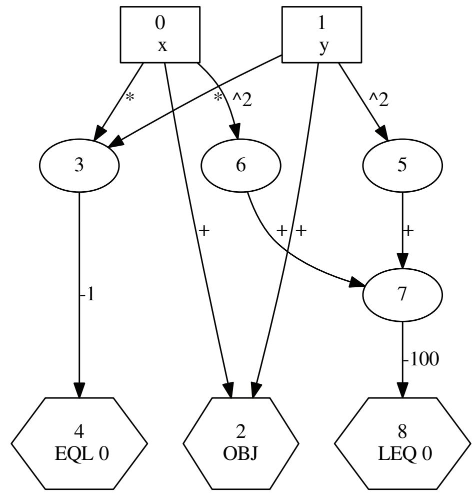


Figure 2-1: DAG representation of the instance of Problem (P) in Example 2.3.41.

involves the solution of auxiliary (convex) optimization problems. A lower bound on the i^{th} variable x_i may be obtained by solving the following convex problem that uses feasibility arguments to exclude regions of the search space (cf. Problem (P^{cv})):

$$\begin{aligned} \min_{\mathbf{x}} \ & x_i \\ \text{s.t. } & g_j^{\text{cv}}(\mathbf{x}) \leq 0, \quad \forall j \in \{1, \cdots, m_I\}, \\ & h_k^{\text{cv}}(\mathbf{x}) \leq 0, \ h_k^{\text{cc}}(\mathbf{x}) \geq 0, \quad \forall k \in \{1, \cdots, m_E\}, \\ & \mathbf{x} \in \text{conv}(X), \end{aligned} \tag{FBBT}$$

where $\mathbf{g}^{\text{cv}}$ denotes a (continuous) convex relaxation of $\mathbf{g}$ on $\text{conv}(X)$, and $\mathbf{h}^{\text{cv}}$ and $\mathbf{h}^{\text{cc}}$ denote (continuous) convex and concave relaxations, respectively, of $\mathbf{h}$ on $\text{conv}(X)$. An upper bound on x_i can be obtained by maximizing the objective in Problem (FBBT) instead of minimizing it. We note that tighter bounds can potentially be obtained by imposing the integrality restrictions on the discrete variables in Problem (P) and solving a convex MINLP. Additionally, a computationally less expensive technique to obtain lower bounds on x_i is to further relax Problem (FBBT) to a linear program. In this thesis, we only solve Problem (FBBT) (or its variants) to try and tighten the bounds on the variables that participate in the construction of relaxations for the functions in Problem (P).

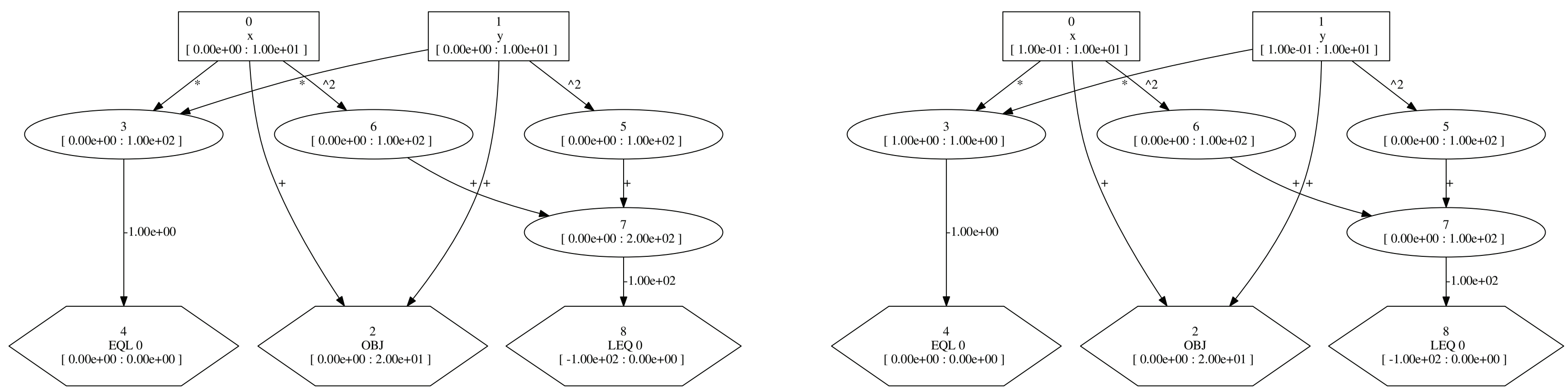


Figure 2-2: Interval bounds after one round of forward-backward interval propagation for Example 2.3.41. Left plot: interval bounds on the node expressions after one round of forward interval propagation. Right plot: interval bounds on the node expressions after the first round of backward interval propagation.

If a valid upper bound UBD on the optimal objective value of Problem (P) is available, then tighter bounds on the problem variables can potentially be obtained by appending the constraint $f^{\text{cv}}(\mathbf{x}) \leq UBD$ to Problem (FBBT), where f^{cv} denotes a (continuous) convex relaxation of the objective function f on $\text{conv}(X)$. We refer to the above technique (and its variants, see Section 2.3.3.1.1) as 'optimality-based bounds tightening' (OBBT).

2.3.2.3 Full-space vs reduced-space branch-and-bound algorithms

As outlined in Section 2.3.2, deterministic global optimization algorithms for nonconvex problems usually involve the concept of partitioning ('branching on') the domain of the decision variables [101]. Since the worst-case running time of all known branch-and-bound algorithms is exponential in the dimension of the variables partitioned, it may be advantageous to utilize 'reduced-space' algorithms which only require branching on a subset of the variables (as opposed to 'full-space' branch-and-bound algorithms which may branch on all of the variables) to guarantee convergence. Despite the potential advantages of reduced-space algorithms for nonconvex problems [20, 69, 76, 169, 217, 237], such methods have not been widely adopted in the literature and in commercial software. One potential reason is that widely-applicable reduced-space branch-and-bound algorithms often do not seem to exhibit favorable convergence rates compared to their full-space counterparts (see Chapter 6). The reader is directed to references [42, 76, 217] for a list of reduced-space branch-and-bound algorithms in the literature.

In this section, we consider the following problem formulation:

$$\begin{aligned} \min_{\mathbf{x},\mathbf{y}} \ & f(\mathbf{x},\mathbf{y}) \\ \text{s.t.} \ & \mathbf{g}(\mathbf{x},\mathbf{y}) \leq \mathbf{0}, \\ & \mathbf{h}(\mathbf{x},\mathbf{y}) = \mathbf{0}, \\ & \mathbf{x} \in X, \ \ \mathbf{y} \in Y, \end{aligned} \tag{RS}$$

where $X \subset \mathbb{R}^{n_x}$ and $Y \subset \mathbb{R}^{n_y}$ are nonempty compact intervals, $f : X \times Y \to \mathbb{R}$ and $\mathbf{g} : X \times Y \to \mathbb{R}^{m_I}$ are partially convex with respect to $\mathbf{x}$, i.e., $f(\cdot,\mathbf{y})$ and $\mathbf{g}(\cdot,\mathbf{y})$ are convex on X for each $\mathbf{y} \in Y$, and $\mathbf{h} : X \times Y \to \mathbb{R}^{m_E}$ is affine with respect to $\mathbf{x}$, i.e., $\mathbf{h}(\cdot,\mathbf{y})$ is affine on X for each $\mathbf{y} \in Y$. Additionally, we assume that the functions f, $\mathbf{g}$, and $\mathbf{h}$ are continuous on $X \times Y$.

When the dimension n_y of the Y-space corresponding to the nonconvexities in the functions in Problem (RS) is significantly smaller than the dimension n_x of the X-space, it may be computationally advantageous to partition only the Y-space during the course of a branch-and-bound algorithm (assuming, of course, that the reduced-space algorithm is guaranteed to converge). However, the convergence rate of a reduced-space branch-and-bound algorithm may be different compared to a similar full-space algorithm (see Chapter 6), which makes it difficult to judge *a priori* whether using a reduced-space branch-and-bound approach would be advantageous. Chapter 6 investigates the convergence orders of a convex relaxation-based reduced-space lower bounding scheme for a subclass of Problem (RS) [76] and a Lagrangian duality-based reduced-space lower bounding scheme [69, Section 3.3] for Problem (RS), and determines necessary and sufficient conditions under which these reduced-space lower bounding schemes have favorable convergence properties.

Algorithm 2.2 outlines a generic reduced-space branch-and-bound algorithm for Problem (RS). It should be noted that Algorithm 2.2 merely provides the backbone of a generic reduced-space branch-and-bound algorithm. In practice, the order in which the subproblems are solved may vary and additional subproblems may be solved to speed up the convergence of the algorithm. The reader can compare and contrast Algorithm 2.2 with the full-space branch-and-bound Algorithm 2.1. A key distinction between the two algorithms is that the branching step in Algorithm 2.2 only partitions the domains of the $\mathbf{y}$ variables, whereas the branching step in Algorithm 2.1 may partition the domains of all of the variables. We direct the reader to references [76] and [69] for two widely-applicable instances of Algorithm 2.2, and for examples of their application.

2.3.2.4 Convergence of branch-and-bound algorithms

In this section, we review some background definitions and results, primarily from Chapter IV of Horst and Tuy [101], that form the basis of a convergence theory for a wide class of B&B algorithms (such as Algorithm 2.1 for Problem (P) and Algorithm 2.2 for Problem (RS)). The definitions and results in this section will be used to establish the convergence of a (reduced-space) branch-and-bound algorithm in Chapter 3 for solving a class of two-stage stochastic MINLPs. For ease of exposition, the results in this section will be presented in the context of the reduced-space formulation Problem (RS). The first definition imposes a requirement on the branching step.

Algorithm 2.2 Outline of a generic reduced-space B&B algorithm for Problem (RS)

Initialize:

a. Iteration counter $k = 0$, interval bounds X^0 and Y^0 on $\mathbf{x}$ and $\mathbf{y}$, respectively, after the optional application of reduced-space preprocessing techniques to the input data, and tolerances $\varepsilon > 0, \varepsilon^l > 0$, and $\varepsilon^u > 0$ such that $\varepsilon^l + \varepsilon^u \leq \varepsilon$.

b. Domain of the root node $M^0 := X^0 \times Y^0$, and the initial partition $\mathcal{P}^0 = \{M^0\}$.

c. Objective function value of the best found feasible point, $UBD = +\infty$; lower bound on the optimal objective value on the root node, $LBD^0 = -\infty$; and the best found feasible point, $\{\mathbf{x}^*, \mathbf{y}^*\} = \emptyset$.

repeat

1. (Node Selection) Pick an active node $n \in \left\{n \in \mathbb{N} \cup \{0\} : M^n \in \mathcal{P}^k\right\}$ using some node selection heuristic and set $\mathcal{P}^{k+1} = \mathcal{P}^k \backslash \{M^n\}$.
2. (Optional Upper Bounding Step) Solve an upper bounding problem on M^n with a termination tolerance of ε^u to try and determine a feasible point. Update UBD, $(\mathbf{x}^*, \mathbf{y}^*)$ if a feasible solution better than the current best solution is obtained.
3. (Optional Bounds Tightening Step) Apply finite reduced-space bounds tightening techniques to obtain $\bar{X}^n \subset X^n$ and $\bar{Y}^n \subset Y^n$, and set $X^n = \bar{X}^n$, $Y^n = \bar{Y}^n$. If either X^n or Y^n is empty, **goto** Step 6. Otherwise, update $M^n = X^n \times Y^n$.
4. (Lower Bounding Step) Solve a reduced-space lower bounding problem on M^n to ε^l-optimality to obtain the lower bound LBD^n (if the lower bounding problem on M^n is infeasible, set $LBD^n = +\infty$). If node n can be fathomed, **goto** Step 6.
5. (Branching Step) Partition M^n into M^{n_1} and M^{n_2} by branching only on the Y-space. Set $\mathcal{P}^{k+1} = \mathcal{P}^{k+1} \cup \{M^{n_1}\} \cup \{M^{n_2}\}$, $LBD^{n_1} = LBD^{n_2} = LBD^n$.
6. Set $\mathcal{P}^{k+1} = \mathcal{P}^{k+1} \backslash \left\{M^p \in \mathcal{P}^{k+1} : LBD^p \geq UBD - \varepsilon\right\}$, $k = k + 1$.

until $\mathcal{P}^k = \emptyset$

If $UBD < +\infty$, then $(\mathbf{x}^*, \mathbf{y}^*)$ provides an ε-optimal solution to Problem (P).

Definition 2.3.42. [Exhaustive Partitioning] Given sets $X \subset \mathbb{R}^{n_x}$ and $Y \in \mathbb{IR}^{n_y}$, a subdivision of $X \times Y$ is said to be exhaustive on Y if every infinite decreasing sequence $\{M^n\} := \{(X^n \times Y^n)\}$ of successively refined partition elements produced by the subdivision process satisfies $\lim\limits_{n\to\infty} w(Y^n) = 0$.

The above definition is borrowed from [76], and is an extension of Definition IV.10 in [101]. It is clear that $\lim\limits_{n\to\infty} Y^n = \bigcap\limits_{n=1}^{\infty} Y^n = \{\bar{\mathbf{y}}\}$ for some $\bar{\mathbf{y}} \in Y$ when the subdivision process is exhaustive on Y (see Theorem 3.10 in [192]). The next result follows.

Lemma 2.3.43. If the subdivision process is exhaustive on Y, then $\forall \delta > 0, \exists N_\delta \in \mathbb{N}$ such that $n \geq N_\delta \implies w(Y^n) < \delta$.

Proof. The claim follows from Definition 2.3.42 and the notion of convergence. $\square$

The next couple of definitions introduce conditions on B&B bounding schemes that will later be leveraged to establish convergence of B&B algorithms for Problem (RS).

Definition 2.3.44. [Consistent Bounding Operation] A bounding operation is called consistent if, at every step, any unfathomed partition element can be further refined, and if any infinite decreasing sequence $\{M^n\} := \{(X^n \times Y^n)\}$ of successively refined partition elements satisfies

$$\lim_{n\to\infty} (UBD^n - LBD^n) = 0,$$

where UBD^n is the overall upper bound after step n of the B&B procedure, and LBD^n is the lower bound over the domain M^n.

We include the following remark to help unpack the above definition.

Remark 2.3.45. Any infinite decreasing sequence $\{M^n\}$ of successively refined partition elements has to obey $LBD^n < UBD^n$, $\forall n \in \mathbb{N}$, since any partition element that does not satisfy this condition would otherwise had to have been fathomed by value dominance.

Definition 2.3.46. [Strongly Consistent Bounding Operation] Suppose we are given an infinite decreasing sequence of successively refined partition elements $\{M^n\} := \{(X^n \times Y^n)\}$ produced by a subdivision of Y that is exhaustive on Y with $\lim_{n\to\infty} Y^n = \{\bar{\mathbf{y}}\}$, and satisfying $X^{n+1} \subset X^n$, $\forall n \in \mathbb{N}$. Define $X^\infty := \cap_{n=1}^{\infty} X^n$. A lower bounding procedure for the sequence of restricted optimization problems (RS)

$$\begin{aligned} \min_{(\mathbf{x},\mathbf{y})\in M^n} \quad & f(\mathbf{x},\mathbf{y}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{x},\mathbf{y}) \leq \mathbf{0}, \\ & \mathbf{h}(\mathbf{x},\mathbf{y}) = \mathbf{0} \end{aligned}$$

that yields a sequence of bounds $\{LBD^n\}$ is said to be strongly consistent on Y if there exists a subsequence of partition elements $\{M^{n_q}\} := \{(X^{n_q} \times Y^{n_q})\}$ that satisfies

$$\begin{aligned} \lim_{q\to\infty} LBD^{n_q} = \min_{\mathbf{x}\in X^\infty} \quad & f(\mathbf{x},\bar{\mathbf{y}}) \\ \text{s.t.} \quad & \mathbf{g}(\mathbf{x},\bar{\mathbf{y}}) \leq \mathbf{0}, \\ & \mathbf{h}(\mathbf{x},\bar{\mathbf{y}}) = \mathbf{0}. \end{aligned}$$

Remark 2.3.47. If the subdivision process is exhaustive on X as well with $X^{\infty} = \{\bar{\mathbf{x}}\}$ in Definition 2.3.46, we recover, in effect, the conventional definition of a strongly consistent bounding operation, see Definition IV.7 in [101].

Definition 2.3.48. [Bound Improving Selection] A node selection operation is said to be bound improving if, at least each time after a finite number of steps, the partition element selected for further partitioning at the current step k, $\mathcal{Q}^k$, corresponds to the domain of a node on which the current overall lower bound for the B&B tree is attained, i.e.,

$$\mathcal{Q}^k \cap \underset{M^n \in \mathcal{P}^k}{\arg\min}\, LBD^n \neq \emptyset,$$

where $\mathcal{P}^k$ denotes the unfathomed partition elements at the start of step k of the B&B algorithm.

The convergence proofs of the algorithms we consider in this thesis will require that every infinite decreasing sequence of successively refined partition elements $\{M^n\}$ generated by the algorithm should satisfy $M^n \cap \mathcal{F}(X \times Y) \neq \emptyset$, $\forall n \in \mathbb{N}$, where $\mathcal{F}(X \times Y)$ denotes the feasible set of Problem (RS) on the initial domain $X \times Y$. We call this requirement 'deletion by infeasibility is certain in the limit' (see [101, Definition IV.8]).

The following two results, proofs of which can be found in [101, Chapter IV], will prove useful towards establishing the convergence of generic B&B algorithms.

Theorem 2.3.49. Consider the sequence of optimization problems in Definition 2.3.46, and suppose the B&B subdivision procedure is exhaustive (on Y). Furthermore, suppose every infinite decreasing sequence $\{M^n\}$ $(= \{(X^n \times Y^n)\})$ of successively refined partition elements satisfies $M^n \cap \mathcal{F}(X \times Y) \neq \emptyset$, $\forall n \in \mathbb{N}$ (i.e., the deletion by infeasibility rule is certain in the limit). Then every strongly consistent pair of lower and upper bounding operations yields a consistent bounding operation.

Proof. See Lemma IV.5. of [101]. □

Theorem 2.3.50. Consider the sequence of optimization problems in Definition 2.3.46. Suppose that for a B&B procedure, the bounding operation is consistent and the selection operation is bound improving. Then the procedure is convergent, i.e.,

$$UBD^* := \lim_{n\to\infty} UBD^n = \min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}(X\times Y)} f(\mathbf{x},\mathbf{y}) = \lim_{n\to\infty} LBD^n_{\text{overall}} =: LBD^*,$$

where $\mathcal{F}(X \times Y)$ is the feasible set of the optimization problem on $X \times Y$, UBD^n and LBD^n_{overall} are the overall upper and lower bounds after step n of the B&B procedure, and UBD^* and LBD^* are the respective values to which they converge (see [192, Theorem 3.14] for the proof of existence of UBD^* and LBD^*).

Proof. See Theorem IV.3. of [101]. □

2.3.3 Optimization under uncertainty

Real-life optimization models often contain uncertain model parameters. Optimal solutions to models which simply consider a single realization of the uncertain parameters can be quite sensitive to the chosen parameter values, potentially rendering these solutions economically worthless and even disastrous in safety-critical applications. While there are apparent advantages of considering uncertainties in optimization models rigorously, accounting for their effects using off-the-shelf global optimization software is computationally prohibitive due to the inherent nonlinear and combinatorial nature of the associated models. This motivates the development of efficient algorithms and software for the solution of optimization problems with parametric uncertainty for the applications of interest. We refer the interested reader to Section 1.1.3 of Chapter 1 for popular approaches to model optimization problems under uncertainty.

Stochastic programming [35, 181] and robust optimization [21] are two widely adopted approaches for solving optimization problems with parametric uncertainty. The stochastic programming framework of interest assumes that the uncertain parameters can take on one of a finite number of values, each with a known probability (these 'scenarios' are obtained in practice by sampling an estimated distribution of the uncertain parameters). Robust optimization-based approaches, on the other hand, require guaranteed satisfaction of constraints for all possible realizations of the uncertain parameters. In the next section, we present the stochastic programming formulation of interest and review prior decomposition approaches for the scalable solution of two-stage stochastic mixed-integer nonlinear programs. The reader is directed to [65] and the references therein for solution techniques for 'static' robust optimization formulations (referred to therein as semi-infinite programs).

2.3.3.1 Two-stage stochastic programming

In this section, we consider the following general class of nonconvex two-stage stochastic mixed-integer nonlinear programs with recourse:

$$\begin{aligned} \inf_{\mathbf{y},\mathbf{z}} \; & f^{(1)}(\mathbf{y},\mathbf{z}) + \mathbb{E}_{\boldsymbol{\omega}\in\Omega}(R(\mathbf{y},\mathbf{z},\boldsymbol{\omega})) \\ \text{s.t. } & (\mathbf{y},\mathbf{z}) \in \mathcal{F}_{\mathrm{FS}}, \end{aligned} \tag{P}$$

where

$$\begin{aligned} R(\mathbf{y},\mathbf{z},\boldsymbol{\omega}) := \inf_{\mathbf{x}} \; & f^{(2)}(\mathbf{x},\mathbf{y},\mathbf{z},\boldsymbol{\omega}) \\ \text{s.t. } & \mathbf{g}(\mathbf{x},\mathbf{y},\mathbf{z},\boldsymbol{\omega}) \leq \mathbf{0}, \\ & \mathbf{x} \in X(\boldsymbol{\omega}), \end{aligned} \tag{R}$$

$X(\boldsymbol{\omega}) := \{\mathbf{x} \in \{0,1\}^{n_{x_b}} \times \Pi_x(\boldsymbol{\omega}) : \mathbf{r}_x(\mathbf{x},\boldsymbol{\omega}) \leq \mathbf{0}\}$, $\forall \boldsymbol{\omega} \in \Omega$, $\Pi_x : \Omega \to \mathcal{P}(\mathbb{R}^{n_{x_c}})$ is assumed to be defined such that $\Pi_x(\boldsymbol{\omega})$ is convex $\forall \boldsymbol{\omega} \in \Omega$, $\mathcal{P}(S)$ denotes the power set of S, $Y = \{\mathbf{y} \in \{0,1\}^{n_y} : \mathbf{r}_y(\mathbf{y}) \leq \mathbf{0}\}$, $Z = \{\mathbf{z} \in \Pi_z : \mathbf{r}_z(\mathbf{z}) \leq \mathbf{0}\}$, $\Pi_z \subset \mathbb{R}^{n_z}$ is convex, $f^{(1)} : [0,1]^{n_y} \times \Pi_z \to \mathbb{R}$, $f^{(2)} : [0,1]^{n_{x_b}} \times \bar{\Pi}_x(\Omega) \times [0,1]^{n_y} \times \Pi_z \times \Omega \to \mathbb{R}$, $\mathbf{g} : [0,1]^{n_{x_b}} \times \bar{\Pi}_x(\Omega) \times [0,1]^{n_y} \times \Pi_z \times \Omega \to \mathbb{R}^m$, $\mathbf{r}_x : [0,1]^{n_{x_b}} \times \bar{\Pi}_x(\Omega) \times \Omega \to \mathbb{R}^{m_x}$, $\bar{\Pi}_x(\Omega) \subset \mathbb{R}^{n_{x_c}}$ is such that $\Pi_x(\boldsymbol{\omega}) \subset \bar{\Pi}_x(\Omega)$, $\forall \boldsymbol{\omega} \in \Omega$, $\mathbf{r}_y : [0,1]^{n_y} \to \mathbb{R}^{m_y}$, $\mathbf{r}_z : \Pi_z \to \mathbb{R}^{m_z}$, $\boldsymbol{\omega}$ is a random variable from a probability space $(\Omega, \mathcal{F}_e, \mathbb{P})$ with $\Omega \subset \mathbb{R}^u$, $\mathbb{E}$ denotes the expected value operator,

$$\begin{aligned} \mathcal{F}_{\mathrm{FS}} := \big\{ (\mathbf{y},\mathbf{z}) \in (Y \times Z) \; : \; \mathbf{r}_{y,z}(\mathbf{y},\mathbf{z}) \leq \mathbf{0} \text{ and, for all possible events } \boldsymbol{\omega} \in \Omega, \\ \exists \mathbf{x}(\boldsymbol{\omega}) \in X(\boldsymbol{\omega}) : \mathbf{g}(\mathbf{x}(\boldsymbol{\omega}),\mathbf{y},\mathbf{z},\boldsymbol{\omega}) \leq \mathbf{0} \big\} \end{aligned}$$

is the projected feasible set in the domain of the complicating variables, $\mathbf{r}_{y,z} : [0,1]^{n_y} \times \Pi_z \to \mathbb{R}^{m_{y,z}}$, and the functions $\mathbf{r}_y, \mathbf{r}_z, \mathbf{r}_{y,z}$ are assumed to be continuous. The variables $\mathbf{y}$ and $\mathbf{z}$ denote the discrete and continuous first-stage/complicating decisions, respectively, that are made before the realization of the uncertainties, while the mixed-integer variables $\mathbf{x}$ denote the second-stage/recourse decisions made after the realization of the uncertain parameters $\boldsymbol{\omega}$. The term $f^{(1)}(\mathbf{y},\mathbf{z})$ represents the cost associated with the first-stage decisions, and the term $R(\mathbf{y},\mathbf{z},\boldsymbol{\omega})$ represents the optimal recourse cost for a given first-stage decision vector $(\mathbf{y},\mathbf{z})$ and a realization of the uncertain parameters $\boldsymbol{\omega}$.

As discussed previously, typical approaches to solve Problem (P) rely on a scenario representation of the uncertainties, where a finite number, s, of possible realizations of $\boldsymbol{\omega}$ are considered. We also make such an assumption in this thesis.

Assumption 2.3.51. The random variable $\boldsymbol{\omega}$ has finite support, i.e., $\Omega = \{\boldsymbol{\omega}_1, \cdots, \boldsymbol{\omega}_s\}$ with $\mathbb{P}(\boldsymbol{\omega} = \boldsymbol{\omega}_h) = p_h > 0$, $\forall h \in \{1, \cdots, s\}$.

Based on the above assumption, the two-stage stochastic program with recourse, Problem (P), can be equivalently written in extensive form as a so-called deterministic equivalent program (DEP):

$$\begin{aligned}
\inf_{\mathbf{x}_1, \cdots, \mathbf{x}_s, \mathbf{y}, \mathbf{z}} \quad & \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}) \\
\text{s.t.} \quad & \mathbf{g}_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\}, \\
& \mathbf{r}_{y,z}(\mathbf{y}, \mathbf{z}) \leq \mathbf{0}, \\
& \mathbf{x}_h \in X_h, \quad \forall h \in \{1, \cdots, s\}, \\
& \mathbf{y} \in Y, \ \ \mathbf{z} \in Z,
\end{aligned} \tag{DEP}$$

where $\mathbf{x}_h$ denotes $\mathbf{x}(\boldsymbol{\omega}_h)$, $X_h = \{\mathbf{x}_h \in \{0,1\}^{n_{x_b}} \times \Pi_{x,h} : \mathbf{r}_{x,h}(\mathbf{x}_h) \leq \mathbf{0}\}$ denotes $X(\boldsymbol{\omega}_h)$, $\Pi_{x,h}$ denotes $\Pi_x(\boldsymbol{\omega}_h)$, $f_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}) := f^{(1)}(\mathbf{y}, \mathbf{z}) + f_h^{(2)}(\mathbf{x}_h, \mathbf{y}, \mathbf{z})$, $\forall (\mathbf{x}_h, \mathbf{y}, \mathbf{z}) \in [0,1]^{n_{x_b}} \times \Pi_{x,h} \times [0,1]^{n_y} \times \Pi_z$, is defined for notational convenience, $f_h^{(2)}(\mathbf{x}_h, \mathbf{y}, \mathbf{z})$, $\mathbf{g}_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z})$, and $\mathbf{r}_{x,h}(\mathbf{x}_h)$ are used to denote $f^{(2)}(\mathbf{x}(\boldsymbol{\omega}_h), \mathbf{y}, \mathbf{z}, \boldsymbol{\omega}_h)$, $\mathbf{g}(\mathbf{x}(\boldsymbol{\omega}_h), \mathbf{y}, \mathbf{z}, \boldsymbol{\omega}_h)$, and $\mathbf{r}_x(\mathbf{x}(\boldsymbol{\omega}_h), \boldsymbol{\omega}_h)$, respectively, and the functions $f_h : [0,1]^{n_{x_b}} \times \Pi_{x,h} \times [0,1]^{n_y} \times \Pi_z \to \mathbb{R}$, $\mathbf{g}_h : [0,1]^{n_{x_b}} \times \Pi_{x,h} \times [0,1]^{n_y} \times \Pi_z \to \mathbb{R}^m$, and $\mathbf{r}_{x,h} : [0,1]^{n_{x_b}} \times \Pi_{x,h} \to \mathbb{R}^{m_x}$ are assumed to be continuous $\forall h \in \{1, \cdots, s\}$. Equality constraints in the formulation are assumed to be modeled using a pair of inequalities and bounded general integer variables are assumed to be equivalently reformulated using binary variables in Problem (DEP) purely for ease of exposition. Additionally, the following assumptions are made throughout this section.

Assumption 2.3.52. The sets Y and Z are nonempty, and the set Z is compact.

Assumption 2.3.53. The set X_h is nonempty and compact for each $h \in \{1, \cdots, s\}$.

Remark 2.3.54. Assumptions 2.3.52 and 2.3.53 along with the assumption of continuity of the functions in Problem (DEP) imply, by Weierstrass' theorem, that Problem (DEP) either has a finite optimal objective value or is infeasible.

Due to their special decomposable structure, two-stage stochastic programs have traditionally been solved using duality-based decomposition techniques with the advantage that the solution time typically scales linearly with the number of scenarios. One class of such methods, which is applicable to a class of separable stochastic programs with recourse problems that are linear programs (i.e., Problem (DEP) reduces to a linear program (LP) when the first-stage variables $\mathbf{y}$ and $\mathbf{z}$ are fixed; note that the above assumption necessitates $n_{x_b} = 0$), is Benders decomposition (BD) [25], also called the L-shaped method in the stochastic programming literature [231]. Over the years, several modifications and extensions to Benders' method have been proposed, including techniques to accelerate the convergence of BD by strengthening the formulation of Problem (DEP) and generating strong cutting planes [150], a multi-cut version [36] that seeks to decrease the number of iterations of the algorithm, the integer L-shaped method [132] and a reformulation-linearization technique-based (RLT-based) procedure [211] that extend the applicability of BD to a subclass of problems with integer recourse variables, and combinatorial Benders' cuts [58] that avoid the use of 'big M' coefficients in problems involving logical constraints. The reader is directed to references [35, 125, 196] for an overview of algorithms and software for stochastic mixed-integer linear programs (MILPs).

Geoffrion [85] generalized Benders' decomposition technique to a class of stochastic programs with nonlinear convex recourse programs (using nonlinear duality theory) and called the resulting method 'generalized Benders decomposition' (GBD). Note that the GBD algorithm for Problem (DEP) essentially boils down to Benders' decomposition algorithm when the assumptions for GBD are satisfied and all of the functions in Problem (DEP) are affine functions of the recourse variables. The above methods, however, usually rely on strong duality for convergence, which is not guaranteed for most problems that involve nonconvex functions in their formulation.

Another class of duality-based decomposition methods, termed Lagrangian relaxation (LR) [48, 91] (also sometimes called Lagrangian decomposition), is based on the solution of a Lagrangian dual problem. These methods (or variants [112, 119]) are applicable to a more general class of stochastic programs, but typically involve the solution of expensive nonsmooth convex optimization problems in a branch-and-bound setting, which may be computationally intensive for large-scale problems. A few modifications to the conventional Lagrangian relaxation method have been proposed in the literature, including cross-decomposition techniques that integrate Benders decomposition and Lagrangian relaxation approaches [166, 230] and techniques that integrate LR within branch-and-cut approaches [112].

It is noteworthy that the solution of two-stage stochastic programs using general-purpose state-of-the-art deterministic global optimization methods such as branch-and-reduce [224, 225] is usually not practical for problems with a large number of scenarios since these methods do not exploit their decomposable structure. This is evidenced by the case studies in Chapter 4. When all of the first-stage decision variables in Problem (DEP) are integer variables with finite bounds, Problem (DEP) can also be solved using nonconvex outer-approximation [118], which is a nearly-decomposable extension of the well-known outer-approximation algorithm for convex MINLPs [72, 79]. Other approaches that can potentially exploit the structure of Problem (DEP) include: a B&B-based global optimization approach [82] that exploits (weak) Lagrangian duality for a subclass of Problem (DEP) when it only contains continuous variables, a Lagrangian dual-based reduced-space B&B algorithm [69] for a subclass of Problem (DEP) with convex recourse programs, and a convex relaxation-based reduced-space B&B algorithm [76] for a subclass of Problem (DEP) with convex recourse programs.

Recently, an efficient decomposition algorithm termed nonconvex generalized Benders decomposition (NGBD), which extends BD and GBD, has been proposed to solve two-stage stochastic programs of the form (DEP) whose first-stage decisions are only bounded integers [138, 139]. This algorithm exploits the decomposable structure of stochastic programs, resulting in a computation that typically grows linearly with the number of scenarios as opposed to classical global optimization algorithms that are worst-case exponential. NGBD and its variants have been shown to be effective in solving several large-scale stochastic programs including the design and operation of natural gas production networks [136], pharmaceutical capacity planning under clinical trial uncertainty [219], the design and operation of flexible energy polygeneration systems [54], and integrated crude selection and refinery operation [241], where thousands of scenarios may be required to model the uncertainty accurately.

In the next two sections, we outline the NGBD and LR algorithms for solving (a subclass of) Problem (DEP). The developments in these sections will prove useful in Chapter 3, where we develop the first fully decomposable procedure for solving Problem (DEP) that provably converges to an ε-optimal solution in finite time by integrating LR, NGBD, and scalable bounds tightening techniques.

2.3.3.1.1 Nonconvex generalized Benders decomposition

GBD [85] employs the concepts of projection, restriction, and dualization to solve Problem (DEP) when the sets X_h are convex (this entails $n_{x_b} = 0$), $\forall h \in \{1, \cdots, s\}$, the participating functions f_h and $\mathbf{g}_h$ are partially convex with respect to the recourse variables $\mathbf{x}_h$ (on X_h), and the so-called Property (P) holds for Problem (DEP) (see [85, p. 251]). NGBD is an extension of GBD that can handle nonconvexities and the nonsatisfaction of Property (P) by the functions participating in Problem (DEP) [138]; however, NGBD can be guaranteed to converge only when Problem (DEP) has just bounded discrete first-stage decisions. Therefore, we consider the following more restrictive formulation in this section that drops the continuous first-stage variables $\mathbf{z}$ from Problem (DEP):

$$\begin{aligned}
\min_{\mathbf{x}_1,\cdots,\mathbf{x}_s,\mathbf{y}} \quad & \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h, \mathbf{y}) \qquad\qquad \text{(NGBD-DEP)}\\
\text{s.t.} \quad & \mathbf{g}_h(\mathbf{x}_h, \mathbf{y}) \le \mathbf{0}, \quad \forall h \in \{1, \cdots, s\},\\
& \mathbf{x}_h \in X_h, \quad \forall h \in \{1, \cdots, s\},\\
& \mathbf{y} \in Y,
\end{aligned}$$

where $X_h = \{\mathbf{x}_h \in \{0,1\}^{n_{x_b}} \times \Pi_{x,h} : \mathbf{r}_{x,h}(\mathbf{x}_h) \le \mathbf{0}\}$, $\Pi_{x,h} \subset \mathbb{R}^{n_{x_c}}$ is convex for each $h \in \{1, \cdots, s\}$, X_h is nonempty and compact $\forall h \in \{1, \cdots, s\}$, $Y = \{\mathbf{y} \in \{0,1\}^{n_y} : \mathbf{r}_y(\mathbf{y}) \le \mathbf{0}\}$ is nonempty, and the functions $f_h : [0,1]^{n_{x_b}} \times \Pi_{x,h} \times [0,1]^{n_y} \to \mathbb{R}$, $\mathbf{g}_h : [0,1]^{n_{x_b}} \times \Pi_{x,h} \times [0,1]^{n_y} \to \mathbb{R}^m$, $\mathbf{r}_{x,h} : [0,1]^{n_{x_b}} \times \Pi_{x,h} \to \mathbb{R}^{m_x}$, $\forall h \in \{1, \cdots, s\}$, and $\mathbf{r}_y : [0,1]^{n_y} \to \mathbb{R}^{m_y}$ are assumed to be continuous.

NGBD considers a convexified, separable version of Problem (NGBD-DEP) that is subject to the manipulations and solution strategies employed by GBD to yield a valid lower bound. Valid upper bounds for Problem (NGBD-DEP) are obtained by fixing the binary complicating variables in Problem (NGBD-DEP) to points visited by the convexified problem in an efficient and systematic manner. The finiteness of the NGBD algorithm is guaranteed by the finite number of values that the binary complicating variables, $\mathbf{y}$, can take in Problem (NGBD-DEP), as implied by Theorem 2.3.61. An outline of the NGBD algorithm is provided in Section 2.3.3.1.1.5. The following sections outline the subproblems used by the NGBD algorithm to solve Problem (NGBD-DEP). The details in these sections are adapted from references [138, 139].

2.3.3.1.1.1 Convexification

In this section, we outline techniques that can be used to construct a convex relaxation of Problem (NGBD-DEP) such that the resulting convex relaxation can be solved in a decomposable manner using GBD.

Definition 2.3.55. [Convex Relaxation of an Optimization Problem] Given a convex set $\Pi_{x,h}$, continuous functions $f_h^{\mathrm{cv}} : [0,1]^{n_{x_b}} \times \Pi_{x,h} \times [0,1]^{n_y} \times \Theta_h \to \mathbb{R}, \mathbf{g}^{\mathrm{cv}} : [0,1]^{n_{x_b}} \times \Pi_{x,h} \times [0,1]^{n_y} \times \Theta_h \to \mathbb{R}^m, \mathbf{r}_{x,h}^{\mathrm{cv}} : [0,1]^{n_{x_b}} \times \Pi_{x,h} \times \Theta_h \to \mathbb{R}^{m_x}, \mathbf{r}_y^{\mathrm{cv}} : [0,1]^{n_y} \times \Gamma \to \mathbb{R}^{m_y}$, $\mathbf{q}_h^{\mathrm{cv}} : [0,1]^{n_{x_b}} \times \Pi_{x,h} \times [0,1]^{n_y} \times \Theta_h \to \mathbb{R}^{m_q}$, and $\mathbf{v}^{\mathrm{cv}} : [0,1]^{n_y} \times \Gamma \to \mathbb{R}^{m_v}$, whose domains involve the auxiliary convex sets $\Theta_h \subset \mathbb{R}^{n_q}$ and $\Gamma \subset \mathbb{R}^{n_v}$, are said to constitute a convex relaxation of Problem (NGBD-DEP), which involves the functions $f_h, \mathbf{g}_h$, $\mathbf{r}_{x,h}$, and $\mathbf{r}_y$, if:

1. f_h^{cv}, $\mathbf{g}^{\mathrm{cv}}$, $\mathbf{r}_{x,h}^{\mathrm{cv}}$, $\mathbf{r}_y^{\mathrm{cv}}$, $\mathbf{q}_h^{\mathrm{cv}}$, and $\mathbf{v}^{\mathrm{cv}}$ are convex on their domains;
2. for any $(\mathbf{x}_h, \mathbf{y}) \in [0,1]^{n_{x_b}} \times \Pi_{x,h} \times [0,1]^{n_y}$, there exists $(\mathbf{q}_h, \mathbf{v}) \in \Theta_h \times \Gamma$ such that

$$\begin{aligned}
f_h^{\mathrm{cv}}(\mathbf{x}_h, \mathbf{y}, \mathbf{q}_h) &\leq f_h(\mathbf{x}_h, \mathbf{y}),\\
\mathbf{g}^{\mathrm{cv}}(\mathbf{x}_h, \mathbf{y}, \mathbf{q}_h) &\leq \mathbf{g}_h(\mathbf{x}_h, \mathbf{y}),\\
\mathbf{r}_{x,h}^{\mathrm{cv}}(\mathbf{x}_h, \mathbf{q}_h) &\leq \mathbf{r}_{x,h}(\mathbf{x}_h),\\
\mathbf{r}_y^{\mathrm{cv}}(\mathbf{y}, \mathbf{v}) &\leq \mathbf{r}_y(\mathbf{y}),\\
\mathbf{q}_h^{\mathrm{cv}}(\mathbf{x}_h, \mathbf{y}, \mathbf{q}_h) &\leq \mathbf{0},\\
\mathbf{v}^{\mathrm{cv}}(\mathbf{y}, \mathbf{v}) &\leq \mathbf{0}.
\end{aligned}$$

The newly defined convex functions involve additional variables ($\mathbf{q}_h$ and $\mathbf{v}$) and constraints ($\mathbf{q}_h^{\mathrm{cv}}(\mathbf{x}_h, \mathbf{y}, \mathbf{q}_h) \leq \mathbf{0}$ and $\mathbf{v}^{\mathrm{cv}}(\mathbf{y}, \mathbf{v}) \leq \mathbf{0}$) that may be required to construct differentiable relaxations when standard McCormick relaxations are employed [84, 213] (we remark that the recently-developed differentiable McCormick relaxation framework [124] provides an alternative approach that does not introduce auxiliary variables and constraints). Note that the domains over which the relaxations are constructed are not written out explicitly to keep the notation (relatively) simple.

By replacing the nonconvex functions in Problem (NGBD-DEP) with the above convex relaxations, we obtain the following problem which provides a lower bound on its optimal objective function value (cf. Proposition 2.3.62):

$$\begin{aligned}\min_{\substack{\mathbf{x}_1,\cdots,\mathbf{x}_s,\mathbf{y},\\ \mathbf{q}_1,\cdots,\mathbf{q}_s}} \quad & \sum_{h=1}^{s} p_h f_h^{\mathrm{cv}}(\mathbf{x}_h,\mathbf{y},\mathbf{q}_h) \\ \text{s.t.} \quad & \tilde{\mathbf{g}}_h^{\mathrm{cv}}(\mathbf{x}_h,\mathbf{y},\mathbf{q}_h) \le \mathbf{0}, \quad \forall h \in \{1,\cdots,s\}, \\ & (\mathbf{x}_h,\mathbf{q}_h) \in D_h, \quad \forall h \in \{1,\cdots,s\}, \\ & \mathbf{y} \in \bar{Y},\end{aligned} \tag{NGBD-LBP-NS}$$

where $\tilde{\mathbf{g}}_h^{\mathrm{cv}} : [0,1]^{n_{x_b}} \times \Pi_{x,h} \times [0,1]^{n_y} \times \Theta_h \to \mathbb{R}^{\tilde{m}}$, with $\tilde{m} := m + m_q$, is defined as $\tilde{\mathbf{g}}_h^{\mathrm{cv}}(\mathbf{x}_h,\mathbf{y},\mathbf{q}_h) := (\mathbf{g}^{\mathrm{cv}}(\mathbf{x}_h,\mathbf{y},\mathbf{q}_h), \mathbf{q}_h^{\mathrm{cv}}(\mathbf{x}_h,\mathbf{y},\mathbf{q}_h))$, $\forall (\mathbf{x}_h,\mathbf{y},\mathbf{q}_h) \in [0,1]^{n_{x_b}} \times \Pi_{x,h} \times [0,1]^{n_y} \times \Theta_h$, for ease of exposition, $D_h := \left\{(\mathbf{x}_h,\mathbf{q}_h) \in [0,1]^{n_{x_b}} \times \Pi_{x,h} \times \Theta_h : \mathbf{r}_{x,h}^{\mathrm{cv}}(\mathbf{x}_h,\mathbf{q}_h) \le \mathbf{0}\right\}$, and $\bar{Y} := \left\{\mathbf{y} \in \{0,1\}^{n_y} : \exists \mathbf{v} \in \Gamma \,:\, \mathbf{r}_y^{\mathrm{cv}}(\mathbf{y},\mathbf{v}) \le \mathbf{0}, \mathbf{v}^{\mathrm{cv}}(\mathbf{y},\mathbf{v}) \le \mathbf{0}\right\}$. Note that the binary recourse variables have been relaxed to their continuous counterparts in the definition of the sets D_h.

Problem (NGBD-LBP-NS) is a convex MINLP that cannot be solved using GBD (in general) unless Property (P) can be guaranteed to hold. To ensure satisfaction of Property (P), we further relax Problem (NGBD-LBP-NS) into the following form (if necessary):

$$\begin{aligned}\min_{\substack{\mathbf{x}_1,\cdots,\mathbf{x}_s,\mathbf{y},\\ \mathbf{q}_1,\cdots,\mathbf{q}_s}} \quad & \sum_{h=1}^{s} p_h \left[f_{1,h}^{\mathrm{cv}}(\mathbf{x}_h,\mathbf{q}_h) + f_{2,h}^{\mathrm{cv}}(\mathbf{y})\right] \\ \text{s.t.} \quad & \tilde{\mathbf{g}}_{1,h}^{\mathrm{cv}}(\mathbf{x}_h,\mathbf{q}_h) + \tilde{\mathbf{g}}_{2,h}^{\mathrm{cv}}(\mathbf{y}) \le \mathbf{0}, \quad \forall h \in \{1,\cdots,s\}, \\ & (\mathbf{x}_h,\mathbf{q}_h) \in D_h, \quad \forall h \in \{1,\cdots,s\}, \\ & \mathbf{y} \in \bar{Y},\end{aligned} \tag{NGBD-LBP}$$

where $f_{1,h}^{\mathrm{cv}} : [0,1]^{n_{x_b}} \times \Pi_{x,h} \times \Theta_h \to \mathbb{R}$, $f_{2,h}^{\mathrm{cv}} : [0,1]^{n_y} \to \mathbb{R}$, we define $\tilde{\mathbf{g}}_{1,h}^{\mathrm{cv}} : [0,1]^{n_{x_b}} \times \Pi_{x,h} \times \Theta_h \to \mathbb{R}^{\tilde{m}}$ and $\tilde{\mathbf{g}}_{2,h}^{\mathrm{cv}} : [0,1]^{n_y} \to \mathbb{R}^{\tilde{m}}$ for each $(\mathbf{x}_h,\mathbf{y},\mathbf{q}_h) \in [0,1]^{n_{x_b}} \times \Pi_{x,h} \times [0,1]^{n_y} \times \Theta_h$ as:

$$\begin{aligned}\tilde{\mathbf{g}}_{1,h}^{\mathrm{cv}}(\mathbf{x}_h,\mathbf{q}_h) &:= \left(\mathbf{g}_{1,h}^{\mathrm{cv}}(\mathbf{x}_h,\mathbf{q}_h), \mathbf{q}_{1,h}^{\mathrm{cv}}(\mathbf{x}_h,\mathbf{q}_h)\right), \\ \tilde{\mathbf{g}}_{2,h}^{\mathrm{cv}}(\mathbf{y}) &:= \left(\mathbf{g}_{2,h}^{\mathrm{cv}}(\mathbf{y}), \mathbf{q}_{2,h}^{\mathrm{cv}}(\mathbf{y})\right),\end{aligned}$$

$\mathbf{g}_{1,h}^{\mathrm{cv}} : [0,1]^{n_{x_b}} \times \Pi_{x,h} \times \Theta_h \to \mathbb{R}^m$, $\mathbf{g}_{2,h}^{\mathrm{cv}} : [0,1]^{n_y} \to \mathbb{R}^m$, $\mathbf{q}_{1,h}^{\mathrm{cv}} : [0,1]^{n_{x_b}} \times \Pi_{x,h} \times \Theta_h \to \mathbb{R}^{m_q}$, $\mathbf{q}_{2,h}^{\mathrm{cv}} : [0,1]^{n_y} \to \mathbb{R}^{m_q}$, and the functions $f_{1,h}^{\mathrm{cv}}, f_{2,h}^{\mathrm{cv}}, \tilde{\mathbf{g}}_{1,h}^{\mathrm{cv}}$, and $\tilde{\mathbf{g}}_{2,h}^{\mathrm{cv}}$ are continuous and convex on their domains. In addition, $\forall(\mathbf{x}_h, \mathbf{y}, \mathbf{q}_h) \in [0,1]^{n_{x_b}} \times \Pi_{x,h} \times [0,1]^{n_y} \times \Theta_h$, we require the following conditions to be satisfied to ensure that Problem (NGBD-LBP) yields a valid lower bound for Problem (NGBD-DEP):

$$\begin{aligned}
&f_{1,h}^{\mathrm{cv}}(\mathbf{x}_h, \mathbf{q}_h) + f_{2,h}^{\mathrm{cv}}(\mathbf{y}) \leq f_h^{\mathrm{cv}}(\mathbf{x}_h, \mathbf{y}, \mathbf{q}_h), \text{ and} \\
&\tilde{\mathbf{g}}_{1,h}^{\mathrm{cv}}(\mathbf{x}_h, \mathbf{q}_h) + \tilde{\mathbf{g}}_{2,h}^{\mathrm{cv}}(\mathbf{y}) \leq \tilde{\mathbf{g}}_h^{\mathrm{cv}}(\mathbf{x}_h, \mathbf{y}, \mathbf{q}_h).
\end{aligned}$$

One way to obtain Problem (NGBD-LBP) from Problem (NGBD-LBP-NS) is via outer-linearization [169]. Problem (NGBD-LBP) is a convex MINLP which can be solved in a decomposable manner using GBD to obtain a lower bound on the optimal objective value of Problem (NGBD-DEP). We make the following two additional assumptions.

Assumption 2.3.56. The sets D_h, $\forall h \in \{1, \cdots, s\}$, are compact.

Remark 2.3.57. From Assumption 2.3.56 and the assumption of continuity of the functions in Problem (NGBD-LBP), Weierstrass' theorem implies that Problem (NGBD-LBP) either has a finite optimal objective value, or is infeasible.

Assumption 2.3.58. Problem (NGBD-LBP) satisfies Slater's condition for $\mathbf{y}$ fixed to those points in $\bar{Y}$ for which Problem (NGBD-LBP) is feasible.

Remark 2.3.59. A consequence of Assumption 2.3.58 is that strong duality holds for Problem (NGBD-LBP) for $\mathbf{y}$ fixed to those points in $\bar{Y}$ for which Problem (NGBD-LBP) is feasible. This validates the dualization manipulation of Problem (NGBD-LBP) to a master problem, which is presented in the following section.

2.3.3.1.1.2 Master problem

Problem (NGBD-LBP) is potentially a large-scale convex MINLP since the number of variables in its formulation is an affine function of the number of scenarios considered. Using the principle of projection in [85], Problem (NGBD-LBP) can be projected onto the space of the first-stage variables and using duality theory, any subproblem with a fixed $\mathbf{y}$ can be reformulated into its dual. This yields the following (semi-infinite) master problem which is equivalent to Problem (NGBD-LBP) as stated in Proposition 2.3.63:

$$\begin{aligned}
&\min_{\eta,\mathbf{y}}\ \eta && \text{(NGBD-MP)}\\
&\text{s.t. } \eta \geq \sum_{h=1}^{s} \min_{(\mathbf{x}_h,\mathbf{q}_h)\in D_h} \left[p_h f_{1,h}^{\mathrm{cv}}(\mathbf{x}_h,\mathbf{q}_h) + \boldsymbol{\lambda}_h^{\mathrm{T}} \tilde{\mathbf{g}}_{1,h}^{\mathrm{cv}}(\mathbf{x}_h,\mathbf{q}_h)\right] \\
&\qquad\quad + \sum_{h=1}^{s} \left(p_h f_{2,h}^{\mathrm{cv}}(\mathbf{y}) + \boldsymbol{\lambda}_h^{\mathrm{T}} \tilde{\mathbf{g}}_{2,h}^{\mathrm{cv}}(\mathbf{y})\right), \quad \forall (\boldsymbol{\lambda}_1,\cdots,\boldsymbol{\lambda}_s) \in \mathbb{R}_+^{\tilde{m}\times s},\\
&\quad\ \ 0 \geq \sum_{h=1}^{s} \min_{(\mathbf{x}_h,\mathbf{q}_h)\in D_h} \left[\boldsymbol{\mu}_h^{\mathrm{T}} \tilde{\mathbf{g}}_{1,h}^{\mathrm{cv}}(\mathbf{x}_h,\mathbf{q}_h)\right] + \sum_{h=1}^{s} \boldsymbol{\mu}_h^{\mathrm{T}} \tilde{\mathbf{g}}_{2,h}^{\mathrm{cv}}(\mathbf{y}), \quad \forall (\boldsymbol{\mu}_1,\cdots,\boldsymbol{\mu}_s) \in \mathbb{R}_+^{\tilde{m}\times s},\\
&\quad\ \ \mathbf{y} \in \bar{Y}, \eta \in \mathbb{R}.
\end{aligned}$$

In Section 2.3.3.1.1.4, the above master problem will be relaxed to yield (in effect) a 'lower bounding problem' for Problem (NGBD-LBP).

2.3.3.1.1.3 Restricted subproblems

The primal problem for NGBD is obtained by restricting $\mathbf{y}$ in Problem (NGBD-DEP) to a point $\bar{\mathbf{y}} \in Y$ and, if feasible, provides an upper bound on the dual function value as stated in Proposition 2.3.64.

$$\begin{aligned}
obj_{\text{NGBD-PP}}(\bar{\mathbf{y}}) := \min_{\mathbf{x}_1,\cdots,\mathbf{x}_s}\ & \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h,\bar{\mathbf{y}}) && \text{(NGBD-PP)}\\
\text{s.t.}\ \ & \mathbf{g}_h(\mathbf{x}_h,\bar{\mathbf{y}}) \leq \mathbf{0}, \quad \forall h \in \{1,\cdots,s\},\\
& \mathbf{x}_h \in X_h, \quad \forall h \in \{1,\cdots,s\}.
\end{aligned}$$

Problem (NGBD-PP) can be decomposed into s independent scenario subproblems of the form[7]:

$$\begin{aligned}
obj_{\text{NGBD-PP}_h}(\bar{\mathbf{y}}) := \min_{\mathbf{x}_h \in X_h}\ & p_h f_h(\mathbf{x}_h,\bar{\mathbf{y}}) && (\text{NGBD-PP}_h)\\
\text{s.t.}\ \ & \mathbf{g}_h(\mathbf{x}_h,\bar{\mathbf{y}}) \leq \mathbf{0},
\end{aligned}$$

[7]We note that the probability p_h is post-multiplied with the optimal objective value of the (rest of the) minimization in Problem (NGBD-PP$_h$) in numerical implementations to avoid ill-conditioned subproblems (that may occur especially when the number of scenarios is large).

with $obj_{\text{NGBD-PP}}(\bar{\mathbf{y}}) = \sum_{h=1}^{s} obj_{\text{NGBD-PP}_h}(\bar{\mathbf{y}})$. Note that Problem (NGBD-PP$_h$), which is potentially a nonconvex MINLP, can be solved to ε-optimality in finite time by state-of-the-art global optimization solvers, provided it can be expressed in a form that can be handled by those solvers.

The primal bounding problem for NGBD is obtained by restricting the binary complicating variables $\mathbf{y}$ in Problem (NGBD-LBP) to a point $\bar{\mathbf{y}} \in \bar{Y}$.

$$
\begin{aligned}
obj_{\text{NGBD-PBP}}(\bar{\mathbf{y}}) := \min_{\substack{\mathbf{x}_1,\cdots,\mathbf{x}_s,\\ \mathbf{q}_1,\cdots,\mathbf{q}_s}} \; & \sum_{h=1}^{s} p_h \left[f_{1,h}^{\text{cv}}(\mathbf{x}_h, \mathbf{q}_h) + f_{2,h}^{\text{cv}}(\bar{\mathbf{y}}) \right] \\
\text{s.t.} \quad & \tilde{\mathbf{g}}_{1,h}^{\text{cv}}(\mathbf{x}_h, \mathbf{q}_h) + \tilde{\mathbf{g}}_{2,h}^{\text{cv}}(\bar{\mathbf{y}}) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\}, \\
& (\mathbf{x}_h, \mathbf{q}_h) \in D_h, \quad \forall h \in \{1, \cdots, s\}.
\end{aligned} \tag{NGBD-PBP}
$$

Problem (NGBD-PBP), which provides an upper bound on the optimal objective value of Problem (NGBD-LBP), can be decomposed into s independent scenario subproblems, the solution of each of which provides a lower bound to the optimal objective value of the corresponding Problem (NGBD-PP$_h$) as stated in Proposition 2.3.65:

$$
\begin{aligned}
obj_{\text{NGBD-PBP}_h}(\bar{\mathbf{y}}) := \min_{(\mathbf{x}_h, \mathbf{q}_h) \in D_h} \; & p_h \left[f_{1,h}^{\text{cv}}(\mathbf{x}_h, \mathbf{q}_h) + f_{2,h}^{\text{cv}}(\bar{\mathbf{y}}) \right] \\
\text{s.t.} \quad & \tilde{\mathbf{g}}_{1,h}^{\text{cv}}(\mathbf{x}_h, \mathbf{q}_h) + \tilde{\mathbf{g}}_{2,h}^{\text{cv}}(\bar{\mathbf{y}}) \leq \mathbf{0}.
\end{aligned} \tag{NGBD-PBP$_h$}
$$

Note that $obj_{\text{NGBD-PBP}}(\bar{\mathbf{y}}) = \sum_{h=1}^{s} obj_{\text{NGBD-PBP}_h}(\bar{\mathbf{y}})$. If Problem (NGBD-PBP$_h$) is infeasible for any scenario $h \in \{1, \cdots, s\}$, Problem (NGBD-PBP) (and consequently, Problem (NGBD-PP)) is infeasible as well and the following feasibility problem is solved:

$$
\begin{aligned}
obj_{\text{NGBD-FP}}(\bar{\mathbf{y}}) := \min_{\substack{\mathbf{x}_1,\cdots,\mathbf{x}_s,\\ \mathbf{q}_1,\cdots,\mathbf{q}_s,\\ \mathbf{w}_1,\cdots,\mathbf{w}_s}} \; & \sum_{h=1}^{s} p_h \|\mathbf{w}_h\| \\
\text{s.t.} \quad & \tilde{\mathbf{g}}_{1,h}^{\text{cv}}(\mathbf{x}_h, \mathbf{q}_h) + \tilde{\mathbf{g}}_{2,h}^{\text{cv}}(\bar{\mathbf{y}}) \leq \mathbf{w}_h, \quad \forall h \in \{1, \cdots, s\}, \\
& (\mathbf{x}_h, \mathbf{q}_h) \in D_h, \;\; \mathbf{w}_h \in W_h, \quad \forall h \in \{1, \cdots, s\},
\end{aligned} \tag{NGBD-FP}
$$

where $\|\mathbf{w}_h\|$ (used here to denote an arbitrary norm of the slack variable vector $\mathbf{w}_h$) measures the violation of the constraints in Problem (NGBD-PBP$_h$), and $W_h \subset \mathbb{R}_+^{\tilde{m}}$ is chosen such that it possesses the following properties:

1. W_h is a convex set;

2. W_h is a pointed cone, i.e., $\mathbf{0} \in W_h$, and $\forall \alpha > 0, \mathbf{w} \in W_h \implies \alpha \mathbf{w} \in W_h$;

3. there exists $\mathbf{w} \in W_h$ with $\mathbf{w} > \mathbf{0}$ (therefore, the cone W_h is unbounded in each dimension).

The sets W_h are typically simply chosen to be $\mathbb{R}_+^{\tilde{m}}$. Since any norm is convex, we have that Problem (NGBD-FP) is convex (we usually use either the 1-norm formulation, or the ∞-norm formulation). Furthermore, Problem (NGBD-FP) can naturally be decomposed into s independent scenario subproblems of the form:

$$
\begin{aligned}
obj_{\text{NGBD-FP}_h}(\bar{\mathbf{y}}) := \min_{\mathbf{x}_h, \mathbf{q}_h, \mathbf{w}_h} \quad & p_h \|\mathbf{w}_h\| \\
\text{s.t.} \quad & \tilde{\mathbf{g}}_{1,h}^{\text{cv}}(\mathbf{x}_h, \mathbf{q}_h) + \tilde{\mathbf{g}}_{2,h}^{\text{cv}}(\bar{\mathbf{y}}) \le \mathbf{w}_h, \\
& (\mathbf{x}_h, \mathbf{q}_h) \in D_h, \ \ \mathbf{w}_h \in W_h,
\end{aligned}
\tag{NGBD-FP$_h$}
$$

with $obj_{\text{NGBD-FP}}(\bar{\mathbf{y}}) = \sum_{h=1}^{s} obj_{\text{NGBD-FP}_h}(\bar{\mathbf{y}})$.

If the convex subproblems (NGBD-PBP$_h$) and (NGBD-FP$_h$) are smooth (or smooth reformulations are possible), they can be solved by gradient-based (local optimization) solvers. Otherwise, they could be solved using nonsmooth optimization methods such as bundle methods [100].

2.3.3.1.1.4 Relaxed master problem

The master problem (NGBD-MP) can be difficult to solve since it is a semi-infinite program; therefore, it is relaxed to the following relaxed master problem (see Proposition 2.3.67):

$$
\begin{aligned}
\min_{\eta, \mathbf{y}} \ & \eta \\
\text{s.t. } & \eta \ge obj_{\text{NGBD-PBP}}(\mathbf{y}^j) + \sum_{h=1}^{s} p_h \left[f_{2,h}^{\text{cv}}(\mathbf{y}) - f_{2,h}^{\text{cv}}(\mathbf{y}^j) \right] + \sum_{h=1}^{s} \left(\boldsymbol{\lambda}_h^j\right)^{\mathrm{T}} \left[\tilde{\mathbf{g}}_{2,h}^{\text{cv}}(\mathbf{y}) - \tilde{\mathbf{g}}_{2,h}^{\text{cv}}(\mathbf{y}^j) \right], \ \ \forall j \in T, \\
& 0 \ge obj_{\text{NGBD-FP}}(\mathbf{y}^i) + \sum_{h=1}^{s} \left(\boldsymbol{\mu}_h^i\right)^{\mathrm{T}} \left[\tilde{\mathbf{g}}_{2,h}^{\text{cv}}(\mathbf{y}) - \tilde{\mathbf{g}}_{2,h}^{\text{cv}}(\mathbf{y}^i) \right], \ \ \forall i \in S, \\
& \left| \left\{ r \in \{1, \cdots, n_y\} : y_r^t = 1 \right\} \right| - 1 \ge \sum_{r(t) \in \{r \in \{1,\cdots,n_y\} : y_r^t = 1\}} y_{r(t)} - \sum_{r(t) \in \{r \in \{1,\cdots,n_y\} : y_r^t = 0\}} y_{r(t)}, \ \ \forall t \in T \cup S, \\
& \mathbf{y} \in \bar{Y}, \eta \in \mathbb{R},
\end{aligned}
\tag{NGBD-RMP}
$$

where T and S are index sets that store whether the binary complicating variable values 'visited thus far' (see Algorithm 2.3 in Section 2.3.3.1.1.5) by the NGBD algorithm are feasible for the primal bounding problem, i.e., at any particular iteration $k \geq 1$ of the NGBD algorithm:

$$T = \left\{ j \in \{1, \cdots, k\} : \text{Problem (NGBD-PBP) is feasible for } \bar{\mathbf{y}} = \mathbf{y}^j \right\},$$
$$S = \left\{ i \in \{1, \cdots, k\} : \text{Problem (NGBD-PBP) is infeasible for } \bar{\mathbf{y}} = \mathbf{y}^i \right\},$$

$\mathbf{y}^i$ is the binary complicating variable value visited by the NGBD algorithm during iteration i (with y_r^i denoting its r^{th} component), the first set of constraints in Problem (NGBD-RMP) correspond to optimality cuts for the iterations T in which Problem (NGBD-PBP) is feasible, the second set of constraints correspond to feasibility cuts for the iterations S in which the corresponding Problem (NGBD-PBP) is infeasible, the third set of constraints correspond to a set of canonical integer cuts that exclude previously examined binary complicating variable values from the feasible region [10], $\boldsymbol{\lambda}_h^j$ are Lagrange multipliers for Problem (NGBD-PBP$_h$) with $\bar{\mathbf{y}}$ fixed to $\mathbf{y}^j$ when $j \in T$, and $\boldsymbol{\mu}_h^i$ are Lagrange multipliers for Problem (NGBD-FP$_h$) with $\bar{\mathbf{y}}$ fixed to $\mathbf{y}^i$ when $i \in S$.

When $T = \emptyset$ (i.e., none of the binary complicating variable points visited thus far are feasible for Problem (NGBD-PBP)), Problem (NGBD-RMP) is unbounded and the following feasibility relaxed master problem is solved instead:

$$\begin{aligned}
\min_{\mathbf{y}} \ & \|\mathbf{y}\| && \text{(NGBD-FRMP)}\\
\text{s.t. } & 0 \geq obj_{\text{NGBD-FP}}(\mathbf{y}^i) + \sum_{h=1}^{s} \left(\boldsymbol{\mu}_h^i\right)^{\mathrm{T}} \left[\tilde{\mathbf{g}}_{2,h}^{\text{cv}}(\mathbf{y}) - \tilde{\mathbf{g}}_{2,h}^{\text{cv}}(\mathbf{y}^i)\right], \quad \forall i \in S,\\
& \left|\left\{r \in \{1, \cdots, n_y\} : y_r^t = 1\right\}\right| - 1 \geq \sum_{r(t) \in \{r \in \{1,\cdots,n_y\} : y_r^t = 1\}} y_{r(t)} - \sum_{r(t) \in \{r \in \{1,\cdots,n_y\} : y_r^t = 0\}} y_{r(t)}, \quad \forall t \in S,\\
& \mathbf{y} \in \bar{Y},
\end{aligned}$$

where $\|\mathbf{y}\|$ is used above to denote an arbitrary norm of $\mathbf{y}$. Note that the sizes of Problems (NGBD-RMP) and (NGBD-FRMP), which are convex MINLPs, are independent of the number of scenarios in the original problem.

An overview of the subproblems used in the (basic) NGBD algorithm for solving Problem (NGBD-DEP) is presented in Figure 2-3. In the next section, we provide an informal summary of the workings of the (basic) NGBD algorithm along with a formal algorithmic

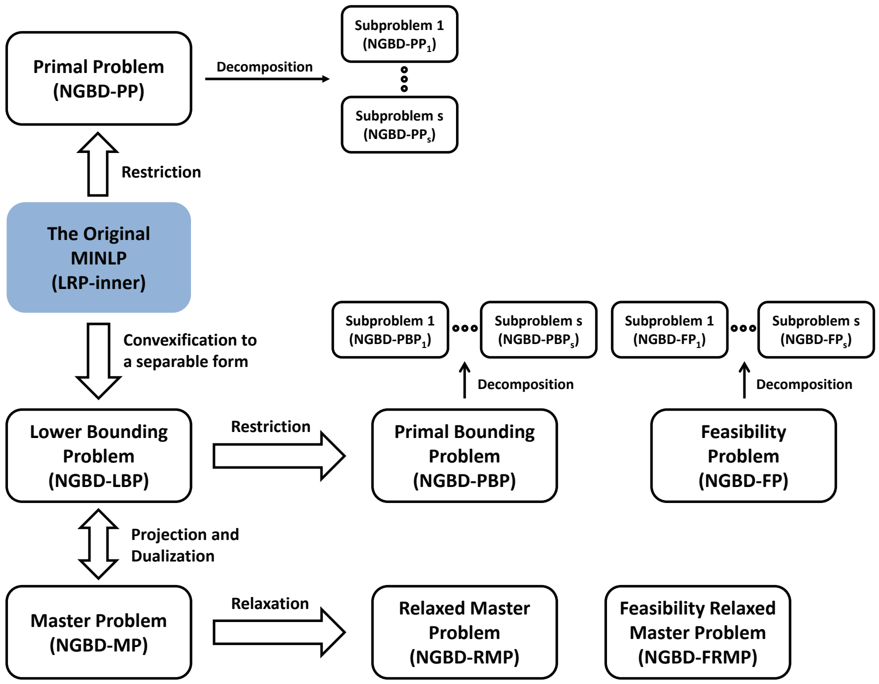


Figure 2-3: Overview of the reformulation and the subproblems in the basic NGBD algorithm (based on [139, Fig. 1]).

description and a result stating its convergence.

2.3.3.1.1.5 The NGBD algorithm

In the inner loop of the NGBD algorithm, Problem (NGBD-LBP) is solved using GBD, which utilizes as its subproblems Problem (NGBD-PBP), Problem (NGBD-FP), Problem (NGBD-RMP), and Problem (NGBD-FRMP) (note that the relaxed master problems, Problem (NGBD-RMP) and Problem (NGBD-FRMP), differ from their GBD counterparts in that they include canonical integer cuts to exclude previously visited binary complicating variable values from their feasible regions). Once the inner GBD loop has converged, Problem (NGBD-PP) is solved at candidate binary complicating variable values (values for which the optimal objective value of the corresponding Problem (NGBD-PBP) is no worse than: i. the optimal objective value of Problem (NGBD-RMP) at termination of the inner GBD loop, and ii. the objective value of the current best known feasible point for Problem (NGBD-PP)) to try and determine feasible points for Problem (NGBD-DEP) in the outer loop of the NGBD algorithm. If all of the candidate binary complicating variable values have been exhausted without satisfying the termination criteria for the NGBD algorithm, the inner loop is revisited and the above steps are repeated until the termination criteria are met.

In the previous section, we have only detailed the subproblems employed by the basic NGBD algorithm to solve Problem (NGBD-DEP). Techniques that can potentially enhance the performance of the basic NGBD algorithm include: feasibility and optimality-based bounds tightening techniques [134, 241, 242] and piecewise convex relaxations [52, 137] that could help improve the strength of the relaxations used in the subproblems and potentially reduce the number of primal problems solved, and adding multiple cuts to the master problem [36, 52] at each iteration in a bid to speed up the convergence of the inner GBD loop.

Below, we present a formal algorithmic outline of the NGBD algorithm for solving Problem (NGBD-DEP) based on [139, Section 4.1] and [138, Section 1.4.1].

Algorithm 2.3 Nonconvex Generalized Benders Decomposition

Initialize:

- Iteration counter $k = 0$, and index sets $T = \emptyset$, $S = \emptyset$, $U = \emptyset$.
- Upper bound on Problem (NGBD-DEP), $UBD = +\infty$; upper bound on Problem (NGBD-LBP), $UBDPB = +\infty$; and lower bound on Problem (NGBD-LBP), $LBD = -\infty$.
- Tolerances $\varepsilon_h > 0$, $\forall h \in \{1, \cdots, s\}$, and $\varepsilon > 0$ such that $\sum_{h=1}^{s} \varepsilon_h = \varepsilon$.
- Current best feasible point, $\{(\mathbf{x}_1^*, \cdots, \mathbf{x}_s^*, \mathbf{y}^*)\} = \emptyset$, for Problem (NGBD-DEP).
- Initial binary complicating variable realization $\mathbf{y}^1$.
- Candidate binary complicating variable value at which Problem (NGBD-PP) is to be solved, $\tilde{\mathbf{y}} = \mathbf{y}^1$, and corresponding NGBD inner iteration counter, $\tilde{k} = 1$.

repeat

  if $k = 0$ **or** (Problem (NGBD-RMP) is feasible **and** $LBD < UBDPB$ **and** $LBD < UBD - \varepsilon$) **then**

    repeat

1. Set $k = k + 1$.
2. Solve Problem (NGBD-PBP$_h$) with $\mathbf{y}$ fixed to $\mathbf{y}^k$ for each $h \in \{1, \cdots, s\}$ sequentially. If Problem (NGBD-PBP$_h$) is feasible for every $h \in \{1, \cdots, s\}$ with Lagrange multipliers $\boldsymbol{\lambda}_h^k$, add an optimality cut to Problem (NGBD-RMP) and set $T = T \cup \{k\}$. If $obj_{\text{NGBD-PBP}}(\mathbf{y}^k) < UBDPB$, update $UBDPB = obj_{\text{NGBD-PBP}}(\mathbf{y}^k)$, $\tilde{\mathbf{y}} = \mathbf{y}^k$, $\tilde{k} = k$.

Algorithm 2.3 Nonconvex Generalized Benders Decomposition (continued)

3. If Problem (NGBD-PBP$_h$) is infeasible for any scenario $\bar{h} \in \{1, \cdots, s\}$, do not solve it for scenarios $\bar{h}+1, \cdots, s$. Instead, set $S = S \cup \{k\}$, $\boldsymbol{\mu}_h^k = \mathbf{0}$ for scenarios $h = 1, \cdots, \bar{h}-1$, and solve Problem (NGBD-FP$_h$) with $\mathbf{y}$ fixed to $\mathbf{y}^k$ for scenarios $h = \bar{h}, \cdots, s$ to obtain the corresponding Lagrange multipliers $\boldsymbol{\mu}_h^k$. Add a feasibility cut to Problem (NGBD-RMP).
4. If $T = \emptyset$, solve Problem (NGBD-FRMP); otherwise solve Problem (NGBD-MP). In the latter case, set LBD to the optimal objective value of Problem (NGBD-RMP) if it is feasible. In both cases, set $\mathbf{y}^{k+1}$ to the $\mathbf{y}$ value corresponding to the solution of either problem.

until $LBD \geq UBDPB$ **or** Problem (NGBD-RMP) is infeasible **or** Problem (NGBD-FRMP) is infeasible

end if

if $UBDPB < UBD - \varepsilon$ **then**

1. Solve Problem (NGBD-PP$_h$) with $\mathbf{y}$ fixed to $\tilde{\mathbf{y}}$ to ε_h-optimality for each scenario $h \in \{1, \cdots, s\}$ sequentially. Set $U = U \cup \{\tilde{k}\}$. If Problem (NGBD-PP$_h$) with $\mathbf{y}$ fixed to $\tilde{\mathbf{y}}$ has an optimal solution $\tilde{\mathbf{x}}_h$ for each $h \in \{1, \cdots, s\}$ and $obj_{\text{NGBD-PP}}(\tilde{\mathbf{y}}) < UBD$, update $UBD = obj_{\text{NGBD-PP}}(\tilde{\mathbf{y}})$ and set $\mathbf{y}^* = \tilde{\mathbf{y}}$ and $\mathbf{x}_h^* = \tilde{\mathbf{x}}_h$ for each $h \in \{1, \cdots, s\}$.
2. If $T \backslash U = \emptyset$, set $UBDPB = +\infty$.
3. Else if $T \backslash U \neq \emptyset$, pick index $i \in T \backslash U$ such that $obj_{\text{NGBD-PBP}}(\mathbf{y}^i) = \min_{j \in T \backslash U} obj_{\text{NGBD-PBP}}(\mathbf{y}^j)$. Update $UBDPB = obj_{\text{NGBD-PBP}}(\mathbf{y}^i)$, $\tilde{\mathbf{y}} = \mathbf{y}^i$, $\tilde{k} = i$.

end if

until $UBDPB \geq UBD - \varepsilon$ **and** (Problem (NGBD-RMP) is infeasible **or** Problem (NGBD-FRMP) is infeasible **or** $LBD \geq UBD - \varepsilon$)

If $UBD < +\infty$, then $(\mathbf{x}_1^*, \cdots, \mathbf{x}_s^*, \mathbf{y}^*)$ is an ε-optimal solution to Problem (NGBD-DEP).

The following (reasonable) assumption is made to simplify the proof of convergence of the NGBD algorithm.

Assumption 2.3.60. We can solve Problems (NGBD-PBP), (NGBD-FP), (NGBD-RMP), and (NGBD-FRMP) to tolerances much tighter than ε in finite time.

The above assumption implies that the tolerances of the above subproblems can be neglected in comparison to the tolerance to which Problem (NGBD-PP) is solved. The following result states that the NGBD algorithm converges finitely to an ε-optimal solution to Problem (NGBD-DEP).

Theorem 2.3.61. (Finite convergence of the NGBD algorithm) If, for each $h \in \{1, \cdots, s\}$, Problem (NGBD-PP$_h$) can be solved to $\frac{\varepsilon}{s}$-optimality in a finite number of steps and Assumption 2.3.60 holds, then the NGBD algorithm either terminates in a finite number of steps with an ε-optimal solution to Problem (NGBD-DEP), or an indication that it is infeasible.

Proof. See Theorem 4.1 in [139] and Theorem 1 in [138]. □

2.3.3.1.1.6 Bounds tightening techniques

This section describes a couple of decomposable bounds tightening techniques that can be used to accelerate the solution of the NGBD algorithm for Problem (NGBD-DEP). The reader is directed to Section 2.3.2.2 for details of the forward-backward interval propagation technique that provides a computationally inexpensive and scalable method for tightening the bounds on the variables in the NGBD subproblems using interval arithmetic. The first approach we detail adapts the feasibility-based bounds tightening subproblem in Section 2.3.2.2, Problem (FBBT), to tighten the bounds on the continuous recourse variables in Problem (NGBD-DEP).

Let x_h^i denote the i^{th} recourse variable corresponding to scenario $h \in \{1, \cdots, s\}$ for any $i \in \{1, \cdots, n_{x_b} + n_{x_c}\}$. Lower bounds on the continuous recourse variable x_h^i, $\forall i \in \{n_{x_b} + 1, \cdots, n_{x_b} + n_{x_c}\}$, can be obtained from the solution of the following convex MINLP which uses feasibility arguments to exclude regions of the search space:

$$
\begin{aligned}
\min_{\mathbf{x}_h, \mathbf{y}, \mathbf{q}_h} \quad & x_h^i \\
\text{s.t.} \quad & \tilde{\mathbf{g}}_{1,h}^{\text{cv}}(\mathbf{x}_h, \mathbf{q}_h) + \tilde{\mathbf{g}}_{2,h}^{\text{cv}}(\mathbf{y}) \leq \mathbf{0}, \\
& (\mathbf{x}_h, \mathbf{q}_h) \in D_h, \;\; \mathbf{y} \in \bar{Y}.
\end{aligned} \tag{FBBT$_x$}
$$

Note that Problem (FBBT$_x$) is a convex MINLP that can either be solved directly, or further relaxed to a mixed-integer linear program (MILP) and solved subsequently (we note that: i. the integrality restrictions on the binary recourse variables in Problem (FBBT$_x$) could also be enforced to derive tighter bounds, and ii. the integrality restrictions on the binary complicating variables in Problem (FBBT$_x$) can be relaxed to keep the computational cost of solving Problem (FBBT$_x$) in check at the expense of weaker bounds). Problem (FBBT$_x$) only considers the relaxations of the constraints in scenario h of Problem (NGBD-DEP) to determine the lower bound on x_h^i; therefore, bounds on the continuous recourse variables can be determined in a decomposable manner. Furthermore, the set D_h is updated after each bound tightening solve so that subsequent iterations of Problem (FBBT$_x$) can exploit the updated variable bounds (and the potentially tighter relaxations that can be constructed as a result). Upper bounds on x_h^i can be determined in an analogous manner by maximizing the objective in Problem (FBBT$_x$) instead of minimizing it. Bounds on the continuous recourse variables $\mathbf{q}_h$ can also be obtained in an analogous manner. Our implementation of the above technique in Chapters 3 and 4 only tightens the bounds on the continuous recourse variables that explicitly participate in the construction of the relaxations of the functions in Problem (NGBD-DEP).

Next, we describe a nearly-decomposable optimality-based bounds tightening technique that integrates the primal bounding and (a multi-cut version of the) relaxed master problems of NGBD, Problems (NGBD-PBP$_h$) and (NGBD-RMP), along with an optimality cut to accelerate the convergence of the NGBD algorithm [242, Section 3.3.2]. For instance, to tighten the lower bound on x_h^i, for any $i \in \{n_{x_b}+1, \cdots, n_{x_b}+n_{x_c}\}$, during any given iteration of the NGBD algorithm for Problem (NGBD-DEP), we can either solve the following convex MINLP, or relax the integrality restrictions on its $\mathbf{y}$ variables and solve a convex program:

$$\begin{aligned}
\min_{\substack{\mathbf{x}_h,\mathbf{y},\mathbf{q}_h,\\ \eta_1,\cdots,\eta_s}} \quad & x_h^i && \text{(OBBT}_x\text{)}\\
\text{s.t.} \quad & \tilde{\mathbf{g}}_{1,h}^{\text{cv}}(\mathbf{x}_h,\mathbf{q}_h) + \tilde{\mathbf{g}}_{2,h}^{\text{cv}}(\mathbf{y}) \leq \mathbf{0},\\
& \eta_{\bar{h}} \geq obj_{\text{NGBD-PBP}_{\bar{h}}}(\mathbf{y}^j) + p_{\bar{h}}\left[f_{2,\bar{h}}^{\text{cv}}(\mathbf{y}) - f_{2,\bar{h}}^{\text{cv}}(\mathbf{y}^j)\right] + \\
& \qquad \left(\boldsymbol{\lambda}_{\bar{h}}^j\right)^{\text{T}}\left[\mathbf{g}_{2,\bar{h}}^{\text{cv}}(\mathbf{y}) - \mathbf{g}_{2,\bar{h}}^{\text{cv}}(\mathbf{y}^j)\right], \quad \forall \bar{h} \in \{1,\cdots,s\}\backslash\{h\},\ \forall j \in T,\\
& 0 \geq obj_{\text{NGBD-FP}}(\mathbf{y}^i) + \sum_{\bar{h}=1}^{s}\left(\boldsymbol{\mu}_{\bar{h}}^i\right)^{\text{T}}\left[\mathbf{g}_{2,\bar{h}}^{\text{cv}}(\mathbf{y}) - \mathbf{g}_{2,\bar{h}}^{\text{cv}}(\mathbf{y}^i)\right], \quad \forall i \in S,\\
& \left|\left\{r \in \{1,\cdots,n_y\} : y_r^t = 1\right\}\right| - 1 \geq \sum_{r(t)\in\{r\in\{1,\cdots,n_y\}:y_r^t=1\}} y_{r(t)} - \\
& \qquad \sum_{r(t)\in\{r\in\{1,\cdots,n_y\}:y_r^t=0\}} y_{r(t)}, \quad \forall t \in T \cup S,\\
& p_h\left[f_{1,h}^{\text{cv}}(\mathbf{x}_h,\mathbf{q}_h) + f_{2,h}^{\text{cv}}(\mathbf{y})\right] + \sum_{\bar{h}\neq h}\eta_{\bar{h}} \leq UBD,\\
& (\mathbf{x}_h,\mathbf{q}_h) \in D_h,\ \mathbf{y} \in \bar{Y},\ (\eta_1,\cdots,\eta_s) \in \mathbb{R}^s,
\end{aligned}$$

where the last inequality corresponds to an optimality cut that could possibly exclude binary complicating variable realizations that have not yet been visited by the NGBD algorithm.

We note that upper bounds on x_h^i can be tightened in a similar fashion, and refer the reader to [134] for related domain reduction ideas that could accelerate the convergence of the NGBD algorithm. Since the size of Problem (OBBT$_x$) can increase significantly with the number of iterations of the inner GBD loop of the NGBD algorithm, our implementation in Chapter 4 adds a limit on the maximum time for its solution.

2.3.3.1.1.7 Properties of the subproblems

The following results from [138, 139] validate the workings of the NGBD algorithm.

Proposition 2.3.62. The optimal objective value of Problem (NGBD-LBP) is a lower bound on the optimal objective value of Problem (NGBD-DEP).

Proof. See Proposition 1 in [138]. □

Proposition 2.3.63. Problems (NGBD-LBP) and (NGBD-MP) are equivalent in the sense that:

1. Problem (NGBD-LBP) is feasible if and only if Problem (NGBD-MP) is feasible;
2. Problems (NGBD-LBP) and (NGBD-MP) have the same optimal objective function values;
3. The optimal objective value of Problem (NGBD-LBP) is attained with an integer realization if and only if the optimal objective value of Problem (NGBD-MP) is attained with the same integer realization.

Proof. See Propositions 2 and 3 in [138]. □

Proposition 2.3.64. If Problem (NGBD-PP) is feasible for $\bar{\mathbf{y}}$ fixed to any element of Y, then its optimal objective value is no less than the optimal objective value of Problem (NGBD-DEP).

Proof. See Proposition 4 in [138]. □

Proposition 2.3.65. If for any $h \in \{1, \cdots, s\}$, we have that Problem (NGBD-PP$_h$) is feasible, then Problem (NGBD-PBP$_h$) is feasible as well. Additionally, the optimal objective value of Problem (NGBD-PP$_h$) is no less than the optimal objective value of Problem (NGBD-PBP$_h$).

Proof. See Proposition 5 in [138]. □

Proposition 2.3.66. For any $h \in \{1, \cdots, s\}$, Problem (NGBD-FP$_h$) satisfies Slater's condition and it always has a minimum, say $(\hat{\mathbf{x}}_h, \hat{\mathbf{q}}_h, \hat{\mathbf{w}}_h)$. Furthermore, $\|\hat{\mathbf{w}}_h\| > 0$ for those scenarios in which Problem (NGBD-PBP$_h$) is infeasible.

Proof. See Proposition 6 in [138]. □

Proposition 2.3.67. The optimal objective value of Problem (NGBD-RMP) is a valid lower bound for Problems (NGBD-LBP) and (NGBD-DEP) augmented with the relevant canonical integer cuts.

Proof. See Corollary 2 in [138]. □

Proposition 2.3.68. The optimal objective value of Problem (FBBT$_x$) is lesser than any feasible value of x_h^i for Problem (NGBD-DEP).

Proof. Any feasible x_h^i for Problem (NGBD-DEP) satisfies $\tilde{\mathbf{g}}_h(\mathbf{x}_h, \mathbf{y}) \leq \mathbf{0}$ for associated vectors $(\mathbf{x}_h, \mathbf{y}) \in \{0,1\}^{n_{x_b}} \times \Pi_{x,h} \times \{0,1\}^{n_y}$. Since $\tilde{\mathbf{g}}_{1,h}^{\text{cv}} + \tilde{\mathbf{g}}_{2,h}^{\text{cv}}$ provides a convex relaxation of $\tilde{\mathbf{g}}_h$ on $[0,1]^{n_{x_b}} \times \Pi_{x,h} \times [0,1]^{n_y}$, there exists $\mathbf{q}_h \in \Theta_h$ such that $\tilde{\mathbf{g}}_{1,h}^{\text{cv}}(\mathbf{x}_h, \mathbf{q}_h) + \tilde{\mathbf{g}}_{2,h}^{\text{cv}}(\mathbf{y}) \leq \tilde{\mathbf{g}}_h(\mathbf{x}_h, \mathbf{y}) \leq \mathbf{0}$ and, therefore, x_h^i is feasible for Problem (FBBT$_x$). □

2.3.3.1.2 Lagrangian relaxation

This section presents the details of a Lagrangian relaxation algorithm for solving Problem (DEP). First, Problem (DEP) is reformulated by explicitly writing out the non-anticipativity constraints corresponding to the (binary and continuous) complicating variables and dualizing them [48, 91] to obtain the Lagrangian dual lower bounding problem (cf. Section 2.3.2.1.3). Next, some techniques that can be used to solve the above dual problem are noted. Finally, we close this section with a few structure-specific techniques that can be used to try and generate feasible points for Problem (DEP).

The Lagrangian relaxation approach [48, 91] begins by creating copies of the first-stage variables, $\mathbf{y}$ and $\mathbf{z}$, for each of the s scenarios, and adding the so-called non-anticipativity constraints that restrict these copies to take the same value across all scenarios.

$$\min_{\substack{\mathbf{x}_1,\cdots,\mathbf{x}_s,\\ \mathbf{y}_1,\cdots,\mathbf{y}_s,\\ \mathbf{z}_1,\cdots,\mathbf{z}_s}} \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h, \mathbf{y}_h, \mathbf{z}_h) \tag{RP}$$

$$\text{s.t. } \mathbf{g}_h(\mathbf{x}_h, \mathbf{y}_h, \mathbf{z}_h) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\},$$

$$\mathbf{r}_{y,z}(\mathbf{y}_h, \mathbf{z}_h) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\},$$

$$\mathbf{y}_h = \mathbf{y}_{h+1}, \ \ \mathbf{z}_h = \mathbf{z}_{h+1}, \quad \forall h \in \{1, \cdots, s-1\}, \tag{NA}$$

$$\mathbf{x}_h \in X_h, \ \ \mathbf{y}_h \in Y, \ \ \mathbf{z}_h \in Z, \quad \forall h \in \{1, \cdots, s\}.$$

The non-anticipativity constraints, Equations (NA), are dualized to obtain the Lagrangian dual problem, Problem (LRP), which is a valid lower bounding problem because of the weak duality theorem (see Theorem 2.3.39).

$$\sup_{\substack{\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_{s-1}\\ \boldsymbol{\gamma}_1,\cdots,\boldsymbol{\gamma}_{s-1}}} \min_{\substack{\mathbf{x}_1,\cdots,\mathbf{x}_s,\\ \mathbf{y}_1,\cdots,\mathbf{y}_s,\\ \mathbf{z}_1,\cdots,\mathbf{z}_s}} \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h, \mathbf{y}_h, \mathbf{z}_h) + \sum_{h=1}^{s-1} \boldsymbol{\gamma}_h^{\mathrm{T}}(\mathbf{y}_h - \mathbf{y}_{h+1}) + \sum_{h=1}^{s-1} \boldsymbol{\beta}_h^{\mathrm{T}}(\mathbf{z}_h - \mathbf{z}_{h+1}) \tag{LRP}$$

$$\text{s.t. } \mathbf{g}_h(\mathbf{x}_h, \mathbf{y}_h, \mathbf{z}_h) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\},$$

$$\mathbf{r}_{y,z}(\mathbf{y}_h, \mathbf{z}_h) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\},$$

$$\mathbf{x}_h \in X_h, \ \ \mathbf{y}_h \in Y, \ \ \mathbf{z}_h \in Z, \quad \forall h \in \{1, \cdots, s\}.$$

For each $h \in \{1, \cdots, s-1\}$, the vectors $\boldsymbol{\gamma}_h \in \mathbb{R}^{n_y}$ and $\boldsymbol{\beta}_h \in \mathbb{R}^{n_z}$ in Problem (LRP) are dual variable/Lagrange multiplier vectors corresponding to the non-anticipativity constraints $\mathbf{y}_h - \mathbf{y}_{h+1} = \mathbf{0}$ and $\mathbf{z}_h - \mathbf{z}_{h+1} = \mathbf{0}$. Therefore, even though Problem (DEP) is potentially a large-scale nonconvex MINLP containing $(sn_x + n_y + n_z)$ decision variables, the inner minimization of the above lower bounding problem decomposes into s independent scenario MINLP subproblems each of which consists of $(n_x + n_y + n_z)$ variables; however, a B&B strategy, which solves a Lagrangian dual problem at each node of the B&B tree, has to be adopted in the $(Y \times Z)$-space to converge to a solution of Problem (DEP). This approach has the advantage that the number of variables to be subdivided in the B&B procedure is independent of the number of scenarios. For ease of exposition, we shall define the

(Lagrangian) dual function as:

$$d(\boldsymbol{\gamma},\boldsymbol{\beta}) := \min_{\substack{\mathbf{x}_1,\cdots,\mathbf{x}_s,\\ \mathbf{y}_1,\cdots,\mathbf{y}_s,\\ \mathbf{z}_1,\cdots,\mathbf{z}_s}} \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h,\mathbf{y}_h,\mathbf{z}_h) + \sum_{h=1}^{s-1} \boldsymbol{\gamma}_h^{\mathrm{T}}(\mathbf{y}_h - \mathbf{y}_{h+1}) + \sum_{h=1}^{s-1} \boldsymbol{\beta}_h^{\mathrm{T}}(\mathbf{z}_h - \mathbf{z}_{h+1}) \tag{LRP-inner}$$

$$\begin{aligned} \text{s.t.}\quad & \mathbf{g}_h(\mathbf{x}_h,\mathbf{y}_h,\mathbf{z}_h) \leq \mathbf{0}, \quad \forall h \in \{1,\cdots,s\},\\ & \mathbf{r}_{y,z}(\mathbf{y}_h,\mathbf{z}_h) \leq \mathbf{0}, \quad \forall h \in \{1,\cdots,s\},\\ & \mathbf{x}_h \in X_h,\ \ \mathbf{y}_h \in Y,\ \ \mathbf{z}_h \in Z, \quad \forall h \in \{1,\cdots,s\}, \end{aligned}$$

where $\boldsymbol{\gamma} := (\boldsymbol{\gamma}_1,\cdots,\boldsymbol{\gamma}_{s-1})$ and $\boldsymbol{\beta} := (\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_{s-1})$.

Although the Lagrangian dual problem, Problem (LRP), is a convex optimization problem, the dual function is typically nondifferentiable and its solution requires the use of a nonsmooth optimization techniques [100], making it tedious to solve to optimality in general. Moreover, evaluation of the dual function at a point involves the solution of several (scenario) nonconvex MINLPs which are in general NP-hard (however, they are likely to be significantly less computationally intensive to solve than Problem (DEP) in practice, especially when the number of scenarios is large). Consequently, typical implementations of the conventional Lagrangian relaxation algorithm only carry out a small predefined number of iterations of a nonsmooth optimization algorithm to obtain lower bounds [112, 119]. The next section lists some widely used approaches for solving the outer dual problem.

2.3.3.1.2.1 Solving the Lagrangian dual problem

Obtaining tight lower bounds for Problem (DEP) using Problem (LRP) might require several iterations of an algorithm in the space of the dual variables which, in turn, may require several calls to a global optimization solver to evaluate the dual function for different values of the dual variables. Note, however, that it is not necessary to maximize the dual function precisely to guarantee a valid lower bound for Problem (DEP) since intermediate iterations of an algorithm applied to the dual provide valid lower bounds as a consequence of the weak duality theorem (see Theorem 2.3.39). Furthermore, the number of variables in the dual problem increases affinely with the number of scenarios considered which might makes precise maximization of the dual function computationally prohibitive for large-scale problems. Consequently, our implementation of the LR algorithm in this thesis (see Chapter 4) considers only a few iterations of an algorithm applied to the dual function to generate lower bounds for Problem (DEP). Below, we list some algorithms that can be used to solve the (outer) dual problem, Problem (LRP).

The dual function is typically nondifferentiable which necessitates the use of nonsmooth convex optimization techniques to solve the (outer) dual problem. A popular technique for solving the Lagrangian dual problem is the subgradient method [212] and its variants [126, 130, 175]. Another approach is to use the cutting plane method for convex optimization [55, 117]; note, however, that pure cutting-plane approaches can be unstable (see [100, Section XV.1.1]). The approach we adopt in this thesis, based on the empirical comparison of nonsmooth optimization software in [110], falls in the broad category of bundle methods [100, 151].

Finally, we note that the solution of the dual problem can be warm started by initializing the dual variables/Lagrange multipliers for Problem (LRP) at a child node using the dual variables corresponding to the lower bounding solution at the parent node in the B&B tree. Therefore, with an initialization of the Lagrange multipliers at the root node of the B&B tree, initial Lagrange multiplier values at the subsequent nodes in the tree are automatically determined.

2.3.3.1.2.2 Upper bounding techniques for solving Problem (DEP)

In this section, we list some structure-specific upper bounding techniques for Problem (DEP) for the scalable generation of feasible points. The first approach that we consider fixes the integer variables (both first- and second-stage variables) in Problem (DEP) to the lower bounding solution obtained by (partially) solving Problem (LRP) at any particular node of the B&B tree, and solving the resulting nonconvex NLP locally (using solvers such as IPOPT [235, 247] or PIPS-NLP [56] that can exploit the decomposable structure of Problem (DEP)) to try and generate feasible solutions. In our computational experience, the above approach usually provides good feasible solutions relatively early on in the B&B tree since the lower bounds obtained by solving Problem (LRP) are usually quite tight (compared to the lower bounds generated using convex relaxation-based B&B approaches for Problem (DEP), which are usually weak early on in the tree especially when the number of scenarios is large).

An alternative and computationally more intensive approach, which guarantees the generation of a feasible point in the limit of the B&B procedure when Problem (DEP) is feasible (see Chapter 3), can also be used to try and generate feasible points and is described below. This upper bounding problem on node n is obtained by fixing the complicating variables $\mathbf{y}$ and $\mathbf{z}$ in Problem (DEP) to the point $(\mathbf{y}^n_{\text{avg}}, \mathbf{z}^n_{\text{avg}}) := \left(\lceil \frac{1}{s}\sum_{h=1}^{s} \mathbf{y}^n_h \rceil, \frac{1}{s}\sum_{h=1}^{s} \mathbf{z}^n_h\right)$, which roughly corresponds to the scenario average of the complicating variable solutions from (an incomplete solution of) the lower bounding problem, Problem (LRP). The resulting Problem (UBP^n) is solved (globally) in a decomposable manner using a global optimization solver:

$$\begin{aligned} \min_{\mathbf{x}_1,\cdots,\mathbf{x}_s} \quad & \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h, \mathbf{y}^n_{\text{avg}}, \mathbf{z}^n_{\text{avg}}) \\ \text{s.t.} \quad & \mathbf{g}_h(\mathbf{x}_h, \mathbf{y}^n_{\text{avg}}, \mathbf{z}^n_{\text{avg}}) \leq \mathbf{0}, \quad \forall h \in \{1,\cdots,s\}, \\ & \mathbf{r}_{y,z}(\mathbf{y}^n_{\text{avg}}, \mathbf{z}^n_{\text{avg}}) \leq \mathbf{0}, \\ & \mathbf{x}_h \in X^n_h, \quad \forall h \in \{1,\cdots,s\}. \end{aligned} \tag{UBP^n}$$

Note that if the constraints $\mathbf{r}_{y,z}(\mathbf{y}^n_{\text{avg}}, \mathbf{z}^n_{\text{avg}}) \leq \mathbf{0}$ trivially hold, they can be eliminated from Problem (UBP^n); otherwise, Problem (UBP^n) is trivially infeasible. If Problem (UBP^n) is feasible when $n = 0$, a feasible point for Problem (DEP) is generated at the root node; otherwise, if the partitioning procedure is exhaustive, either a feasible point will be obtained from the solution of Problem (UBP^n) along a (possibly infinite) decreasing sequence of successively refined partition elements of the B&B tree, or the sequence of successively refined partition elements of the B&B tree will be deemed to be converging to an infeasible point (deletion by infeasibility rule is certain in the limit, see Section 2.3.2.4). A formal proof of a related result is provided in Chapter 3 (also see [47]).

Chapter 3

A modified Lagrangian relaxation algorithm for nonconvex two-stage stochastic mixed-integer nonlinear programs

Mixed-integer nonlinear programs (MINLPs) provide a powerful framework for modeling applications in the chemical process industries. All known global optimization algorithms for solving general classes of MINLPs, however, suffer from a worst-case exponential growth in the runtime with the problem size, which makes the solution of large-scale two-stage stochastic MINLPs using general-purpose software impractical for applications of interest. This chapter develops a Lagrangian relaxation-type decomposition algorithm by integrating two existing decomposition approaches along with efficient domain reduction techniques for solving a broad class of two-stage stochastic MINLPs.

3.1 Introduction

Mixed-integer nonlinear programs provide a powerful framework for modeling many diverse applications, including airline scheduling, concrete structure design, environmental planning, portfolio optimization, supply chain optimization, traffic planning, and process systems engineering [18, 226]. Over the past few decades, several local and global optimization algorithms have been proposed for the solution of MINLPs [44, 80, 226]. These techniques have been successfully implemented in the general-purpose state-of-the-art deterministic global optimization software ANTIGONE [162], BARON [225], Couenne [19], LINDOGlobal [145], and SCIP [233]. With significant advancements in optimization techniques and software, there has been a growing interest in rigorously accounting for uncertainties in optimization models [21, 35, 181, 196], especially when their impact on the decision process is significant. In this chapter, we propose a decomposition strategy for the solution of the following class of nonconvex two-stage stochastic programs with recourse (see Problem (DEP) in Section 2.3.3.1 of Chapter 2):

$$\begin{aligned}
\min_{\mathbf{x}_1,\cdots,\mathbf{x}_s,\mathbf{y},\mathbf{z}} \quad & \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h,\mathbf{y},\mathbf{z}) \\
\text{s.t.} \quad & \mathbf{g}_h(\mathbf{x}_h,\mathbf{y},\mathbf{z}) \leq \mathbf{0}, \quad \forall h \in \{1,\cdots,s\}, \\
& \mathbf{r}_{y,z}(\mathbf{y},\mathbf{z}) \leq \mathbf{0}, \\
& \mathbf{x}_h \in X_h, \quad \forall h \in \{1,\cdots,s\}, \\
& \mathbf{y} \in Y, \;\; \mathbf{z} \in Z,
\end{aligned} \tag{DEP}$$

where $X_h := \{\mathbf{x}_h \in \{0,1\}^{n_{x_b}} \times \Pi_{x,h} : \mathbf{r}_{x,h}(\mathbf{x}_h) \leq \mathbf{0}\}$, $\Pi_{x,h} \subset \mathbb{R}^{n_{x_c}}$ is a convex set, $Y := \{\mathbf{y} \in \{0,1\}^{n_y} : \mathbf{r}_y(\mathbf{y}) \leq \mathbf{0}\}$, $Z := \{\mathbf{z} \in \Pi_z : \mathbf{r}_z(\mathbf{z}) \leq \mathbf{0}\}$, $\Pi_z \subset \mathbb{R}^{n_z}$ is convex, $f_h : [0,1]^{n_{x_b}} \times \Pi_{x,h} \times [0,1]^{n_y} \times \Pi_z \to \mathbb{R}$, $\mathbf{g}_h : [0,1]^{n_{x_b}} \times \Pi_{x,h} \times [0,1]^{n_y} \times \Pi_z \to \mathbb{R}^m$, $\mathbf{r}_{x,h} : [0,1]^{n_{x_b}} \times \Pi_{x,h} \to \mathbb{R}^{m_x}$, $\mathbf{r}_y : [0,1]^{n_y} \to \mathbb{R}^{m_y}$, $\mathbf{r}_z : \Pi_z \to \mathbb{R}^{m_z}$, $\mathbf{r}_{y,z} : [0,1]^{n_y} \times \Pi_z \to \mathbb{R}^{m_{y,z}}$, and $p_h > 0$ denotes the probability of occurrence of scenario $h \in \{1,\cdots,s\}$. Along the lines of Chapter 2, we assume that the set X_h is nonempty and compact for each $h \in \{1,\cdots,s\}$, the set Y is nonempty, the set Z is nonempty and compact, and the functions f_h, $\mathbf{g}_h$, and $\mathbf{r}_{x,h}$, $\forall h \in \{1,\cdots,s\}$, $\mathbf{r}_y$, $\mathbf{r}_z$, and $\mathbf{r}_{y,z}$ are continuous. The above assumptions imply that Problem (DEP) either has a finite optimal objective value, or is infeasible. The reader is directed to Section 2.3.3.1 of Chapter 2 for an overview of decomposition algorithms for solving (subclasses of) Problem (DEP).

This chapter is organized as follows. Section 3.2 details the lower bounding problem for Problem (DEP) based on the Lagrangian relaxation and NGBD algorithms that were reviewed in Section 2.3.3.1 of Chapter 2. Section 3.3 details the techniques used to determine feasible solutions and reduce the search space of the decision variables (Section 3.4 highlights relevant properties of the subproblems described in Sections 3.2 and 3.3). Section 3.5 details the modified Lagrangian relaxation algorithm along with proofs of convergence. Finally, Section 3.6 presents two numerical case studies which demonstrate some advantages of the proposed decomposition approach over state-of-the-art deterministic global optimization software and the conventional Lagrangian relaxation algorithm, and Section 3.7 provides some concluding remarks.

3.2 Reformulation and the lower bounding problem

This section presents the reformulation and relaxation procedures applied to Problem (DEP) to derive a valid lower bounding problem. We begin by reformulating Problem (DEP) by explicitly writing out the non-anticipativity constraints corresponding to the continuous complicating variables. Then, these non-anticipativity constraints are dualized in a manner similar to the conventional Lagrangian relaxation technique [48, 91] to obtain the lower bounding problem (the inner minimization of which can be solved in a decomposable manner using NGBD) for the modified Lagrangian relaxation algorithm.

The conventional Lagrangian relaxation approach [48, 91], see Section 2.3.3.1.2, involves creating copies of the first-stage variables, $\mathbf{y}$ and $\mathbf{z}$, for each of the s scenarios, adding the so-called non-anticipativity constraints that restrict these copies to take the same value across all scenarios, dualizing these non-anticipativity constraints to yield a decomposable lower bounding problem for Problem (DEP), and adopting a B&B strategy in the $(Y \times Z)$-space to converge to a solution of Problem (DEP). In this work, copies of only the continuous first-stage variables $\mathbf{z}$ are created along with the corresponding non-anticipativity constraints (Equation (NA)). This results in Problem (RP), which is equivalent to Problem (DEP) as stated in Proposition 3.4.1. It is worth noting that the form of the non-anticipativity constraints used in Problem (RP) is merely one of several ways in which the continuous first-stage variables corresponding to the different scenarios can be linked together.

$$\begin{aligned}
\min_{\substack{\mathbf{x}_1,\cdots,\mathbf{x}_s,\mathbf{y},\\ \mathbf{z}_1,\cdots,\mathbf{z}_s}} \quad & \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}_h) && \text{(RP)} \\
\text{s.t.} \quad & \mathbf{g}_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}_h) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\}, \\
& \mathbf{r}_{y,z}(\mathbf{y}, \mathbf{z}_h) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\}, \\
& \mathbf{z}_h = \mathbf{z}_{h+1}, \ \ \forall h \in \{1, \cdots, s-1\}, && \text{(NA)} \\
& \mathbf{x}_h \in X_h, \ \ \mathbf{z}_h \in Z, \quad \forall h \in \{1, \cdots, s\}, \\
& \mathbf{y} \in Y.
\end{aligned}$$

As in the conventional Lagrangian relaxation approach, the non-anticipativity constraints, Equation (NA), are dualized to obtain Problem (LRP), which is a valid lower bounding problem as stated in Proposition 3.4.2:

$$\begin{aligned}
\sup_{\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_{s-1}} \min_{\substack{\mathbf{x}_1,\cdots,\mathbf{x}_s,\mathbf{y},\\ \mathbf{z}_1,\cdots,\mathbf{z}_s}} \quad & \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}_h) + \sum_{h=1}^{s-1} \boldsymbol{\beta}_h^{\mathrm{T}} (\mathbf{z}_h - \mathbf{z}_{h+1}) && \text{(LRP)} \\
\text{s.t.} \quad & \mathbf{g}_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}_h) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\}, \\
& \mathbf{r}_{y,z}(\mathbf{y}, \mathbf{z}_h) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\}, \\
& \mathbf{x}_h \in X_h, \ \ \mathbf{z}_h \in Z, \quad \forall h \in \{1, \cdots, s\}, \\
& \mathbf{y} \in Y,
\end{aligned}$$

where, for each $h \in \{1, \cdots, s-1\}$, $\boldsymbol{\beta}_h \in \mathbb{R}^{n_z}$ is a Lagrange multiplier vector corresponding to the constraints $\mathbf{z}_h - \mathbf{z}_{h+1} = \mathbf{0}$. The above lower bounding problem provides lower bounds that are at least as tight as the corresponding lower bounding problem in the conventional Lagrangian relaxation algorithm (see Proposition 3.4.3). Furthermore, if the above lower bounding problem is used in a branch-and-bound scheme, it is sufficient to branch on the continuous complicating variables to converge. This is in contrast to the conventional Lagrangian relaxation-based branch-and-bound procedure where it is in general necessary to also branch on the binary complicating variables to guarantee convergence. For ease of exposition, we shall define the (Lagrangian) dual function as:

$$
\begin{aligned}
d(\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_{s-1}) := \min_{\substack{\mathbf{x}_1,\cdots,\mathbf{x}_s,\mathbf{y},\\ \mathbf{z}_1,\cdots,\mathbf{z}_s}} \quad & \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}_h) + \sum_{h=1}^{s-1} \boldsymbol{\beta}_h^{\mathrm{T}}(\mathbf{z}_h - \mathbf{z}_{h+1}) \qquad \text{(LRP-inner)} \\
\text{s.t.} \quad & \mathbf{g}_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}_h) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\}, \\
& \mathbf{r}_{y,z}(\mathbf{y}, \mathbf{z}_h) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\}, \\
& \mathbf{x}_h \in X_h, \ \ \mathbf{z}_h \in Z, \quad \forall h \in \{1, \cdots, s\}, \\
& \mathbf{y} \in Y.
\end{aligned}
$$

Problem (LRP-inner), which evaluates the dual function for given values of the Lagrange multipliers (also referred to as dual variables), is not in a decomposable form since the binary first-stage variables, $\mathbf{y}$, couple the scenario recourse problems together; however, it is in a form that can be solved in a decomposable manner using NGBD [138], see Section 2.3.3.1.1 of Chapter 2[1]. If the assumptions of Theorem 2.3.61 in Section 2.3.3.1.1.5 are satisfied, we have that the NGBD algorithm converges finitely to an ε-optimal solution to Problem (LRP-inner), i.e., to a feasible point of Problem (LRP-inner) with an objective value that is accurate to within ε of the value of the Lagrangian dual function (for any given value of the dual variables). We refer the reader to Section 2.3.3.1.2.1 of Chapter 2 for some widely used approaches for solving the (outer) Lagrangian dual problem, Problem (LRP).

3.3 Upper bounding and bounds tightening techniques

The efficiency of a branch-and-bound implementation not only depends on generating a rapidly converging sequence of tight lower bounds, but also on the effectiveness of the upper bounding scheme and the ability to reduce the search space using bounds tightening techniques. In this section, we outline a few upper bounding and bounds tightening techniques for Problem (DEP), most of which exploit its nearly decomposable structure. Before we proceed, we establish some additional notation than the ones listed in Chapter 2 that will be useful for the rest of this chapter.

[1] We will freely adapt the definitions and results in Section 2.3.3.1.1 to Problem (LRP-inner) in the remainder of this chapter.

Notation

Positive parameters ε, ε_l, and ε_u are chosen as termination tolerances for solving Problem (DEP), Problem (LRP-inner), and the upper bounding problems in Section 3.3.1, respectively, such that $\varepsilon_l + \varepsilon_u \leq \varepsilon$. Each node in the B&B tree is assigned (and referred to using) a unique index n from the set $\mathbb{N} \cup \{0\}$, with the index '0' assigned to the root node of the tree. The level of node n in the B&B tree is denoted by $\mathcal{L}^n$ with $\mathcal{L}^0 := 0$ and $\mathcal{L}^{n_c} := \mathcal{L}^{n_p} + 1$ if n_c is a child node of n_p.

Sets $X_h^n \subset \{0,1\}^{n_{x_b}} \times \Pi_{x,h}$, $Y^n \subset \{0,1\}^{n_y}$, and $Z^n \subset \Pi_z$ (not to be confused with the n-fold Cartesian product of the Z sets; we never use this symbol to denote such a product in this chapter) denote the domains of variables $\mathbf{x}_h$, $\forall h \in \{1, \cdots, s\}$, $\mathbf{y}$, and $\mathbf{z}$, respectively, on node n. The sets D_h^n and $\bar{Y}^n$, defined by

$$D_h^n := \left\{(\mathbf{x}_h, \mathbf{z}_h, \mathbf{q}_h) \in \operatorname{conv}(X_h^n) \times Z^n \times \Theta_h^n : \mathbf{r}_{x,h}^{\mathrm{cv}}(\mathbf{x}_h, \mathbf{q}_h) \leq \mathbf{0}, \mathbf{r}_z^{\mathrm{cv}}(\mathbf{z}_h, \mathbf{q}_h) \leq \mathbf{0}\right\} \text{ and}$$
$$\bar{Y}^n := \left\{\mathbf{y} \in Y^n : \exists \mathbf{v} \in \Gamma^n \text{ s.t. } \mathbf{r}_y^{\mathrm{cv}}(\mathbf{y}, \mathbf{v}) \leq \mathbf{0}, \mathbf{v}^{\mathrm{cv}}(\mathbf{y}, \mathbf{v}) \leq \mathbf{0}\right\},$$

denote the sets D_h and $\bar{Y}$, respectively, in Problem (NGBD-LBP-NS) that are constructed on node n when the NGBD algorithm is employed to solve Problem (LRP-inner). Note that although the B&B algorithm we propose does not branch on the variables $\mathbf{x}_h$, $\forall h \in \{1, \cdots, s\}$, and $\mathbf{y}$, their corresponding domains on node n, X_h^n and Y^n, may be updated using the bounds tightening techniques described in Section 2.3.3.1.1.6 of Chapter 2 and Section 3.3.2 and propagated to child nodes since doing so could help generate tighter relaxations for the NGBD subproblems and tighter lower bounds for Problem (DEP). We denote the 'domain' of (the variables in) node n by the set $M^n := X_1^n \times \cdots \times X_s^n \times Y^n \times Z^n$.

For any $i \in \{1, \cdots, n_{x_b} + n_{x_c}\}$, the i^{th} recourse variable corresponding to scenario h is denoted by x_h^i, and for any $i \in \{1, \cdots, n_z\}$, z^i is used to denote the i^{th} continuous complicating variable. We let $z^{i,n,L}$ and $z^{i,n,U}$ denote the lower and upper bounds for z^i on node n, i.e., $z^i \in [z^{i,n,L}, z^{i,n,U}]$ on node n.

We let $UBD \in \mathbb{R} \cup \{+\infty\}$ denote the best found objective value for Problem (DEP) in the B&B tree, and $(\mathbf{x}_1^*, \cdots, \mathbf{x}_s^*, \mathbf{y}^*, \mathbf{z}^*)$ denote a corresponding (feasible) solution when $UBD < +\infty$. The lower bound on node n in the B&B tree is denoted by $LBD^n \in \mathbb{R} \cup \{-\infty, +\infty\}$. The initial value of the Lagrange multipliers for Problem (LRP) on node n is denoted by $(\boldsymbol{\beta}_1^{n,0}, \cdots, \boldsymbol{\beta}_{s-1}^{n,0})$, and $(\boldsymbol{\beta}_1^{n,*}, \cdots, \boldsymbol{\beta}_{s-1}^{n,*})$ denotes the Lagrange multipliers corresponding to the lower bound LBD^n with $(\mathbf{x}_1^n, \cdots, \mathbf{x}_s^n, \mathbf{y}^n, \mathbf{z}_1^n, \cdots, \mathbf{z}_s^n)$ denoting a ε_l-optimal solution of the dual function evaluated at $(\boldsymbol{\beta}_1^{n,*}, \cdots, \boldsymbol{\beta}_{s-1}^{n,*})$, i.e., $d(\boldsymbol{\beta}_1^{n,*}, \cdots, \boldsymbol{\beta}_{s-1}^{n,*})$. Finally, we define the indicator function $\texttt{fathom} : (\mathbb{R} \cup \{+\infty\}) \times (\mathbb{R} \cup \{+\infty\}) \times \mathbb{R}_+ \to \{0, 1\}$ such that $\texttt{fathom}(LBD, UBD, \varepsilon)$ evaluates to one if a node with the lower bound LBD can be fathomed based on the current best upper bound UBD relative to the termination tolerance ε, and evaluates to zero otherwise.

3.3.1 Upper bounding problems

An efficient B&B algorithm for Problem (DEP) would entail the generation of good feasible solutions, if they exist, that are close in objective function value to a global optimal solution early on in the B&B tree. Good feasible solutions help fathom nodes of the branch-and-bound tree by value dominance, thereby accelerating the convergence of branch-and-bound algorithms that otherwise converge using only a combination of consistent bounding operations and exhaustive partitioning [101]. While several heuristic approaches that look to generate good feasible solutions have been proposed in the literature (see [28, 60], for instance), only a few simple (but scalable) techniques are used in this work.

The simplest heuristics that can be used to try and generate feasible solutions involves solving Problem (DEP) using a convex MINLP solver (such as BONMIN [40] or DICOPT [90]) or a nonconvex MINLP solver with a predefined limit on the solution time. Alternatively, the integer variables in Problem (DEP) can be fixed to the lower bounding solution obtained by (partially) solving Problem (LRP) at any particular node of the B&B tree, and the resulting nonconvex NLP can be solved locally (using solvers such as IPOPT [235, 247] or PIPS-NLP [56] that can exploit the decomposable structure of Problem (DEP)) to try and generate feasible solutions, see Section 2.3.3.1.2.2. It is empirically observed from the computational studies (see Section 3.6 and Chapter 4) that the latter approach provides good feasible solutions in the cases when the former approach fails, since the lower bounds obtained by solving Problem (LRP) are usually quite tight.

A third approach, which guarantees the generation of a feasible point in the limit when Problem (DEP) is feasible, can also be used to try and generate feasible points and is described below (also see the related approach in Section 2.3.3.1.2.2). This upper bounding problem on node n is obtained by fixing the continuous complicating variables $\mathbf{z}$ in Problem (DEP) to the point $\mathbf{z}_{\text{avg}}^n := \frac{1}{s} \sum_{h=1}^{s} \mathbf{z}_h^n$, which corresponds to the scenario average of the continuous complicating variable solutions for the lower bounding problem[2], and the resulting Problem (UBP-NGBD^n) is solved (globally) in a decomposable manner using NGBD:

$$\begin{aligned}
\min_{\mathbf{x}_1,\cdots,\mathbf{x}_s,\mathbf{y}} \quad & \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}^n_{\text{avg}}) \\
\text{s.t.} \quad & \mathbf{g}_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}^n_{\text{avg}}) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\}, \\
& \mathbf{r}_{y,z}(\mathbf{y}, \mathbf{z}^n_{\text{avg}}) \leq \mathbf{0}, \\
& \mathbf{x}_h \in X^n_h, \quad \forall h \in \{1, \cdots, s\}, \\
& \mathbf{y} \in Y^n.
\end{aligned} \tag{UBP-NGBD^n}$$

Note that Problem (UBP-NGBD^n) only contains binary complicating variables $\mathbf{y}$; therefore, it can be solved in a decomposable manner to ε_u-optimality using NGBD. If Problem (UBP-NGBD^n) is feasible when $n = 0$, a feasible point for Problem (DEP) is generated at the root node; otherwise, if the partitioning procedure is exhaustive, either a feasible point will be obtained from the solution of Problem (UBP-NGBD^n) along a (possibly infinite) decreasing sequence of successively refined partition elements of the branch-and-bound tree, or the sequence of successively refined partition elements of the branch-and-bound tree will be deemed to be converging to an infeasible point (deletion by infeasibility rule is certain in the limit) [101]. A formal proof of the above claim is provided in Section 3.5.2.

Finally, feasible solutions can also potentially be generated by attempting to solve Problem (DEP) using NGBD (although the NGBD algorithm does not necessarily converge to an optimal solution of Problem (DEP) because of the presence of continuous complicating variables in its formulation, it can potentially identify feasible points for Problem (DEP)). In this implementation, the inner GBD-like loop of NGBD determines a solution to a relaxation of Problem (DEP) (one which is similar to Problem (NGBD-LBP), but without copies of the continuous complicating variables and sans the use of Lagrangian duality) by iteratively restricting the (binary and continuous) complicating variables to various points in their domains. Problem (DEP) is then solved at iterates of the complicating variables for which the above relaxation was feasible to try and generate upper bounds along with associated feasible points in a decomposable manner.

[2]It will become clear from the proofs of Theorems 3.5.5 and 3.5.8 in Section 3.5.2 that the point $\mathbf{z}^n_{\text{avg}}$ is not the only suitable choice in this context.

3.3.2 Bounds tightening techniques

This section lists a couple of decomposable bounds tightening subproblems in addition to the ones listed in Section 2.3.3.1.1.6 for reducing the domains of the continuous complicating variables in Problem (DEP).

The first bounds tightening approach that we consider is a feasibility-based approach that involves the solution of auxiliary (scenario) optimization problems. Lower bounds on the continuous complicating variable z^i, for any $i \in \{1, \cdots, n_z\}$, on node n can be obtained from the solution of the following set of convex MINLPs:

$$\begin{aligned} \max_{h \in \{1,\cdots,s\}} \min_{\mathbf{x}_h, \mathbf{y}, \mathbf{z}_h, \mathbf{q}_h} \quad & z^i_h \\ \text{s.t.} \quad & \tilde{\mathbf{g}}^{\text{cv}}_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}_h, \mathbf{q}_h) \leq \mathbf{0}, \\ & (\mathbf{x}_h, \mathbf{z}_h, \mathbf{q}_h) \in D^n_h, \;\; \mathbf{y} \in \bar{Y}^n. \end{aligned} \tag{FBBT^n_z}$$

Note that Problem (FBBT^n_z) involves the solution of s convex MINLPs each of which can either be solved directly, or further relaxed to a MILP, if necessary, and solved subsequently. Problem (FBBT^n_z) creates copies of the continuous complicating variables for each scenario h, minimizes z^i by only considering the constraints in scenario h of Problem (NGBD-LBP-NS), and imposes the restriction that the lower bounds for z^i have to be the same over all scenarios $h = 1, \cdots, s$. The number of bounding problems solved to determine lower bounds on all the continuous complicating variables is sn_z, which is linear in the number of scenarios. Upper bounds on z^i can be determined in an analogous manner by replacing the outer max with a min and the inner min with a max in Problem (FBBT^n_z). In what follows, we describe an optimality-based bounds tightening technique that is used the tighten the bounds on the continuous complicating variables based on aggressive bounds tightening (ABT) [19].

Suppose UBD is the current best upper bound for Problem (DEP) and bounds on the i^{th} continuous complicating variable on node n are given by $z^i \in [z^{i,n,L}, z^{i,n,U}]$. If, for some candidate point $\bar{z}^{i,n} \in (z^{i,n,L}, z^{i,n,U})$ and candidate Lagrange multiplier vector $(\bar{\boldsymbol{\beta}}^n_1, \cdots, \bar{\boldsymbol{\beta}}^n_{s-1}) \in \mathbb{R}^{(s-1)n_z}$, either the feasibility-based bounds tightening techniques described above determine that restricting z^i to $[z^{i,n,L}, \bar{z}^{i,n}]$ makes Problem (DEP) infeasible on node n, or the solution of the following (lower bounding) problem is greater than UBD, then $z^i \in [\bar{z}^{i,n}, z^{i,n,U}]$ is a valid bound tightening for node n:

$$\begin{aligned}
\min_{\substack{\mathbf{x}_1,\cdots,\mathbf{x}_s,\mathbf{y},\\ \mathbf{z}_1,\cdots,\mathbf{z}_s}} \quad & \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}_h) + \sum_{h=1}^{s-1} \left(\bar{\boldsymbol{\beta}}_h^n\right)^{\mathrm{T}} (\mathbf{z}_h - \mathbf{z}_{h+1}) \\
\text{s.t.} \quad & \mathbf{g}_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}_h) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\}, \\
& \mathbf{r}_{y,z}(\mathbf{y}, \mathbf{z}_h) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\}, \\
& \mathbf{x}_h \in X_h^n, \ \ \mathbf{z}_h \in Z^n \cap \left\{\mathbf{z} : z^i \leq \bar{z}^{i,n}\right\}, \quad \forall h \in \{1, \cdots, s\}, \\
& \mathbf{y} \in Y^n.
\end{aligned} \tag{ABT$_z^n$}$$

Problem (ABT$_z^n$) is similar to Problem (LRP-inner) on node n for a given value of the Lagrange multipliers except for the domain of the continuous complicating variable z^i. Moreover, it is also a nonconvex MINLP that can be solved in a decomposable manner using NGBD. However, a key difference is that the solution of Problem (ABT$_z^n$) need not be carried out to completion since we only wish to determine if a valid bound tightening can be achieved; if, during the course of the NGBD algorithm, a feasible solution for Problem (ABT$_z^n$) is found that has a smaller objective value than UBD (relative to the tolerance ε), the NGBD algorithm can be terminated because the region $z^i \in [z^{i,n,L}, \bar{z}^{i,n}]$ cannot be fathomed based on optimality arguments. Furthermore, if, during the course of the NGBD algorithm, the lower bound for the objective value of Problem (ABT$_z^n$) is determined to be larger than UBD, the NGBD algorithm can be terminated and $z^i \in [\bar{z}^{i,n}, z^{i,n,U}]$ can be declared a valid tightening for the domain of z^i on node n. Upper bounds for the continuous complicating variables are potentially tightened on node n by restricting $z^i \in [\bar{z}^{i,n}, z^{i,n,U}]$ and solving a problem similar to Problem (ABT$_z^n$). We also remark that the application of ABT to the binary first-stage variables results in a procedure similar to probing [198] for those variables.

Problem (ABT$_z^n$) can be extended to involve the iterative solution of several bounding problems for different values of Lagrange multipliers similar to Problem (LRP); however, only one value for the Lagrange multipliers is used in this work since it is desired to minimize computational time for bounding problems. However, several rounds of ABT may be carried out on a per-variable basis at each node of the tree to obtain good bounds. It should be noted that a good upper bound UBD is essential for the ABT step to tighten the bounds on the continuous complicating variables effectively.

3.4 Properties of the subproblems

The following results, in conjunction with the results in Section 2.3.3.1.1.7 of Chapter 2, validate the workings of the proposed modified Lagrangian relaxation algorithm.

Proposition 3.4.1. Problems (RP) and (DEP) are equivalent in the sense that:

1. Problem (RP) is feasible if and only if Problem (DEP) is feasible;

2. The optimal objective function values of Problem (RP) and Problem (DEP) are the same.

3. The point $(\mathbf{x}_1, \cdots, \mathbf{x}_s, \mathbf{y}, \mathbf{z})$ is an optimal solution to Problem (DEP) if and only if the point $(\mathbf{x}_1, \cdots, \mathbf{x}_s, \mathbf{y}, \mathbf{z}, \cdots, \mathbf{z})$ is an optimal solution to Problem (RP).

Proof. The result follows from the fact that $(\mathbf{x}_1, \cdots, \mathbf{x}_s, \mathbf{y}, \mathbf{z})$ is a feasible solution to Problem (DEP) with objective value UBD if and only if $(\mathbf{x}_1, \cdots, \mathbf{x}_s, \mathbf{y}, \mathbf{z}, \cdots, \mathbf{z})$ is a feasible solution to Problem (RP) with objective value UBD. □

Proposition 3.4.2. The optimal objective value of Problem (LRP) is a lower bound on the optimal objective value of Problem (RP). Moreover, the dual function (defined by (LRP-inner)) evaluated at any given value of the Lagrange multipliers/dual variables provides a valid lower bound to the optimal objective value of Problem (RP).

Proof. This follows from weak duality. □

Proposition 3.4.3. Problem (LRP) provides lower bounds that are at least as strong as those provided by the lower bounding problem of the conventional Lagrangian relaxation algorithm.

Proof. The lower bounding problem of conventional Lagrangian relaxation is given by

$$\sup_{\substack{\boldsymbol{\beta}_1,\cdots,\boldsymbol{\beta}_{s-1},\\ \boldsymbol{\gamma}_1,\cdots,\boldsymbol{\gamma}_{s-1}}} \min_{\substack{\mathbf{x}_1,\cdots,\mathbf{x}_s,\\ \mathbf{y}_1,\cdots,\mathbf{y}_s,\\ \mathbf{z}_1,\cdots,\mathbf{z}_s}} \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h, \mathbf{y}_h, \mathbf{z}_h) + \sum_{h=1}^{s-1} \boldsymbol{\beta}_h^{\mathrm{T}}(\mathbf{z}_h - \mathbf{z}_{h+1}) + \sum_{h=1}^{s-1} \boldsymbol{\gamma}_h^{\mathrm{T}}(\mathbf{y}_h - \mathbf{y}_{h+1}) \tag{LRP-conv}$$

$$\begin{aligned} \text{s.t. } & \mathbf{g}_h(\mathbf{x}_h, \mathbf{y}_h, \mathbf{z}_h) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\}, \\ & \mathbf{r}_{y,z}(\mathbf{y}_h, \mathbf{z}_h) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\}, \\ & \mathbf{x}_h \in X_h, \;\; \mathbf{y}_h \in Y, \;\; \mathbf{z}_h \in Z, \quad \forall h \in \{1, \cdots, s\}, \end{aligned}$$

where for each $h \in \{1, \cdots, s-1\}$, $\boldsymbol{\gamma}_h \in \mathbb{R}^{n_y}$ is a Lagrange multiplier corresponding to the relaxed constraint $(\mathbf{y}_h - \mathbf{y}_{h+1}) = \mathbf{0}$.

Since, for each $(\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_{s-1}) \in \mathbb{R}^{(s-1)n_z}$ and $(\boldsymbol{\gamma}_1, \cdots, \boldsymbol{\gamma}_{s-1}) \in \mathbb{R}^{(s-1)n_y}$, the inner minimization of Problem (LRP-conv) provides a lower bound to the inner minimization of Problem (LRP) (by weak duality), the desired result follows. $\square$

Corollary 3.4.4. Consider the solution of Problem (LRP-inner) using NGBD for a given value of the Lagrange multipliers on node n of the branch-and-bound tree. Suppose, at a particular iteration of the NGBD algorithm, the optimal objective value of the adaptation of Problem (NGBD-RMP) for solving Problem (LRP-inner) is greater than UBD, and the minimum of all of the objective values of the adaptations of Problem (NGBD-PP) with $\mathbf{y}$ fixed to each of the feasible integer realizations visited thus far by (adaptations of) Problem (NGBD-PBP) are greater than UBD. Then, the optimal objective value of Problem (LRP-inner) is greater than UBD and node n can be fathomed.

Proof. The result follows from Propositions 3.4.2, 2.3.62, 2.3.63, 2.3.65, and 2.3.67. $\square$

Proposition 3.4.5. The optimal objective value of Problem (FBBT_z^n) is lesser than any feasible value of z^i on node n for Problem (DEP).

Proof. Using arguments similar to the proof of Proposition 2.3.68, the inner minimization of Problem (FBBT_z^n) can be shown to provide a valid lower bound for z_h^i.

Suppose $z_1^{i,n,L}, \cdots, z_s^{i,n,L}$ are the corresponding lower bounds for $z_1^i, \cdots, z_s^i$, respectively, on node n of the branch-and-bound tree, and assume not all $z_1^{i,n,L}, \cdots, z_s^{i,n,L}$ are equal (otherwise the result follows immediately). Consider any point in $X_1^n \times \cdots \times X_s^n \times Y^n \times Z^n \times \cdots \times Z^n$, and let $(z_1^i, \cdots, z_s^i)$ take on values at this point such that $z_h^i \geq z_h^{i,n,L}$, $\forall h \in \{1, \cdots, s\}$. Suppose $\exists \bar{h} \in \{1, \cdots, s\}$ such that $z_{\bar{h}}^i < \max\limits_h \; z_h^{i,n,L}$. Then $(z_1^i, \cdots, z_s^i)$ cannot correspond to a feasible for Problem (RP) since it violates at least one of the constraints $\mathbf{z}_{h'} - \mathbf{z}_{h'+1} = \mathbf{0}$, for some $h' \in \{1, \cdots, s-1\}$. Therefore, for any $h \in \{1, \cdots, s\}$, the lower bound $z^{i,n,L}$ on node n can be redefined to $\max\limits_h \; z_h^{i,n,L}$ without excluding any feasible points for Problem (RP). The desired result follows from Proposition 3.4.1. $\square$

Proposition 3.4.6. Given $\bar{z}^{i,n} \in (z^{i,n,L}, z^{i,n,U})$ and $(\bar{\boldsymbol{\beta}}_1^n, \cdots, \bar{\boldsymbol{\beta}}_{s-1}^n) \in \mathbb{R}^{(s-1)n_z}$, if the solution of Problem (ABT_z^n) is greater than UBD, then $z^i \in [z^{i,n,L}, \bar{z}^{i,n}]$ does not contain a feasible solution with objective value lower than UBD, i.e., $z^i \in [\bar{z}^{i,n}, z^{i,n,U}]$ is a valid bound tightening.

Proof. This follows from Propositions 3.4.1 and 3.4.2, and the fact that Problem (ABT_z^n) is a restriction of Problem (LRP-inner). $\square$

Corollary 3.4.7. If, at any point during the solution of Problem (ABT_z^n) using NGBD, the optimal objective value of the relaxed master problem of NGBD becomes greater than UBD and the inner GBD-like loop of NGBD has not converged, then the optimal objective value of Problem (ABT_z^n) is greater than UBD and $z^i \in [\bar{z}^{i,n}, z^{i,n,U}]$ is a valid bound tightening. Moreover, if a feasible solution to the primal problem of NGBD with an objective value less than UBD is determined for Problem (ABT_z^n), then the optimal objective value of Problem (ABT_z^n) is less than UBD and Problem (ABT_z^n) cannot be used to show $z^i \in [\bar{z}^{i,n}, z^{i,n,U}]$ is a valid bound tightening.

Proof. The first part follows from Corollary 3.4.4 and Proposition 3.4.6. The second part follows from Proposition 2.3.64. $\square$

Proposition 3.4.8. Let n_p and n_c be a parent-child node pair in the branch-and-bound tree. Suppose $(\boldsymbol{\beta}_1^{n_p,*}, \cdots, \boldsymbol{\beta}_{s-1}^{n_p,*})$ are Lagrange multipliers corresponding to the lower bound LBD^{n_p} obtained by (partially) solving Problem (LRP) on node n_p. Then, the optimal objective value of Problem (LRP-inner) on node n_c with the Lagrange multipliers fixed to $(\boldsymbol{\beta}_1^{n_p,*}, \cdots, \boldsymbol{\beta}_{s-1}^{n_p,*})$ is greater than or equal to the lower bound on node n_p.

Proof. This follows from the fact that for $(\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_{s-1})$ fixed to $(\boldsymbol{\beta}_1^{n_p,*}, \cdots, \boldsymbol{\beta}_{s-1}^{n_p,*})$, Problem (LRP-inner) on node n_c is a restriction of Problem (LRP-inner) on node n_p with the same objective function. $\square$

3.5 A modified Lagrangian relaxation algorithm

In this section, an outline of a modified Lagrangian relaxation (hereafter abbreviated to MLR) algorithm for Problem (DEP) is provided based on the general algorithmic structure of branch-and-bound algorithms [101] and an application of the Lagrangian relaxation technique [48, 91] for Problem (DEP). In addition, proofs of convergence of the algorithm to an ε-optimal solution of Problem (DEP) are provided. Before we present a formal outline of the algorithm, we first provide an informal description of the workings of the algorithm below.

The MLR algorithm is a reduced-space branch-and-bound algorithm in which only the domains of the continuous complicating variables $\mathbf{z}$ are partitioned during the course of the algorithm. On any particular node of the branch-and-bound tree, the feasibility-based bounds tightening techniques described in Section 2.3.3.1.1.6 and Section 3.3.2 are first applied to tighten the bounds on the variables in Problem (DEP). Next, Problem (DEP) is solved (locally) using the upper bounding problems described in Section 3.3.1 to try and determine a better upper bound and associated feasible point (while this step is listed as optional in Algorithm 3.1, a finite upper bound is necessary to be able to apply the aggressive bounds tightening technique; hence, we carry out this step at the root node of the branch-and-bound tree). If a good upper bound has been determined, the expensive aggressive bounds tightening technique detailed in Section 3.3.2 is employed to try and reduce the domains of the continuous complicating variables. Next, a few iterations of an algorithm applied to the Lagrangian dual problem, Problem (LRP), are employed to determine a lower bound on the node under consideration. As an aside, we note that Cao and Zavala [47] essentially prescribe using a 'single iteration' of a conventional Lagrangian relaxation algorithm for Problem (LRP) (for the case when it does not contain any binary variables in its formulation) with all of the Lagrange multipliers fixed to zero, which may not work well for applications in which the (samples of the) uncertain parameters affect the optimal objective value of Problem (DEP) significantly (see the integrated crude selection and refinery operation case study in Section 3.6). Problem (DEP) is then solved (locally) using those upper bounding techniques in Section 3.3.1 that utilize the lower bounding solution to either fix the binary variables or the continuous complicating variables in Problem (DEP), or as a suitable initial guess for 'local optimization approaches'. Finally, if the node has not yet been fathomed, we branch on one of the continuous complicating variables and continue with the branch-and-bound procedure by selecting an unfathomed node from the tree.

Throughout this section, we assume, without loss of generality, that the domain of the continuous complicating variables, $\mathbf{z}$, is an interval subset Z of $\mathbb{R}^{n_z}$, i.e., $Z \in \mathbb{IR}^{n_z}$. Additionally, it will be implicitly assumed, without loss of generality, that the sets X_h, $\forall h \in \{1, \cdots, s\}$, and Y satisfy: $X_h = \{0,1\}^{n_{x_b}} \times \Pi_{x,h}$ with $\Pi_{x,h} \in \mathbb{IR}^{n_{x_c}}$, $\forall h \in \{1, \cdots, s\}$, and $Y = \{0,1\}^{n_{x_b}}$.

3.5.1 Outline of the algorithm

Algorithm 3.1 Modified Lagrangian relaxation algorithm

Initialize:

- Iteration counter $k = 0$ and tolerances $\varepsilon, \varepsilon_l, \varepsilon_u > 0$ such that $\varepsilon_l + \varepsilon_u \leq \varepsilon$.
- Bounds $X_1^0, \cdots, X_s^0, Y^0$, and $Z^0 := \prod_{i=1}^{n_z} [z^{i,0,L}, z^{i,0,U}]$ on $\mathbf{x}_1, \cdots, \mathbf{x}_s, \mathbf{y}, \mathbf{z}$, respectively, on the root node (after the optional application of preprocessing techniques to the input data); domain of the root node $M^0 := X_1^0 \times \cdots \times X_s^0 \times Y^0 \times Z^0$, and the initial partition $\mathcal{P}^0 = \{M^0\}$.
- The maximum number of iterations of an algorithm applied to the Lagrangian dual function, $D_{\max} \geq 1$, and the maximum number of ABT iterations (for each variable) on any node, $B_{\max} \geq 0$.
- Boolean indicator variables $b_{\text{ABT}}^{i,L} = 1, b_{\text{ABT}}^{i,U} = 1$, $\forall i \in \{1, \cdots, n_z\}$, which indicate whether the ABT technique should be employed to try and tighten the lower and upper bounds, respectively, for variable z^i on any given node.
- An integer $1 \leq N < \infty$ and a fraction $0 < \lambda < 1$ to be used in determining the branching strategy.
- Initial values for the Lagrange multipliers, corresponding to the nonanticipativity constraints in Problem (RP), at the root node, $\boldsymbol{\beta}_1^{0,0}, \cdots, \boldsymbol{\beta}_{s-1}^{0,0}$.
- Upper bound $UBD = +\infty$, lower bound at the root node $LBD^0 = -\infty$, and the current best feasible point for Problem (DEP), $\{(\mathbf{x}_1^*, \cdots, \mathbf{x}_s^*, \mathbf{y}^*, \mathbf{z}^*)\} = \emptyset$.

repeat

1. (Node Selection) Pick $n \in \underset{\{n \in \mathbb{N} \cup \{0\} : M^n \in \mathcal{P}^k\}}{\arg\min} LBD^n$ and set $\mathcal{P}^{k+1} = \mathcal{P}^k \backslash \{M^n\}$.
2. (Optional FBBT Step) Apply the feasibility-based bounds tightening techniques described in Section 3.3.2 to update the sets $X_1^n, \cdots, X_s^n$, Z^n. If all sets $X_1^n, \cdots, X_s^n$, and Z^n are nonempty, update M^n and **goto** Step 3. Otherwise, **goto** Step 9.

Algorithm 3.1 Modified Lagrangian relaxation algorithm (continued)

3. (Optional Upper Bounding Step) Solve the upper bounding problem with a termination tolerance of ε_u using any of the techniques described in Section 3.3.1. Update UBD and $(\mathbf{x}_1^*, \cdots, \mathbf{x}_s^*, \mathbf{y}^*, \mathbf{z}^*)$ if a feasible solution better than the current best solution is found.

4. (Optional ABT Step) If $\{(\mathbf{x}_1^*, \cdots, \mathbf{x}_s^*, \mathbf{y}^*, \mathbf{z}^*)\} = \emptyset$, **goto** Step 5. Otherwise, set $b_{\text{ABT}}^{i,L} = b_{\text{ABT}}^{i,U} = 1$, $\forall i \in \{1, \cdots, n_z\}$, and solve Problem ($\text{ABT}_z^n$) as follows:

for $m = 1$ to $B_{\max}$ **do**

for $i \in \left\{j \in \{1, \cdots, n_z\} : b_{\text{ABT}}^{j,L} = 1\right\}$ **do**

i. Pick $\bar{z}^{i,n,m} \in (z^{i,n,L}, z^{i,n,U})$ and solve Problem (ABT_z^n), with the Lagrange multipliers $(\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_{s-1})$ fixed to $(\boldsymbol{\beta}_1^{n,0}, \cdots, \boldsymbol{\beta}_{s-1}^{n,0})$, to ε_l-optimality using NGBD to determine if the lower bound $z^{i,n,L}$ can be tightened.

ii. If the solution of Problem (ABT_z^n) is terminated because the lower bound $z^{i,n,L}$ cannot be tightened, set $b_{\text{ABT}}^{i,L} = 0$. Else, if $z^{i,n,L}$ can be tightened, set $z^{i,n,L} = \bar{z}^{i,n,m}$ and update M^n.

end for

for $i \in \left\{j \in \{1, \cdots, n_z\} : b_{\text{ABT}}^{j,U} = 1\right\}$ **do**

i. Pick $\bar{z}^{i,n,m} \in (z^{i,n,L}, z^{i,n,U})$ and solve Problem (ABT_z^n), with the Lagrange multipliers $(\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_{s-1})$ fixed to $(\boldsymbol{\beta}_1^{n,0}, \cdots, \boldsymbol{\beta}_{s-1}^{n,0})$, to ε_l-optimality using NGBD to determine if the upper bound $z^{i,n,U}$ can be tightened.

ii. If the solution of Problem (ABT_z^n) is terminated because the upper bound $z^{i,n,U}$ cannot be tightened, set $b_{\text{ABT}}^{i,U} = 0$. Else, if $z^{i,n,U}$ can be tightened, set $z^{i,n,U} = \bar{z}^{i,n,m}$ and update M^n.

end for

end for

5. (Lower Bounding Step) Set $(\boldsymbol{\beta}_1^{n,*}, \cdots, \boldsymbol{\beta}_{s-1}^{n,*}) = (\boldsymbol{\beta}_1^{n,0}, \cdots, \boldsymbol{\beta}_{s-1}^{n,0})$, and solve the lower bounding problem on node n as follows:

Algorithm 3.1 Modified Lagrangian relaxation algorithm (continued)

for $l = 0$ to $(D_{\max} - 1)$ **do**

i. Solve Problem (LRP-inner), with the Lagrange multipliers $(\boldsymbol{\beta}_1, \cdots, \boldsymbol{\beta}_{s-1})$ fixed to $(\boldsymbol{\beta}_1^{n,l}, \cdots, \boldsymbol{\beta}_{s-1}^{n,l})$, over M^n to ε_l-optimality using NGBD (see Section 2.3.3.1.1) to obtain a lower bound, say $LBD^{n,l}$, and corresponding solution, say $(\mathbf{x}_1^{n,l}, \cdots, \mathbf{x}_s^{n,l}, \mathbf{y}^{n,l}, \mathbf{z}_1^{n,l}, \cdots, \mathbf{z}_s^{n,l})$.

ii. If the solution of Problem (LRP-inner) is terminated because all of the conditions in Corollary 3.4.4 are satisfied (i.e., we detect that the node can be fathomed either by infeasibility, or by value dominance while solving Problem (LRP-inner) using NGBD), **goto** Step 9. Else, if $LBD^{n,l} > LBD^n$, set $LBD^n = LBD^{n,l}$, $(\mathbf{x}_1^n, \cdots, \mathbf{x}_s^n, \mathbf{y}^n, \mathbf{z}_1^n, \cdots, \mathbf{z}_s^n) = (\mathbf{x}_1^{n,l}, \cdots, \mathbf{x}_s^{n,l}, \mathbf{y}^{n,l}, \mathbf{z}_1^{n,l}, \cdots, \mathbf{z}_s^{n,l})$, and $(\boldsymbol{\beta}_1^{n,*}, \cdots, \boldsymbol{\beta}_{s-1}^{n,*}) = (\boldsymbol{\beta}_1^{n,l}, \cdots, \boldsymbol{\beta}_{s-1}^{n,l})$.

iii. If $l < D_{\max} - 1$, determine the Lagrange multipliers for the next iteration, $\boldsymbol{\beta}_1^{n,l+1}, \cdots, \boldsymbol{\beta}_{s-1}^{n,l+1}$, using an algorithm applied to the Lagrangian dual function.

end for

6. (Optional Upper Bounding Step) Solve Problem (DEP) to ε_u-optimality either by restricting $\mathbf{z}$ to the point $\mathbf{z}_{\text{avg}}^n$, where $z_{\text{avg}}^{i,n}$ denotes the average of $z_1^{i,n}, \cdots, z_s^{i,n}$, using NGBD, or by restricting the binary variables in Problem (DEP) to their lower bounding solutions and solving the resulting problem using a local solver (see Section 3.3.1). Update UBD and $(\mathbf{x}_1^*, \cdots, \mathbf{x}_s^*, \mathbf{y}^*, \mathbf{z}^*)$ if a feasible solution better than the current best solution is found.

7. (Branching)

if $\texttt{fathom}(LBD^n, UBD, \varepsilon) = 0$ **then**

Partition the domain M^n as follows:

if $\mathcal{L}^n \equiv 0 \mod N$ **then**

Determine the continuous complicating variable with the largest relative diameter (relative to the root node)

$$i^* \in \underset{\{i \in \{1,\cdots,n_z\}: z^{i,0,U} \neq z^{i,0,L}\}}{\arg\max} \frac{\left(z^{i,n,U} - z^{i,n,L}\right)}{z^{i,0,U} - z^{i,0,L}},$$

Algorithm 3.1 Modified Lagrangian relaxation algorithm (continued)

and bisect the domain of that variable to determine the domains of the child nodes n_1 and n_2:

$$\begin{aligned}
z^{i,n_j,L} &= z^{i,n,L}, \;\; z^{i,n_j,U} = z^{i,n,U}, \qquad \forall j \in \{1,2\}, i \neq i^*, \\
z^{i^*,n_1,L} &= z^{i^*,n,L}, \;\; z^{i^*,n_1,U} = \frac{1}{2}\left(z^{i^*,n,L} + z^{i^*,n,U}\right), \\
z^{i^*,n_2,L} &= \frac{1}{2}\left(z^{i^*,n,L} + z^{i^*,n,U}\right), \;\; z^{i^*,n_2,U} = z^{i^*,n,U}.
\end{aligned}$$

else

Determine the continuous complicating variable with the largest 'dispersion'

$$i^* \in \underset{\left\{i \in \{1,\cdots,n_z\} : \max\limits_h z_h^{i,n} \neq \min\limits_h z_h^{i,n}\right\}}{\arg\max} \sum_{h=1}^{s} \frac{\left|z_h^{i,n} - z_{\text{avg}}^{i,n}\right|}{\max\limits_h z_h^{i,n} - \min\limits_h z_h^{i,n}},$$

and branch on the domain of that variable (using a convex combination of the scenario-averaged lower bounding solution & the midpoint of that variable's domain) to determine the domains of the child nodes n_1 and n_2:

$$\begin{aligned}
z^{i,n_j,L} &= z^{i,n,L}, \;\; z^{i,n_j,U} = z^{i,n,U}, \qquad \forall j \in \{1,2\}, i \neq i^*, \\
z^{i^*,n_1,L} &= z^{i^*,n,L}, \;\; z^{i^*,n_1,U} = \lambda z_{\text{avg}}^{i^*,n} + (1-\lambda)\left(\frac{z^{i^*,n,L} + z^{i^*,n,U}}{2}\right), \\
z^{i^*,n_2,L} &= \lambda z_{\text{avg}}^{i^*,n} + (1-\lambda)\left(\frac{z^{i^*,n,L} + z^{i^*,n,U}}{2}\right), \;\; z^{i^*,n_2,U} = z^{i^*,n,U}.
\end{aligned}$$

end if

else

goto Step 9.

end if

8. Set $X_h^{n_1} = X_h^{n_2} = X_h^n$, $\forall h \in \{1,\cdots,s\}$, $Y^{n_1} = Y^{n_2} = Y^n$, $\mathcal{P}^{k+1} = \mathcal{P}^{k+1} \cup M^{n_1} \cup M^{n_2}$, $\boldsymbol{\beta}_h^{n_1,0} = \boldsymbol{\beta}_h^{n_2,0} = \boldsymbol{\beta}_h^{n,*}$, $\forall h \in \{1,\cdots,s-1\}$, $\mathcal{L}^{n_1} = \mathcal{L}^{n_2} = \mathcal{L}^n + 1$, and $LBD^{n_1} = LBD^{n_2} = LBD^n$.

Algorithm 3.1 Modified Lagrangian relaxation algorithm (continued)

9. Set $\mathcal{P}^{k+1} = \mathcal{P}^{k+1} \backslash \left\{ n \in \mathbb{N} \cup \{0\} : M^n \in \mathcal{P}^{k+1},\ \texttt{fathom}(LBD^n, UBD, \varepsilon) = 1 \right\}$ and $k = k + 1$.

until $\mathcal{P}^{k+1} = \emptyset$

If $UBD < +\infty$, then $(\mathbf{x}_1^*, \cdots, \mathbf{x}_s^*, \mathbf{y}^*, \mathbf{z}^*)$ is an ε-optimal solution to Problem (DEP).

A flowchart of the MLR algorithm is presented in Figure 3-1.

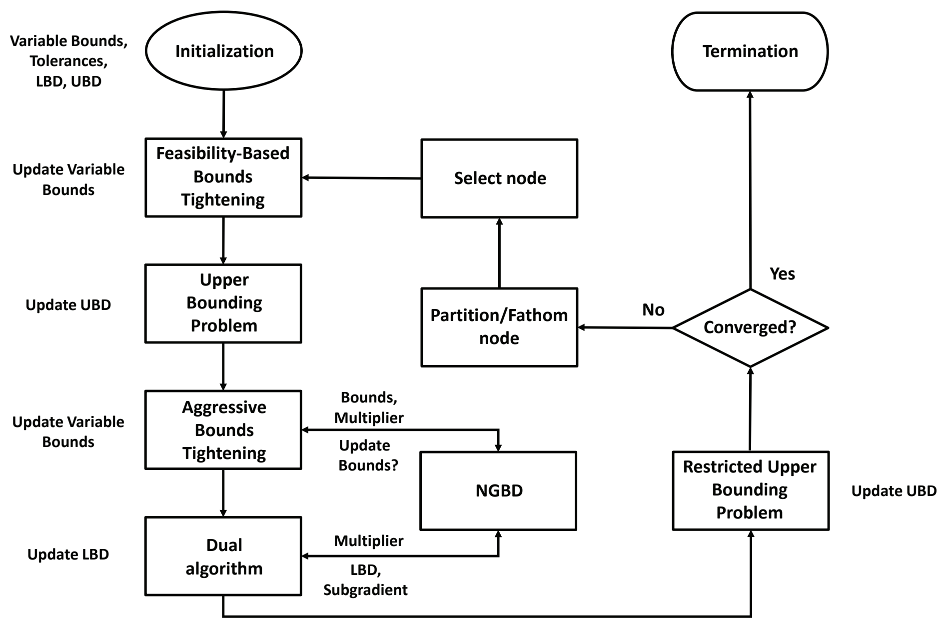


Figure 3-1: Flowchart of the modified Lagrangian relaxation algorithm.

3.5.2 Proof of convergence

In this section, we establish convergence of the above modified Lagrangian relaxation algorithm using the convergence machinery for B&B algorithms reviewed in Section 2.3.2.4 of Chapter 2. Finite convergence of the algorithm is also established under additional assumptions on the set of near-optimal global minimizers for Problem (DEP).

We begin by verifying the requisite conditions of Theorems 2.3.49 and 2.3.50 for the proposed algorithm in a bid to establish its convergence.

Lemma 3.5.1. The subdivision process employed by the MLR algorithm is exhaustive on Z^0.

Proof. This follows from the fact that the branching rule (see Step 7 of Algorithm 3.1) bisects the longest (relative) edge of a partition element if it lies on a level that is a multiple of the user-defined integer parameter N. Therefore, if the longest (relative) edge of Z^0 is of (relative) length Δ_0, the longest (relative) edge of Z^{nNn_z}, which corresponds to the domain of the $\mathbf{z}$ variables at iteration nNn_z of the MLR branch-and-bound algorithm, is at most equal to $\dfrac{\Delta_0}{2^n}$. □

Lemma 3.5.2. The branch-and-bound selection procedure employed by the MLR algorithm is bound improving.

Proof. This is seen to be true upon comparing the node selection procedure in Step 1 of Algorithm 3.1 with Definition 2.3.48. □

Infeasibility of a partition element is detected either using the bounds tightening techniques described in Section 3.3.2, or from the infeasibility of Problem (LRP-inner) while computing lower bounds.

Lemma 3.5.3. Deletion by infeasibility is certain in the limit in the MLR algorithm.

Proof. It suffices to show that every infinite decreasing sequence of successively refined partition elements $\{M^n\} := \{(X_1^n \times \cdots \times X_s^n \times Y^n \times Z^n)\}$ satisfies $M^n \cap \mathcal{F} \neq \emptyset$, where

$$\mathcal{F} := \left\{ (\mathbf{x}_1, \cdots, \mathbf{x}_s, \mathbf{y}, \mathbf{z}) \in \left(\prod_{h=1}^{s} X_h^0 \right) \times Y^0 \times Z^0 : \mathbf{r}_{y,z}(\mathbf{y}, \mathbf{z}) \leq \mathbf{0}, \mathbf{g}_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}) \leq \mathbf{0},\ \forall h \right\},$$

to show that the lower and upper bounding operations result in a consistent bounding operation. Since any sequence in a compact set has a convergent subsequence, it is sufficient to show that any point $(\tilde{\mathbf{x}}_1, \cdots, \tilde{\mathbf{x}}_s, \tilde{\mathbf{y}}, \tilde{\mathbf{z}})$ corresponding to an accumulation point $(\tilde{\mathbf{x}}_1, \cdots, \tilde{\mathbf{x}}_s, \tilde{\mathbf{y}}, \tilde{\mathbf{z}}, \cdots, \tilde{\mathbf{z}})$ of the sequence $\{(\mathbf{x}_1^n, \cdots, \mathbf{x}_s^n, \mathbf{y}^n, \mathbf{z}_1^n, \cdots, \mathbf{z}_s^n)\}$ of lower bounding solutions over $\{M^n\}$ is an element of $\mathcal{F}$. Note that $(\tilde{\mathbf{z}}, \cdots, \tilde{\mathbf{z}})$, where $\{\tilde{\mathbf{z}}\} = \lim_{n \to \infty} Z^n$, is the unique accumulation point of the sequence $\{(\mathbf{z}_1^n, \cdots, \mathbf{z}_s^n)\}$ since the subdivision process is exhaustive on Z. Additionally, the sequence of lower bounding solutions $\{(\mathbf{x}_1^n, \cdots, \mathbf{x}_s^n, \mathbf{y}^n, \mathbf{z}_1^n, \cdots, \mathbf{z}_s^n)\}$ over $\{M^n\}$ corresponds to a sequence of lower bounds $\{LBD^n\}$ with $LBD^n < +\infty$, $\forall n \in \mathbb{N}$, since the corresponding partition elements have not been fathomed. Since $\forall n \in \mathbb{N}$, $X_1^n, \cdots, X_s^n, Y^n, Z^n$ are closed sets, we have that $(\tilde{\mathbf{x}}_1, \cdots, \tilde{\mathbf{x}}_s, \tilde{\mathbf{y}}, \tilde{\mathbf{z}}) \in X_1^n \times \cdots \times X_s^n \times Y^n \times Z^n$ for each $n \in \mathbb{N}$.

Suppose, by way of contradiction, $(\tilde{\mathbf{x}}_1, \cdots, \tilde{\mathbf{x}}_s, \tilde{\mathbf{y}}, \tilde{\mathbf{z}}) \notin \mathcal{F}$. Then the maximum constraint violation at $(\tilde{\mathbf{x}}_1, \cdots, \tilde{\mathbf{x}}_s, \tilde{\mathbf{y}}, \tilde{\mathbf{z}})$ is strictly positive, i.e., there exists $\delta > 0$ such that

$$\max\left\{\max_{h\in\{1,\cdots,s\}} \max_{i\in\{1,\cdots,m\}} g_h^i(\tilde{\mathbf{x}}_h, \tilde{\mathbf{y}}, \tilde{\mathbf{z}}), \max_{j\in\{1,\cdots,m_{y,z}\}} r_{y,z}^j(\tilde{\mathbf{y}}, \tilde{\mathbf{z}})\right\} = \delta > 0,$$

where g_h^i and $r_{y,z}^j$ denote the i^{th} and j^{th} components of $\mathbf{g}_h$ and $\mathbf{r}_{y,z}$, respectively. Suppose the above maximum is attained for the pair of indices (h^*, i^*) with $g_{h^*}^{i^*}(\tilde{\mathbf{x}}_{h^*}, \tilde{\mathbf{y}}, \tilde{\mathbf{z}}) = \delta$ (a similar proof holds if the maximum is attained for some index j^* with $r_{y,z}^{j^*}(\tilde{\mathbf{y}}, \tilde{\mathbf{z}}) = \delta$). Since the constraint functions are all assumed to be continuous, there exists a subsequence $\{M^{n_q}\}$ of $\{M^n\}$, and a corresponding subsequence of lower bounding solutions $\{(\mathbf{x}_1^{n_q}, \cdots, \mathbf{x}_s^{n_q}, \mathbf{y}^{n_q}, \mathbf{z}_1^{n_q}, \cdots, \mathbf{z}_s^{n_q})\}$ such that $(\mathbf{x}_1^{n_q}, \cdots, \mathbf{x}_s^{n_q}) \in X_1^{n_q} \times \cdots \times X_s^{n_q}$ with the limit $\lim\limits_{q\to\infty}(\mathbf{x}_1^{n_q}, \cdots, \mathbf{x}_s^{n_q}) = (\tilde{\mathbf{x}}_1, \cdots, \tilde{\mathbf{x}}_s)$, $\mathbf{y}^{n_q} \in Y^{n_q}$ and $\lim\limits_{q\to\infty}\mathbf{y}^{n_q} = \tilde{\mathbf{y}}$, $\mathbf{z}_h^{n_q} \in Z^{n_q}$, $\forall h$, and $\lim\limits_{q\to\infty}(\mathbf{z}_1^{n_q}, \cdots, \mathbf{z}_s^{n_q}) = (\tilde{\mathbf{z}}, \cdots, \tilde{\mathbf{z}})$, over which the constraint $g_{h^*}^{i^*}$ in Problem (LRP) converges to its limiting value

$$\lim_{q\to\infty} g_{h^*}^{i^*}(\mathbf{x}_{h^*}^{n_q}, \mathbf{y}^{n_q}, \mathbf{z}_{h^*}^{n_q}) = g_{h^*}^{i^*}(\tilde{\mathbf{x}}_{h^*}, \tilde{\mathbf{y}}, \tilde{\mathbf{z}}).$$

Thus, $\exists q_\delta \in \mathbb{N}$ such that $\forall q \geq q_\delta$,

$$\left| g_{h^*}^{i^*}(\mathbf{x}_{h^*}^{n_q}, \mathbf{y}^{n_q}, \mathbf{z}_{h^*}^{n_q}) - g_{h^*}^{i^*}(\tilde{\mathbf{x}}_{h^*}, \tilde{\mathbf{y}}, \tilde{\mathbf{z}}) \right| < \delta.$$

Therefore for $q \geq q_\delta$, $g_{h^*}^{i^*}(\mathbf{x}_{h^*}^{n_q}, \mathbf{y}^{n_q}, \mathbf{z}_{h^*}^{n_q}) > 0$ implying that the lower bounding problem is infeasible for partition elements in the sequence $\{M^{n_q}\}$ beyond $M^{n_{q_\delta}}$, a contradiction.

Since the choice of the accumulation point $(\tilde{\mathbf{x}}_1, \cdots, \tilde{\mathbf{x}}_s, \tilde{\mathbf{y}}, \tilde{\mathbf{z}}, \cdots, \tilde{\mathbf{z}})$ was arbitrary, the desired result follows. $\square$

Lemma 3.5.4. The lower bounding problem employed by the MLR algorithm is strongly consistent.

Proof. Let $\{M^n\} = \{(X_1^n \times \cdots \times X_s^n \times Y^n \times Z^n)\}$ be an infinite nested sequence of successively refined partition elements produced by a subdivision of $X_1 \times \cdots \times X_s \times Y \times Z$ that is exhaustive on Z with $\lim\limits_{n\to\infty} Z^n = \{\dot{\mathbf{z}}\}$. We can assume that Problem (DEP) is feasible for $\mathbf{z}$ restricted to $\dot{\mathbf{z}}$, and consequently, for each $n \in \mathbb{N}$, there exists $(\mathbf{x}_1, \cdots, \mathbf{x}_s, \mathbf{y}) \in X_1^n \times \cdots \times X_s^n \times Y^n$ such that $\mathbf{g}_h(\mathbf{x}_h, \mathbf{y}, \dot{\mathbf{z}}) \leq \mathbf{0}$, $\mathbf{r}_{y,z}(\mathbf{y}, \dot{\mathbf{z}}) \leq \mathbf{0}$ since otherwise, by Lemma 3.5.3, the sequence of partitions $\{M^n\}$ would have been deemed to converge to an infeasible point.

Let $X_h^\infty := \cap_{n=1}^\infty X_h^n \neq \emptyset, \forall h \in \{1, \cdots, s\}$, $Y^\infty := \cap_{n=1}^\infty Y^n \neq \emptyset$, and note that these sets are closed.

Let $(\dot{\mathbf{x}}_1, \cdots, \dot{\mathbf{x}}_s, \dot{\mathbf{y}}, \dot{\mathbf{z}}, \cdots, \dot{\mathbf{z}})$ be any accumulation point of the sequence $\{(\mathbf{x}_1^n, \cdots, \mathbf{x}_s^n, \mathbf{y}^n, \mathbf{z}_1^n, \cdots, \mathbf{z}_s^n)\}$ of lower bounding solutions over $\{M^n\}$. Then $(\dot{\mathbf{x}}_1, \cdots, \dot{\mathbf{x}}_s, \dot{\mathbf{y}}) \in (X_1^\infty \times \cdots \times X_s^\infty \times Y^\infty)$ and from Lemma 3.5.3, $(\dot{\mathbf{x}}_1, \cdots, \dot{\mathbf{x}}_s, \dot{\mathbf{y}}, \dot{\mathbf{z}})$ is feasible for Problem (DEP). Let $\{(\boldsymbol{\beta}_1^{n,*}, \cdots, \boldsymbol{\beta}_{s-1}^{n,*})\}$ be a corresponding sequence of Lagrange multipliers at which the sequence of lower bounds $\{LBD^n\}$ is achieved over $\{M^n\}$ via a (partial) solution of Problem (LRP). We assume without loss of generality that the sequence $\{(\boldsymbol{\beta}_1^{n,*}, \cdots, \boldsymbol{\beta}_{s-1}^{n,*})\}$ is bounded in norm (with a bound that is chosen, for instance, by looking at the first few terms of the sequence); if this assumption were violated for some term of the sequence $(\boldsymbol{\beta}_1^{j,*}, \cdots, \boldsymbol{\beta}_{s-1}^{j,*})$, $j \geq 1$, we can (recursively) replace $(\boldsymbol{\beta}_1^{j,*}, \cdots, \boldsymbol{\beta}_{s-1}^{j,*})$ with $(\boldsymbol{\beta}_1^{j-1,*}, \cdots, \boldsymbol{\beta}_{s-1}^{j-1,*})$ and proceed with the proof (the proof for the original sequence of Lagrange multipliers would follow by sandwich arguments). It suffices to show that there exists a subsequence $\{M^{n_q}\}$ over which

$$
\begin{aligned}
\lim_{q \to \infty} \min_{\substack{(\mathbf{x}_1, \cdots, \mathbf{x}_s, \mathbf{y}) \in (X_1^{n_q} \times \cdots \times X_s^{n_q} \times Y^{n_q}) \\ \mathbf{z}_1, \cdots, \mathbf{z}_s}} \quad & \left[\sum_{h=1}^{s} p_h f_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}_h) + \sum_{h=1}^{s-1} \left(\boldsymbol{\beta}_h^{n_q,*}\right)^{\mathrm{T}} (\mathbf{z}_h - \mathbf{z}_{h+1}) \right] \\
\text{s.t.} \quad & \mathbf{g}_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}_h) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\}, \\
& \mathbf{r}_{y,z}(\mathbf{y}, \mathbf{z}_h) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\}, \\
& \mathbf{z}_h \in Z^{n_q}, \quad \forall h \in \{1, \cdots, s\} \\
= \min_{(\mathbf{x}_1, \cdots, \mathbf{x}_s, \mathbf{y}) \in (X_1^\infty \times \cdots \times X_s^\infty \times Y^\infty)} \quad & \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h, \mathbf{y}, \dot{\mathbf{z}}) \\
\text{s.t.} \quad & \mathbf{g}_h(\mathbf{x}_h, \mathbf{y}, \dot{\mathbf{z}}) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\}, \\
& \mathbf{r}_{y,z}(\mathbf{y}, \dot{\mathbf{z}}) \leq \mathbf{0},
\end{aligned}
\tag{L}
$$

to prove that the lower bounding problem is strongly consistent. Let

$$
\begin{aligned}
\mathcal{F}_{\mathrm{LRP}}^n := \big\{ & (\mathbf{x}_1, \cdots, \mathbf{x}_s, \mathbf{y}, \mathbf{z}_1, \cdots, \mathbf{z}_s) \in \left(X_1^n \times \cdots \times X_s^n \times Y^n \times Z^n \times \cdots \times Z^n\right) : \\
& \mathbf{r}_{y,z}(\mathbf{y}, \mathbf{z}_h) \leq \mathbf{0},\ \mathbf{g}_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}_h) \leq \mathbf{0}, \forall h \in \{1, \cdots, s\} \big\}
\end{aligned}
$$

denote the feasible set of Problem (LRP-inner) on the partition element M^n. Clearly $\mathcal{F}_{\mathrm{LRP}}^{n+1} \subset \mathcal{F}_{\mathrm{LRP}}^n, \forall n \in \mathbb{N}$, and $\mathcal{F}_{\mathrm{LRP}}^\infty = \cap_{n=1}^\infty \mathcal{F}_{\mathrm{LRP}}^n \neq \emptyset$. By virtue of the definition of

$\{(\boldsymbol{\beta}_1^{n,*}, \cdots, \boldsymbol{\beta}_{s-1}^{n,*})\}$ and because of the way in which the Lagrange multipliers are propagated from a parent node to a child node (see the discussion at the end of Section 2.3.3.1.2.1 and Proposition 3.4.8), we have

$$\begin{aligned}
&\min_{(\mathbf{x}_1,\cdots,\mathbf{x}_s,\mathbf{y},\mathbf{z}_1,\cdots,\mathbf{z}_s)\in\mathcal{F}_{\mathrm{LRP}}^{n+1}} \left[\sum_{h=1}^{s} p_h f_h(\mathbf{x}_h,\mathbf{y},\mathbf{z}_h) + \sum_{h=1}^{s-1}\left(\boldsymbol{\beta}_h^{n+1,*}\right)^{\mathrm{T}}(\mathbf{z}_h-\mathbf{z}_{h+1})\right] \\
\geq &\min_{(\mathbf{x}_1,\cdots,\mathbf{x}_s,\mathbf{y},\mathbf{z}_1,\cdots,\mathbf{z}_s)\in\mathcal{F}_{\mathrm{LRP}}^{n}} \left[\sum_{h=1}^{s} p_h f_h(\mathbf{x}_h,\mathbf{y},\mathbf{z}_h) + \sum_{h=1}^{s-1}\left(\boldsymbol{\beta}_h^{n,*}\right)^{\mathrm{T}}(\mathbf{z}_h-\mathbf{z}_{h+1})\right], \forall n\in\mathbb{N}.
\end{aligned}$$

Furthermore, since Problem (DEP) is assumed to be feasible for $\mathbf{z}$ restricted to $\dot{\mathbf{z}}$, we have by weak duality

$$\begin{aligned}
&\min_{(\mathbf{x}_1,\cdots,\mathbf{x}_s,\mathbf{y},\mathbf{z}_1,\cdots,\mathbf{z}_s)\in\mathcal{F}_{\mathrm{LRP}}^{n}} \left[\sum_{h=1}^{s} p_h f_h(\mathbf{x}_h,\mathbf{y},\mathbf{z}_h) + \sum_{h=1}^{s-1}\left(\boldsymbol{\beta}_h^{n,*}\right)^{\mathrm{T}}(\mathbf{z}_h-\mathbf{z}_{h+1})\right] \\
\leq &\min_{(\mathbf{x}_1,\cdots,\mathbf{x}_s,\mathbf{y})\in(X_1^\infty\times\cdots\times X_s^\infty\times Y^\infty)} \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h,\mathbf{y},\dot{\mathbf{z}}) \\
&\qquad\text{s.t.}\qquad \mathbf{g}_h(\mathbf{x}_h,\mathbf{y},\dot{\mathbf{z}})\leq\mathbf{0},\quad \forall h\in\{1,\cdots,s\}, \\
&\qquad\qquad\qquad \mathbf{r}_{y,z}(\mathbf{y},\dot{\mathbf{z}})\leq\mathbf{0},
\end{aligned}$$

$\forall n\in\mathbb{N}$. Therefore, the limit defined in the left hand side of Equation (L) exists. Choose $\{M^{n_q}\}$ to be a subsequence of $\{M^n\}$ such that $(\dot{\mathbf{x}}_1,\cdots,\dot{\mathbf{x}}_s,\dot{\mathbf{y}},\dot{\mathbf{z}},\cdots,\dot{\mathbf{z}})$, which is a feasible point for Problem (DEP), is the limit of $\{(\mathbf{x}_1^{n_q},\cdots,\mathbf{x}_s^{n_q},\mathbf{y}^{n_q},\mathbf{z}_1^{n_q},\cdots,\mathbf{z}_s^{n_q})\}$. From the continuity of f and the boundedness of $\{(\boldsymbol{\beta}_1^{n,*},\cdots,\boldsymbol{\beta}_{s-1}^{n,*})\}$, we have

$$\lim_{q\to\infty}\left[\sum_{h=1}^{s} p_h f_h(\mathbf{x}_h^{n_q},\mathbf{y}^{n_q},\mathbf{z}_h^{n_q}) + \sum_{h=1}^{s-1}\left(\boldsymbol{\beta}_h^{n_q,*}\right)^{\mathrm{T}}(\mathbf{z}_h^{n_q}-\mathbf{z}_{h+1}^{n_q})\right] = \sum_{h=1}^{s} p_h f_h(\dot{\mathbf{x}}_h,\dot{\mathbf{y}},\dot{\mathbf{z}}).$$

Putting all of the above together, we obtain

$$\begin{aligned}
&\sum_{h=1}^{s} p_h f_h(\mathbf{x}_h^{n_q},\mathbf{y}^{n_q},\mathbf{z}_h^{n_q}) + \sum_{h=1}^{s-1}\left(\boldsymbol{\beta}_h^{n_q,*}\right)^{\mathrm{T}}(\mathbf{z}_h^{n_q}-\mathbf{z}_{h+1}^{n_q}) \\
\leq &\min_{(\mathbf{x}_1,\cdots,\mathbf{x}_s,\mathbf{y})\in(X_1^\infty\times\cdots\times X_s^\infty\times Y^\infty)} \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h,\mathbf{y},\dot{\mathbf{z}}) \\
&\qquad\text{s.t.}\qquad \mathbf{g}_h(\mathbf{x}_h,\mathbf{y},\dot{\mathbf{z}})\leq\mathbf{0},\quad \forall h\in\{1,\cdots,s\}, \\
&\qquad\qquad\qquad \mathbf{r}_{y,z}(\mathbf{y},\dot{\mathbf{z}})\leq\mathbf{0}
\end{aligned}$$

$$\leq \sum_{h=1}^{s} p_h f_h(\dot{\mathbf{x}}_h, \dot{\mathbf{y}}, \dot{\mathbf{z}}).$$

Since the choice of the accumulation point was arbitrary, letting $q \to \infty$ gives us (L).

Therefore, the lower bounding problem, which corresponds to a (partial) solution of Problem (LRP), is strongly consistent. □

Theorem 3.5.5. [Convergence] Let $\varepsilon_l > 0, \varepsilon_u > 0$, and $\varepsilon > 0$ be termination tolerances for the solution of Problem (LRP-inner), the upper bounding problem, and Problem (DEP), respectively, such that $\varepsilon \geq \varepsilon_l + \varepsilon_u$. If Problem (LRP-inner) (respectively, the upper bounding problem) can be solved to ε_l-optimality (respectively, ε_u-optimality) in a finite number of steps, then the algorithm either terminates in a finite number of steps with an ε-optimal solution of Problem (DEP), or an indication that Problem (DEP) is infeasible, or, if a feasible point is not found finitely, any accumulation point of the sequence $\{\mathbf{x}_1^n, \cdots, \mathbf{x}_s^n, \mathbf{y}^n, \mathbf{z}_{\text{avg}}^n\}$, where $\mathbf{z}_{\text{avg}}^n := \frac{1}{s}\sum_{h=1}^{s} \mathbf{z}_h^n$, corresponding to an infinite decreasing sequence of successively refined partition elements of the branch-and-bound tree solves Problem (DEP).

Proof. From Lemmata 3.5.1 and 3.5.4, we know that the subdivision process is exhaustive and the lower bounding operation is strongly consistent. Furthermore, from Lemma 3.5.3, we know that every infinite decreasing sequence $\{M^n\}$ of successively refined partition elements satisfies $M^n \cap \mathcal{F} \neq \emptyset$, $\forall n$ (deletion by infeasibility is certain in the limit). Therefore, from Theorem 2.3.49, we have that the lower bounding technique results in a consistent bounding operation. From Lemma 3.5.2 and Theorem 2.3.50, we have that the B&B procedure is convergent for $\varepsilon = 0$ (assuming that all the subproblems involved can be solved to optimality in finite time (cf. Theorem 2.3.61) and the sequence of upper bounds, determined using a corresponding sequence of feasible points, converges to the optimal objective value).

Consider the case when a feasible point has not been found in finite time by the upper bounding problem(s). Let $\{M^n\}$ be an infinite decreasing sequence of successively refined partition elements generated by the B&B procedure with $\{LBD^n\}$ denoting a corresponding sequence of lower bounds. Since the lower bounding problem is strongly consistent and the subdivision procedure is exhaustive on Z, the sequence $\{LBD^n\}$ (and the overall lower bound) will approach to within $(\varepsilon_l + \bar{\varepsilon})$ tolerance of the optimal objective value finitely, where $\bar{\varepsilon} > 0$ is any positive offset tolerance. Note that the inner minimization of the lower bounding problem, Problem (LRP-inner), can be solved to ε_l-optimality in finite time (see Theorem 2.3.61), and at most a finite number of iterations of an algorithm applied to the dual are carried out. Because the selection procedure is bound improving, any subsequence of the sequence of partition elements $\{M^n\}$ explored by the B&B procedure will contain an infinite number of partition elements which correspond to the lowest lower bound. Consider any accumulation point $(\tilde{\mathbf{x}}_1, \cdots, \tilde{\mathbf{x}}_s, \tilde{\mathbf{y}}, \tilde{\mathbf{z}}, \cdots, \tilde{\mathbf{z}})$ of the sequence $\{(\mathbf{x}_1^n, \cdots, \mathbf{x}_s^n, \mathbf{y}^n, \mathbf{z}_1^n, \cdots, \mathbf{z}_s^n)\}$ of lower bounding solutions over $\{M^n\}$. Lemma 3.5.3 guarantees the feasibility of $(\tilde{\mathbf{x}}_1, \cdots, \tilde{\mathbf{x}}_s, \tilde{\mathbf{y}}, \tilde{\mathbf{z}}, \cdots, \tilde{\mathbf{z}})$. Moreover, the sequence $\{\mathbf{z}_{\text{avg}}^n\}$ of points to which the $\mathbf{z}$ variables in Problem (UBP-NGBD^n) is restricted converges to $\tilde{\mathbf{z}}$. Strong consistency of the lower bounding problem therefore implies that the solution of Problem (UBP-NGBD^n) with $\mathbf{z}$ restricted to $\tilde{\mathbf{z}}$ will yield an optimal solution to Problem (DEP).

The desired result follows for given tolerances $\varepsilon_l, \varepsilon_u, \varepsilon > 0$ such that $\varepsilon \geq \varepsilon_l + \varepsilon_u$. □

The following result shows that under relatively mild additional assumptions, finite convergence of the MLR algorithm (relative to a given tolerance) can be guaranteed. To establish this result, we assume that node n is fathomed only if its lower bound $LBD^n \geq UBD$, where UBD is the overall upper bound of the branch-and-bound tree.

Assumption 3.5.6. There exists an ε_S-optimal ideal Slater point in $X_1 \times \cdots \times X_s \times Y \times Z$ for Problem (DEP), i.e., there exists an ε_S-optimal point $(\mathbf{x}_1^S, \cdots, \mathbf{x}_s^S, \mathbf{y}^S, \mathbf{z}^S) \in X_1 \times \cdots \times X_s \times Y \times Z$ such that $\mathbf{g}_h(\mathbf{x}_h^S, \mathbf{y}^S, \mathbf{z}^S) < \mathbf{0}$, $\forall h \in \{1, \cdots, s\}$, and $\mathbf{r}_{y,z}(\mathbf{y}^S, \mathbf{z}^S) < \mathbf{0}$.

Remark 3.5.7. Assumption 3.5.6 along with the assumptions of compactness of the sets and the continuity of the functions in Problem (DEP) imply, by Weierstrass' theorem, that Problem (DEP) has a finite optimal objective value. The inspiration for the above assumption arose from reading [168]. Note that Assumption 3.5.6 is slightly weaker than [47, Assumption 3], which the authors therein use to prove convergence of the sequence of upper bounds (for a subclass of Problem (DEP) that does not contain any discrete variables).

Theorem 3.5.8. [Finite-ε convergence] Let $\varepsilon_l > 0, \varepsilon_u > 0$, and $\varepsilon > 0$ be termination tolerances for the solution of Problem (LRP-inner), the upper bounding problem (Problem (UBP-NGBD^n)), and Problem (DEP), respectively. If Problem (LRP-inner) (respectively, the upper bounding problem) used by the modified Lagrangian relaxation algorithm can be solved to ε_l-optimality (respectively, ε_u-optimality) in a finite number of steps, Assumption 3.5.6 holds with $\varepsilon_S \geq 0$, and the termination tolerance $\varepsilon > \varepsilon_l + \varepsilon_u + \varepsilon_S$, the modified Lagrangian relaxation algorithm terminates in a finite number of steps with an ε-optimal solution of Problem (DEP).

Proof. We first provide the intuition behind the proof before formally stating it.

We know from Theorem 3.5.5 that the overall lower bound of the branch-and-bound algorithm is guaranteed to approach to within ε_l of the optimal objective value (from below) in the limit. Because we assume the existence of an ε_S-optimal Slater point, say $(\mathbf{x}_1^S, \cdots, \mathbf{x}_s^S, \mathbf{y}^S, \mathbf{z}^S)$, for Problem (DEP), there exists a neighborhood of $\mathbf{z}^S$ (relative to Z) such that restricting $\mathbf{z}$ to any point in this neighborhood in Problem (UBP-NGBD^n) is guaranteed to result in the generation of a near-optimal feasible point and a corresponding upper bound. This leads to two possibilities: either this neighborhood of $\mathbf{z}^S$ was part of a node that had been fathomed early on by value dominance, or, because the partitioning of Z is exhaustive and the node selection procedure is bound improving, the $\mathbf{z}$-component of the domain of some branch-and-bound node that is visited finitely is guaranteed to lie within this neighborhood. The former case can only happen if the branch-and-bound algorithm has already found a good enough upper bound, in which case the algorithm terminates finitely anyway. If the latter scenario occurs, then solving Problem (UBP-NGBD^n) with $\mathbf{z}$ restricted to any point in this neighborhood (for instance, $\mathbf{z}^n_{\text{avg}}$) will yield a good enough upper bound that will eventually guarantee finite convergence. In what follows, we state the above arguments formally.

From Assumption 3.5.6, there exists $(\mathbf{x}_1^S, \cdots, \mathbf{x}_s^S, \mathbf{y}^S, \mathbf{z}^S) \in X_1 \times \cdots \times X_s \times Y \times Z$ such that

$$\mathbf{g}_h(\mathbf{x}_h^S, \mathbf{y}^S, \mathbf{z}^S) < \mathbf{0}, \ \ \forall h \in \{1, \cdots, s\}, \quad \mathbf{r}_{y,z}(\mathbf{y}^S, \mathbf{z}^S) < \mathbf{0},$$

and

$$\sum_{h=1}^{s} p_h f_h(\mathbf{x}_h^S, \mathbf{y}^S, \mathbf{z}^S) \leq \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h^*, \mathbf{y}^*, \mathbf{z}^*) + \varepsilon_S,$$

where $(\mathbf{x}_1^*, \cdots, \mathbf{x}_s^*, \mathbf{y}^*, \mathbf{z}^*)$ is an optimal solution to Problem (DEP).

From the continuity of the associated functions, there exists $\delta > 0$ such that for all points $(\mathbf{x}_1^B, \cdots, \mathbf{x}_s^B, \mathbf{y}^B, \mathbf{z}^B) \in \mathcal{N}_\delta^\infty((\mathbf{x}_1^S, \cdots, \mathbf{x}_s^S, \mathbf{y}^S, \mathbf{z}^S))$, where $\mathcal{N}_\delta^\infty((\mathbf{x}_1, \cdots, \mathbf{x}_s, \mathbf{y}, \mathbf{z}))$ is the δ-neighborhood of $(\mathbf{x}_1, \cdots, \mathbf{x}_s, \mathbf{y}, \mathbf{z})$ relative to $X_1 \times \cdots \times X_s \times Y \times Z$ with respect to the ∞-norm,

$$\left| \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h^B, \mathbf{y}^B, \mathbf{z}^B) - \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h^S, \mathbf{y}^S, \mathbf{z}^S) \right| < (\varepsilon - \varepsilon_l - \varepsilon_u - \varepsilon_S)$$

and

$$\mathbf{g}_h(\mathbf{x}_h^B, \mathbf{y}^B, \mathbf{z}^B) < \mathbf{0}, \ \forall h \in \{1, \cdots, s\}, \quad \mathbf{r}_{y,z}(\mathbf{y}^B, \mathbf{z}^B) < \mathbf{0}.$$

For any such $\mathbf{z}^B$, we have that Problem (DEP) is feasible when $\mathbf{z}$ is restricted to $\mathbf{z}^B$. Moreover, it can be easily seen that solving the corresponding Problem (UBP-NGBD^n) to a termination tolerance of ε_u provides an $(\varepsilon - \varepsilon_l)$-optimal solution to Problem (DEP).

Thus, if there exists a node M^{n_δ} in the B&B tree such that $\forall \mathbf{z} \in Z^{n_\delta}$, $\exists(\mathbf{x}_1, \cdots, \mathbf{x}_s, \mathbf{y}) \in X_1^{n_\delta} \times \cdots \times X_s^{n_\delta} \times Y^{n_\delta}$ such that $(\mathbf{x}_1, \cdots, \mathbf{x}_s, \mathbf{y}, \mathbf{z}) \in \mathcal{N}_\delta^\infty((\mathbf{x}_1^S, \cdots, \mathbf{x}_s^S, \mathbf{y}^S, \mathbf{z}^S))$, we have that $\mathbf{z}$ restricted to any point in Z^{n_δ} in Problem (UBP-NGBD^n) generates an $(\varepsilon - \varepsilon_l)$-optimal solution to Problem (DEP).

Suppose, by way of contradiction, the B&B procedure does not converge finitely. Let $\{M^n\}$ be the sequence of successively refined partitioned elements generated by the B&B subdivision process that is exhaustive on Z. From Lemmata 2.3.43 and 3.5.1, $\exists N_\delta \in \mathbb{N}$ such that for all partition elements M^n with $n \geq N_\delta$, $w(Z^n) < \frac{\delta}{2}$. One such partition, say M^{n_δ}, has to contain $(\mathbf{x}_1^S, \cdots, \mathbf{x}_s^S, \mathbf{y}^S, \mathbf{z}^S)$ (otherwise, a feasible point that is ε_S-optimal has already been found by the B&B algorithm). By the definition of δ, we have $\forall \mathbf{z} \in Z^{n_\delta}, \exists(\mathbf{x}_1, \cdots, \mathbf{x}_s, \mathbf{y}) \in X_1^{n_\delta} \times \cdots \times X_s^{n_\delta} \times Y^{n_\delta}$ such that $(\mathbf{x}_1, \cdots, \mathbf{x}_s, \mathbf{y}, \mathbf{z}) \in \mathcal{N}_{\frac{\delta}{2}}^\infty((\mathbf{x}_1^S, \cdots, \mathbf{x}_s^S, \mathbf{y}^S, \mathbf{z}^S))$. Lemma 2.3.43, Theorems 2.3.61 and 3.5.5, and the B&B selection procedure guarantee that the node M^{n_δ} is visited finitely, resulting in a finite generation of an $(\varepsilon - \varepsilon_l)$-optimal point for Problem (DEP).

Since, from Theorem 3.5.5, the branch-and-bound algorithm is convergent with the sequence of lower bounds $\{LBD^n\}$ converging to within an ε_l tolerance of the optimal objective value and, by the above argument, an $(\varepsilon - \varepsilon_l)$-optimal point has been generated finitely, we have that the branch-and-bound procedure is finitely convergent for the given tolerance ε, a contradiction. $\square$

The above result, as presented, does not apply to problems with equality constraints that are reformulated to the form of Problem (DEP) using pairs of inequality constraints (since the proof roughly assumes that the feasible set has a nonempty interior). We note that Theorem 3.5.8 can be easily extended to the case when Problem (DEP) contains affine equality constraints if the upper bounding problem, Problem (UBP-NGBD^n), is solved at each node of the branch-and-bound tree by fixing the continuous complicating variables $\mathbf{z}$ to the scenario-averaged lower bounding continuous complicating variable solution $\mathbf{z}^n_{\text{avg}}$. Furthermore, Theorem 3.5.8 can also be extended to the case when (some) affine inequality constraints are active at the ε_S-optimal Slater point.

3.6 Computational studies

In this section, we compare the performance of the modified Lagrangian relaxation (MLR) algorithm with the performance of the conventional Lagrangian relaxation (LR) algorithm and four general-purpose state-of-the-art global optimization solvers ANTIGONE 1.1 [162], BARON 16.3.4 [225], Couenne 0.5 [19], and SCIP 3.2 [233], which are accessed via GAMS 24.7.1 [83], on two case studies from the literature (also see Section 4.4 of Chapter 4). The LR and MLR algorithms used for comparison are implemented as part of our software GOSSIP (see Chapter 4), which we plan on making available to the academic community soon on our lab's website `http://yoric.mit.edu/software`. Our case studies include: a tank sizing and scheduling model for a multi-product plant with uncertainties in the product demands developed by Rebennack, Kallrath, and Pardalos [186], and a model for integrated crude selection and refinery operation with uncertainties in the crude qualities and yields, developed by Yang and Barton [241], that is modified to include continuous first-stage variables. Multiple instances of each of the above models, with a different number of scenarios in each instance, were solved using the six different algorithms/solvers. Sections 3.6.1 to 3.6.2 briefly present details of these models and instances. We redirect the interested reader to Sections 4.2.4 and 4.4 of Chapter 4 for the implementation details for these computational studies.

3.6.1 Tank sizing and scheduling for a multi-product plant

This case study considers a chemical plant with a single reactor that is used to produce three products, which are then stored in three distinct tanks. The objective of this model is to: i. determine the optimal sizes of the tanks for each product, ii. determine an optimal production schedule that satisfies the uncertain demands for each of the products, and iii.

determine the optimal campaign lengths and production sizes for the three products. This model includes bilinear and univariate signomial terms and binary recourse variables that make it challenging to solve. We consider one uncertain product demand with instances ranging from one to twenty five scenarios as part of our case study in Section 3.6.3. The details of the deterministic and two-stage stochastic programming models can be found in [186]. The authors of [186] observe that general-purpose state-of-the-art global optimization solvers are unable to prove optimality of their best found solutions even for the stochastic programming models with a relatively small number of scenarios, which is in line with our observations in Section 3.6.3

3.6.2 Integrated crude selection and refinery operation

This case study integrates a simplified refinery model with a pooling model while considering uncertainties in the crude oil qualities and yields. The purpose of this model is to determine the optimal crude purchase while taking the above uncertainties into account so that the expected revenue of the refinery can be maximized while satisfying key refinery product quality constraints. This model includes a choice of ten crudes, a fixed-yield crude distillation unit model, mass balances, market demands, quality constraints, and capacity and supply restrictions. Similar to previous work [241], we assume that the vacuum residue yields and the sulfur fractions of gas oil from the CDU for each crude are uncertain parameters; however, in a departure from previous work [241], which assumes that crudes can be purchased in certain predefined discrete amounts, we assume that the crude purchase quantities are 'semi-continuous' variables, i.e., the crude purchase quantity can either be zero if the decision maker decides not to purchase that particular crude, or it must lie between prespecified positive lower and upper bounds. This update to the model implies that the NGBD algorithm employed by Yang and Barton can no longer be applied to solve the resulting two-stage stochastic program to guaranteed global optimality, which necessitates the use of alternative decomposition techniques such as LR and MLR. We generate case studies with one, five, ten, twenty, forty, and one hundred and twenty scenarios of the uncertain parameters as part of our case study in Section 3.6.3. The essential details of the two-stage stochastic programming model can be found in [241].

3.6.3 Results and discussion

This section presents computational results for the tank sizing and integrated crude selection problems under uncertainty. For each of these case studies, we list the solution time in seconds (rounded to the nearest 0.1 second) and percentage relative termination gap separated by '/', with the gap defined as

$$\text{gap} = \min\left\{100, 100 \times \frac{\texttt{upperbound} - \texttt{lowerbound}}{\max\left\{|\texttt{lowerbound}|, |\texttt{upperbound}|\right\} + \delta}\right\} [\%],$$

where $\delta \in (0, 1)$ such that $\delta \ll 1$, for each tested solver for varying numbers of scenarios. We note that the reported percentage termination gap is rounded to the nearest 0.1 percent if the instance was not solved within the time limit, and set to be equal to 0.1 percent (which is the desired optimality level) otherwise. A blank entry ('-') for the solution time indicates that the solution reached the time limit of 10,000 seconds. The entry '`t`' for the solution time indicates that the solver terminated prematurely due to failure (likely due to insufficient memory to continue). The entry '`i`' for the solution time indicated that the solver wrongly concluded that the model is infeasible. A blank entry ('-') for the termination gap either indicates that a feasible point wasn't found within the time limit, or that the termination gap is not relevant for this entry. Empty entries for both the solution time and termination gap indicates that the computational experiment was not carried out.

The results for the tank sizing and scheduling problem are presented in Tables 3.1 and 3.2, and Figure 3-2 compares the performance of the different solvers for instances with varying numbers of scenarios for this case study. Couenne wrongly concludes that the single scenario problem is infeasible, and is unable to solve multiple scenario instances of this problem with 10,000 seconds. ANTIGONE and BARON fail to solve problems with more than one scenario within the time limit of 10,000 seconds, whereas SCIP can only solve the one and two-scenario instances within the time limit. The solution times of both LR and MLR appear to scale affinely with the number of scenarios as expected; however, MLR can be seen to be more effective for solving large-scenario instances and can solve instances with up to 21 scenarios within the prescribed time limit. Since this case study does not include any binary complicating variables, the lower bounding problem of the MLR algorithm effectively reduces to the lower bounding problem of the conventional Lagrangian relaxation algorithm. Consequently, the significant computational advantage

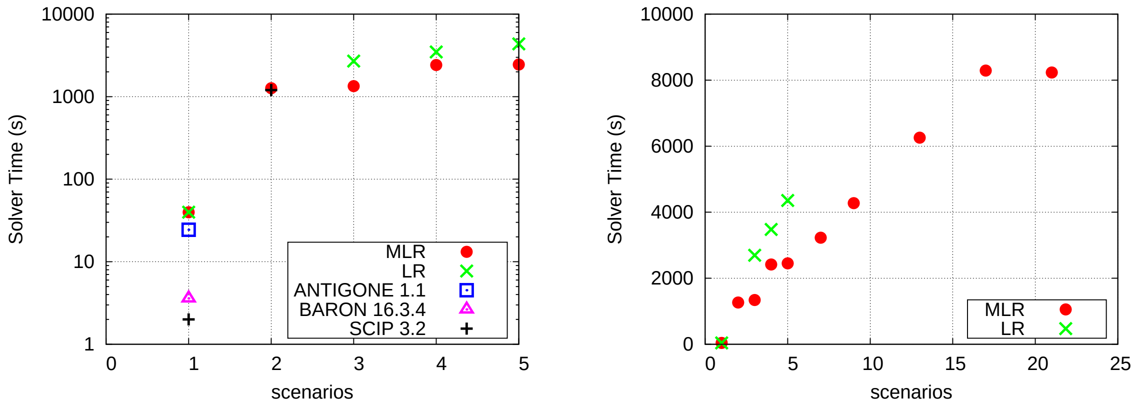


Figure 3-2: Comparison of the different solvers on the tank sizing and scheduling problem [186] (see Tables 3.1 and 3.2). The general-purpose global optimization solvers are generally ineffective in solving multiple scenario instances of this problem. The LR and MLR algorithms, on the other hand, perform favorably compared to the general-purpose solvers. The implementation of the aggressive bounds tightening technique within the MLR algorithm improves its performance significantly compared to the conventional LR algorithm.

Table 3.1: Comparison of the different solvers on the tank sizing and scheduling problem. This case study includes 3 continuous complicating variables, 9 binary recourse variables per scenario, 44 continuous recourse variables per scenario, 76 second-stage constraints per scenario, 32 bilinear terms per scenario, and 3 univariate signomial terms per scenario.

# Scenarios	1	2	3	4	5
ANTIGONE 1.1	24.4/0.1	-/13.9	-/7.4	-/14.3	-/14.3
BARON 16.3.4	3.6/0.1	t/23.1	-/10.6	-/59.8	-/49.8
COUENNE 0.5	i/-	-/12.8	-/25.6	-/35.4	-/39.9
SCIP 3.2	2.0/0.1	1201.7/0.1	-/0.2	t/16.6	t/10.5
MLR	39.7/0.1	1265.2/0.1	1340.7/0.1	2416.3/0.1	2451.2/0.1
LR	39.9/0.1	-/0.14	2696.6/0.1	3477.1/0.1	4354.8/0.1

of MLR towards solving larger scenario instances can be attributed to the effectiveness of the aggressive bounds tightening (ABT) technique (detailed in Section 3.3.2) in eliminating suboptimal regions of the search space early on in the B&B tree.

Table 3.3 provides detailed computational results for the integrated crude selection and refinery operation case study, and Figure 3-3 compares the performance of the different solvers on various instances of this problem. This case study provides the opportunity to test all of the modifications of the LR algorithm that are incorporated as part of the MLR algorithm since it includes binary and continuous first-stage decisions. Both the LR and MLR algorithms, however, face difficulties in converging to within 0.1% tolerance for instances with more than one scenario, whereas the solvers ANTIGONE and BARON can

Table 3.2: Extended results for the decomposition methods for continuous tank sizing problem.

# Scenarios	7	9	13	17	21	25
MLR	3228.9/0.1	4274.4/0.1	6256.9/0.1	8289.2/0.1	8230.6/0.1	-/0.12
LR	-/0.11	-/0.12	-/0.12	-/0.14	-/0.14	-/0.14

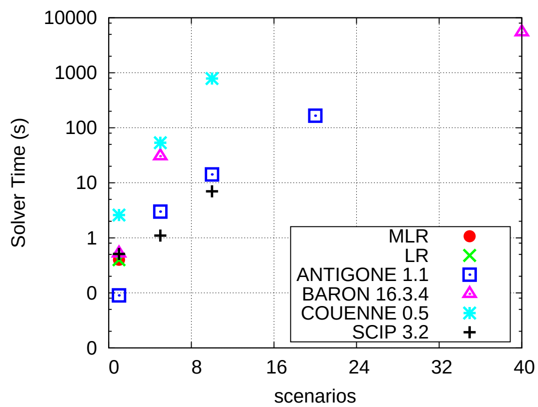


Figure 3-3: Comparison of the different solvers on the integrated crude selection and refinery operation model (see Table 3.3). Both LR and MLR are unable to solve instances with more than one scenario, whereas the commercial solvers ANTIGONE and BARON can solve instances with more than ten scenarios within 10,000 seconds.

solve instances with more than ten scenarios for this case study within the time limit (these solvers, however, do not scale favorably with the number of scenarios). In our experience, a major contributing factor towards the lackluster performances of both the decomposition algorithms is the ineffectiveness of nonsmooth optimization techniques in solving the respective dual problems. As a validation of the theory, we note that the MLR algorithm can indeed provide tighter bounds than the conventional LR algorithm. This is evidenced, for example, by the fact that the initial root node bound of the MLR algorithm corresponds to a relative gap of 14.7% for the instance with ten scenarios, whereas the corresponding root node relative gap of the conventional LR algorithm is 16.2%. We also note that the upper bounding techniques detailed in Section 3.3.1 are able to determine a global optimal solution for instances with up to forty scenarios at the root node of the B&B tree for the LR and MLR implementations.

Table 3.3: Comparison of the different solvers on the integrated crude selection and refinery operation model. This case study includes 10 binary complicating variables, 10 continuous complicating variables, 122 continuous recourse variables per scenario, 21 first-stage constraints, 111 second-stage constraints per scenario, and 26 bilinear terms per scenario.

# Scenarios	1	5	10	20	40	120
ANTIGONE 1.1	0.1/0.1	3.0/0.1	14.2/0.1	166.3/0.1	-/0.2	-/0.7
BARON 16.3.4	0.5/0.1	30.5/0.1	-/0.4	-/0.4	5448.2/0.1	-/0.7
COUENNE 0.5	2.6/0.1	53.4/0.1	785.3/0.1	-/0.2	-/0.7	-/-
SCIP 3.2	0.5/0.1	1.1/0.1	7.0/0.1	-/0.2	-/0.2	-/36.0
MLR	0.4/0.1	-/9.3	-/9.5	-/10.4	-/11.2	-/-
LR	0.4/0.1	-/9.6	-/9.6	-/11.1	-/12.0	-/-

3.7 Conclusion

This chapter presented a modified Lagrangian relaxation algorithm for solving two-stage stochastic MINLPs with mixed-integer variables in both stages by combining NGBD, Lagrangian relaxation, and scalable bounds tightening techniques. To the best of our knowledge, this chapter details the first fully decomposable algorithm for solving this class of problems that provably converges to an ε-optimal solution in finite time. The theoretical basis of the algorithm was established by building on the framework of the NGBD and Lagrangian relaxation algorithms and a general convergence theory of B&B algorithms, and the performance of the algorithm was tested on two case studies from the literature. While the proposed algorithm is able to mitigate some of the issues faced by the conventional Lagrangian relaxation algorithm via the use of decomposable bounds tightening techniques, it appears to still suffer from the following major (numerical) limitation of the conventional LR algorithm: although the (outer) Lagrangian dual problem is convex, the Lagrangian dual function is generally nonsmooth, and there seems to be a divide between theory and practice in solving nonsmooth (convex) optimization problems. Future work on Lagrangian relaxation-type algorithms must therefore seek to mitigate this limitation if one wishes to solve practical applications using such techniques in reasonable times.

Chapter 4

GOSSIP: decomposition software for the Global Optimization of nonconvex two-Stage Stochastic mixed-Integer nonlinear Programs

Stochastic programming provides a natural way of incorporating uncertainty in model parameters, and has been receiving increasing attention in the process systems engineering literature. Despite rapid advances in decomposition techniques for solving nonconvex two-stage stochastic mixed-integer nonlinear programs (MINLPs), there is no publicly available software framework which implements these techniques. Motivated by the above, this chapter presents `GOSSIP`, a decomposition framework for the global optimization of two-stage stochastic MINLPs. `GOSSIP` includes implementations of nonconvex generalized Benders decomposition, Lagrangian relaxation, and a modified Lagrangian relaxation algorithm for solving a broad class of two-stage stochastic MINLPs.

4.1 Introduction

This chapter introduces `GOSSIP`, decomposition software for the global solution of the following class of scenario-based two-stage stochastic MINLPs (see Problem (SP) in Chapter 1):

$$\begin{aligned} \min_{\mathbf{x}_1,\cdots,\mathbf{x}_s,\mathbf{y},\mathbf{z}} \quad & \sum_{h=1}^{s} p_h f_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}) \\ \text{s.t.} \quad & \mathbf{g}_h(\mathbf{x}_h, \mathbf{y}, \mathbf{z}) \leq \mathbf{0}, \quad \forall h \in \{1, \cdots, s\}, \\ & \mathbf{r}_{y,z}(\mathbf{y}, \mathbf{z}) \leq \mathbf{0}, \\ & \mathbf{x}_h \in X_h, \quad \forall h \in \{1, \cdots, s\}, \\ & \mathbf{y} \in Y, \;\; \mathbf{z} \in Z, \end{aligned} \tag{SP}$$

where $X_h = \{0,1\}^{n_{x_b}} \times \Pi_{x,h}$ with $\Pi_{x,h} \in \mathbb{R}^{n_{x_c}}$, $\forall h \in \{1, \cdots, s\}$, $Y = \{0,1\}^{n_y}$, $Z \in \mathbb{R}^{n_z}$, and functions $f_h : [0,1]^{n_{x_b}} \times \Pi_{x,h} \times [0,1]^{n_y} \times Z \to \mathbb{R}$, $\mathbf{g}_h : [0,1]^{n_{x_b}} \times \Pi_{x,h} \times [0,1]^{n_y} \times Z \to \mathbb{R}^m$, $\forall h \in \{1, \cdots, s\}$, and $\mathbf{r}_{y,z} : [0,1]^{n_y} \times Z \to \mathbb{R}^{m_{y,z}}$ are assumed to be continuous. The variables $\mathbf{y}$ and $\mathbf{z}$ in Problem (SP) denote the discrete and continuous first-stage/complicating decisions, respectively, that are made before the realization of the uncertainties, while, for each $h \in \{1, \cdots, s\}$, the mixed-integer variables $\mathbf{x}_h$ denote the second-stage/recourse decisions made after the uncertain model parameters realize their 'scenario h' values. The quantity $p_h > 0$ represents the probability of occurrence of scenario h with $\sum_{h=1}^{s} p_h = 1$. Equality constraints in the formulation are assumed to be modeled using pairs of inequalities purely for ease of exposition. Additionally, bounded general integer variables are assumed to be equivalently reformulated using binary variables in Problem (SP) mainly for ease of exposition (`GOSSIP` automatically reformulates only the first-stage bounded integer variables using binary variables and, if necessary, auxiliary constraints).

We assume that all of the functions in Problem (SP) are factorable, i.e., they can be expressed as finite compositions of binary and unary operations, from predefined libraries of binary (for example, $+$, $-$, $\times$, $/$, and $\hat{}$) and unary functions (for example, univariate powers, $|\cdot|$, exp, and log), respectively, applied to the variables in Problem (SP). Additionally, we will implicitly assume that the functional forms of the subproblems we formulate can be handled by (one of) the solvers linked to by `GOSSIP`. Further details on the functional forms supported by `GOSSIP` will be provided in the ensuing sections. The solution techniques in GOSSIP also roughly assume the existence of global optimization software that can solve, in reasonable times to desired termination tolerances, the single-scenario versions of Problem (SP) corresponding to each possible realization of the uncertain parameters; if this assumption does not hold, we (reasonably) do not expect GOSSIP to be able to solve multiple scenarios instances (possibly even a single scenario instance) of such problems in practical solution times (see Section 4.4).

Decomposition algorithms for Problem (SP) try to exploit its near-decomposable structure so that their solution times scale reasonably with an increase in the number of scenarios. General-purpose deterministic global optimization methods [224, 225] for Problem (SP), on the other hand, face a worst-case exponential increase in solution times with the number of scenarios since they do not exploit its structure, which typically makes the solution of large-scenario instances of Problem (SP) using such techniques impractical for applications of interest. We refer the reader to Section 2.3.3.1 of Chapter 2 for an overview of techniques for solving various subclasses of Problem (SP).

While there are a few commercial and open-source software packages for solving two-stage stochastic MILPs (see [125] for details), there are hardly any software implementations for the scalable solution of nonlinear and mixed-integer nonlinear stochastic programming problems. When Problem (SP) only contains continuous variables, Schur-IPOPT [247] can exploit its block-angular structure to obtain a 'local optimal solution' using structured linear algebra techniques (also see [107, 131]). PIPS-NLP [56] is another C-based library that also implements a structured linear algebra-based filter line-search interior point method that can solve such problems locally. The progessive hedging algorithm of Rockafellar and Wets [190] has been implemented within PySP [236], a Python-based software package for stochastic programming, and can potentially be used to solve Problem (SP), albeit without any convergence guarantees. Finally, SNGO [47] provides a Julia-based implementation of a Lagrangian relaxation-type algorithm for the global solution of Problem (SP), along with bounds tightening and tailored branching & upper bounding techniques, when it only contains continuous variables. Among all of the software listed in this section, SNGO is the most relevant to our work since the other implementations usually cannot guarantee finding a global solution to Problem (SP) when it involves nonconvex functions in its formulation.

GOSSIP includes efficient implementations of NGBD (see Section 2.3.3.1.1 of Chapter 2), Lagrangian relaxation (see Section 2.3.3.1.2 of Chapter 2), and a modified Lagrangian relaxation algorithm (see Chapter 3) for solving Problem (SP). GOSSIP includes several advanced techniques for reformulating and preprocessing user input, automatically constructs and coordinates the solution of the subproblems used by the decomposition algorithms, and integrates state-of-the-art decomposition methods with scalable bounds tightening techniques. To the best of our knowledge, GOSSIP will be the first publicly available decomposition software for the global solution of Problem (SP) (we plan to make GOSSIP available for download to academics without charge soon on our lab website http://yoric.mit.edu/software). This chapter demonstrates the capabilities of GOSSIP on a diverse set of test cases from the literature.

This chapter is organized as follows. Section 4.2 discusses the implementation details pertaining to some of GOSSIP's main features, including reformulating user input, detecting special structures, relaxation techniques, and bounds tightening techniques. Section 4.3 briefly outlines, to the best of our knowledge, the first test library for two-stage stochastic MINLPs[1]. Section 4.4 presents the results of our computational experiments with GOSSIP and demonstrates the advantage of our software implementation in comparison to state-of-the-art deterministic global optimization software. Finally, Section 4.5 concludes the chapter, and Section 4.6 includes detailed results for the computational experiments in Section 4.4.

4.2 Implementing GOSSIP

This section outlines the key components of the GOSSIP codebase for solving Problem (SP), and closes with a wish list for future implementation work. GOSSIP consists of more than a hundred thousand lines of source code[2], written primarily in C++ (with a few links to C and FORTRAN-based libraries), and includes subroutines for: reformulating user input and detecting special structures, automatic construction of the subproblems required by the decomposition techniques (which includes construction of convex relaxations, see Section 2.3.3.1.1), domain reduction, linking to state-of-the-art MILP, NLP, and

[1] We do, however, note that http://minlp.org/ includes some instances of stochastic programs.

[2] Dr. Achim Wechsung, a former member of our group at MIT, contributed about 5000 lines of C++ code to GOSSIP from his previous work [237, 239], including subroutines for constructing (and parsing) computational graphs based on user-defined models, an interval-based bounds tightening method that leverages the computational graph representation, and a framework for implementing B&B algorithms (part of Dr. Wechsung's work that provides a flexible computational graph framework is available to academics for download at no charge via our group's website: http://yoric.mit.edu/software/compgraph). All of the remaining code in GOSSIP was written by the author.

MINLP solvers, and a generic B&B framework. The following sections describe the major implementation-related aspects of `GOSSIP`.

4.2.1 Reformulating user input and detecting special structures

The reformulation strategies in `GOSSIP` aim to eliminate redundant variables, terms, and constraints in the user-defined model, reformulate it into a form that can be handled by the implemented methods, and elucidate any structure within the resulting (simplified) model in order to construct tight relaxations for use within the subproblems of the implemented decomposition techniques. Many of the reformulation strategies in `GOSSIP` are based on similar implementations in the commercial software ANTIGONE and BARON that are described in the articles [11, 161–163].

In the first step, the user defines their two-stage stochastic programming model using `GOSSIP`'s modeling language in C++ and sets the relevant solution options, including the solution method and termination tolerances, to be used by `GOSSIP`. Once the user-defined model is converted to a computational graph format [239], `GOSSIP` stores the problem variables and constraints within an `InputProblem` object using variable and constraint classes, respectively. The model stored in the `InputProblem` object is then subject to various preprocessing techniques, including the 'elimination' of redundant variables and constraints, reformulating bounded integer complicating variables using binary complicating variables (and adding auxiliary constraints, if necessary), distributing/disaggregating products of linear and nonlinear terms (see [223, Chapter 3]), simplifying/flattening the computational graph (see [163, Section 3.1.1]), and extracting (and storing) the coefficients of linear constraints in the formulation. Next, the deterministic equivalent/extensive form of the user-defined two-stage stochastic program is constructed either if the user wishes to solve the extensive form directly using a supported solver, or if the chosen solution approach requires the solution of the (restricted) deterministic equivalent form using local optimization techniques for generating feasible points.

Once the initial reformulations on the user-defined model are carried out, the above `InputProblem` object is then decomposed into individual scenario `PrimalProblem` objects, which are used either to generate feasible points for the NGBD algorithm, or to generate lower bounds for the LR and MLR algorithms (note that each `PrimalProblem` object model, by way of its definition, is usually solved using a global solver irrespective of whether the solution method is NGBD, LR, or MLR). Each of the scenario models stored in the `PrimalProblem` objects are then temporarily reformulated using the auxiliary variable technique [213] in order to detect simple common nonlinear subexpressions in those models (we do not permanently reformulate the `PrimalProblem` object models because we wish to leave the choice of reformulation steps to the global solvers that are used to solve these models). `GOSSIP` tracks a few well-studied nonlinear functional forms of (auxiliary) variables, including: products, quotients, univariate and multivariate signomial terms, trilinear and quadrilinear terms, exponentials, logarithms, composite exponential and logarithmic terms, and absolute value terms. The above temporarily-reformulated `PrimalProblem` object models are relaxed using the techniques listed in Section 4.2.2 to construct the corresponding `PrimalBoundingProblem` object models, which are either used within the GBD loop of the NGBD algorithm, or are used to construct the subproblems employed by the bounds tightening techniques listed in Section 4.2.3, or both. Next, depending on the choice of the solution method, `FeasibilityProblem` and `RelaxedMasterProblem` objects are constructed based on the definitions of Problems (NGBD-FP) and (NGBD-RMP), respectively. Finally, we construct feasibility-based bounds tightening and optimality-based bounds tightening model objects based on the subproblems listed in Section 4.2.3.

4.2.2 Relaxation strategies

Table 4.1 lists the various term-specific relaxation strategies employed by `GOSSIP` for the construction of its subproblems. We note that the convexity/concavity of univariate and multivariate signomial terms (see [163, Appendix B]), and logarithmic, exponential, and composite logarithmic and exponential terms are (potentially) detected by `GOSSIP` and utilized during the construction of relaxations. The dependence of the relaxations on the (relevant) variable bounds are stored so that the relaxations can be efficiently updated if and when the domains of the variables are reduced using bounds tightening techniques. Unlike the outer-linearization approaches used by state-of-the-art global optimization software [224, 225], `GOSSIP` currently employs nonlinear convex relaxations wherever relevant, since generating tighter relaxations could dramatically improve the performance of the implemented algorithms. Once additional advanced relaxation strategies, similar to those used by state-of-the-art global optimization software [162, 225], are implemented within `GOSSIP`, we plan on switching to a polyhedral relaxation strategy to leverage the robustness of com-

Table 4.1: Summary of the relaxation strategies for the different 'simple terms' detected by GOSSIP.

Term	Relaxation strategies
xy	McCormick envelope [7, 154]
$\frac{x}{y}$	Bilinear reformulation, Zamora & Grossmann [246] envelope
x^c	Secant, Liberti & Pantelides [142] linearization
$\log(x)$	Secant
$\exp(x)$	Secant
x^y	Reformulate as $\exp(y\log(x))$
$\lvert x\rvert$	MILP reformulation (see [163, Section 3.2.3])
$\min(x,y)$	Reformulate as $\frac{1}{2}(x+y-\lvert x-y\rvert)$
$\max(x,y)$	Reformulate as $\frac{1}{2}(x+y+\lvert x-y\rvert)$
$x\log(x)$	Secant
$x\exp(x)$	Bilinear reformulation, Secant
xyz	Meyer & Floudas [155, 156] envelope
$xyzw$	Facets of the convex hull (see [163, Section 3.3.1])
${x_1}^{c_1}\cdot{x_2}^{c_2}\cdots{x_n}^{c_n}$	Bilinear reformulation, Secant, Transformation-based [149, 152]

mercial LP solvers and the warm-start capabilities of such a framework. Future work with regards to relaxation strategies also involves the incorporation of advanced reformulation-linearization technique (RLT) cuts [143, 161, 163, 164, 209, 210] and a piecewise convex relaxation framework [27, 52, 137, 158] within GOSSIP to improve the strength of constructed relaxations (we currently manually incorporate some relevant RLT cuts as part of our models for the computational studies in Section 4.4).

4.2.3 Bounds tightening techniques

While several bounds tightening techniques have been proposed in the literature [19, 182, 224], based on feasibility [17, 94] and optimality [194, 246] arguments, that are directly applicable to nonconvex MINLPs in the form of Problem (SP), we chose to implement tailored bounds tightening techniques within GOSSIP so that the computational effort expended in the bounds tightening steps scales linearly with the number of scenarios while still providing tight bounds. The domain reduction techniques that are part of GOSSIP include: forward-backward interval propagation (see Section 2.3.2.2 of Chapter 2), optimization-based FBBT techniques (see Section 2.3.3.1.1.6 of Chapter 2 and Section 3.3.2 of Chapter 3), and OBBT techniques (see Problem (OBBT_x) in Section 2.3.3.1.1.6 and Problem (ABT_z^n) in Section 3.3.2).

4.2.4 Implementing the decomposition techniques within `GOSSIP`

This section briefly summarizes the implementation aspects of the three decomposition algorithms (NGBD, LR, and MLR; see Chapters 2 and 3) implemented within `GOSSIP` and lists the default settings for key solver options for those algorithms.

All of the implemented decomposition techniques use the forward-backward interval propagation technique at various stages in their solution process with the termination criterion requiring that the improvement in each of the variables' bounds fall below 0.1% (with respect to their bounds during the previous iteration of the technique). `GOSSIP` interfaces with a few state-of-the-art software for the solution of subproblems, including a C++ interface to ANTIGONE 1.1 [162], a C++ interface to IPOPT 3.12.8 [235], a C++ interface to SNOPT 7.2-4 [86], a C interface to CPLEX 12.6.1 [103], and FORTRAN interfaces to MPBNGC 2.0 [151] and Solvopt 1.1 [130]. The default solver choices are: ANTIGONE 1.1 for the global optimization of convex MINLPs, nonconvex NLPs and nonconvex MINLPs, IPOPT 3.12.8 for solving convex NLPs and for generating feasible points for LR and MLR (for the case of nonconvex NLPs and MINLPs), CPLEX 12.6.1 for solving LPs and MILPs, and MPBNGC 2.0 for solving the dual problems. The subproblems used by the decomposition techniques are solved using interfaces to the above solvers via a `Solver` class. `GOSSIP` also interfaces with the BOOST C++ interval arithmetic library [62] to calculate interval bounds on factorable expressions for bounds tightening techniques and during the construction of relaxations, and interfaces with the FADBAD++ C++ library [26] for computing first- and second-derivative information.

The relative termination tolerances ε_h, for each $h \in \{1, \cdots, s\}$, for the NGBD primal problems (see Algorithm 2.3 in Chapter 2) are all set to be equal to $\frac{\varepsilon}{s}$. The default choice for the initial binary complicating variable realization $\mathbf{y}^1$ in Algorithm 2.3 is an optimal solution to Problem (NGBD-FRMP) with the initialization $S = \emptyset$; in practice, $\mathbf{y}^1$ is set to the best found feasible solution to Problem (NGBD-FRMP) if the time limit of five seconds is reached. The NGBD algorithm uses five initial iterations of the optimization-based FBBT techniques described in Section 2.3.3.1.1.6 that are solved either using a supported LP solver, or using a supported convex NLP solver depending on the type of formulation used for Problem (NGBD-PBP) (we relax the integrality restrictions on the integer variables in the FBBT subproblems by default). The default choice of norm for Problem (NGBD-FP) is the ∞-norm, whereas the default choice for Problem (NGBD-FRMP) is the 1-norm. We note that the solution of the primal problems for NGBD, Problems (NGBD-PP$_h$), are potentially accelerated by providing 'cutoff objective values' that the global solvers can use to fathom nodes of their B&B trees (see Proposition 3.6 and Remark 3.2 in [139]). If the upper bound in the NGBD algorithm is updated after a sequence of primal problems is solved, we use a single pass of the OBBT technique described in Problem (OBBT$_x$) of Chapter 2 to try and tighten the variables' bounds using optimality arguments (once again, the default setting is to relax the integrality restrictions on the integer variables). Since the size of the OBBT subproblems can increase significantly with the number of iterations of the inner (GBD) loop of the NGBD algorithm, we limit the solution time for each OBBT subproblem (for each considered variable) to $\frac{1}{s}$ seconds to ensure that the total OBBT solution time scales well with the number of scenarios. The current implementation within `GOSSIP` only employs the OBBT step when the OBBT subproblems reduce to LPs after the integrality restrictions are relaxed; if Problem (FBBT$_x$) is not a MILP, then Problem (OBBT$_x$) is not employed.

The termination tolerances for the (overall) upper and lower bounding problems for LR and MLR (see Algorithm 3.1 in Chapter 3) are set to be $\frac{\varepsilon}{10}$ and $\frac{\varepsilon}{1.2}$, respectively. The default maximum number of 'dual iterations' for these algorithms is $D_{\max} = 5$, and the default branching strategies for these algorithms employs the parameter values $N = 2$ and $\lambda = 0.75$. The Lagrange multipliers for the non-anticipativity constraints in LR and MLR are all initially set to zero. At each node of the B&B tree, two rounds of the optimization-based FBBT techniques described in Sections 2.3.3.1.1.6 and 3.3.2 are solved using a supported LP/NLP solver depending on the type of formulation used. Both the LR and MLR algorithms use ANTIGONE and IPOPT with a maximum time limit of five seconds to try and generate feasible points at each node of the B&B tree (see Section 2.3.3.1.2.2 of Chapter 2 and Section 3.3.1 of Chapter 3). The default setting for the modified Lagrangian relaxation algorithm is to use at most three rounds of ABT per variable per lower/upper bound at each node with the Lagrange multipliers fixed to the initial values at those nodes; on the other hand, the default setting for the conventional Lagrangian relaxation algorithm is to not use the ABT steps.

4.2.5 Future work

We list some important avenues for future implementation work within `GOSSIP`, organized into a few categories, below:

- Algorithms: nonconvex outer-approximation [118].
- Relaxation techniques: polyhedral relaxation framework [224, 225], piecewise linear relaxations [160, 165], reformulation-linearization cuts [143, 161, 163, 164, 209, 210], and other specialized relaxation strategies [11, 12, 120, 121, 161–164, 227, 250, 251].
- Finding feasible points: multi-start methods [228], variable neighborhood search methods and their variants [95, 141], and the use of software for two-stage stochastic NLPs [56, 247] for the efficient generation of feasible points.
- Software links: GUROBI [92], BONMIN [40], SCIP [233].
- Miscellaneous: exploiting parallelizability.

4.3 The `GOSSIP` test library

The `GOSSIP` test library currently consists of twenty test cases from eight different applications, including the stochastic pooling problem [136], continuous and discrete-versions of a tank sizing problem [186], a trim loss minimization problem [96], and continuous and discrete-versions of a refinery model [241]. Multiple instances of each of these test cases involving different numbers of scenarios are generated. Out of the twenty test cases, two involve continuous complicating variables and require the use of the MLR algorithm in place of the NGBD algorithm. Additionally, we attempt to solve all twenty test cases using ANTIGONE 1.1, BARON 16.3.4, Couenne 0.5, SCIP 3.2, and the LR algorithm. A summary of the contents of the `GOSSIP` test library is provided in Table 4.2. We plan to make the test library available for download on the group website `http://yoric.mit.edu/software`.

4.4 Computational experiments

To test the capabilities of `GOSSIP`, we compare the performance of the decomposition techniques (NGBD/MLR and LR) in `GOSSIP` against performance of four general-purpose

Table 4.2: The GOSSIP test library. Check marks in the columns for n_y, n_z, n_{x_b}, and n_{x_c} indicate the presence of binary complicating variables, continuous complicating variables, binary recourse variables, and continuous recourse variables, respectively. The entries in the last column indicate the types of nonlinear terms present in each model, with BLIN denoting bilinear terms, TRILIN denoting trilinear terms, USIG denoting univariate signomial terms, MSIG denoting multivariate signomial terms, and LOG denoting logarithmic terms.

Problem Type	**# Cases**	$\mathbf{n_y}$	$\mathbf{n_z}$	$\mathbf{n_{x_b}}$	$\mathbf{n_{x_c}}$	**Nonlinear term types**
Pooling [2, 136, 159]	5	✓			✓	BLIN
Sarawak [136, 139, 140]	3	✓			✓	BLIN
Pump network [137]	2	✓			✓	BLIN, USIG
Software reliability [139]	2	✓			✓	BLIN, TRILIN, MSIG, LOG
Continuous tank sizing [186]	1		✓	✓	✓	BLIN, USIG
Discrete tank sizing [186]	2	✓		✓	✓	BLIN
Discrete refinery model [241]	1	✓			✓	BLIN
Continuous refinery model [241]	1	✓	✓		✓	BLIN
Trim loss [96]	2	✓		✓		BLIN
Knapsack [9]	1	✓		✓		

state-of-the-art deterministic global optimization software ANTIGONE 1.1, BARON 16.3.4, Couenne 0.5, and SCIP 3.2, accessed through GAMS 24.7.1 [83], on the GOSSIP test library summarized in Table 4.2 (in the last knapsack case study, we test the performance of NGBD and LR against SCIP 3.2 and MILP software CBC 2.9 and CPLEX 12.6.3.0). We note that SNGO [47] cannot be used to solve any of the problems in Table 4.2 since all of the problems therein involve discrete variables in their formulation. We constructed multiple instances of each of the test cases in Table 4.2 with a varying number of scenarios to analyze the rates at which the solution times of these techniques increases with an increase in the number of scenarios. The maximum number of scenarios considered in each case study is determined by the best-performing global solver.

All of the case study instances were solved using a Dell Precision T5810 workstation with a 3.5 GHz Intel Xeon E5-1650 v3 processor on a VMWare 11.0 Workstation running a Ubuntu 14.04 virtual machine with 6 GB of memory. An absolute tolerance of 10^{-9}, a relative tolerance of 10^{-3}, and a time limit of 10,000 seconds (about 2.8 hours) were imposed for all solvers for each problem instance. A maximum iteration limit of 10^9 was set for each global solver, with the remaining parameter values set to their defaults. To minimize the effect of fluctuating machine load, each of the four global solvers was successively used to solve each problem instance (see [66, Section 3]); the decomposition techniques in `GOSSIP` were used to solve the case studies at a later time. The detailed results of the computational experiments are presented as part of the supplementary information in Section 4.6. In the rest of this section, we will present graphical comparisons of the performance of the six different solution techniques as the number of scenarios is varied for the problems in the test library. This is in contrast to conventional comparison plots made among global optimization software (see, for instance, [162, Section 5]), where performance profiles of the CPU time and the final relative termination gap are compared across solvers. Plots that provide insights on the empirical scaling of the solution times of the implemented decomposition algorithms (and considered general-purpose software) as a function of the number of scenarios are also provided wherever relevant. Plots in which data points corresponding to any one of the solvers is missing indicates that the corresponding solver(s) was unable to solve any of the scenario problems for that test case.

Stochastic pooling problems

Figures 4-1 to 4-4 compare the performance of the different solvers on the first four stochastic pooling problem instances. Since only ANTIGONE was able to solve even the single scenario problem of stochastic pooling problem #5 (in 1879 seconds), we do not plot a comparison plot for that case. The advantages of NGBD and LR over the four general-purpose global optimization software for the first three instances of the stochastic pooling problem are evident from Figures 4-1 to 4-3. The solution times of both the NGBD and LR implementations appear to scale affinely with the number of scenarios for these instances, whereas the solution time of even the best-performing general-purpose solver appears to scale exponentially. Only ANTIGONE and BARON are able to solve even the single scenario problem for the fourth instance, and ANTIGONE is the only solver that is able to solve the single scenario case for the fifth instance (see Tables 4.9 and 4.10). Although the Lagrangian relaxation algorithm uses ANTIGONE to generate its lower (and upper) bounds, it is unable to converge within the time limit for the single scenario instances of stochastic pooling problems #4 and #5. This suggests that the version of ANTIGONE within GAMS 24.7.1 works differently than the version linked within `GOSSIP`. Tables 4.3 to 4.10 present detailed computational results for the five stochastic pooling problems.

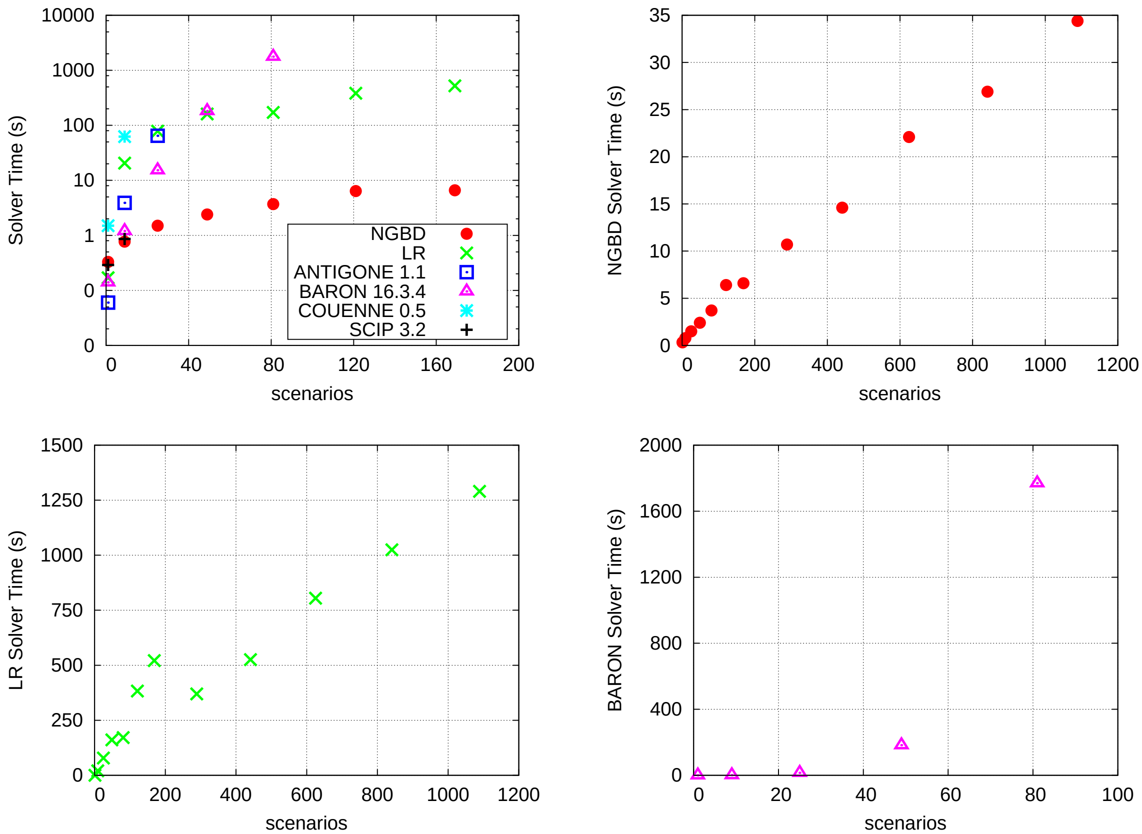


Figure 4-1: Comparison of the different solvers on stochastic pooling problem #1 [136] (see Tables 4.3 and 4.4). The solution times of NGBD and LR seem to scale affinely with the number of scenarios, whereas the solution time of the best-performing general-purpose solver for this instance, BARON, appears to scale less favorably.

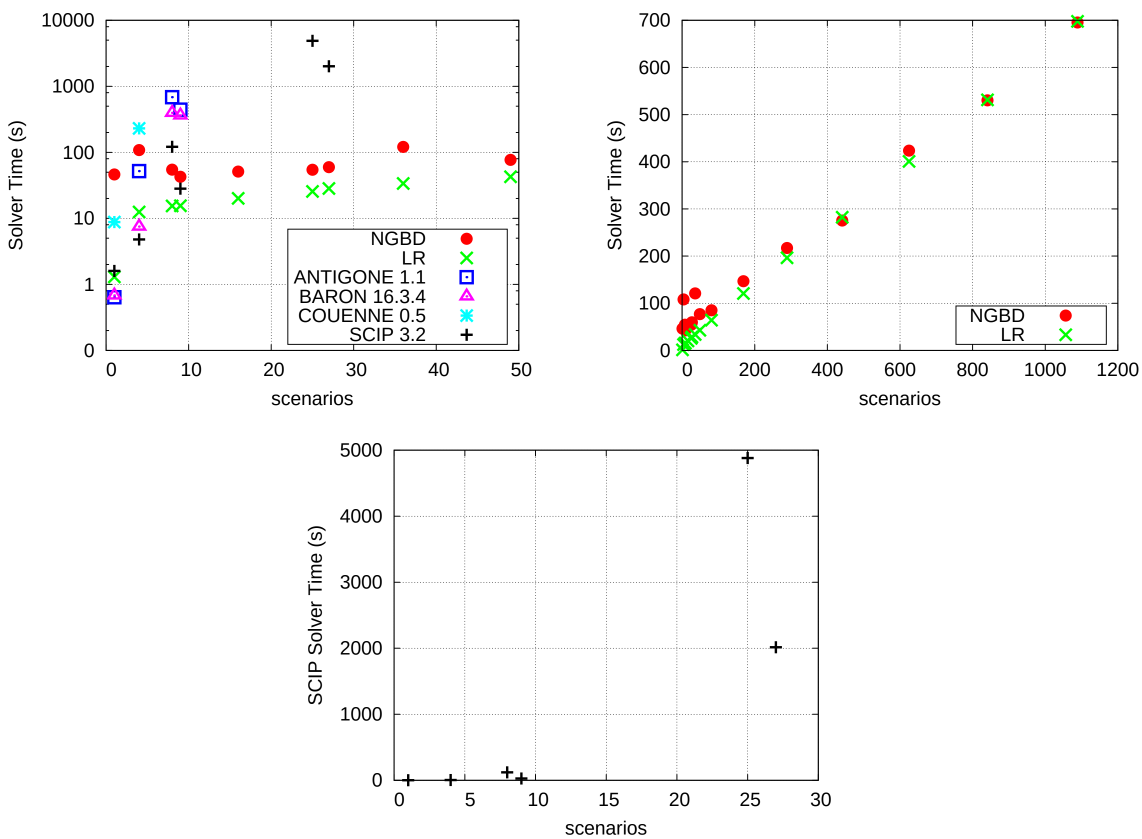


Figure 4-2: Comparison of the different solvers on stochastic pooling problem #2 based on [2] (see Tables 4.5 and 4.6). The solution times of NGBD and LR seem to scale affinely with the number of scenarios, whereas the solution time of the best-performing general-purpose solver for this instance, SCIP, increases significantly with the number of scenarios.

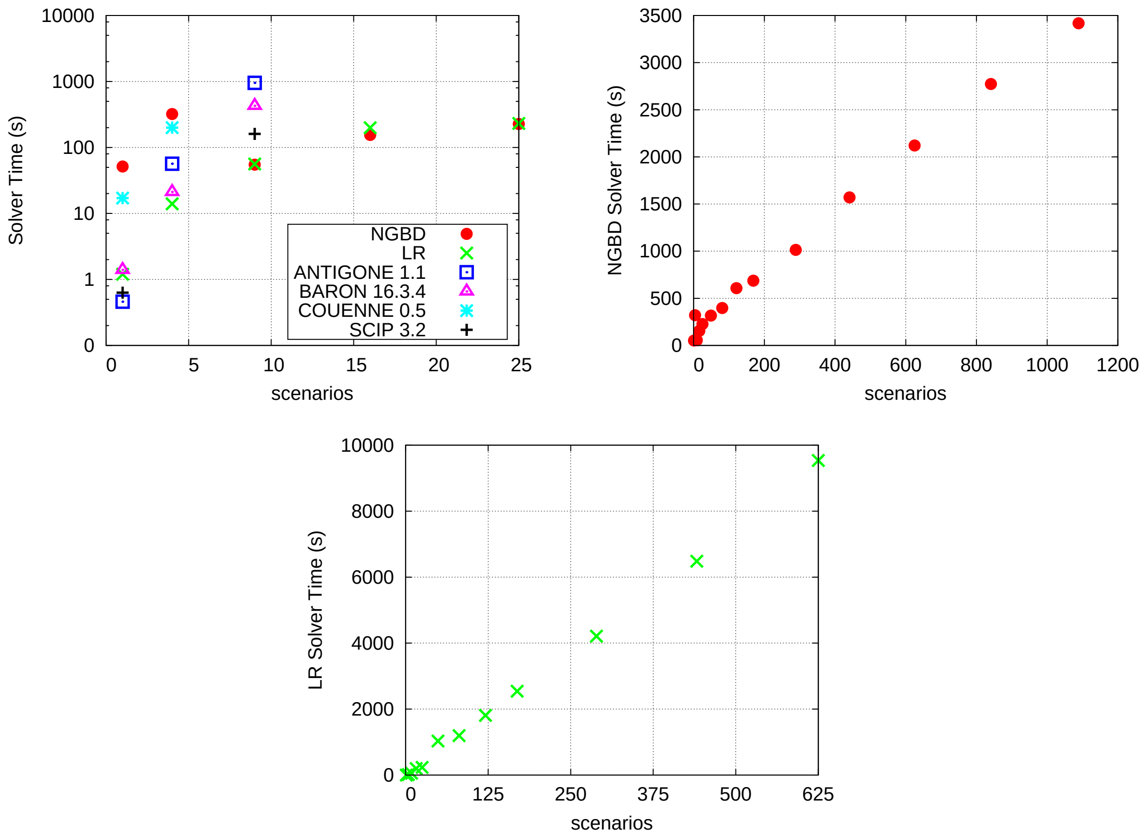


Figure 4-3: Comparison of the different solvers on stochastic pooling problem #3 based on [2] (see Tables 4.7 and 4.8). The solution times of NGBD and LR appear to scale affinely with the number of scenarios.

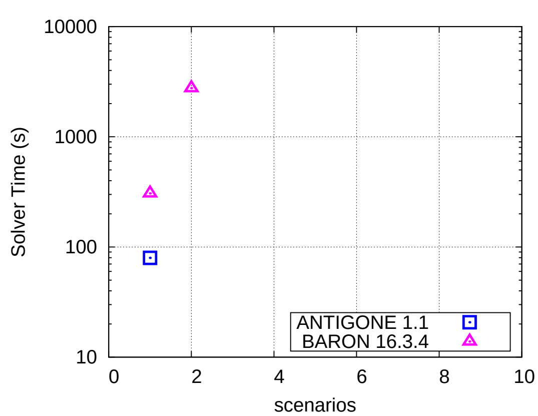


Figure 4-4: Comparison of the different solvers on stochastic pooling problem #4 based on [159] (see Table 4.9). Only ANTIGONE and BARON manage to solve even the single scenario instance within 10,000 seconds.

Sarawak gas production network problems

Figures 4-5 to 4-7 compare the performance of the different solvers on the Sarawak gas production system problem instances. Once again, these case studies demonstrate the advantages of the implemented decomposition techniques over general-purpose state-of-the-art global optimization software for large-scenario instances of Problem (SP). While both NGBD and LR outperform the general-purpose software for large-scenario instances of the first Sarawak case study, we observe from Figure 4-5 that their solution times grows faster than affinely with the number of scenarios. This decline in the performance of these decomposition algorithms for large-scenario instances is due to the difficulty faced by the version of ANTIGONE within GOSSIP for solving the corresponding problems to near global optimality. In particular, this undesirable scaling may be a consequence of the fact that both of these algorithms attempt to solve subproblems of the larger scenario instances of the gas network problem to increasingly tighter termination tolerances using ANTIGONE (see Section 4.2.4 for the details of the default termination tolerances for the subproblems). For the remaining two instances of the Sarawak gas network problem, the solution times of both NGBD and LR exhibit the expected affine scaling behavior with the number of scenarios. Tables 4.11 to 4.16 present detailed computational results for the three Sarawak gas production system-based case studies.

Pump network problems

Figures 4-8 and 4-9 compare the performance of the different solvers on the pump network problem instances. The nonaffine scaling of the solution time of NGBD and LR with the number of scenarios for the first pump network problem (see Figure 4-8) is once again due to the nonaffine scaling of the time taken to solve the primal problems in NGBD using ANTIGONE. While the general-purpose global solvers seem to be scaling better than worst-case exponential for the above example (as seen from Figure 4-8), they are unable to solve problems with more than 27 scenarios within 10,000 seconds of solver time. Tables 4.17 to 4.20 present detailed computational results for the two pump network problems.

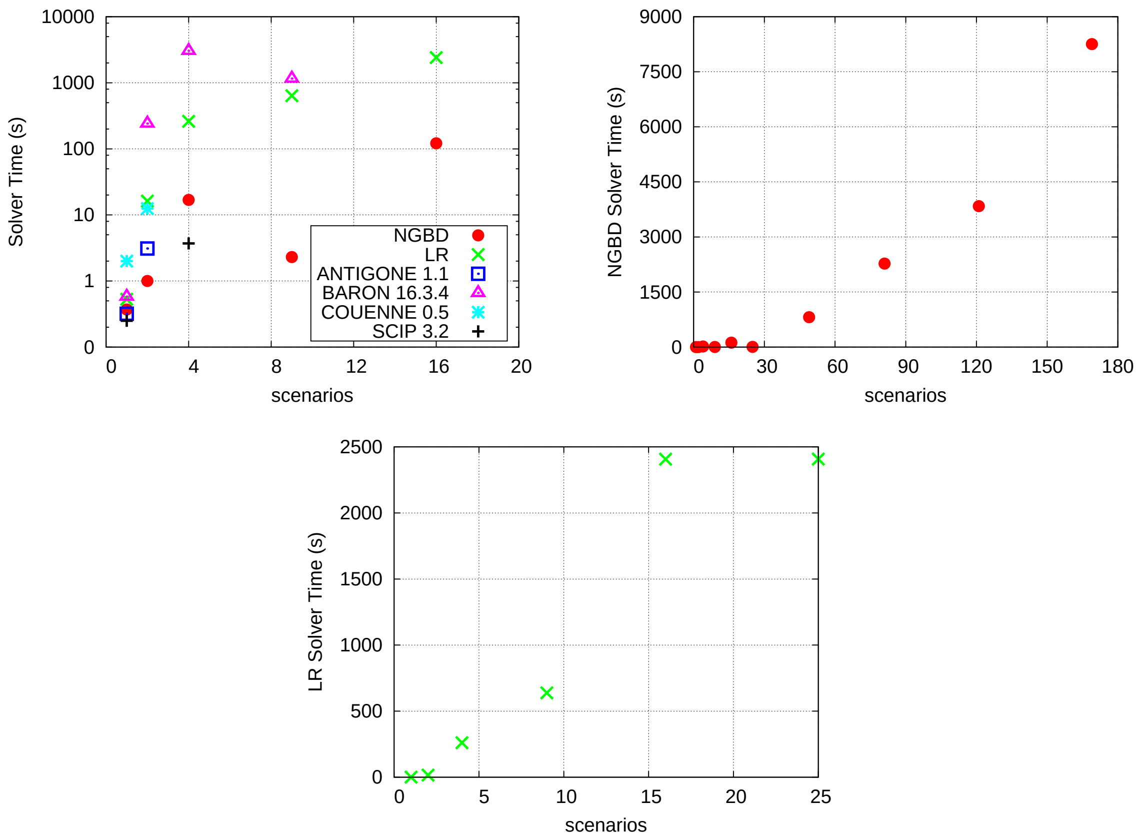


Figure 4-5: Comparison of the different solvers on the Sarawak gas production problem #1 [136] (see Tables 4.11 and 4.12). The solution times of NGBD and LR appear to scale worse than affinely with the number of scenarios; however, these decomposition algorithms perform better than the general-purpose solvers for instances with more than eight scenarios.

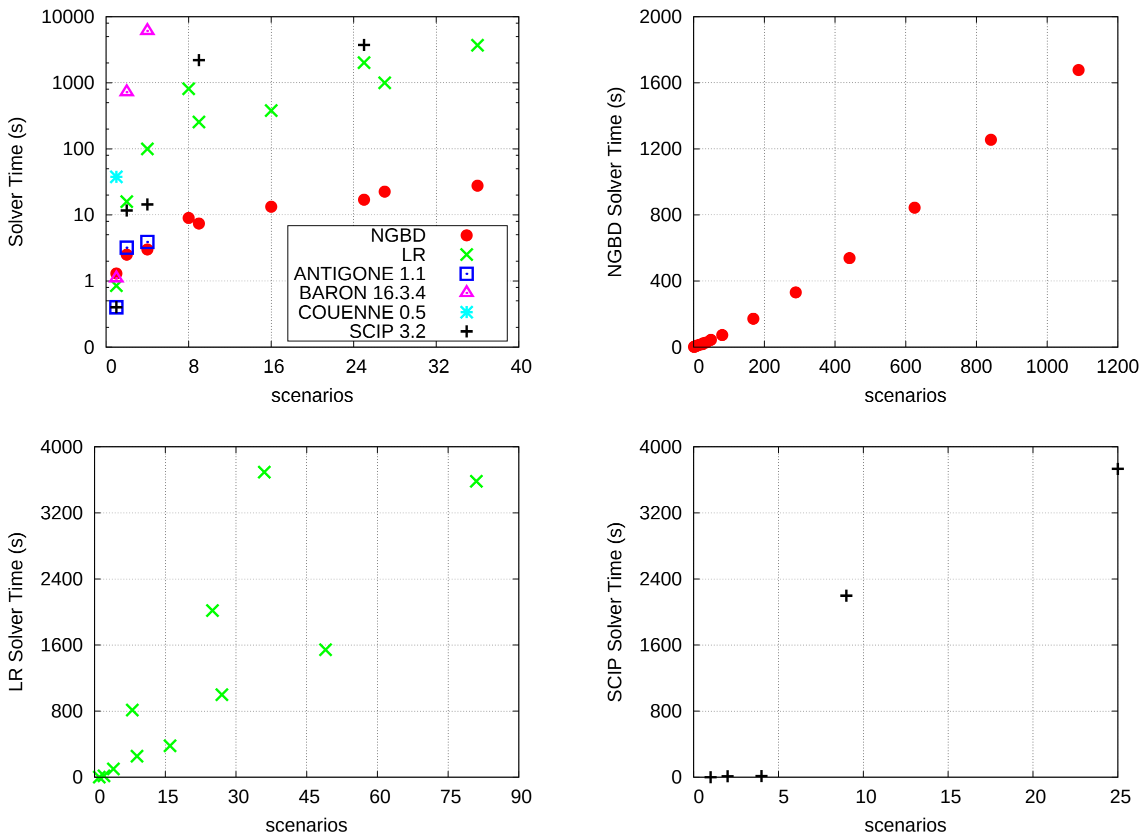


Figure 4-6: Comparison of the different solvers on the Sarawak gas production problem #2 [136] (see Tables 4.13 and 4.14). The solution times of NGBD and LR scale favorably with the number of scenarios, whereas the solution time of the best-performing general-purpose solver for this instance, SCIP, increases significantly with the number of scenarios.

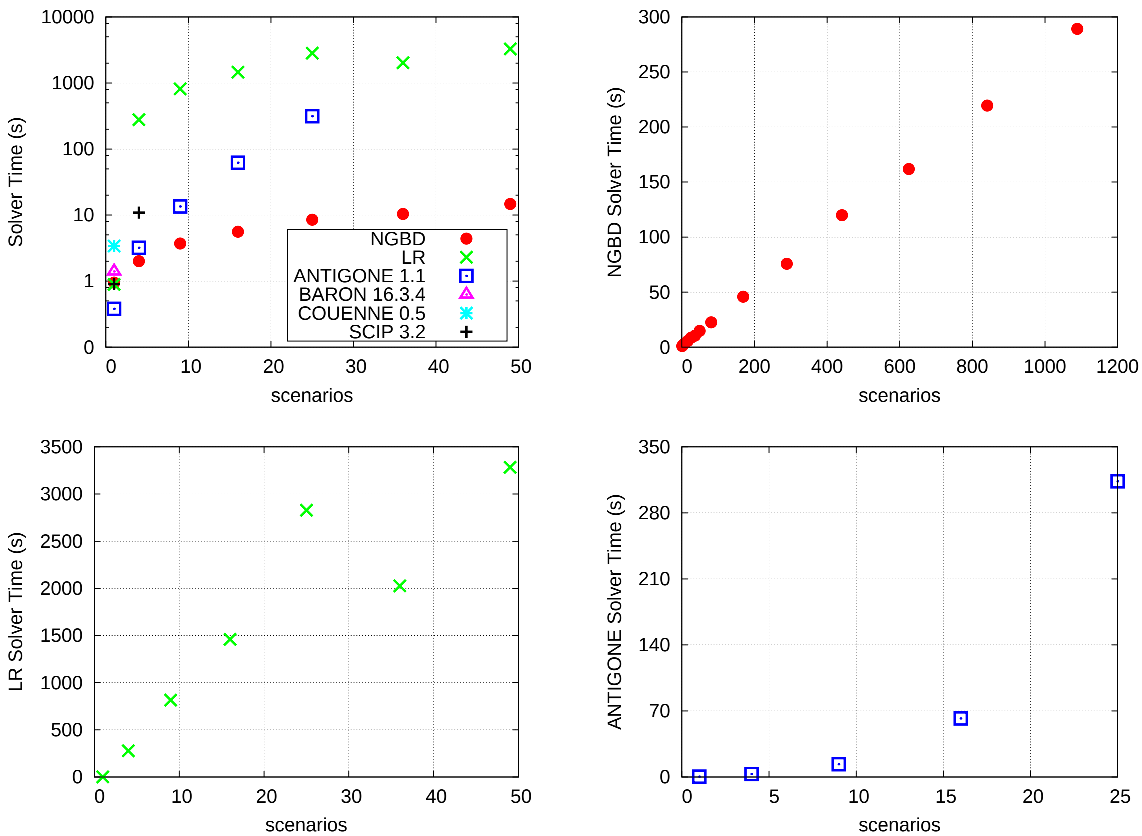


Figure 4-7: Comparison of the different solvers on the Sarawak gas production problem #3 [140] (see Tables 4.15 and 4.16). The solution times of NGBD and LR scale favorably with the number of scenarios, whereas the solution time of the best-performing general-purpose solver for this instance, ANTIGONE, seems to increase asymptotically exponentially with the number of scenarios.

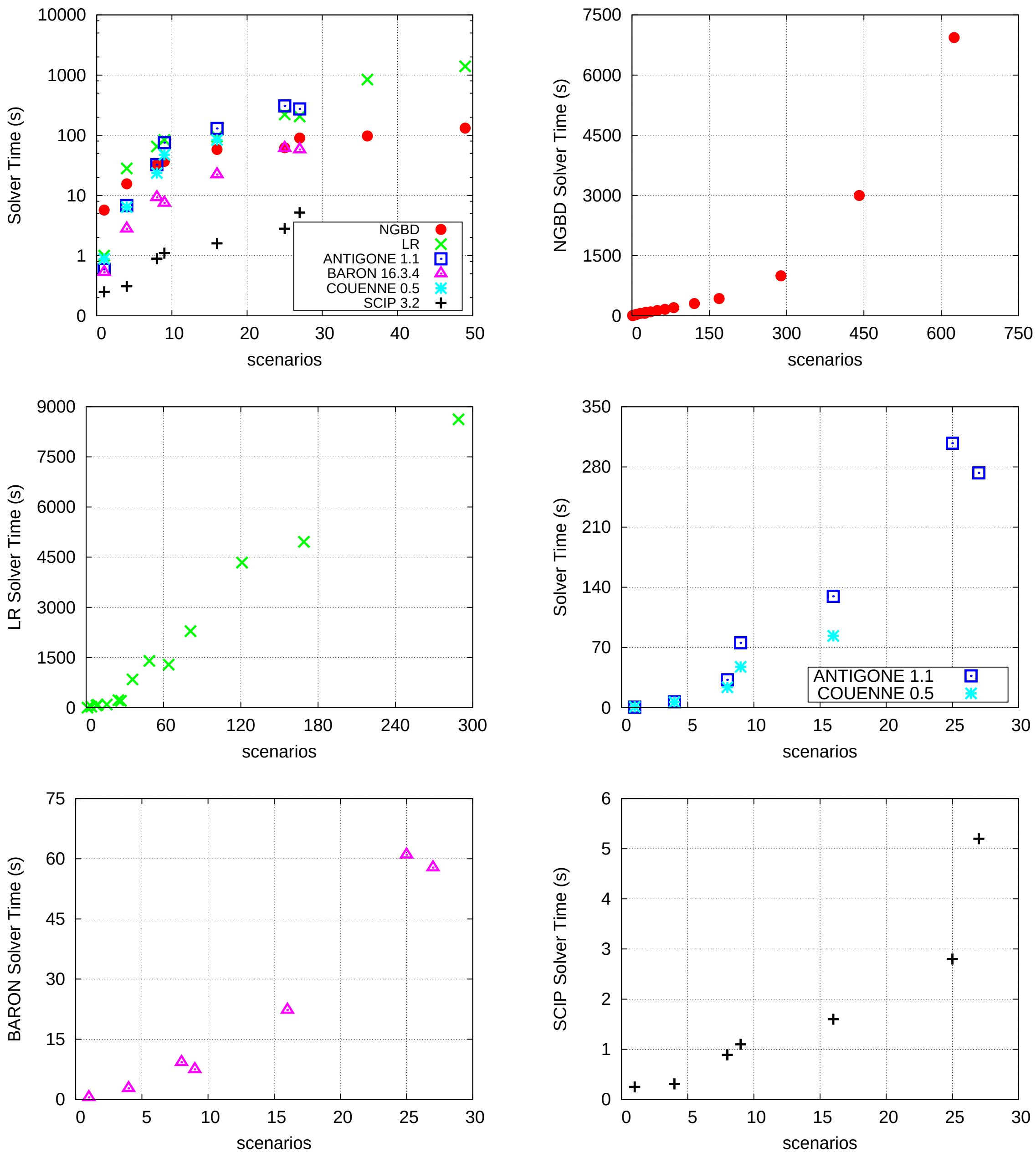


Figure 4-8: Comparison of the different solvers on pump network problem #1 [137] (see Tables 4.17 and 4.18). Although the solution times of NGBD and LR appear to scale worse than affinely with the number of scenarios, they can solve problems with larger number of scenarios within the time limit.

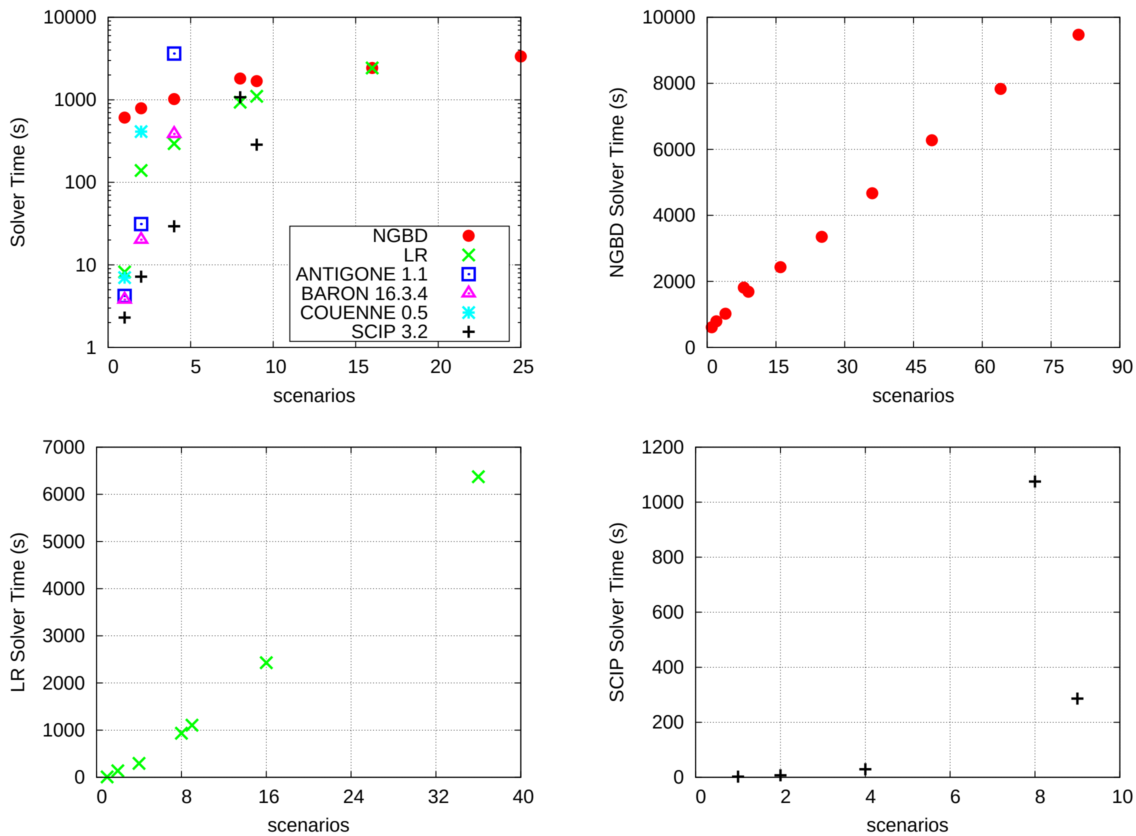


Figure 4-9: Comparison of the different solvers on pump network problem #2 [137] (see Tables 4.19 and 4.20). The solution times of NGBD and LR scale favorably with the number of scenarios, whereas the solution time of the best-performing general-purpose solver for this instance, SCIP, seems to increase significantly with the number of scenarios.

Software reliability problems

Figures 4-10 to 4-11 compare the performance of the different solvers on the software reliability problem instances. Despite the fact that the solvers BARON and SCIP appear to be scaling sub-exponentially with the number of scenarios for the first software reliability case study, the reader can verify that the decomposition techniques outperform these general-purpose software for this example (we note that LR does not converge for the instances with more than 729 scenarios because of a memory-related bug within the `GOSSIP` interface to IPOPT that will hopefully be fixed in the future). For the second instance of the software reliability problem, the solvers BARON and SCIP once again seem to be scaling sub-exponentially with the number of scenarios and perform better than NGBD and LR within the time limit; however, NGBD seems to enjoy a favorable affine scaling for this example, and could be expected to perform better than the general-purpose software for larger-scenario instances. Tables 4.21 and 4.22 present detailed computational results for the two software reliability case studies.

Tank sizing Problems

Continuous tank sizing

Figure 4-12 compares the performance of the different solvers on a continuous tank sizing problem instance. Tables 4.23 and 4.24 provides detailed computational results for the continuous tank sizing problem (also see Section 3.6.3 of Chapter 3). While all of the general-purpose solvers struggle to solve the cases with more than two scenarios to within the termination tolerance in 10,000 seconds, the two applicable decomposition techniques fare better for this example. LR is able to solve instances with up to five scenarios for this case, whereas MLR improves upon the performance of LR by virtue of the aggressive bounds tightening technique (see Section 3.3.2 of Chapter 3) and solves problems with up to 21 scenarios within the time limit.

Discrete tank sizing

Figures 4-13 and 4-14 compares the performance of the different solvers on two discrete tank sizing problem instances. Tables 4.25 to 4.27 present detailed computational results for the discrete tank sizing instances. NGBD fails to converge for instances with more

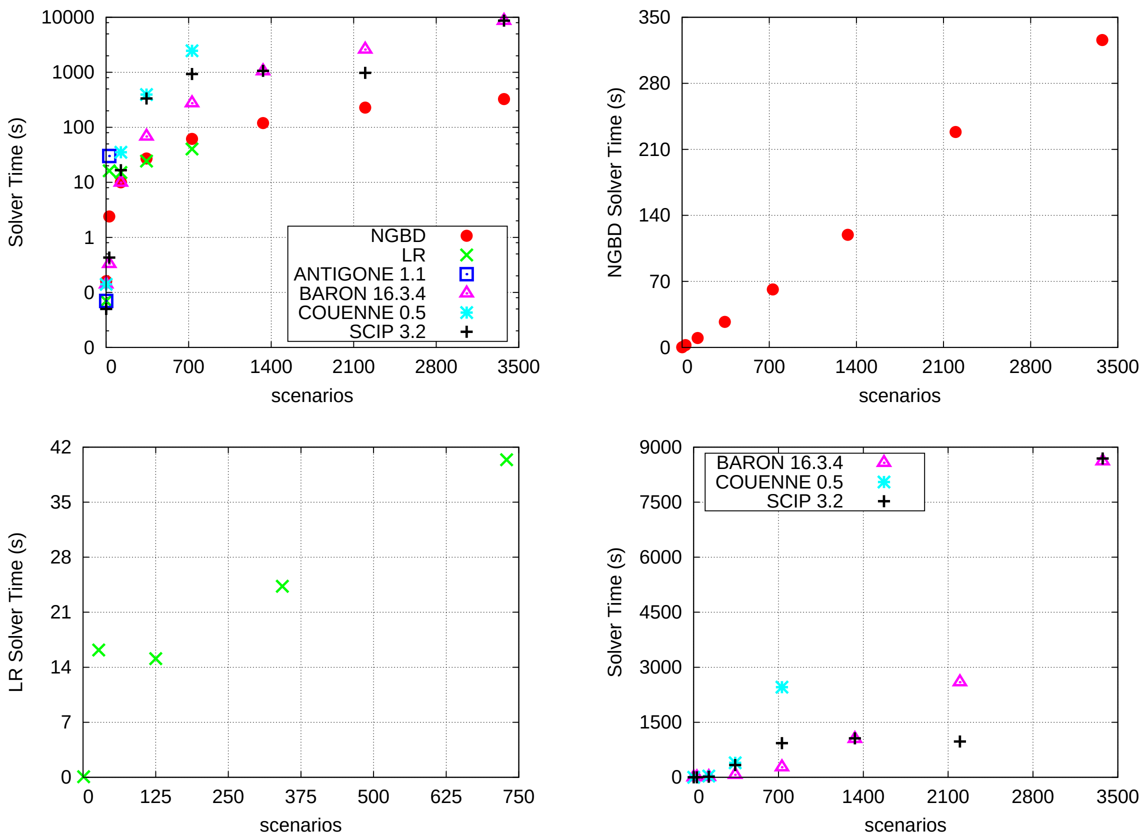


Figure 4-10: Comparison of the different solvers on software reliability problem #1 [139] (see Table 4.21). NGBD outperforms all the other solution techniques for large-scenario instances. Interestingly, BARON, Couenne, and SCIP appear to be scaling sub-exponentially with the number of scenarios.

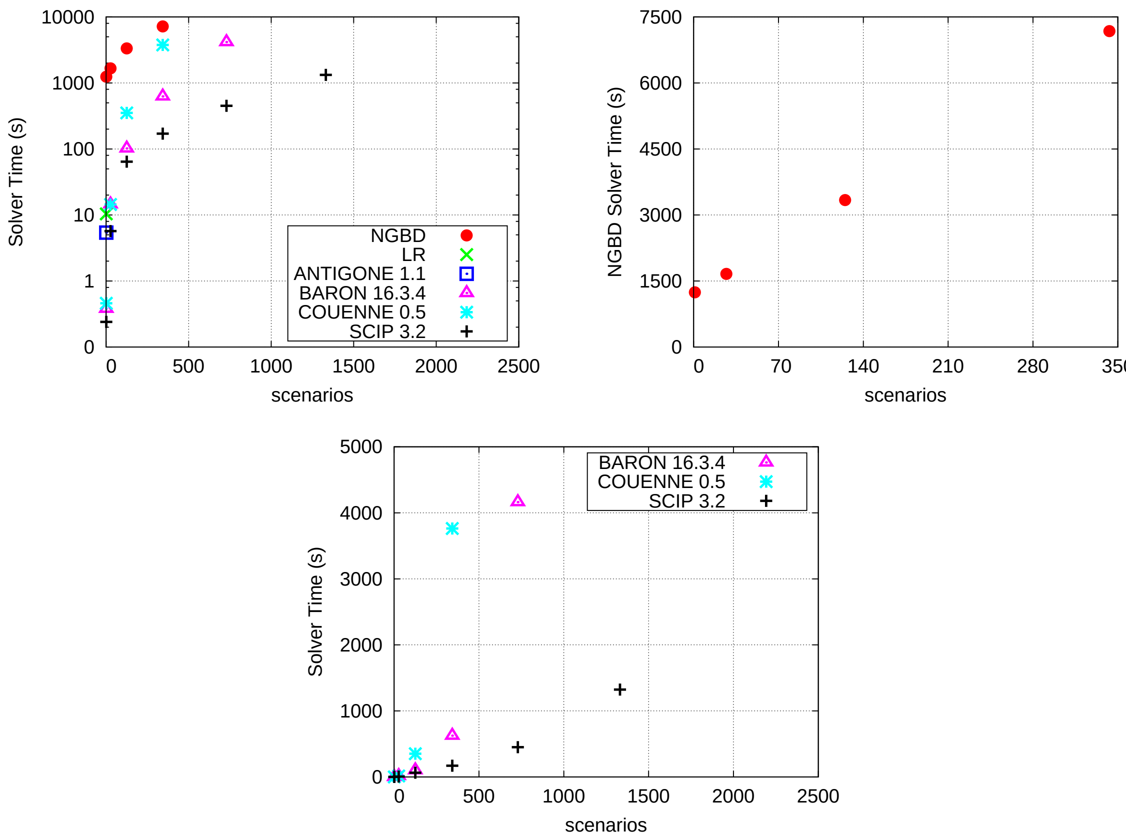


Figure 4-11: Comparison of the different solvers on software reliability problem #2 (see Table 4.22). BARON, Couenne, and SCIP (which is the best-performing solver for this case study) appear to be scaling sub-exponentially with the number of scenarios, whereas NGBD empirically scales affinely with the number of scenarios.

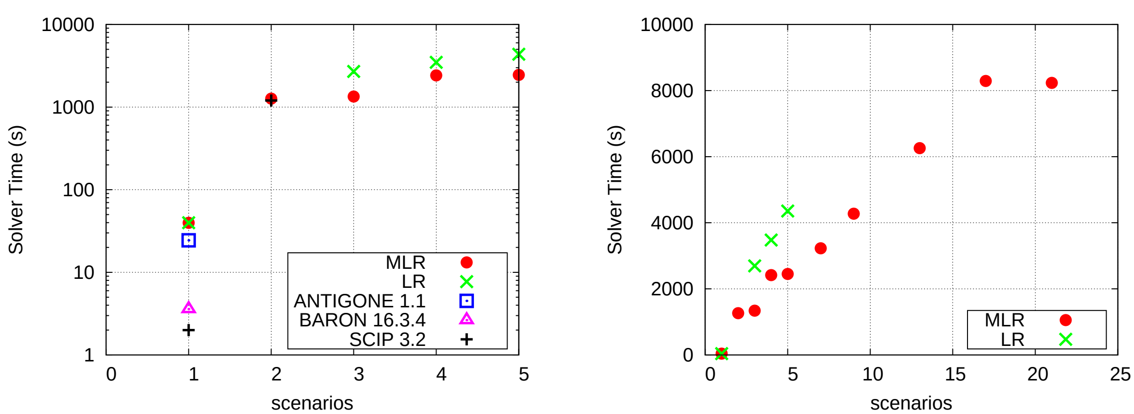


Figure 4-12: Comparison of the different solvers on the continuous tank sizing problem [186] (see Tables 4.23 and 4.24). The decomposition algorithms perform favorably compared to the general-purpose solvers for this case study. The implementation of the aggressive bounds tightening technique within the MLR algorithm improves its performance significantly compared to the LR algorithm.

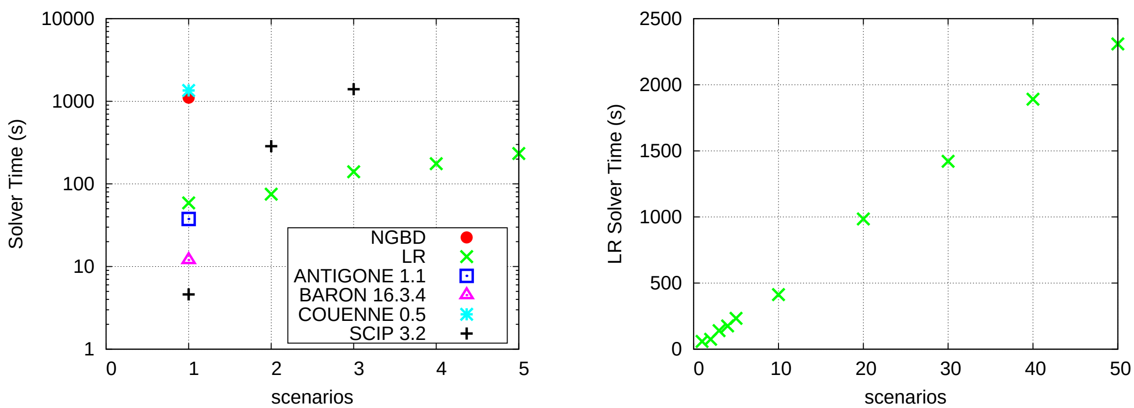


Figure 4-13: Comparison of the different solvers on the discrete tank sizing problem #1 [186] (see Tables 4.25 and 4.26). Only the LR algorithm, which empirically scales affinely with the number of scenarios, is able to solve large-scenario problems for this case study.

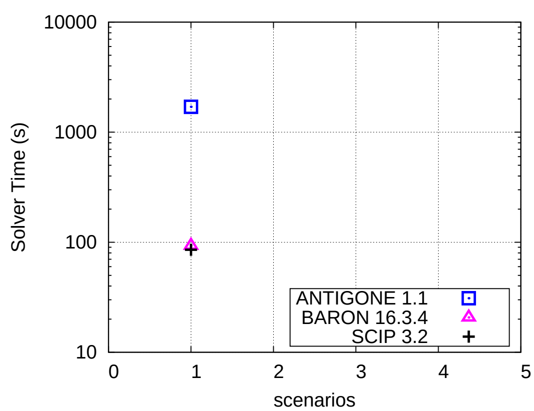


Figure 4-14: Comparison of the different solvers on discrete tank sizing problem #2 [186] (see Table 4.27). None of the tested solvers can solve instances with more than one scenario, with the decomposition methods unable to even solve the single scenario instance within the time limit.

than one scenario for the first case study (and even for the single scenario instance of the second case study) within the time limit because of the large number of binary variable realizations explored during the course of the algorithm due to the significant underestimation/relaxation gap. Lagrangian relaxation, on the other hand, outperforms all of the tested solution techniques for the first case study, and its solution time scales linearly with the number of scenarios for this case. Both NGBD and LR fail to solve even the single scenario instance of the second case study, and the state-of-the-art global solvers struggle to solve multi-scenario instances of this case study within the time limit.

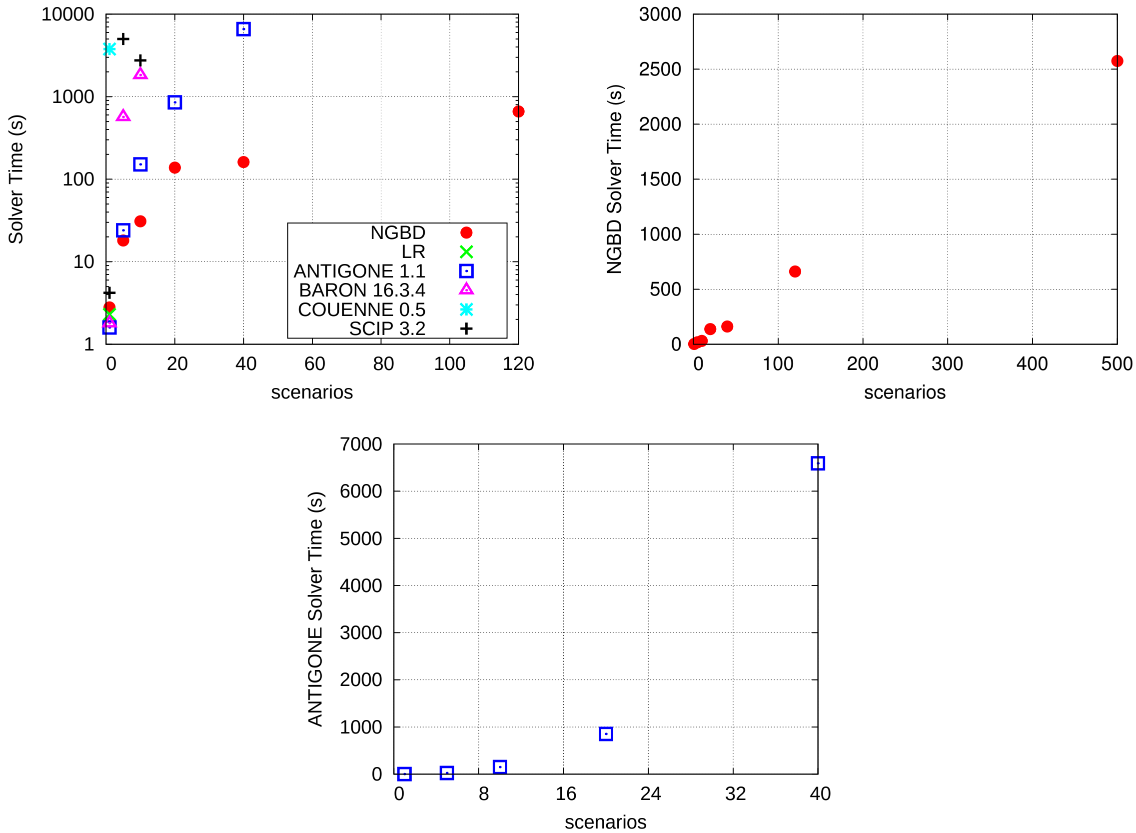


Figure 4-15: Comparison of the different solvers on the discrete refinery model [241, Example 2] (see Table 4.28). NGBD is the best-performing algorithm for this case study with its solution time scaling affinely with the number of scenarios. The best-performing general purpose solver, ANTIGONE, on the other hand, appears to scale unfavorably with the number of scenarios.

Refinery model problems

Discrete refinery model

Figure 4-15 compares the performance of the different solvers on the refinery model instance, and Table 4.28 provides detailed computational results for the discrete refinery model case study. ANTIGONE is the best-performing general-purpose solver for this case study; however, its solution time appears to grow significantly with the number of scenarios. The solution time of NGBD, on the other hand, empirically grows affinely with the number of scenarios, and it is able to solve instances with up to 500 scenarios in less than one hour.

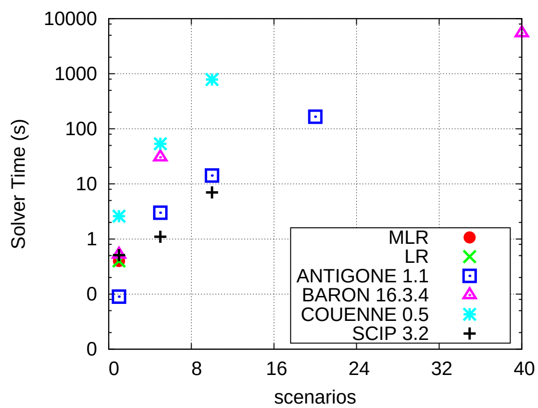


Figure 4-16: Comparison of the different solvers on the continuous refinery model (see Table 4.29). Both LR and MLR are unable to solve multi-scenario instances, whereas the commercial solvers ANTIGONE and BARON can solve instances with more than ten scenarios within 10,000 seconds.

Continuous refinery model

Figure 4-16 compares the performance of the different solvers on the continuous refinery model instance (also see Section 3.6.3 of Chapter 3). Both the MLR and LR algorithms face difficulties in converging to an optimal solution for multi-scenario instances. The solvers ANTIGONE and BARON, on the other hand, can solve instances with more than ten scenarios for this case study within the time limit despite not scaling favorably with the number of scenarios. Table 4.29 provides detailed computational results for the continuous refinery model case study.

Trim loss minimization problems

Figures 4-17 and 4-18 compare the performance of the different solvers on the two (challenging) trim loss minimization problem instances. NGBD fails to converge for even the single scenario instances of these case studies within the time limit because of the significant underestimation/relaxation gap induced by the challenging discrete nature of the problem. SCIP outperforms all of the tested solution techniques for the first case study by virtue of its strengths in handling integer programs. None of the solution techniques can solve instances with more than one scenario for the second example. Tables 4.30 and 4.31 present detailed computational results for these instances.

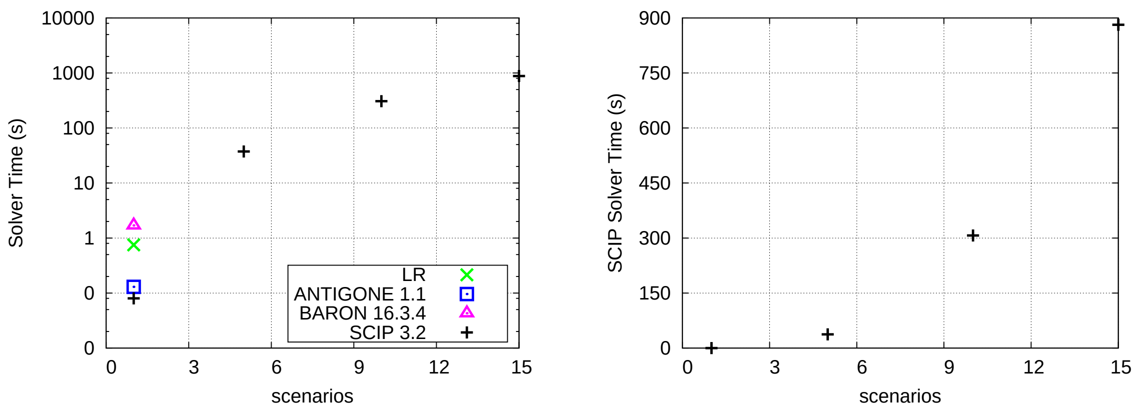


Figure 4-17: Comparison of the different solvers on trim loss model #1 (see Table 4.30). NGBD fails to solve even the single scenario instance for this case, while SCIP outperforms all of the tested solvers and can solve problems with up to 15 scenarios within 10,000 seconds.

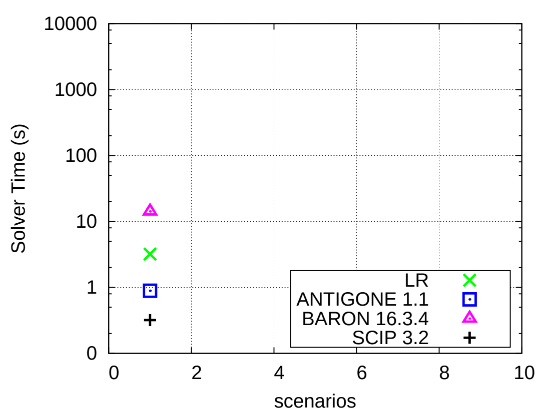


Figure 4-18: Comparison of the different solvers on trim loss model #2 (see Table 4.31). None of the tested solvers can solve multi-scenario instances of this problem, and NGBD once again fails to solve even the single scenario instance within the time limit.

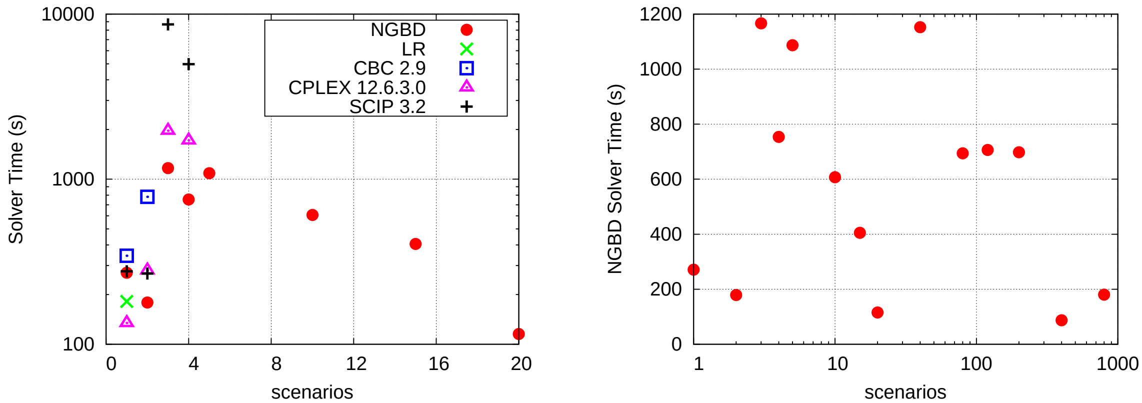


Figure 4-19: Comparison of the different solvers on a knapsack problem (see Tables 4.32 and 4.33). NGBD outperforms all of the general-purpose MILP solvers and the LR algorithm for this case study. Note that the horizontal axis of the plot on the right is in logarithmic scale.

Knapsack problem

Figure 4-19 compares the performance of the different solvers on instances of a knapsack problem. The solution time of NGBD seems to be quite independent of the number of scenarios for this problem, as seen from Figure 4-19 (note that the horizontal axis is in logarithmic scale), while sophisticated MILP software such as CPLEX fail to solve instances with five or more scenarios within 10,000 seconds. Tables 4.32 and 4.33 provide detailed computational results for this case study.

4.5 Conclusion

This chapter introduces GOSSIP, decomposition software for the global optimization of a broad class of two-stage stochastic MINLPs. GOSSIP includes implementations of state-of-the-art decomposition techniques such as nonconvex generalized Benders decomposition (NGBD), Lagrangian relaxation (LR), and a modified Lagrangian relaxation (MLR) algorithm, and will be the first publicly available decomposition software for the global solution of two-stage stochastic MINLPs. This chapter also instituted the first soon-to-be publicly available test library for the above challenging class of problems.

Computational experiments demonstrated that the decomposition techniques implemented within GOSSIP generally outperformed four state-of-the-art software for MINLPs. Additionally, the decomposition techniques typically exhibited a favorable affine scaling with the number of scenarios (on a serial computer) compared to general-purpose software whose solution times usually increased significantly with the number of scenarios. In particular, at least one of the decomposition techniques exhibited the best performance on large-scenario instances of thirteen of the twenty case studies. On four of the seven of the case studies in which all of the decomposition techniques were not the best-performing 'solver', none of the solvers could solve instances with more than two scenarios, highlighting the challenging nature of these problems. `GOSSIP` can therefore (justifiably) be viewed as a decomposition toolkit that provides a framework for the scalable solution of a challenging class of two-stage stochastic MINLPs.

4.6 Detailed results of the computational experiments

This section tabulates the detailed computational results for the computational studies summarized in Section 4.4. For each case study, we list the solution time in seconds (rounded to the nearest 0.1 second) and percentage relative termination gap separated by '/', defined as

$$\text{gap} = \min\left\{100, 100 \times \frac{\texttt{upperbound} - \texttt{lowerbound}}{\max\left\{|\texttt{lowerbound}|, |\texttt{upperbound}|\right\} + \delta}\right\} [\%],$$

where $\delta \in (0, 1)$ such that $\delta \ll 1$, for each tested solver for varying numbers of scenarios. We note that the reported percentage termination gap is rounded to the nearest 0.1 percent if the instance was not solved within the time limit, and set to be equal to 0.1 percent (which is the desired solution accuracy) otherwise. A blank entry ('`-`') for the solution time indicates that the solution reached the time limit of 10,000 seconds. The entry '`t`' for the solution time indicates that the solver terminated prematurely due to failure (likely due to insufficient memory to continue). The entry '`wd`' for the solution time indicates that the solver returned an incorrect solution as optimal because of an incorrect dual (lower bound) value. The entry '`i`' for the solution time indicated that the solver wrongly concluded that the model is infeasible. A blank entry ('`-`') for the termination gap either indicates that a feasible point wasn't found within the time limit, or that the termination gap is not relevant for this entry. An entry of '100' for the termination gap indicates that a feasible solution was found, but the relative termination gap is large (nearly 100%). Empty entries for both the solution time and termination gap indicates that the computational experiment was not carried out.

Stochastic pooling problems

Table 4.3: Comparison of the different solvers on stochastic pooling problem #1. This case study includes 16 binary complicating variables, 21 continuous recourse variables per scenario, 26 first-stage constraints, 55 second-stage constraints per scenario, and 8 bilinear terms per scenario.

# Scenarios	1	9	25	49	81	121	169
ANTIGONE 1.1	0.1/0.1	3.9/0.1	64.7/0.1	-/28.9	-/33.3	-/34.0	-/34.2
BARON 16.3.4	0.1/0.1	1.2/0.1	15.3/0.1	182.4/0.1	1769.6/0.1	-/27.0	-/22.3
COUENNE 0.5	1.5/0.1	62.3/0.1	-/32.0	-/33.9	-/37.8	-/37.7	-/100
SCIP 3.2	0.3/0.1	0.9/0.1	t,wd/-	-,wd/-	-,wd/-	-/10.9	-/28.4
NGBD	0.3/0.1	0.8/0.1	1.5/0.1	2.4/0.1	3.7/0.1	6.4/0.1	6.6/0.1
LR	0.2/0.1	20.6/0.1	78.5/0.1	160.8/0.1	171.5/0.1	382.8/0.1	521.4/0.1

Table 4.4: Extended results for the decomposition methods for stochastic pooling problem #1.

# Scenarios	289	441	625	841	1089
NGBD	10.7/0.1	14.6/0.1	22.1/0.1	26.9/0.1	34.4/0.1
LR	369.8/0.1	525.4/0.1	804.9/0.1	1024.4/0.1	1290.2/0.1

Table 4.5: Comparison of the different solvers on stochastic pooling problem #2. This case study includes 24 binary complicating variables, 89 continuous recourse variables per scenario, 39 first-stage constraints, 134 second-stage constraints per scenario, and 20 bilinear terms per scenario.

# Scenarios	1	4	8	9	16	25	27	36	49
ANTIGONE 1.1	0.6/0.1	52.0/0.1	683.6/0.1	442.3/0.1	-/4.7	-/3.3	-/3.0	-/3.6	-/3.6
BARON 16.3.4	0.7/0.1	7.5/0.1	396.2/0.1	362.0/0.1	-/5.6	-/4.6	-/6.4	-/5.0	-/12.0
COUENNE 0.5	8.8/0.1	230.4/0.1	-/2.4	-/4.5	-/12.3	-/6.1	-/4.8	-/100	-/100
SCIP 3.2	1.6/0.1	4.8/0.1	121.1/0.1	28.2/0.1	-,wd/-	4881.2/0.1	2014.7/0.1	-/100	-,wd/-
NGBD	46.2/0.1	108.1/0.1	54.6/0.1	42.6/0.1	51.1/0.1	54.4/0.1	59.7/0.1	121.0/0.1	76.9/0.1
LR	1.3/0.1	12.5/0.1	15.4/0.1	15.5/0.1	20.1/0.1	25.6/0.1	28.4/0.1	33.7/0.1	42.7/0.1

Table 4.6: Extended results for the decomposition methods for stochastic pooling problem #2.

# Scenarios	81	169	289	441	625	841	1089
NGBD	85.0/0.1	146.7/0.1	217.2/0.1	275.4/0.1	423.6/0.1	530.2/0.1	695.3/0.1
LR	64.2/0.1	120.7/0.1	196.5/0.1	282.4/0.1	400.7/0.1	531.2/0.1	697.8/0.1

Table 4.7: Comparison of the different solvers on stochastic pooling problem #3. This case study includes 33 binary complicating variables, 108 continuous recourse variables per scenario, 53 first-stage constraints, 212 second-stage constraints per scenario, and 40 bilinear terms per scenario.

# Scenarios	1	4	9	16	25
ANTIGONE 1.1	0.5/0.1	56.8/0.1	953.8/0.1	-/2.7	-/5.1
BARON 16.3.4	1.4/0.1	21.2/0.1	429.0/0.1	-/0.7	-/1.7
COUENNE 0.5	17.1/0.1	200.0/0.1	-/1.4	-/3.2	-/7.2
SCIP 3.2	0.6/0.1	`wd/-`	160.5/0.1	-/0.2	-/0.7
NGBD	51.4/0.1	321.0/0.1	54.9/0.1	154.8/0.1	226.9/0.1
LR	1.2/0.1	14.0/0.1	56.1/0.1	199.1/0.1	231.5/0.1

Table 4.8: Extended results for the decomposition methods for stochastic pooling problem #3.

# Scenarios	49	81	121	169	289	441	625	841	1089
NGBD	316.1/0.1	397.1/0.1	607.2/0.1	686.6/0.1	1013.9/0.1	1570.0/0.1	2121.1/0.1	2773.6/0.1	3417.9/0.1
LR	1032.8/0.1	1194.1/0.1	1807.3/0.1	2541.2/0.1	4210.3/0.1	6480.4/0.1	9534.9/0.1	/	/

Table 4.9: Comparison of the different solvers on stochastic pooling problem #4. This case study includes 55 binary complicating variables, 63 continuous recourse variables per scenario, 100 first-stage constraints, 138 second-stage constraints per scenario, and 48 bilinear terms per scenario.

# Scenarios	1	2	4	8	9
ANTIGONE 1.1	79.7/0.1	-/3.5	-/10.4	-/21.2	-/21.5
BARON 16.3.4	306.7/0.1	2763.7/0.1	-/25.6	-/33.3	-/22.6
COUENNE 0.5	-/21.6	-/28.4	-/32.7	-/42.0	-/39.0
SCIP 3.2	t/50.3	t/55.9	t/58.3	t/58.3	t/-
NGBD	-/45.4	-/45.4	-/45.4	-/43.2	-/45.4
LR	-/7.7	-/27.6	-/100	-/100	-/100

Table 4.10: Comparison of the different solvers on stochastic pooling problem #5. This case study includes 187 binary complicating variables, 207 continuous recourse variables per scenario, 1090 first-stage constraints, 426 second-stage constraints per scenario, and 300 bilinear terms per scenario.

# Scenarios	1	2	4	9
ANTIGONE 1.1	1879.3/0.1	-/8.8	-/34.7	-/54.4
BARON 16.3.4	-/32.6	-/37.1	-/37.7	-/53.6
COUENNE 0.5	-/40.3	-/47.0	-/52.6	-/-
SCIP 3.2	t/69.5	t/71.9	t/67.1	t/61.8
NGBD	-/-	-/-	-/-	-/-
LR	-/32.4	-/100	-/100	-/-

Sarawak gas production network problems

Table 4.11: Comparison of the different solvers on Sarawak gas production problem #1. This case study includes 38 binary complicating variables, 93 continuous recourse variables per scenario, 76 first-stage constraints, 205 second-stage constraints per scenario, and 34 bilinear terms per scenario.

# Scenarios	1	2	4	9	16
ANTIGONE 1.1	0.3/0.1	3.1/0.1	-/0.2	-/1.5	-/2.1
BARON 16.3.4	0.6/0.1	241.3/0.1	3043.1/0.1	1158.9/0.1	-/2.0
COUENNE 0.5	2.0/0.1	12.5/0.1	-/0.2	-/1.8	-/3.2
SCIP 3.2	0.2/0.1	wd/-	3.7/0.1	-/0.5	-/1.4
NGBD	0.4/0.1	1.0/0.1	16.9/0.1	2.3/0.1	121.2/0.1
LR	0.5/0.1	16.2/0.1	261.1/0.1	638.6/0.1	2407.1/0.1

Table 4.12: Extended results for the decomposition methods for Sarawak gas production problem #1.

# Scenarios	25	49	81	121	169	225
NGBD	6.2/0.1	813.5/0.1	2274.3/0.1	3839.0/0.1	8254.4/0.1	-/0.2
LR	2407.9/0.1	-/0.7	-/100	/	/	/

Table 4.13: Comparison of the different solvers on Sarawak gas production problem #2. This case study includes 38 binary complicating variables, 93 continuous recourse variables per scenario, 76 first-stage constraints, 205 second-stage constraints per scenario, and 34 bilinear terms per scenario.

# Scenarios	1	2	4	8	9	16	25	27	36
ANTIGONE 1.1	0.4/0.1	3.2/0.1	3.9/0.1	-/0.7	-/0.6	-/0.6	-/0.5	-/1.2	-/0.4
BARON 16.3.4	1.1/0.1	722.5/0.1	6093.1/0.1	-/2.7	-/1.1	-/0.9	-/1.1	-/1.4	-/0.4
COUENNE 0.5	37.7/0.1	-/2.5	-/1.8	-/2.3	-/0.6	-/1.1	-/0.7	-/2.9	-/0.4
SCIP 3.2	0.4/0.1	11.7/0.1	14.4/0.1	-/0.2	2198.7/0.1	-/0.3	3734.5/0.1	-/0.8	-/1.2
NGBD	1.3/0.1	2.5/0.1	3.0/0.1	9.0/0.1	7.4/0.1	13.3/0.1	17.0/0.1	22.5/0.1	27.7/0.1
LR	0.8/0.1	15.9/0.1	100.0/0.1	812.6/0.1	254.7/0.1	379.8/0.1	2018.3/0.1	999.5/0.1	3694.5/0.1

Table 4.14: Extended results for the decomposition methods for Sarawak gas production problem #2.

# Scenarios	49	81	169	289	441	625	841	1089
NGBD	43.6/0.1	72.9/0.1	171.5/0.1	330.7/0.1	538.5/0.1	843.9/0.1	1255.1/0.1	1677.1/0.1
LR	1543.2/0.1	3583.4/0.1	t/29.7	/	/	/	/	/

Table 4.15: Comparison of the different solvers on Sarawak gas production problem #3. This case study includes 38 binary complicating variables, 93 continuous recourse variables per scenario, 76 first-stage constraints, 205 second-stage constraints per scenario, and 34 bilinear terms per scenario.

# Scenarios	1	4	9	16	25	36	49
ANTIGONE 1.1	0.4/0.1	3.2/0.1	13.5/0.1	62.1/0.1	313.5/0.1	-/0.5	-/0.9
BARON 16.3.4	1.4/0.1	-/2.5	-/1.7	t/1.3	-/1.1	-/1.0	-/0.9
COUENNE 0.5	3.4/0.1	-/2.5	-/1.8	-/1.3	-/4.1	-/1.0	-/0.9
SCIP 3.2	0.9/0.1	10.9/0.1	-/0.3	-/0.4	-/0.4	-/0.6	-/0.7
NGBD	1.0/0.1	2.0/0.1	3.7/0.1	5.6/0.1	8.5/0.1	10.4/0.1	14.7/0.1
LR	0.9/0.1	278.2/0.1	815.5/0.1	1459.1/0.1	2828.1/0.1	2026.1/0.1	3283.3/0.1

Table 4.16: Extended results for the decomposition methods for the Sarawak gas production problem #3.

# Scenarios	81	169	289	441	625	841	1089
NGBD	22.5/0.1	45.8/0.1	75.7/0.1	119.9/0.1	161.8/0.1	219.5/0.1	289.2/0.1
LR	-/100	/	/	/	/	/	/

Pump network problems

Table 4.17: Comparison of the different solvers on pump network problem #1. This case study includes 18 binary complicating variables, 38 continuous recourse variables per scenario, 33 first-stage constraints, 95 second-stage constraints per scenario, 6 bilinear terms per scenario, and 6 univariate signomial terms per scenario.

# Scenarios	1	4	8	9	16	25	27	36	49
ANTIGONE 1.1	0.6/0.1	6.8/0.1	32.4/0.1	75.4/0.1	129.5/0.1	307.6/0.1	273.1/0.1	-/1.7	-/9.8
BARON 16.3.4	0.5/0.1	2.8/0.1	9.3/0.1	7.5/0.1	22.3/0.1	61.0/0.1	57.8/0.1	-/1.3	-/1.8
COUENNE 0.5	0.9/0.1	6.4/0.1	23.7/0.1	47.4/0.1	83.5/0.1	-,wd/-	-/2.1	-/1.3	-/1.5
SCIP 3.2	0.2/0.1	0.3/0.1	0.9/0.1	1.1/0.1	1.6/0.1	2.8/0.1	5.2/0.1	-/0.7	-/1.0
NGBD	5.7/0.1	15.6/0.1	32.2/0.1	36.8/0.1	58.1/0.1	61.8/0.1	89.6/0.1	97.3/0.1	131.4/0.1
LR	1.0/0.1	28.1/0.1	65.0/0.1	83.5/0.1	95.1/0.1	221.9/0.1	204.8/0.1	841.4/0.1	1397.6/0.1

Table 4.18: Extended results for the decomposition methods for pump network problem #1.

# Scenarios	64	81	121	169	289	441	625
NGBD	161.6/0.1	202.5/0.1	306.1/0.1	430.2/0.1	996.0/0.1	2998.7/0.1	6934.3/0.1
LR	1287.1/0.1	2288.2/0.1	4338.6/0.1	4959.4/0.1	8624.4/0.1	/	/

Table 4.19: Comparison of the different solvers on pump network problem #2. This case study includes 27 binary complicating variables, 57 continuous recourse variables per scenario, 49 first-stage constraints, 142 second-stage constraints per scenario, 9 bilinear terms per scenario, and 9 univariate signomial terms per scenario.

# Scenarios	1	2	4	8	9	16	25
ANTIGONE 1.1	4.2/0.1	31.2/0.1	3637.5/0.1	-/18.7	-/20.3	-/57.0	-/52.7
BARON 16.3.4	3.8/0.1	20.2/0.1	386.0/0.1	-/9.0	-/30.9	-/47.0	-/50.7
COUENNE 0.5	7.0/0.1	411.7/0.1	10000/0.1	10000/0.1	10000/0.1	-/26.0	-/30.5
SCIP 3.2	2.3/0.1	7.2/0.1	29.4/0.1	1075.1/0.1	286.2/0.1	-/0.4	-/6.4
NGBD	609.5/0.1	790.0/0.1	1020.6/0.1	1810.9/0.1	1685.1/0.1	2430.3/0.1	3351.0/0.1
LR	8.2/0.1	139.2/0.1	294.4/0.1	935.0/0.1	1104.4/0.1	2430.4/0.1	-/2.5

Table 4.20: Extended results for the decomposition methods for pump network problem #2.

# Scenarios	36	49	64	81
NGBD	4668.9/0.1	6274.9/0.1	7830.7/0.1	9469.8/0.1
LR	6373.3/0.1	-/2.4	-/10.0	-/-

Software reliability problems

Table 4.21: Comparison of the different solvers on software reliability problem #1. This case study includes 8 binary complicating variables, 3 continuous recourse variables per scenario, 4 first-stage constraints, 3 second-stage constraints per scenario, 1 trilinear term per scenario, and 3 logarithmic terms per scenario.

# Scenarios	1	27	125	343	729	1331	2197	3375
ANTIGONE 1.1	0.1/0.1	30.1/0.1	-/1.9	-/3.4	-/0.2	-/2.0	-/2.6	-/-
BARON 16.3.4	0.1/0.1	0.3/0.1	9.9/0.1	68.0/0.1	274.2/0.1	1046.8/0.1	2593.7/0.1	8615.8/0.1
COUENNE 0.5	0.1/0.1	wd/-	35.2/0.1	394.1/0.1	2456.7/0.1	-/7.6	-/11.2	-/11.2
SCIP 3.2	0.1/0.1	0.4/0.1	16.6/0.1	334.4/0.1	931.0/0.1	1062.9/0.1	977.4/0.1	8691.2/0.1
NGBD	0.2/0.1	2.4/0.1	10.0/0.1	27.1/0.1	61.5/0.1	119.4/0.1	228.4/0.1	325.9/0.1
LR	0.1/0.1	16.2/0.1	15.1/0.1	24.3/0.1	40.4/0.1	t/-	t/-	t/-

Table 4.22: Comparison of the different solvers on software reliability problem #2. This case study includes 18 binary complicating variables, 5 continuous recourse variables per scenario, 6 first-stage constraints, 5 second-stage constraints per scenario, 1 multivariate signomial term per scenario, and 5 logarithmic terms per scenario.

# Scenarios	1	27	125	343	729	1331	2197
ANTIGONE 1.1	5.4/0.1	-/26.0	-/35.9	-/37.3	-/37.3	-/37.4	-/-
BARON 16.3.4	0.4/0.1	14.5/0.1	102.0/0.1	625.7/0.1	4162.0/0.1	-/5.2	-/32.4
COUENNE 0.5	0.5/0.1	14.4/0.1	352.0/0.1	3762.4/0.1	-/38.1	-/50.0	-/50.0
SCIP 3.2	0.2/0.1	5.7/0.1	64.1/0.1	170.7/0.1	451.2/0.1	1322.6/0.1	t/8.0
NGBD	1242.3/0.1	1661.0/0.1	3339.9/0.1	7177.5/0.1	-/5.3	-/7.3	-/13.3
LR	10.4/0.1	t/0.9	-/8.2	-/10.3	t/-	t/-	t/-

Tank sizing Problems

Continuous tank sizing

Table 4.23: Comparison of the different solvers on the continuous tank sizing problem. This case study includes 3 continuous complicating variables, 9 binary recourse variables per scenario, 44 continuous recourse variables per scenario, 76 second-stage constraints per scenario, 32 bilinear terms per scenario, and 3 univariate signomial terms per scenario.

# Scenarios	1	2	3	4	5
ANTIGONE 1.1	24.4/0.1	-/13.9	-/7.4	-/14.3	-/14.3
BARON 16.3.4	3.6/0.1	t/23.1	-/10.6	-/59.8	-/49.8
COUENNE 0.5	i/-	-/12.8	-/25.6	-/35.4	-/39.9
SCIP 3.2	2.0/0.1	1201.7/0.1	-/0.2	t/16.6	t/10.5
MLR	39.7/0.1	1265.2/0.1	1340.7/0.1	2416.3/0.1	2451.2/0.1
LR	39.9/0.1	-/0.14	2696.6/0.1	3477.1/0.1	4354.8/0.1

Table 4.24: Extended results for the decomposition methods for the continuous tank sizing problem.

# Scenarios	7	9	13	17	21	25
MLR	3228.9/0.1	4274.4/0.1	6256.9/0.1	8289.2/0.1	8230.6/0.1	-/0.12
LR	-/0.11	-/0.12	-/0.12	-/0.14	-/0.14	-/0.14

Discrete tank sizing

Table 4.25: Comparison of the different solvers on discrete tank sizing problem #1. This case study includes 48 binary complicating variables, 9 binary recourse variables per scenario, 44 continuous recourse variables per scenario, 3 first-stage constraints, 76 second-stage constraints per scenario, and 32 bilinear terms per scenario.

# Scenarios	1	2	3	4	5
ANTIGONE 1.1	37.7/0.1	-/9.0	-/6.4	-/15.2	-/14.0
BARON 16.3.4	12.0/0.1	-/32.0	-/38.4	-/64.5	-/66.0
COUENNE 0.5	1351.5/0.1	-/44.7	i/-	-/65.7	-/65.1
SCIP 3.2	4.6/0.1	286.4/0.1	1401.3/0.1	t/10.0	t/19.2
NGBD	1106.4/0.1	-/80.4	-/84.3	-/86.9	-/87.3
LR	58.9/0.1	75.1/0.1	140.5/0.1	175.7/0.1	233.6/0.1

Table 4.26: Extended results for the decomposition methods for discrete tank sizing problem #1.

# Scenarios	10	20	30	40	50
NGBD	/	/	/	/	/
LR	413.2/0.1	984.8/0.1	1421.2/0.1	1890.7/0.1	2308.6/0.1

Table 4.27: Comparison of the different solvers on discrete tank sizing problem #2. This case study includes 64 binary complicating variables, 16 binary recourse variables per scenario, 73 continuous recourse variables per scenario, 4 first-stage constraints, 128 second-stage constraints per scenario, and 74 bilinear terms per scenario.

# Scenarios	1	2	3	4	5
ANTIGONE 1.1	1700.0/0.1	-/20.9	-/24.1	-/59.9	-/79.4
BARON 16.3.4	93.0/0.1	-/65.8	-/86.4	-/91.6	-/91.9
COUENNE 0.5	i/-	-/62.9	i/-	i/-	i/-
SCIP 3.2	85.4/0.1	t/3.6	t/17.7	t/60.4	t/78.8
NGBD	-/86.0	-/88.0	-/90.9	-/91.5	-/91.7
LR	-/0.13	-/100	-/100	-/100	-/100

Refinery model problems

Discrete refinery model

Table 4.28: Comparison of the different solvers on the discrete refinery model. This case study includes 100 binary complicating variables, 122 continuous recourse variables per scenario, 101 first-stage constraints, 111 second-stage constraints per scenario, and 26 bilinear terms per scenario.

# Scenarios	1	5	10	20	40	120	500
ANTIGONE 1.1	1.6/0.1	24.0/0.1	151.4/0.1	853.1/0.1	6590.6/0.1	-/0.9	/
BARON 16.3.4	1.8/0.1	567.2/0.1	1830.0/0.1	-/0.5	-/0.8	-/0.9	/
COUENNE 0.5	3770.0/0.1	-/1.4	-/0.8	-/1.1	-/-	-/-	/
SCIP 3.2	4.2/0.1	4998.2/0.1	2752.8/0.1	-/0.2	-/0.4	-/0.4	/
NGBD	2.8/0.1	18.1/0.1	30.9/0.1	138.2/0.1	161.4/0.1	661.5/0.1	2574.2/0.1
LR	2.3/0.1	-/15.4	-/13.4	-/17.3	-/18.1	-/-	/

Continuous refinery model

Table 4.29: Comparison of the different solvers on the continuous refinery model. This case study includes 10 binary complicating variables, 10 continuous complicating variables, 122 continuous recourse variables per scenario, 21 first-stage constraints, 111 second-stage constraints per scenario, and 26 bilinear terms per scenario.

# Scenarios	1	5	10	20	40	120
ANTIGONE 1.1	0.1/0.1	3.0/0.1	14.2/0.1	166.3/0.1	-/0.2	-/0.7
BARON 16.3.4	0.5/0.1	30.5/0.1	-/0.4	-/0.4	5448.2/0.1	-/0.7
COUENNE 0.5	2.6/0.1	53.4/0.1	785.3/0.1	-/0.2	-/0.7	-/-
SCIP 3.2	0.5/0.1	1.1/0.1	7.0/0.1	-/0.2	-/0.2	-/36.0
MLR	0.4/0.1	-/9.3	-/9.5	-/10.4	-/11.2	-/-
LR	0.4/0.1	-/9.6	-/9.6	-/11.1	-/12.0	-/-

Trim loss minimization problems

Table 4.30: Comparison of the different solvers on trim loss minimization problem #1. This case study includes 4 binary complicating variables, 16 integer complicating variables, 20 integer recourse variables per scenario, 12 first-stage constraints, 40 second-stage constraints per scenario, and 48 bilinear terms per scenario.

# Scenarios	1	5	10	15	20
ANTIGONE 1.1	0.1/0.1	-/1.1	-/7.9	-/17.9	-/9.2
BARON 16.3.4	1.7/0.1	-/1.1	-/13.0	-/12.4	-/15.8
COUENNE 0.5	-/4.4	-/13.2	-/13.0	-/13.3	-/14.0
SCIP 3.2	0.1/0.1	37.4/0.1	307.1/0.1	881.3/0.1	t/2.1
NGBD	-/-	-/-	-/-	-/-	-/-
LR	0.8/0.1	-/2.3	-/2.0	-/3.4	-/2.6

Table 4.31: Comparison of the different solvers on trim loss minimization problem #2. This case study includes 6 binary complicating variables, 36 integer complicating variables, 42 integer recourse variables per scenario, 18 first-stage constraints, 72 second-stage constraints per scenario, and 108 bilinear terms per scenario.

# Scenarios	1	3	5	7	9
ANTIGONE 1.1	0.9/0.1	-/4.8	-/5.9	-/2.6	-/4.4
BARON 16.3.4	14.2/0.1	-/6.0	-/8.2	-/7.2	-/29.7
COUENNE 0.5	-/3.3	-/11.2	-/11.4	-/9.0	-/19.8
SCIP 3.2	0.3/0.1	t/2.9	t/3.6	t/1.1	-/5.4
NGBD	-/-	-/-	-/-	-/-	-/-
LR	3.2/0.1	-/7.0	-/7.4	-/6.1	-/9.0

Knapsack problem

Table 4.32: Comparison of the different solvers on the knapsack problem. This case study includes 240 binary complicating variables, 120 binary recourse variables per scenario, 50 first-stage constraints, and 5 second-stage constraints per scenario.

# Scenarios	1	2	3	4	5	10	15	20
CBC 2.9	343.6/0.1	782.0/0.1	-/0.3	-/0.2	-/0.4	-/0.6	-/1.0	-/0.6
CPLEX 12.6.3.0	134.6/0.1	280.5/0.1	1968.7/0.1	1718.3/0.1	t/0.12	t/0.14	t/0.2	t/0.2
SCIP 3.2	276.8/0.1	268.3/0.1	8661.7/0.1	4974.5/0.1	t/0.2	t/0.2	t/0.2	t/0.2
NGBD	271.2/0.1	179.1/0.1	1166.5/0.1	753.6/0.1	1087.1/0.1	607.1/0.1	405.1/0.1	115.4/0.1
LR	181.6/0.1	-/0.2	-/0.6	-/0.4	-/0.6	-/0.7	-/0.8	-/0.6

Table 4.33: Extended results for the decomposition methods for the knapsack problem.

# Scenarios	40	80	120	200	400	800
NGBD	1152.4/0.1	694.0/0.1	706.2/0.1	697.9/0.1	87.1/0.1	180.3/0.1

Chapter 5

The cluster problem in constrained global optimization

Deterministic branch-and-bound algorithms for continuous global optimization often visit a large number of boxes in the neighborhood of a global minimizer, resulting in the so-called cluster problem [68]. This chapter extends previous analyses of the cluster problem in unconstrained global optimization [68, 238] to the constrained setting based on the notion of convergence order for convex relaxation-based lower bounding schemes developed in Chapter 6 (while the only external material that this chapter significantly relies on is the background provided in Chapter 2, we feel that the results of this chapter will be most appreciated by the reader, if at all, when they are read while keeping the results of Chapter 6 in mind and vice versa). The material in this chapter has been published as the article [108].

5.1 Introduction

One of the key issues faced by deterministic branch-and-bound algorithms for continuous global optimization [101] is the so-called cluster problem, where a large number of boxes may be visited by the algorithm in the vicinity of a global minimizer [68, 178, 238]. Du and Kearfott [68, 116] were the first to analyze this phenomenon in the context of interval branch-and-bound algorithms for unconstrained global optimization. They established that the accuracy with which the bounding scheme estimates the range of the objective function, as determined by the notion of convergence order in Definition 2.3.34, dictates the extent of the cluster problem. Furthermore, they determined that, in the worst case, at least second-order convergence of the bounding scheme is required to mitigate 'clustering' [68]. Next, Neumaier [178, Section 15] provided a similar analysis and concluded that even second-order convergence of the bounding scheme might, in the worst case, result in an exponential number of boxes in the vicinity of an unconstrained global minimizer. In addition, Neumaier claimed that a similar situation holds in a reduced manifold for the constrained case [178, Section 15].

Recently, Wechsung et al. [238] provided a refined analysis of Neumaier's argument for unconstrained global optimization which corroborated the previous analyses. In addition, they showed that the number of boxes visited in the vicinity of a global minimizer may scale differently depending on the convergence order prefactor. As a result, second-order convergent bounding schemes with small-enough prefactors may altogether eliminate the cluster problem, while second-order convergent bounding schemes with large-enough prefactors may result in an exponential number of boxes being visited. Also note the analysis by Wechsung [237, Section 2.3] that shows first-order convergence of the bounding scheme may be sufficient to mitigate the cluster problem in unconstrained optimization when the optimizer sits at a point of nondifferentiability of the objective function.

As highlighted above, the convergence order of the bounding scheme plays a key role in the analysis of the cluster problem. This concept, which is based on the rate at which the notion of excess width from interval extensions [172] shrinks to zero, compares the rate of convergence of an estimated range of a function to its true range. Bompadre and Mitsos [38] developed the notions of Hausdorff and pointwise convergence rates of bounding schemes, and established sharp rules for the propagation of convergence orders of bounding schemes constructed using McCormick's composition rules [154]. In addition, Bompadre and Mitsos [38] demonstrated second-order pointwise convergence of schemes of convex and concave envelopes of twice continuously differentiable functions, second-order pointwise convergence of schemes of αBB relaxations [4], and provided a conservative estimate of the prefactor of αBB relaxation schemes for the case of constant α. Scholz [205] demonstrated second-order convergence of centered forms (also see, for instance, the article by Krawczyk and Nickel [127]). Bompadre and coworkers [39] established sharp rules for the propagation of convergence orders of Taylor and McCormick-Taylor models. Najman and Mitsos [174] established sharp rules for the propagation of convergence orders of the multivariate McCormick relaxations developed in [173, 227]. Finally, Khan and coworkers [124] developed a continuously differentiable variant of McCormick relaxations [154, 173, 227], and established second-order pointwise convergence of schemes of the differentiable McCormick relaxations for twice continuously differentiable functions. The above literature not only helps develop bounding schemes for unconstrained optimization with the requisite convergence order, but also provides conservative estimates for the convergence order prefactor (see Definition 2.3.34). Also note the related definition for the rate of convergence of (lower) bounding schemes for geometric branch-and-bound methods provided by Schöbel and Scholz [203].

This chapter provides an analysis of the cluster problem for constrained global optimization. It is shown that clustering can occur both on feasible and infeasible regions in the neighborhood of a global minimizer. Akin to the case of unconstrained optimization, both the convergence order of a lower bounding scheme and its corresponding prefactor (see Definition 5.2.3) may be crucial towards tackling the cluster problem; however, in contrast to the case of unconstrained optimization, where at least second-order convergent schemes of relaxations are required to mitigate the cluster problem when the minimizer sits at a point of differentiability of the objective function, it is shown that first-order convergent lower bounding schemes with small-enough prefactors may eliminate the cluster problem under certain conditions. Additionally, conditions under which second-order convergence of the lower bounding scheme may be sufficient to mitigate clustering are developed. The analysis in this chapter reduces to previous analyses of the cluster problem for unconstrained optimization under suitable assumptions.

This chapter assumes that boxes can be placed such that global minimizers are always in their relative interior; otherwise, an exponential number of boxes can contain global minimizers. Techniques such as epsilon-inflation [153] or back-boxing [178, 229] can potentially be used to place boxes with global minimizers in their relative interior.

This chapter is organized as follows. Section 5.2 provides the problem formulation, describes the notions of convergence used in this chapter, and sets up the framework for analyzing the cluster problem in Section 5.3. Section 5.3.1 analyzes the cluster problem on the set of nearly-optimal feasible points in a neighborhood of a global minimizer and determines conditions under which first-order and second-order convergent bounding schemes may be sufficient to mitigate clustering in such neighborhoods. Section 5.3.2 analyzes the cluster problem on the set of nearly-feasible points in a neighborhood of a global minimizer that have a 'good-enough' objective function value, and develops conditions under which first-order and second-order convergent bounding schemes may be sufficient to mitigate clustering in such neighborhoods. Finally, Section 5.4 lists the conclusions of this chapter.

5.2 Problem formulation and background

In this chapter, we consider the nonlinear programming formulation

$$\begin{aligned}
\min_{\mathbf{x}} \ & f(\mathbf{x}) \\
\text{s.t. } & \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \\
& \mathbf{h}(\mathbf{x}) = \mathbf{0}, \\
& \mathbf{x} \in X,
\end{aligned} \tag{P}$$

where $X \subset \mathbb{R}^{n_x}$ is a nonempty open bounded convex set, and the functions $f : X \to \mathbb{R}$, $\mathbf{g} : X \to \mathbb{R}^{m_I}$, and $\mathbf{h} : X \to \mathbb{R}^{m_E}$ are continuous on X. The following assumptions are enforced throughout this chapter.

Assumption 5.2.1. The constraints define a nonempty compact set

$$\{\mathbf{x} \in X : \mathbf{g}(\mathbf{x}) \leq \mathbf{0},\ \mathbf{h}(\mathbf{x}) = \mathbf{0}\} \subset X.$$

Assumption 5.2.2. Let $\mathbf{x}^* \in X$ be a global minimum for Problem (P), and assume that the branch-and-bound algorithm has found the upper bound $UBD = f(\mathbf{x}^*)$ sufficiently early on. Let ε be the termination tolerance for the branch-and-bound algorithm, and suppose the algorithm fathoms node k when $UBD - LBD_k \leq \varepsilon$, where LBD_k is the lower bound on node k.

When Assumption 5.2.1 is enforced, Problem (P) attains its optimal solution on X by virtue of the assumption that f is continuous on X. Note that the assumption that X is an open set is made purely for ease of exposition, particularly when differentiability assumptions on the functions in Problem (P) are made, and is not practically implementable in general. As a result, we implicitly assume throughout this chapter that finite bounds on the variables (which define an interval in the interior of X) are available for use in a branch-and-bound setting.

Assumption 5.2.2 essentially assumes that the convergence of the overall lower bound is the limiting factor for the convergence of the branch-and-bound algorithm. This is usually a reasonable assumption in the context of branch-and-bound algorithms for global optimization where most of the effort is typically spent in proving ε-optimality of feasible solutions found using (heuristic) local optimization-based techniques. The cluster problem analysis in this chapter is asymptotic in ε in general; we provide conservative estimates of the worst-case number of boxes visited by the branch-and-bound algorithm in nearly-optimal and nearly-feasible neighborhoods of global minimizers for some sufficiently small $\varepsilon > 0$. The conservatism of the above estimates decreases as $\varepsilon \to 0$. The asymptotic nature of our analysis with respect to ε is not only a result of considering the local behavior of the objective function in the vicinity of a global minimizer (which is also a limitation of the analyses of the cluster problem in unconstrained optimization [68, 178, 237, 238]), but is also a consequence of considering the local behavior of the constraints (and, therefore, the feasible region) in the vicinity of a global minimizer. In practice, values of ε for which the analysis of the cluster problem provides a reasonable overestimate of the number of boxes visited can be much larger than the machine precision (on the order of 10^{-1}). This is evidenced by the examples in Section 5.3. Also note that the fathoming criterion for the branch-and-bound algorithm in this chapter is different from the one considered by Wechsung et al. [238], who assume that node k is fathomed only when $LBD_k > UBD$; however, the worst-case estimates of the number of boxes visited by the branch-and-bound algorithm are not affected by this difference in our assumptions.

Throughout this chapter, we will use $\mathbf{x}^*$ to denote a global minimizer of Problem (P), and Z^{C} to denote the relative complement of a set $Z \subset \mathbb{R}^{n_x}$ with respect to X. The reader is directed to Chapter 2 for other notational definitions, and for the background definitions used in this chapter (in particular, we will use Definitions 2.2.1, 2.2.4, 2.2.5, 2.3.3, 2.3.5, 2.3.8, 2.3.13, 2.3.14, 2.3.23, 2.3.28, 2.3.31, and 2.3.34). The following definition seeks to extend the notion of convergence order of a bounding scheme [38, 39, 238] to constrained problems (also see Definition 6.3.12). Conditions under which specific lower bounding schemes are guaranteed to exhibit a certain convergence order are presented in Chapter 6.

Definition 5.2.3. [Convergence Order of a Lower Bounding Scheme] Consider Problem (P). For any $Z \in \mathbb{I}X$, let $\mathcal{F}(Z) = \{\mathbf{x} \in Z : \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\}$ denote the feasible set of Problem (P) with $\mathbf{x}$ restricted to Z.

Let $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}X}$ and $(\mathbf{g}_Z^{\mathrm{cv}})_{Z\in\mathbb{I}X}$ denote continuous schemes of convex relaxations of f and $\mathbf{g}$, respectively, in X, and let $(\mathbf{h}_Z^{\mathrm{cv}},\mathbf{h}_Z^{\mathrm{cc}})_{Z\in\mathbb{I}X}$ denote a continuous scheme of relaxations of $\mathbf{h}$ in X. For any $Z\in\mathbb{I}X$, let $\mathcal{F}^{\mathrm{cv}}(Z) = \{\mathbf{x}\in Z : \mathbf{g}_Z^{\mathrm{cv}}(\mathbf{x})\le\mathbf{0}, \mathbf{h}_Z^{\mathrm{cv}}(\mathbf{x})\le\mathbf{0}, \mathbf{h}_Z^{\mathrm{cc}}(\mathbf{x})\ge\mathbf{0}\}$ denote the feasible set of the convex relaxation-based lower bounding scheme. The convex relaxation-based lower bounding scheme is said to have convergence of order $\beta>0$ at

1. a feasible point $\mathbf{x}\in X$ if there exists $\tau\ge 0$ such that for every $Z\in\mathbb{I}X$ with $\mathbf{x}\in Z$,

$$\min_{\mathbf{z}\in\mathcal{F}(Z)} f(\mathbf{z}) - \min_{\mathbf{z}\in\mathcal{F}^{\mathrm{cv}}(Z)} f_Z^{\mathrm{cv}}(\mathbf{z}) \le \tau w(Z)^{\beta}.$$

2. an infeasible point $\mathbf{x}\in X$ if there exists $\bar{\tau}\ge 0$ such that for every $Z\in\mathbb{I}X$ with $\mathbf{x}\in Z$,

$$d\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(Z), \mathbb{R}_-^{m_I}\times\{\mathbf{0}\}\right) - d\left(\mathcal{I}_C(Z), \mathbb{R}_-^{m_I}\times\{\mathbf{0}\}\right) \le \bar{\tau} w(Z)^{\beta},$$

where $\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(Z)$ denotes the image of Z under the vector-valued function $\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}$, and $\mathcal{I}_C(Z)$ is defined by

$$(\mathcal{I}_C(Z))_{Z\in\mathbb{I}X} := \big(\{(\mathbf{v},\mathbf{w})\in\mathbb{R}^{m_I}\times\mathbb{R}^{m_E} : \mathbf{v}=\mathbf{g}_Z^{\mathrm{cv}}(\mathbf{z}), \mathbf{h}_Z^{\mathrm{cv}}(\mathbf{z})\le\mathbf{w}\le\mathbf{h}_Z^{\mathrm{cc}}(\mathbf{z}) \\ \text{for some } \mathbf{z}\in Z\}\big)_{Z\in\mathbb{I}X}.$$

The scheme of lower bounding problems is said to have convergence of order $\beta>0$ on X if it has convergence of order (at least) β at each $\mathbf{x}\in X$, with the constants τ and $\bar{\tau}$ independent of $\mathbf{x}$.

Definition 5.2.3 is motivated by the requirements of a lower bounding scheme to fathom feasible and infeasible regions in a branch-and-bound procedure [101]. On nested sequences of intervals converging to a feasible point of Problem (P), we require that the corresponding sequences of lower bounds converge rapidly to the corresponding sequences of minimum objective values. On the other hand, on nested sequences of intervals converging to an infeasible point of Problem (P), we require that the corresponding sequences of lower bounding problems rapidly detect the (eventual) infeasibility of the corresponding sequences of intervals for Problem (P). The latter requirement is enforced by requiring that the measures of infeasibility of the corresponding lower bounding problems, as determined by the distance function d (see Definition 2.3.5), converge rapidly to the measures of infeasibility of the corresponding restricted Problems (P). Note that some intervals that only contain infeasible points may also potentially be fathomed by value dominance if the lower bounds on those intervals obtained by solving the corresponding relaxation-based lower bounding problems is greater than or equal to $UBD - \varepsilon$. This possibility in considered later in this section (see, for instance, Lemma 5.2.6) and in Section 5.3.2.

The following lemmata detail worst-case conditions under which nodes containing a global minimum and infeasible points are fathomed.

Lemma 5.2.4. [Fathoming Nodes Containing Global Minimizers] Let $X^* \in \mathbb{I}X$, with $\mathbf{x}^* \in X^*$, correspond to the domain of node k^* in the branch-and-bound tree. Suppose the convex relaxation-based lower bounding scheme has convergence of order $\beta^* > 0$ at $\mathbf{x}^*$ with a prefactor $\tau^* > 0$ (see Definition 5.2.3). For node k^* to be fathomed, we require, in that worst case, that

$$w(X^*) \leq \left(\frac{\varepsilon}{\tau^*}\right)^{\frac{1}{\beta^*}}.$$

Proof. The condition for node k^* to be fathomed by value dominance is $UBD - LBD_{k^*} = f(\mathbf{x}^*) - LBD_{k^*} \leq \varepsilon$. Since we are concerned about convergence at the feasible point $\mathbf{x}^* \in X$, we have from Definition 5.2.3 that

$$\begin{aligned} \min_{\mathbf{z} \in \mathcal{F}(X^*)} f(\mathbf{z}) - \min_{\mathbf{z} \in \mathcal{F}^{\text{cv}}(X^*)} f^{\text{cv}}_{X^*}(\mathbf{z}) &\leq \tau^* w(X^*)^{\beta^*} \\ \implies LBD_{k^*} = \min_{\mathbf{z} \in \mathcal{F}^{\text{cv}}(X^*)} f^{\text{cv}}_{X^*}(\mathbf{z}) &\geq f(\mathbf{x}^*) - \tau^* w(X^*)^{\beta^*}. \end{aligned}$$

Therefore, in the worst case, node k^* is fathomed only when

$$LBD_{k^*} \geq f(\mathbf{x}^*) - \tau^* w(X^*)^{\beta^*} \geq f(\mathbf{x}^*) - \varepsilon \iff w(X^*) \leq \left(\frac{\varepsilon}{\tau^*}\right)^{\frac{1}{\beta^*}}. \qquad \square$$

Lemma 5.2.5. [Fathoming Infeasible Nodes by Infeasibility] Let $X^I \in \mathbb{I}X$, with

$$X^I \subset \left\{ \mathbf{x} \in X : d\left(\begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix} (\mathbf{x}), \mathbb{R}^{m_I}_{-} \times \{\mathbf{0}\} \right) > \varepsilon^f \right\}$$

for some $\varepsilon^f > 0$, correspond to the domain of node k^I in the branch-and-bound tree. Suppose the convex relaxation-based lower bounding scheme has convergence of order $\beta^I > 0$ at each $\mathbf{x} \in X^I$ with a prefactor $\tau^I > 0$ that is independent of $\mathbf{x}$ (see Definition 5.2.3). For node k^I to be fathomed by infeasibility, we require, in the worst case, that

$$w(X^I) \leq \left(\frac{\varepsilon^f}{\tau^I}\right)^{\frac{1}{\beta^I}}.$$

Proof. For node k^I to be fathomed by infeasibility, we require that the convex relaxation-based lower bounding problem is infeasible on X^I, i.e., $d\left(\mathcal{I}_C(X^I), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\}\right) > 0$. Since we are concerned about convergence at infeasible points, we have from Definition 5.2.3 that

$$\begin{aligned}
&d\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(X^I), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\}\right) - d\left(\mathcal{I}_C(X^I), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\}\right) \leq \tau^I w(X^I)^{\beta^I}\\
\implies &d\left(\mathcal{I}_C(X^I), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\}\right) \geq d\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(X^I), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\}\right) - \tau^I w(X^I)^{\beta^I}.
\end{aligned}$$

Therefore, node k^I is fathomed, in the worst case, only when

$$\begin{aligned}
&d\left(\mathcal{I}_C(X^I), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\}\right) \geq d\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(X^I), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\}\right) - \tau^I w(X^I)^{\beta^I} > 0\\
\iff &\varepsilon^f - \tau^I w(X^I)^{\beta^I} \geq 0\\
\iff &w(X^I) \leq \left(\frac{\varepsilon^f}{\tau^I}\right)^{\frac{1}{\beta^I}}.
\end{aligned}$$

□

Lemma 5.2.6. [Fathoming Infeasible Nodes by Value Dominance] Let $X^I \in \mathbb{I}X$, with

$$X^I \subset \left\{\mathbf{x} \in X : d\left(\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}(\mathbf{x}), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\}\right) > 0\right\},$$

correspond to the domain of node k^I in the branch-and-bound tree. Suppose $\forall \mathbf{x} \in X^I$, $f(\mathbf{x}) \geq f(\mathbf{x}^*)$. Furthermore, suppose the scheme $(f_Z^{\mathrm{cv}})_{Z \in \mathbb{I}X}$ has convergence of order $\beta^f > 0$

at each $\mathbf{x} \in X^I$ with a prefactor $\tau^f > 0$ that is independent of $\mathbf{x}$ (see Definition 2.3.34). If

$$w(X^I) \leq \left(\frac{\varepsilon}{\tau^f}\right)^{\frac{1}{\beta^f}},$$

then node k^I will be fathomed.

Proof. A sufficient condition for node k^I to be fathomed is

$$\min_{\mathbf{z} \in \mathcal{F}^{\mathrm{cv}}(X^I)} f^{\mathrm{cv}}_{X^I}(\mathbf{z}) \geq f(\mathbf{x}^*) - \varepsilon.$$

Since $(f^{\mathrm{cv}}_Z)_{Z \in \mathbb{I}X}$ has convergence of order β^f, we have from Definition 2.3.34 that

$$\begin{aligned}
\min_{\mathbf{z} \in X^I} f^{\mathrm{cv}}_{X^I}(\mathbf{z}) &\geq \min_{\mathbf{z} \in X^I} f(\mathbf{z}) - \tau^f w(X^I)^{\beta^f} \\
&\geq \min_{\mathbf{z} \in X^I} f(\mathbf{z}) - \varepsilon \\
&\geq f(\mathbf{x}^*) - \varepsilon,
\end{aligned}$$

where Step 2 uses $w(X^I) \leq \left(\frac{\varepsilon}{\tau^f}\right)^{\frac{1}{\beta^f}}$, and Step 3 uses $f(\mathbf{x}) \geq f(\mathbf{x}^*)$, $\forall \mathbf{x} \in X^I$. Therefore,

$$\min_{\mathbf{z} \in \mathcal{F}^{\mathrm{cv}}(X^I)} f^{\mathrm{cv}}_{X^I}(\mathbf{z}) \geq \min_{\mathbf{z} \in X^I} f^{\mathrm{cv}}_{X^I}(\mathbf{z}) \geq f(\mathbf{x}^*) - \varepsilon.$$

The desired result follows. $\square$

In what follows, we shall partition the set X into distinct regions with the aim of constructing regions that are either relatively easy to fathom (based on Lemmata 5.2.4 to 5.2.6), or are relatively hard to fathom. Suppose the convex relaxation-based lower bounding scheme has convergence of order $\beta^* > 0$ on $\mathcal{F}(X)$ with prefactor $\tau^* > 0$, and convergence of order $\beta^I > 0$ on $(\mathcal{F}(X))^{\mathrm{C}}$ with prefactor $\tau^I > 0$ (note that it is sufficient for the lower bounding scheme to have the requisite convergence orders on some neighborhood of the global minimizers of Problem (P) for our analysis to hold, as will become clear in Section 5.3). Furthermore, suppose the scheme $(f^{\mathrm{cv}}_Z)_{Z \in \mathbb{I}X}$ has convergence of order $\beta^f > 0$ on X with prefactor $\tau^f > 0$. Pick a feasibility tolerance ε^f and an optimality tolerance ε^o

such that

$$\left(\frac{\varepsilon^f}{\tau^I}\right)^{\frac{1}{\beta^I}} = \left(\frac{\varepsilon^o}{\tau^f}\right)^{\frac{1}{\beta^f}} = \left(\frac{\varepsilon}{\tau^*}\right)^{\frac{1}{\beta^*}}, \tag{TOL}$$

and consider the following partition of X:

$$X_1 := \left\{ \mathbf{x} \in X : d\left(\begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix} (\mathbf{x}), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\} \right) > \varepsilon^f \right\},$$

$$X_2 := \left\{ \mathbf{x} \in X : d\left(\begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix} (\mathbf{x}), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\} \right) \in (0, \varepsilon^f] \text{ and } f(\mathbf{x}) - f(\mathbf{x}^*) > \varepsilon^o \right\},$$

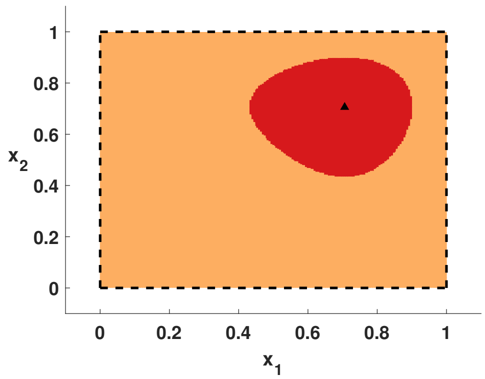
$$X_3 := \left\{ \mathbf{x} \in X : d\left(\begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix} (\mathbf{x}), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\} \right) \in (0, \varepsilon^f] \text{ and } f(\mathbf{x}) - f(\mathbf{x}^*) \leq \varepsilon^o \right\},$$

$$X_4 := \left\{ \mathbf{x} \in X : d\left(\begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix} (\mathbf{x}), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\} \right) = 0 \text{ and } f(\mathbf{x}) - f(\mathbf{x}^*) > \varepsilon \right\}, \text{ and}$$

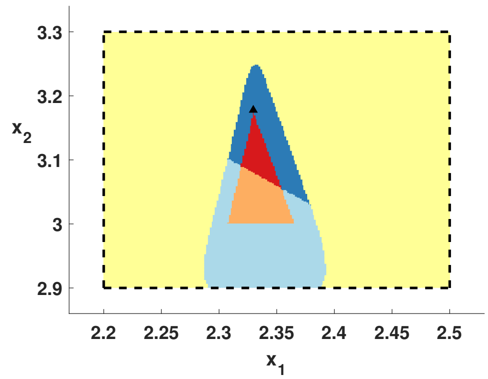
$$X_5 := \left\{ \mathbf{x} \in X : d\left(\begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix} (\mathbf{x}), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\} \right) = 0 \text{ and } f(\mathbf{x}) - f(\mathbf{x}^*) \leq \varepsilon \right\}.$$

The set X_1 corresponds to the set of infeasible points for Problem (P) with the measure of infeasibility greater than ε^f. The set X_2 corresponds to the set of infeasible points for Problem (P) with the measure of infeasibility less than or equal to ε^f and with the objective function value greater than $f(\mathbf{x}^*) + \varepsilon^o$, while the set X_3 corresponds to the set of infeasible points for Problem (P) with the measure of infeasibility less than or equal to ε^f and the objective function value less than or equal to $f(\mathbf{x}^*) + \varepsilon^o$. The set X_4 corresponds to the set of feasible points for Problem (P) with objective value greater than $f(\mathbf{x}^*) + \varepsilon$, while the set X_5 corresponds to the set of feasible points for Problem (P) with objective value less than or equal to $f(\mathbf{x}^*) + \varepsilon$. The sets X_1 through X_5 are illustrated in Figure 5-1 for the three two-dimensional problems presented in Examples 5.2.7 to 5.2.9.

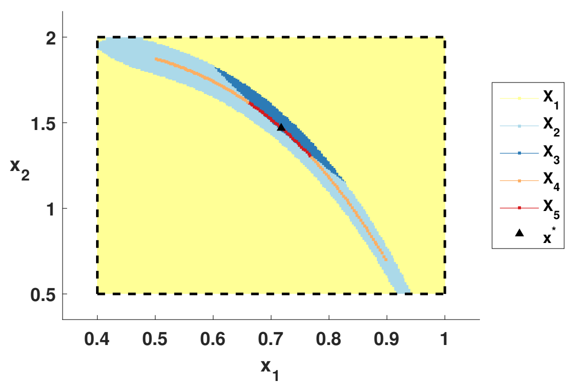
Intuitively, we expect that nodes with domains contained in the sets X_1 and X_2 can be fathomed relatively easily (by infeasibility and value dominance, respectively) compared to nodes with domains contained in the set X_3. Similarly, we expect that nodes with domains contained in the set X_4 can be fathomed relatively easily (by value dominance) compared to nodes with domains contained in the set X_5. This intuition is formalized in Corollary 5.2.10.



(a) Example 5.2.7 (unconstrained)



(b) Example 5.2.8 (inequality-constrained)



(c) Example 5.2.9 (equality-constrained)

Figure 5-1: Plots of the sets X_1 through X_5 for an unconstrained, an inequality-constrained, and an equality-constrained problem. The dashed lines define the sets X, and the filled-in triangles denote the unique global minimizers of the problems on X. All plots use $\varepsilon = \varepsilon^o = \varepsilon^f = 0.1$ for illustration.

Consequently, the extent of clustering is dictated primarily by the number of boxes required to cover the regions X_3 and X_5. Section 5.3 provides conservative estimates of the number of boxes of certain widths that are required to cover X_3 and X_5 under suitable assumptions. As an aside, note that the condition specified by Equation (TOL) is used to roughly enforce that nodes with domains contained in the sets X_1, X_2, and X_4 can, in the worst case, be fathomed using a similar level of effort.

Example 5.2.7. Let $X = (0,1) \times (0,1)$, $m_I = m_E = 0$, and $f(\mathbf{x}) = x_1^4 + x_2^4 - x_1^2 - x_2^2$ with $\mathbf{x}^* = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$. We have:

$$X_1 = X_2 = X_3 = \emptyset,$$

$$X_4 = \left\{\mathbf{x} \in X : x_1^4 + x_2^4 - x_1^2 - x_2^2 > f(\mathbf{x}^*) + \varepsilon\right\}, \text{ and}$$

$$X_5 = \left\{ \mathbf{x} \in X : x_1^4 + x_2^4 - x_1^2 - x_2^2 \leq f(\mathbf{x}^*) + \varepsilon \right\}.$$

The sets X_1 through X_5 are depicted in Figure 5-1a for $\varepsilon = 0.1$.

Example 5.2.8. Let $X = (2.2, 2.5) \times (2.9, 3.3)$, $m_I = 3$, $m_E = 0$, $f(\mathbf{x}) = -x_1 - x_2$, $g_1(\mathbf{x}) = x_2 - 2x_1^4 + 8x_1^3 - 8x_1^2 - 2$, $g_2(\mathbf{x}) = x_2 - 4x_1^4 + 32x_1^3 - 88x_1^2 + 96x_1 - 36$, and $g_3(\mathbf{x}) = 3 - x_2$ with $\mathbf{x}^* \approx (2.33, 3.18)$ (based on Example 4.10 in [81]). We have:

$$\begin{aligned}
X_1 &= \left\{ \mathbf{x} \in X : \sqrt{\sum_{j=1}^{3} (\max\{0, g_j(\mathbf{x})\})^2} > \varepsilon^f \right\}, \\
X_2 &= \left\{ \mathbf{x} \in X : \sqrt{\sum_{j=1}^{3} (\max\{0, g_j(\mathbf{x})\})^2} \in (0, \varepsilon^f],\ -x_1 - x_2 > f(\mathbf{x}^*) + \varepsilon^o \right\}, \\
X_3 &= \left\{ \mathbf{x} \in X : \sqrt{\sum_{j=1}^{3} (\max\{0, g_j(\mathbf{x})\})^2} \in (0, \varepsilon^f],\ -x_1 - x_2 \leq f(\mathbf{x}^*) + \varepsilon^o \right\}, \\
X_4 &= \{ \mathbf{x} \in X : \mathbf{g}(\mathbf{x}) \leq \mathbf{0},\ -x_1 - x_2 > f(\mathbf{x}^*) + \varepsilon \}, \text{ and} \\
X_5 &= \{ \mathbf{x} \in X : \mathbf{g}(\mathbf{x}) \leq \mathbf{0},\ -x_1 - x_2 \leq f(\mathbf{x}^*) + \varepsilon \}.
\end{aligned}$$

The sets X_1 through X_5 are depicted in Figure 5-1b for $\varepsilon = \varepsilon^o = \varepsilon^f = 0.1$.

Example 5.2.9. Let $X = (0.4, 1.0) \times (0.5, 2.0)$, $m_I = 2$, $m_E = 1$, $f(\mathbf{x}) = -12x_1 - 7x_2 + x_2^2$, $g_1(\mathbf{x}) = x_1 - 0.9$, $g_2(\mathbf{x}) = 0.5 - x_1$, and $h(\mathbf{x}) = x_2 + 2x_1^4 - 2$ with $\mathbf{x}^* \approx (0.72, 1.47)$ (based on Example 4.9 in [81]). We have:

$$\begin{aligned}
X_1 &= \left\{ \mathbf{x} \in X : \sqrt{\sum_{j=1}^{2} (\max\{0, g_j(\mathbf{x})\})^2 + |h(\mathbf{x})|^2} > \varepsilon^f \right\}, \\
X_2 &= \left\{ \mathbf{x} \in X : \sqrt{\sum_{j=1}^{2} (\max\{0, g_j(\mathbf{x})\})^2 + |h(\mathbf{x})|^2} \in (0, \varepsilon^f],\ -12x_1 - 7x_2 + x_2^2 > f(\mathbf{x}^*) + \varepsilon^o \right\}, \\
X_3 &= \left\{ \mathbf{x} \in X : \sqrt{\sum_{j=1}^{2} (\max\{0, g_j(\mathbf{x})\})^2 + |h(\mathbf{x})|^2} \in (0, \varepsilon^f],\ -12x_1 - 7x_2 + x_2^2 \leq f(\mathbf{x}^*) + \varepsilon^o \right\}, \\
X_4 &= \left\{ \mathbf{x} \in X : \mathbf{g}(\mathbf{x}) \leq \mathbf{0},\ h(\mathbf{x}) = 0,\ -12x_1 - 7x_2 + x_2^2 > f(\mathbf{x}^*) + \varepsilon \right\}, \text{ and} \\
X_5 &= \left\{ \mathbf{x} \in X : \mathbf{g}(\mathbf{x}) \leq \mathbf{0},\ h(\mathbf{x}) = 0,\ -12x_1 - 7x_2 + x_2^2 \leq f(\mathbf{x}^*) + \varepsilon \right\}.
\end{aligned}$$

The sets X_1 through X_5 are depicted in Figure 5-1c for $\varepsilon = \varepsilon^o = \varepsilon^f = 0.1$.

The following corollary of Lemmata 5.2.4, 5.2.5, and 5.2.6, similar to Lemma 2 in [238], provides sufficient conditions under which nodes with domains contained in X_1, X_2, and X_4 can be fathomed.

Corollary 5.2.10. [Fathoming Nodes Contained in X_1, X_2, and X_4] Let $\delta = \left(\frac{\varepsilon}{\tau^*}\right)^{\frac{1}{\beta^*}}$.

1. Suppose the convex relaxation-based lower bounding scheme has convergence of order $\beta^I > 0$ at each $\mathbf{x} \in X_1$ with a prefactor $\tau^I > 0$ that is independent of $\mathbf{x}$. Consider $\bar{X}_1 \in \mathbb{I}X_1$ corresponding to the domain of node k_1 in the branch-and-bound tree. If $w(\bar{X}_1) \leq \delta$, then node k_1 will be fathomed by infeasibility.

2. Suppose the scheme of convex relaxations $(f_Z^{\mathrm{cv}})_{Z \in \mathbb{I}X}$ has convergence of order $\beta^f > 0$ at each $\mathbf{x} \in X_2$ with a prefactor $\tau^f > 0$ that is independent of $\mathbf{x}$. Consider $\bar{X}_2 \in \mathbb{I}X_2$ corresponding to the domain of node k_2 in the branch-and-bound tree. If $w(\bar{X}_2) \leq \delta$, then node k_2 will be fathomed by value dominance.

3. Suppose the convex relaxation-based lower bounding scheme has convergence of order $\beta^* > 0$ at each $\mathbf{x} \in X_4$ with a prefactor $\tau^* > 0$ that is independent of $\mathbf{x}$. Consider $\bar{X}_4 \in \mathbb{I}X_4$ corresponding to the domain of node k_4 in the branch-and-bound tree. If $w(\bar{X}_4) \leq \delta$, then node k_4 will be fathomed by value dominance.

Corollary 5.2.10 implies that nodes with domains $\bar{X}_1$, $\bar{X}_2$, and $\bar{X}_4$ such that $\bar{X}_1 \in \mathbb{I}X_1$, $\bar{X}_2 \in \mathbb{I}X_2$, and $\bar{X}_4 \in \mathbb{I}X_4$ can be fathomed when or before their widths are δ (in fact, nodes with domains in $\mathbb{I}X_2$ and $\mathbb{I}X_4$ can be fathomed when or before their widths are $\left(\frac{\varepsilon^o+\varepsilon}{\tau^f}\right)^{\frac{1}{\beta^f}}$ and $\left(\frac{2\varepsilon}{\tau^*}\right)^{\frac{1}{\beta^*}}$, respectively). However, nodes $\bar{X}_5 \in \mathbb{I}X_5$ may, in the worst case, need to be covered by boxes of width δ before they are fathomed. Furthermore, nodes $\bar{X}_3 \in \mathbb{I}X_3$ may need to be covered by a large number of boxes depending on the convergence properties of the lower bounding scheme on X_3. The following example presents a case in which clustering may occur on X_3 because the lower bounding scheme does not have a sufficiently-large convergence order at infeasible points.

Example 5.2.11. Let $X = (-2, 2)$, $m_I = 3$, and $m_E = 0$ with $f(x) = x$, $g_1(x) = x^2$, $g_2(x) = x - 1$, and $g_3(x) = -1 - x$. We have $x^* = 0$ (which is the only feasible point). For any $[x^{\mathrm{L}}, x^{\mathrm{U}}] =: Z \in \mathbb{I}X$, let

$$\begin{aligned}
f_Z^{\mathrm{cv}}(x) &= x,\\
g_{1,Z}^{\mathrm{cv}}(x) &= \begin{cases} -(x^{\mathrm{U}} - x^{\mathrm{L}}), & \text{if } 0 \in [x^{\mathrm{L}}, x^{\mathrm{U}}] \\ \min\left(\left(x^{\mathrm{L}}\right)^2, \left(x^{\mathrm{U}}\right)^2\right) - (x^{\mathrm{U}} - x^{\mathrm{L}}), & \text{otherwise} \end{cases},\\
g_{2,Z}^{\mathrm{cv}}(x) &= x - 1,\\
g_{3,Z}^{\mathrm{cv}}(x) &= -1 - x.
\end{aligned}$$

We have $\beta^* = \beta^I = 1$ and β^f arbitarily-large with prefactors τ^*, τ^I, and τ^f, respectively, greater than zero.

Suppose $\varepsilon, \varepsilon^f \ll 1$. Pick $\gamma > 0$ and $\alpha \in (0, \gamma)$ such that $(\gamma + \alpha)^2 = \varepsilon^f$. Let $x^{\mathrm{L}} := -\gamma - \alpha = -\sqrt{\varepsilon^f}$ and $x^{\mathrm{U}} := -\gamma + \alpha < 0$. The width of Z is $w(Z) = 2\alpha$. Note that g_2 and g_3 are feasible on Z; therefore, we need only be concerned with the feasibility of g_1.

We have $\overline{g}_1(Z) = [(\gamma - \alpha)^2, (\gamma + \alpha)^2]$ and $d(\overline{\mathbf{g}}(Z), \mathbb{R}_-^{m_I}) = (\gamma - \alpha)^2$. This implies g_1 is infeasible at each $x \in Z$. Note that $X_3 = [x^{\mathrm{L}}, 0) \cup \left(0, \min\{\varepsilon^o, \sqrt{\varepsilon^f}\}\right]$ (which follows, in part, from each $x \in [x^{\mathrm{L}}, 0)$ being infeasible with $f(x) \leq f(x^*)$ and $d(\{\mathbf{g}(x)\}, \mathbb{R}_-^{m_I}) \leq \varepsilon^f$).

We have $\overline{g}_{1,Z}^{\mathrm{cv}}(Z) = [(\gamma - \alpha)^2 - 2\alpha, (\gamma - \alpha)^2 - 2\alpha]$ and $d(\overline{\mathbf{g}}_Z^{\mathrm{cv}}(Z), \mathbb{R}_-^{m_I}) = \max\{0, (\gamma - \alpha)^2 - 2\alpha\}$. The optimal objective value of the lower bounding problem on Z is $-\gamma - \alpha$ when $d(\overline{\mathbf{g}}_Z^{\mathrm{cv}}(Z), \mathbb{R}_-^{m_I}) = 0$, and is $+\infty$ otherwise. Note that the lower bounding problem is infeasible on Z when $(\gamma - \alpha)^2 - 2\alpha > 0$, which can be achieved by choosing α to be sufficiently-small (and increasing γ accordingly).

The maximum width of the interval Z for which it can be fathomed by infeasibility can be shown to be $w(Z) = 2\alpha^* := 2(1 + \gamma) - 2\sqrt{1 + 2\gamma} = O(\gamma^2) = O(\varepsilon^f)$ (note that $\gamma \ll 1$ because $\varepsilon^f \ll 1$). For $\alpha > \alpha^*$, the interval Z cannot be fathomed by infeasibility and the optimal objective value of the lower bounding problem on Z is $-\gamma - \alpha = -\sqrt{\varepsilon^f} = O(\sqrt{\varepsilon})$. Such an interval Z cannot be fathomed by value dominance either since $\varepsilon \ll 1$.

Therefore, in the worst case, the interval Z can be fathomed only when $w(Z) = O(\gamma^2) = O(\varepsilon^f)$. This causes clustering in the worst case since $w([x^{\mathrm{L}}, 0)) = O(\sqrt{\varepsilon^f})$ and $[x^{\mathrm{L}}, 0) \subset X_3$.

5.3 Analysis of the cluster problem

In this section, conservative estimates for the number of boxes required to cover X_3 and X_5 are provided based on assumptions on Problem (P) (in particular, on its set of global minimizers), and characteristics of the branch-and-bound algorithm.

5.3.1 Estimates for the number of boxes required to cover X_5

This section assumes that Problem (P) has a finite number of global minimizers (which implies each global minimum is a strict local minimum), and ε is small enough that X_5 is guaranteed to be contained in neighborhoods of global minimizers under additional assumptions. An estimate for the number of boxes of width δ required to cover some neighborhood of a minimum $\mathbf{x}^*$ that contains the subset of X_5 around $\mathbf{x}^*$ is provided under suitable assumptions. An estimate for the number of boxes required to cover X_5 can be obtained by summing the above estimates over the set of global minimizers. Throughout this section, we assume that $\mathbf{x}^*$ is a nonisolated feasible point (see Definition 2.3.8); otherwise, $\exists \alpha > 0$ such that $\mathcal{N}^2_\alpha(\mathbf{x}^*) \cap X_5 = \{\mathbf{x}^*\}$, which can be covered using a single box.

Lemma 5.3.1. Consider Problem (P). Suppose $\mathbf{x}^*$ is nonisolated and f is differentiable at $\mathbf{x}^*$. Then $\forall \theta > 0$, $\exists \alpha > 0$ such that

$$\inf_{\left\{\mathbf{d}:\|\mathbf{d}\|_1=1,\, \exists t>0\,:\,(\mathbf{x}^*+t\mathbf{d})\in\mathcal{N}^1_\alpha(\mathbf{x}^*)\cap\mathcal{F}(X)\right\}} \nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} > \min_{\left\{\mathbf{d}:\|\mathbf{d}\|_1=1,\mathbf{d}\in T(\mathbf{x}^*)\right\}} \nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} - \theta.$$

Proof. We proceed by contradiction. Define

$$\begin{aligned} L(\alpha) &:= \inf_{\left\{\mathbf{d}:\|\mathbf{d}\|_1=1,\, \exists t>0\,:\,(\mathbf{x}^*+t\mathbf{d})\in\mathcal{N}^1_\alpha(\mathbf{x}^*)\cap\mathcal{F}(X)\right\}} \nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d}, \\ L^* &:= \min_{\left\{\mathbf{d}:\|\mathbf{d}\|_1=1,\mathbf{d}\in T(\mathbf{x}^*)\right\}} \nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d}, \end{aligned}$$

and note that $L(\alpha)$ is monotonically nonincreasing on $(0, +\infty)$. Suppose $\exists \theta > 0$ such that $\forall \alpha > 0$, we have $L(\alpha) \leq L^* - \theta$. Consider a sequence $\{\alpha_k\} \to 0$ with $\alpha_k > 0$, and a corresponding sequence $\{\mathbf{d}_k\}$ such that

$$\mathbf{d}_k \in \left\{\mathbf{d} : \|\mathbf{d}\|_1 = 1,\, \exists t_k > 0 \,:\, (\mathbf{x}^* + t_k\mathbf{d}) \in \mathcal{N}^1_{\alpha_k}(\mathbf{x}^*) \cap \mathcal{F}(X), \nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} \leq L^* - \frac{\theta}{2}\right\}.$$

The existence of $\mathbf{d}_k$ follows from the assumption that $L(\alpha) \leq L^* - \theta$, $\forall \alpha > 0$. Since $\|\mathbf{d}_k\|_1 = 1$, $\forall k$, we have the existence of $\mathbf{d}^* \in \mathbb{R}^{n_x}$ with $\mathbf{d}^* = \lim_{k_q \to \infty} \mathbf{d}_{k_q}$ and $\|\mathbf{d}^*\|_1 = 1$ for some convergent subsequence $\{\mathbf{d}_{k_q}\}$. Furthermore, $\mathbf{d}^* \in T(\mathbf{x}^*)$ and $\nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d}^* \leq L^* - \frac{\theta}{2}$, since $\forall k_q$ we have $\nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d}_{k_q} \leq L^* - \frac{\theta}{2}$, which contradicts the definition of L^*. □

The following result, inspired by Lemma 2.4 in [237], provides a conservative estimate of the subset of X_5 around a nonisolated $\mathbf{x}^*$ under the assumption that the objective function grows linearly on the feasible region in some neighborhood of $\mathbf{x}^*$. The reader can compare the assumptions of Lemma 5.3.2 with what follows from Lemma 5.3.1 and the necessary optimality conditions in Theorem 2.3.15 (see Remark 5.3.3 for details).

Lemma 5.3.2. Consider Problem (P). Suppose $\mathbf{x}^*$ is nonisolated, f is differentiable at $\mathbf{x}^*$, and $\exists \alpha > 0$ such that

$$L := \inf_{\left\{\mathbf{d}:\|\mathbf{d}\|_1=1,\, \exists t>0\,:\,(\mathbf{x}^*+t\mathbf{d})\in\mathcal{N}^1_\alpha(\mathbf{x}^*)\cap\mathcal{F}(X)\right\}} \nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} > 0.$$

Then, $\exists \hat{\alpha} \in (0, \alpha]$ such that the region $\mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) \cap X_5$ can be conservatively approximated by

$$\hat{X}_5 = \left\{\mathbf{x} \in \mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) : L\|\mathbf{x} - \mathbf{x}^*\|_1 \leq 2\varepsilon\right\}.$$

Proof. Let $\mathbf{x} = \mathbf{x}^* + t\mathbf{d} \in \mathcal{N}^1_\alpha(\mathbf{x}^*) \cap \mathcal{F}(X)$ with $\|\mathbf{d}\|_1 = 1$ and $t = \|\mathbf{x} - \mathbf{x}^*\|_1 > 0$. We have

$$\begin{aligned}
f(\mathbf{x}) &= f(\mathbf{x}^* + t\mathbf{d}) \\
&= f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x} - \mathbf{x}^*) + o(\|\mathbf{x} - \mathbf{x}^*\|_1) \\
&= f(\mathbf{x}^*) + t\nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} + o(t) \\
&\geq f(\mathbf{x}^*) + Lt + o(t),
\end{aligned}$$

where Step 2 follows from the differentiability of f at $\mathbf{x}^*$. Consequently, there exists $\hat{\alpha} \in (0, \alpha]$ such that for all $\mathbf{x} = \mathbf{x}^* + t\mathbf{d} \in \mathcal{F}(X)$ with $\|\mathbf{d}\|_1 = 1$ and $t \in [0, \hat{\alpha})$:

$$f(\mathbf{x}) \geq f(\mathbf{x}^*) + Lt + o(t) \geq f(\mathbf{x}^*) + \frac{L}{2}t.$$

Therefore, $\forall \mathbf{x} \in \mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) \cap X_5$ we have $\mathbf{x} = \mathbf{x}^* + t\mathbf{d} \in \mathcal{F}(X)$ with $\|\mathbf{d}\|_1 = 1$ and $t =$

$\|\mathbf{x} - \mathbf{x}^*\|_1 < \hat{\alpha}$, and

$$\varepsilon \geq f(\mathbf{x}) - f(\mathbf{x}^*) \geq \frac{L}{2}t \implies Lt = L\|\mathbf{x} - \mathbf{x}^*\|_1 \leq 2\varepsilon. \qquad \square$$

A conservative estimate of the number of boxes of width δ required to cover $\mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) \cap X_5$ can be obtained by estimating the number of boxes of width δ required to cover $\hat{X}_5$ (see Theorem 5.3.11). The following remark is in order.

Remark 5.3.3.

1. Lemma 5.3.2 is not applicable when $L = 0$. This can occur, for instance, when $\mathbf{x}^*$ is an unconstrained minimum, in which case other techniques have to be employed to analyze the cluster problem [68, 178, 237, 238] under alternative assumptions. This is because when f is differentiable at an unconstrained minimizer $\mathbf{x}^*$, it grows slower than linearly around $\mathbf{x}^*$ as a result of the first-order necessary optimality condition $\nabla f(\mathbf{x}^*) = \mathbf{0}$ (note that if f is twice-differentiable at $\mathbf{x}^*$ and $\nabla^2 f(\mathbf{x}^*)$ is positive definite, then f grows quadratically around $\mathbf{x}^*$, see Theorems 2.3.9, 2.3.11, and 2.3.12). The assumptions of Lemma 5.3.2 may be satisfied for a constrained problem, however, because they only require that the objective function grow linearly in the set of directions that lead to feasible points in some neighborhood of $\mathbf{x}^*$. An example of $L = 0$ when $\mathbf{x}^*$ is not an unconstrained minimum is: $X = (-2, 2)$, $m_I = 2$, $m_E = 0$, $f(x) = x^3$, $g_1(x) = x - 1$, and $g_2(x) = -x$ with $x^* = 0$. In this example, the objective function only grows cubically around x^* in the direction from x^* that leads to feasible points.

   From Lemma 5.3.1, we have that a sufficient condition for the key assumption of Lemma 5.3.2 to be satisfied is $\min\limits_{\left\{\mathbf{d}:\|\mathbf{d}\|_1=1,\mathbf{d}\in T(\mathbf{x}^*)\right\}} \nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} > 0$. It is not hard to show that this condition is also necessary when f is differentiable at $\mathbf{x}^*$. Proposition 5.3.7 shows that the assumptions of Lemma 5.3.2 will not be satisfied when Problem (P) does not contain any active inequality constraints and the minimizer corresponds to a KKT point (see Definition 2.3.18) for Problem (P).

2. The constant $\hat{\alpha}$ depends on the local behavior of f around $\mathbf{x}^*$, but is independent of ε since it is determined by the subset of $\mathcal{N}^1_{\alpha}(\mathbf{x}^*) \cap \mathcal{F}(X)$ on which the affine function $f(\mathbf{x}^*) + \frac{L}{2}t$ underestimates $f(\mathbf{x})$. Consequently, for sufficiently small ε, $\hat{X}_5 = \{\mathbf{x} \in X : L\|\mathbf{x} - \mathbf{x}^*\|_1 \leq 2\varepsilon\}$ since $\{\mathbf{x} \in X : L\|\mathbf{x} - \mathbf{x}^*\|_1 \leq 2\varepsilon\}$ will then be a subset of

$\mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*)$. Note that the factor '2' in the denominator of '$\frac{L}{2}t$' is arbitrarily chosen; any factor > 1 can instead be chosen with a corresponding $\hat{\alpha}$. Furthermore, $\mathbf{x}^*$ is necessarily the unique global minimizer of Problem (P) on $\mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*)$ since $L > 0$.

3. If, in addition to the assumptions of Lemma 5.3.2, f is assumed to be convex on $\mathcal{N}^1_{\alpha}(\mathbf{x}^*)$, then we can choose $\hat{\alpha} = \alpha$. Additionally, $\mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) \cap X_5$ can be conservatively approximated by $\{\mathbf{x} \in X : L\|\mathbf{x} - \mathbf{x}^*\|_1 \leq \varepsilon\}$ when ε is small enough.

4. The estimate $\hat{X}_5$ becomes less conservative as ε is decreased since the higher order term $o(t) \to 0$ as $\varepsilon \to 0$. Simply put, this is because the affine approximation $f(\mathbf{x}^*) + Lt$ provides a better description of f as $\varepsilon \to 0$.

In fact, under the assumptions of Lemma 5.3.2, a less conservative estimate of X_5 can be obtained by accounting for the fact that not all points $\mathbf{x} \in \left\{\mathbf{x} \in \mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) : L\|\mathbf{x} - \mathbf{x}^*\|_1 \leq 2\varepsilon\right\}$ satisfy $\nabla f(\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x} - \mathbf{x}^*) \geq L\|\mathbf{x} - \mathbf{x}^*\|_1$.

Proposition 5.3.4. Consider Problem (P), and suppose the assumptions of Lemma 5.3.2 are satisfied. Then, $\exists \hat{\alpha} \in (0, \alpha]$ such that the region $\mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) \cap X_5$ can be conservatively approximated by

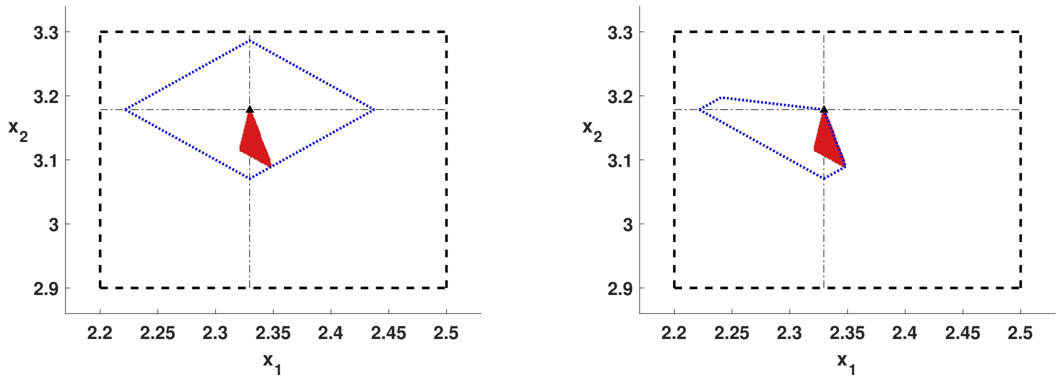
$$\hat{X}_5 = \left\{\mathbf{x} \in \mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) : L\|\mathbf{x} - \mathbf{x}^*\|_1 \leq 2\varepsilon,\ L\|\mathbf{x} - \mathbf{x}^*\|_1 \leq \nabla f(\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x} - \mathbf{x}^*)\right\}.$$

Proof. The desired result follows from Lemma 5.3.2 and the fact that

$$\nabla f(\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x} - \mathbf{x}^*) \geq L\|\mathbf{x} - \mathbf{x}^*\|_1, \quad \forall \mathbf{x} \in \mathcal{N}^1_{\alpha}(\mathbf{x}^*) \cap \mathcal{F}(X),$$

from the assumptions of Lemma 5.3.2. $\square$

As an illustration of the application of Lemma 5.3.2, let us reconsider Example 5.2.8. Recall that $X = (2.2, 2.5) \times (2.9, 3.3)$, $m_I = 3$, $m_E = 0$, $f(\mathbf{x}) = -x_1 - x_2$, $g_1(\mathbf{x}) = x_2 - 2x_1^4 + 8x_1^3 - 8x_1^2 - 2$, $g_2(\mathbf{x}) = x_2 - 4x_1^4 + 32x_1^3 - 88x_1^2 + 96x_1 - 36$, and $g_3(\mathbf{x}) = 3 - x_2$ with $\mathbf{x}^* \approx (2.33, 3.18)$. Let $\varepsilon \leq 0.07$. We have $\mathcal{F}(X) = \{\mathbf{x} \in X : \mathbf{g}(\mathbf{x}) \leq \mathbf{0}\}$, $\nabla f(\mathbf{x}^*) = (-1, -1)$, $\alpha = +\infty$, $L \approx 0.649$, and $X_5 = \{\mathbf{x} \in X : \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, -x_1 - x_2 \leq f(\mathbf{x}^*) + \varepsilon\}$. Choose $\hat{\alpha} = +\infty$ in Lemma 5.3.2. From Lemma 5.3.2 and Remark 5.3.3, we have $\hat{X}_5 = \{\mathbf{x} : 0.649\|\mathbf{x} - \mathbf{x}^*\|_1 \leq \varepsilon\}$ (since f is convex).



(a) X_5 and estimate $\hat{X}_5$ from Lemma 5.3.2 (b) X_5 and estimate $\hat{X}_5$ from Proposition 5.3.4

Figure 5-2: Plots of X_5 (solid regions) and $\hat{X}_5$ (the areas between the dotted lines) for Example 5.2.8 for $\varepsilon = 0.07$ (note that we do not use $\varepsilon = 0.1$ as in Figure 5-1b because the corresponding $\hat{X}_5$ are not contained in X). The dashed lines define the set X, the filled-in triangles correspond to the minimizer $\mathbf{x}^*$, and the dash-dotted lines represent the axes translated to $\mathbf{x}^*$.

Figure 5-2a plots X_5 and $\hat{X}_5$ for $\varepsilon = 0.07$, and Figure 5-2b shows the improvement in the estimate when Proposition 5.3.4 is used, in which case we obtain the estimate $\hat{X}_5 = \{\mathbf{x} : 0.649\|\mathbf{x} - \mathbf{x}^*\|_1 \leq \varepsilon,\ 0.649\|\mathbf{x} - \mathbf{x}^*\|_1 \leq -(x_1 - x_1^*) - (x_2 - x_2^*)\}$. Note that an even better estimate of X_5 may be obtained by using knowledge of the local feasible set $\mathcal{N}^1_\alpha(\mathbf{x}^*) \cap \mathcal{F}(X)$. However, other than in some special cases (see Lemma 5.3.13), we shall stick with the estimate $\hat{X}_5$ from Lemma 5.3.2 since we are mainly concerned with the dependence of the extent of clustering on the convergence rate of the lower bounding scheme.

Before we provide an estimate of the number of boxes of width δ required to cover $\mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) \cap X_5$, we provide a few more examples that satisfy the assumptions of Lemma 5.3.2 and present an approach that could help determine if its assumptions are satisfied. Example 5.3.5 illustrates another inequality-constrained case which satisfies the assumptions of Lemma 5.3.2. Note that the minimizer $\mathbf{x}^*$ does not satisfy the KKT conditions (see Theorem 2.3.17) in this case.

Example 5.3.5. Let $\varepsilon \leq 1$, $X = (-2, 2)$, $m_I = 3$, and $m_E = 0$ with $f(x) = -x$, $g_1(x) = x^3$, $g_2(x) = x - 1$, $g_3(x) = -1 - x$, and $x^* = 0$. We have $\mathcal{F}(X) = [-1, 0]$, $\nabla f(x^*) = -1$, $\alpha = +\infty$, $L = 1$, and $X_5 = [-\varepsilon, 0]$. Choose $\hat{\alpha} = +\infty$ in Lemma 5.3.2. From Lemma 5.3.2 and Remark 5.3.3, we have $\hat{X}_5 = [-\varepsilon, +\varepsilon]$ (since f is convex).

The reader may conjecture, based on Example 5.3.5 and other examples of low dimension, that every nonisolated minimizer $\mathbf{x}^*$ which does not satisfy the KKT conditions will automatically satisfy the main assumption of Lemma 5.3.2. Example 5.3.6, inspired by [99, Section 4.1], however illustrates a case when the assumptions of Lemma 5.3.2 are not satisfied even though $\mathbf{x}^*$ does not satisfy the KKT conditions.

Example 5.3.6. Let $X = (-2, 2)^3$, $m_I = 5$, and $m_E = 0$ with $f(\mathbf{x}) = x_1 + x_3^2$, $g_1(\mathbf{x}) = x_1 - 1$, $g_2(\mathbf{x}) = x_2 - x_1$, $g_3(\mathbf{x}) = x_2^2$, $g_4(\mathbf{x}) = -x_3$, $g_5(\mathbf{x}) = x_3 - 1$, and $\mathbf{x}^* = (0, 0, 0)$. We have $\mathcal{F}(X) = \left\{\mathbf{x} \in [0, 1]^3 : x_2 = 0\right\}$, $\nabla f(\mathbf{x}^*) = (1, 0, 0)$, and $L = 0$ for every $\alpha > 0$ since $(0, 0, 1) \in T(\mathbf{x}^*)$ and $\nabla f(\mathbf{x}^*)^{\mathrm{T}}(0, 0, 1) = 0$.

The next result provides conditions under which the assumptions of Lemma 5.3.2 will not be satisfied. In particular, it is shown that the assumptions of Lemma 5.3.2 will not be satisfied if Problem (P) is purely equality-constrained and all the functions in Problem (P) are differentiable at a nonisolated $\mathbf{x}^*$.

Proposition 5.3.7. Consider Problem (P) with $m_E \geq 1$. Suppose $\mathbf{x}^*$ is nonisolated, f is differentiable at $\mathbf{x}^*$, functions h_k, $k = 1, \cdots, m_E$, are differentiable at $\mathbf{x}^*$, and $\mathcal{A}(\mathbf{x}^*) = \emptyset$. Furthermore, suppose there exist multipliers $\boldsymbol{\lambda}^* \in \mathbb{R}^{m_E}$ corresponding to the equality constraints such that $(\mathbf{x}^*, \mathbf{0}, \boldsymbol{\lambda}^*)$ is a KKT point. Then

$$\min_{\left\{\mathbf{d}: \|\mathbf{d}\|_1 = 1, \mathbf{d} \in T(\mathbf{x}^*)\right\}} \nabla f(\mathbf{x}^*)^{\mathrm{T}} \mathbf{d} = 0.$$

Proof. Since $(\mathbf{x}^*, \mathbf{0}, \boldsymbol{\lambda}^*)$ is a KKT point, we have (see Definition 2.3.18):

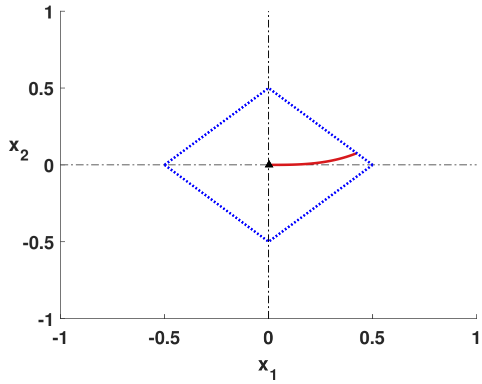
$$\nabla f(\mathbf{x}^*) + \sum_{k=1}^{m_E} \lambda_k^* \nabla h_k(\mathbf{x}^*) = \mathbf{0}.$$

From the assumption that $\mathbf{x}^*$ is a nonisolated feasible point, we have that the set of unit-norm directions $\{\mathbf{d} : \|\mathbf{d}\|_1 = 1, \mathbf{d} \in T(\mathbf{x}^*)\}$ is nonempty. Additionally, we have

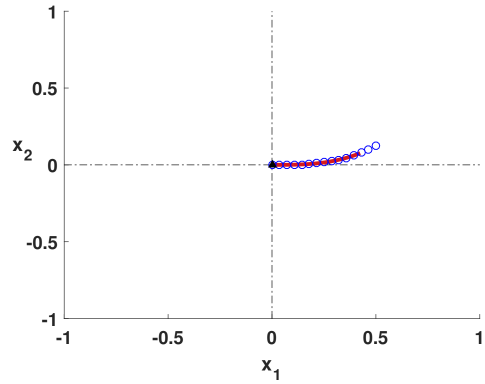
$$T(\mathbf{x}^*) \subset \mathcal{L}(\mathbf{x}^*) := \left\{\mathbf{d} \in \mathbb{R}^{n_x} : \nabla h_k(\mathbf{x}^*)^{\mathrm{T}} \mathbf{d} = 0, \forall k \in \{1, \cdots, m_E\}\right\},$$

where $\mathcal{L}(\mathbf{x}^*)$ denotes the linearized cone at $\mathbf{x}^*$ (see, for instance, [13]). Consequently, for each $\mathbf{d} \in T(\mathbf{x}^*)$ with $\|\mathbf{d}\|_1 = 1$, we have $\nabla f(\mathbf{x}^*)^{\mathrm{T}} \mathbf{d} = 0$. ☐

Note that the above result can naturally be extended to accommodate weakly active inequality constraints (see [13, Section 4.4]). The ensuing examples illustrate that the as-



(a) X_5 and estimate $\hat{X}_5$ from Lemma 5.3.2



(b) X_5 and estimate $\hat{X}_5$ from Lemma 5.3.13

Figure 5-3: Plots of X_5 (solid curves) and $\hat{X}_5$ (left figure: area between the dotted lines, right figure: curve depicted by the circles) for Example 5.3.9 for $\varepsilon = 0.5$. The filled-in triangles correspond to the minimizer $\mathbf{x}^*$, and the dash-dotted lines represent the axes translated to $\mathbf{x}^*$.

sumptions of Lemma 5.3.2 may be satisfied when individual assumptions of Proposition 5.3.7 do not hold.

Example 5.3.8. Let $\varepsilon \leq 0.5$, $X = (-2,2) \times (-2,2)$, $m_I = 1$, and $m_E = 1$ with $f(\mathbf{x}) = x_1 + 10x_2^2$, $g(\mathbf{x}) = x_1 - 1$, $h(\mathbf{x}) = x_1 - |x_2|$, and $\mathbf{x}^* = (0,0)$. We have $\mathcal{F}(X) = \{\mathbf{x} \in X : x_1 = |x_2|, x_1 \leq 1\}$, $\nabla f(\mathbf{x}^*) = (1,0)$, $\alpha = +\infty$, $L = 0.5$, and the set $X_5 = \left\{\mathbf{x} \in [0,\varepsilon] \times [-\varepsilon,\varepsilon] : x_1 = |x_2|, x_1 + 10x_2^2 \leq \varepsilon\right\}$. Choose $\hat{\alpha} = +\infty$ in Lemma 5.3.2. From Lemma 5.3.2 and Remark 5.3.3, we have $\hat{X}_5 = \{\mathbf{x} \in X : \|\mathbf{x}\|_1 \leq 2\varepsilon\}$ (since f is convex).

Example 5.3.9. Let $\varepsilon \leq 0.5$, $X = (-2,2) \times (-2,2)$, $m_I = 4$, and $m_E = 1$ with $f(\mathbf{x}) = x_1 + x_2$, $g_1(\mathbf{x}) = -x_1$, $g_2(\mathbf{x}) = -x_2$, $g_3(\mathbf{x}) = x_1 - 1$, $g_4(\mathbf{x}) = x_2 - 1$, $h(\mathbf{x}) = x_2 - x_1^3$, and $\mathbf{x}^* = (0,0)$. We have $\mathcal{F}(X) = \left\{\mathbf{x} \in [0,1]^2 : x_2 = x_1^3\right\}$, $\nabla f(\mathbf{x}^*) = (1,1)$, $\alpha = +\infty$, $L = 1$, and $X_5 = \left\{\mathbf{x} \in [0,\varepsilon] \times [0,\varepsilon] : x_2 = x_1^3, x_1 + x_2 \leq \varepsilon\right\}$. Choose $\hat{\alpha} = +\infty$ in Lemma 5.3.2. From Lemma 5.3.2 and Remark 5.3.3, we have $\hat{X}_5 = \{\mathbf{x} \in X : \|\mathbf{x}\|_1 \leq \varepsilon\}$ (since f is convex).

Figure 5-3a plots X_5 and $\hat{X}_5$ for Example 5.3.9 for $\varepsilon = 0.5$. It is seen that the estimate $\hat{X}_5$ does not capture the 'one-dimensional nature' of X_5 (which is a consequence of the equality constraint in Example 5.3.9). This issue is addressed in Lemma 5.3.13. Note that X_5 for Example 5.3.8 also resides in a reduced-dimensional manifold, but Lemma 5.3.13 does not apply in this case since h is not differentiable at $\mathbf{x}^*$ (the discussion after Lemma 5.3.13 proposes a modification of the assumptions of Lemma 5.3.13 that addresses this issue).

While Lemma 5.3.2 provides a conservative estimate of $\mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) \cap X_5$ under suitable assumptions, verifying the satisfaction of its assumptions is not straightforward. The following proposition provides a conservative approach for determining whether the assumptions of Lemma 5.3.2 are satisfied.

Proposition 5.3.10. Let $L(\alpha)$ denote the constant L in Lemma 5.3.2 for a given $\alpha > 0$. When the active constraints are differentiable at $\mathbf{x}^*$, a lower bound on $L_0 := \lim\limits_{\alpha \to 0^+} L(\alpha)$ can be obtained by solving

$$\begin{aligned} \min_{\mathbf{d}} \; & \nabla f(\mathbf{x}^*)^{\mathrm{T}} \mathbf{d} \\ \text{s.t.} \; & \|\mathbf{d}\|_1 = 1, \\ & \mathbf{d} \in \mathcal{L}(\mathbf{x}^*), \end{aligned}$$

where

$$\mathcal{L}(\mathbf{x}^*) := \left\{ \mathbf{d} \in \mathbb{R}^{n_x} : \nabla g_j(\mathbf{x}^*)^{\mathrm{T}} \mathbf{d} \leq 0, \forall j \in \mathcal{A}(\mathbf{x}^*), \nabla h_k(\mathbf{x}^*)^{\mathrm{T}} \mathbf{d} = 0, \forall k \in \{1, \cdots, m_E\} \right\}$$

denotes the linearized cone at $\mathbf{x}^*$. If $\mathbf{x}^*$ corresponds to a KKT point, the above formulation provides the exact value of L_0.

So far in this section, we have established conditions under which a conservative estimate of the subset of X_5 around a minimizer $\mathbf{x}^*$ can be obtained, presented examples for which the above conditions hold, and isolated a class of problems for which the above conditions are not satisfied. The following theorem follows from Corollary 2.1 in [237], the proof of which is rederived for completeness. It provides a conservative estimate of the number of boxes of width δ required to cover $\hat{X}_5$ from Lemma 5.3.2. Therefore, from Lemma 5.2.4 and the result below, we can get an upper bound on the worst-case number of boxes required to cover $\mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) \cap X_5$ and estimate the extent of the cluster problem on that region (recall from Remark 5.3.3 that the subset of X_5 around $\mathbf{x}^*$ will be contained in $\mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*)$ for sufficiently small ε).

Theorem 5.3.11. Suppose the assumptions of Lemma 5.3.2 hold. Let $\delta = \left(\dfrac{\varepsilon}{\tau^*}\right)^{\frac{1}{\beta^*}}$, $r = \dfrac{2\varepsilon}{L}$.

1. If $\delta \geq 2r$, let $N = 1$.

2. If $\dfrac{2r}{m-1} > \delta \geq \dfrac{2r}{m}$ for some $m \in \mathbb{N}$ with $m \leq n_x$ and $2 \leq m \leq 5$, then let

$$N = \sum_{i=0}^{m-1} 2^i \binom{n_x}{i} + 2n_x \left\lceil \frac{m-3}{3} \right\rceil.$$

3. Otherwise, let

$$N = \left\lceil 2\,(\tau^*)^{\frac{1}{\beta^*}}\,\varepsilon^{\left(1-\frac{1}{\beta^*}\right)} L^{-1} \right\rceil^{n_x-1} \left(\left\lceil 2\,(\tau^*)^{\frac{1}{\beta^*}}\,\varepsilon^{\left(1-\frac{1}{\beta^*}\right)} L^{-1} \right\rceil + 2n_x \left\lceil (\tau^*)^{\frac{1}{\beta^*}}\,\varepsilon^{\left(1-\frac{1}{\beta^*}\right)} L^{-1} \right\rceil \right).$$

Then, N is an upper bound on the number of boxes of width δ required to cover $\hat{X}_5$.

Proof. This proof is rederived based on Corollary 2.1 in [237] and the proof of Lemma 3 in [238]. Note that the condition in the second case is corrected to '$2 \leq m \leq 5$' as opposed to '$2 \leq m \leq 6$' in [237].

From Lemma 5.3.2, we have

$$\hat{X}_5 = \left\{ \mathbf{x} \in \mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) : L\|\mathbf{x} - \mathbf{x}^*\|_1 \leq 2\varepsilon \right\} \subset \left\{ \mathbf{x} : \|\mathbf{x} - \mathbf{x}^*\|_1 \leq \frac{2\varepsilon}{L} \right\} =: \tilde{B}.$$

Therefore, an upper bound on the number of boxes of width δ required to cover $\hat{X}_5$ can be obtained by conservatively estimating the number of boxes of width δ required to cover $\tilde{B}$. In what follows, we will assume without loss of generality that $\mathbf{x}^* = \mathbf{0}$.

1. Suppose $\delta \geq 2r$. Consider the box B_δ of width δ centered at $\mathbf{x}^* = \mathbf{0}$. We have

$$\mathbf{x} \in \tilde{B} \implies \|\mathbf{x}\|_1 \leq \frac{2\varepsilon}{L} \implies \|\mathbf{x}\|_\infty \leq \frac{2\varepsilon}{L} = r \leq \frac{\delta}{2} \implies \mathbf{x} \in B_\delta,$$

where we have used the fact that $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_1$, $\forall \mathbf{x} \in \mathbb{R}^{n_x}$. Therefore, B_δ is sufficient to cover $\tilde{B}$.

2. Suppose $m \leq n_x$ with $m \in \{2, \cdots, 5\}$ and $\delta \geq \frac{2r}{m}$. Place a box B_δ of width δ centered at $\mathbf{x}^* = \mathbf{0}$ (the condition on δ ensures that B_δ intersects the boundary of $\tilde{B}$). Let

$$E_i := \left\{ \mathbf{e} \in \mathbb{R}^{n_x} : e_j \in \left\{ -\frac{\delta}{2}, 0, \frac{\delta}{2} \right\},\ \forall j \in \{1, \cdots, n_x\},\ \sum_{j=1}^{n_x} I_0(e_j) = i \right\},$$

where $I_0 : \mathbb{R} \to \{0,1\}$ is defined as $I_0(x) := \begin{cases} 0, & \text{if } x = 0 \\ 1, & \text{otherwise} \end{cases}$, denote the set of midpoints of the $(n_x - i)$-dimensional faces of B_δ (each element of E_i has exactly i nonzero components, each of which is $\pm\frac{\delta}{2}$). Note that $|E_i| = 2^i\binom{n_x}{i}$, $\forall i \in \{1, \cdots, n_x\}$. Under the assumption $\delta \geq \frac{2r}{m}$, we will show that, in addition to B_δ, it is sufficient to place one box beside B_δ along the directions in $E_1, \cdots, E_{m-1}$ when $m = 2$ or $m = 3$, and two boxes beside B_δ along the directions in E_1 and one box beside B_δ along the directions in $E_2, \cdots, E_{m-1}$ when $m = 4$ or $m = 5$ in order to cover $\tilde{B}$.

First, we show that we need not place any boxes beside B_δ along the directions in $E_m, \cdots, E_{n_x}$. Let $\mathbf{e} \in E_i$ with $i \in \{m, \cdots, n_x\}$. We have $\|\mathbf{e}\|_1 = \frac{\delta}{2}i \geq \frac{i}{m}r \geq r$, which implies $\mathbf{e} \in \partial\tilde{B} \cup \tilde{B}^{\mathrm{C}}$ (where $\partial\tilde{B}$ denotes the boundary of $\tilde{B}$). Consequently, boxes placed beside B_δ along the directions in $E_m, \cdots, E_{n_x}$ do not intersect the interior of $\tilde{B}$ and are not required to cover $\tilde{B}$.

Suppose $\delta \geq \frac{2r}{m}$, and let $\mathbf{e} \in E_i$ for some $i \in \{1, \cdots, m-1\}$. The distance from $\mathbf{e}$, which is the midpoint of an $(n-i)$-dimensional face of B_δ, to $\frac{2r}{\delta i}\mathbf{e}$, which is a point on the boundary of $\tilde{B}$ in the direction $\mathbf{e}$, in the ∞-norm is $\frac{r}{i} - \frac{\delta}{2} \leq \frac{r}{i} - \frac{r}{m}$. If this distance is less than δ for each $i \in \{1, \cdots, m-1\}$, then one box beside B_δ along the directions in $E_1, \cdots, E_{m-1}$ is sufficient to cover $\tilde{B}$. This amounts to requiring

$$\begin{aligned} & \frac{r}{i} - \frac{r}{m} \leq \frac{2r}{m}, \ \forall i \in \{1, \cdots, m-1\} \\ \iff & m \leq 3i, \ \forall i \in \{1, \cdots, m-1\} \\ \iff & m = 2 \ \text{ or } \ m = 3. \end{aligned}$$

Note that if $m = 4$ or $m = 5$, we still have $m \leq 3i$, $\forall i \in \{2, \cdots, m-1\}$. Additionally, $\frac{r}{1} - \frac{r}{m} \leq \frac{4r}{m} \leq 2\delta$ in such cases. Therefore, when $m = 4$ or $m = 5$, two boxes along the directions in E_1 and one box along the directions in $E_2, \cdots, E_{m-1}$ are sufficient to cover $\tilde{B}$.

3. If the previous assumptions on δ are not satisfied, a box of width δ centered at $\mathbf{x}^*$ may not intersect $\partial\tilde{B}$. To estimate the number of boxes of width δ required to cover $\tilde{B}$, we first estimate the number of boxes, N_r, of width $r = \frac{2\varepsilon}{L}$ required to cover $\tilde{B}$ using the previous analysis, and then estimate the number of boxes of width δ required to cover

the intersection of these N_r boxes with $\tilde{B}$.

The number of boxes of width r required to cover $\tilde{B}$ is $N_r := 1 + 2n_x$, where '1' corresponds to the box centered at $\mathbf{x}^* = \mathbf{0}$, and '$2n_x$' corresponds to the boxes along the directions in E_1. Note that E_1 is now defined as

$$E_1 := \left\{ \mathbf{e} \in \mathbb{R}^{n_x} : e_j \in \left\{ -\frac{r}{2}, 0, \frac{r}{2} \right\}, \forall j \in \{1, \cdots, n_x\}, \sum_{j=1}^{n_x} I_0(e_j) = 1 \right\}$$

since $\tilde{B}$ is first covered using boxes of width r. The box of width r centered at $\mathbf{x}^*$ can be covered using $\left\lceil \frac{r}{\delta} \right\rceil^{n_x}$ boxes of width δ. Note that the entire volume of the $2n_x$ boxes along the directions in E_1 need not be covered using boxes of width δ since parts of those boxes have no intersection with $\tilde{B}$. To estimate the extent to which each of the $2n_x$ boxes need to be covered with boxes of width δ, we compute the distance between any $\mathbf{e} \in E_1$ (which is a midpoint of a one-dimensional face of the box of width r centered at $\mathbf{x}^*$) and $\frac{2r}{r \times 1}\mathbf{e} = 2\mathbf{e}$ (which is a point on the boundary of $\tilde{B}$ in the direction $\mathbf{e}$) in the ∞-norm. This distance turns out to be equal to $\frac{r}{2}$. This implies at most half the volumes of the $2n_x$ boxes need to be covered using boxes of width δ, which yields the estimate of $2n_x \left\lceil \frac{r}{\delta} \right\rceil^{n_x - 1} \left\lceil \frac{r}{2\delta} \right\rceil$ boxes of width δ that are required to cover the $2n_x$ boxes of width r along the directions in E_1. $\square$

Remark 5.3.12. Under the assumptions of Lemma 5.3.2, the dependence of N on ε disappears when the lower bounding scheme has first-order convergence on $\mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) \cap \mathcal{F}(X)$, i.e., $\beta^* = 1$. Therefore, the cluster problem on X_5 may be eliminated even using first-order convergent lower bounding schemes with sufficiently small prefactors. This is in contrast to unconstrained global optimization where at least second-order convergent lower bounding schemes are required to eliminate the cluster problem (see Remark 5.3.3 for an intuitive explanation for this qualitative difference in behavior). Note that the dependence of N on the prefactor τ^* can be detailed in a manner similar to Table 1 in [238].

The above scaling has also been empirically observed by Goldsztejn et al. [87], who reason "$\cdots$ removes the tangency between the feasible set and the objective level set, and therefore should prevent the cluster effect."

The next result refines the analysis of Lemma 5.3.2 when Problem (P) contains equality constraints that can locally be eliminated using the implicit function theorem [192].

Lemma 5.3.13. Consider Problem (P) with $1 \leq m_E < n_x$. Suppose $\mathbf{x}^*$ is nonisolated, f is differentiable at $\mathbf{x}^*$, and $\exists \alpha > 0$ such that $\mathbf{h}$ is continuously differentiable on $\mathcal{N}^1_\alpha(\mathbf{x}^*)$ and

$$L := \inf_{\left\{\mathbf{d} : \|\mathbf{d}\|_1 = 1,\, \exists t > 0 \,:\, (\mathbf{x}^* + t\mathbf{d}) \in \mathcal{N}^1_\alpha(\mathbf{x}^*) \cap \mathcal{F}(X)\right\}} \nabla f(\mathbf{x}^*)^{\mathrm{T}} \mathbf{d} > 0.$$

Furthermore, suppose the variables $\mathbf{x}$ can be reordered and partitioned into dependent variables $\mathbf{z} \in \mathbb{R}^{m_E}$ and independent variables $\mathbf{p} \in \mathbb{R}^{n_x - m_E}$, with $\mathbf{x} \equiv (\mathbf{z}, \mathbf{p})$, such that $\nabla_{\mathbf{z}} \mathbf{h}((\mathbf{z}, \mathbf{p}))$ is nonsingular on $\mathcal{N}^1_\alpha((\mathbf{z}^*, \mathbf{p}^*))$, where $\mathbf{x}^* \equiv (\mathbf{z}^*, \mathbf{p}^*)$. Then, $\exists \alpha_{\mathbf{p}}, \alpha_{\mathbf{z}} \in (0, \alpha]$, a continuously differentiable function $\boldsymbol{\phi} : \mathcal{N}^1_{\alpha_{\mathbf{p}}}(\mathbf{p}^*) \to \mathcal{N}^1_{\alpha_{\mathbf{z}}}(\mathbf{z}^*)$, and $\hat{\alpha} \in (0, \alpha_{\mathbf{p}})$ such that the region $\left(\mathcal{N}^1_{\alpha_{\mathbf{z}}}(\mathbf{z}^*) \times \mathcal{N}^1_{\hat{\alpha}}(\mathbf{p}^*)\right) \cap X_5$ can be conservatively approximated by

$$\hat{X}_5 = \left\{(\mathbf{z}, \mathbf{p}) \in \mathcal{N}^1_{\alpha_{\mathbf{z}}}(\mathbf{z}^*) \times \mathcal{N}^1_{\hat{\alpha}}(\mathbf{p}^*) : \mathbf{z} = \boldsymbol{\phi}(\mathbf{p}),\ L\|\mathbf{p} - \mathbf{p}^*\|_1 \leq 2\varepsilon\right\}.$$

Proof. The result follows from the proof of Lemma 5.3.2 and the implicit function theorem [192, Chapter 9]. □

Lemma 5.3.13 effectively states that, under suitable conditions, the subset of X_5 around $\mathbf{x}^*$ resides in a reduced-dimensional manifold. Figure 5-3b compares the estimate $\hat{X}_5$ obtained from Lemma 5.3.13 (when we assume precise knowledge of the implicit function) with the one obtained from Lemma 5.3.2 for Example 5.3.9. The reason for distinguishing between $\alpha_{\mathbf{p}}$ and $\hat{\alpha}$ is so that we can have $\boldsymbol{\phi}$ to be continuously differentiable on $\mathrm{cl}\left(\mathcal{N}^1_{\hat{\alpha}}(\mathbf{p}^*)\right)$; this fact will be used shortly. Note that the assumptions that $\mathbf{h}$ is continuously differentiable on $\mathcal{N}^1_\alpha(\mathbf{x}^*)$ and $\nabla_{\mathbf{z}} \mathbf{h}((\mathbf{z}, \mathbf{p}))$ is nonsingular on $\mathcal{N}^1_\alpha((\mathbf{z}^*, \mathbf{p}^*))$ can be relaxed based on a nonsmooth variant of the implicit function theorem [57, Chapter 7] (which can be used to derive a less conservative estimate of X_5 for Example 5.3.8, for instance).

The following corollary of Theorem 5.3.11 refines the estimate of the number of boxes of width δ required to cover $\hat{X}_5$ under the assumptions of Lemma 5.3.13. It provides an upper bound on the number of boxes of width δ required to cover X_5 that scales as $O\left(\varepsilon^{(n_x - m_E)\left(1 - \frac{1}{\beta^*}\right)}\right)$ in contrast to the scaling $O\left(\varepsilon^{n_x\left(1 - \frac{1}{\beta^*}\right)}\right)$ from Theorem 5.3.11.

Corollary 5.3.14. Suppose the assumptions of Lemma 5.3.13 hold. Let $\delta = \left(\dfrac{\varepsilon}{\tau^*}\right)^{\frac{1}{\beta^*}}$,

$r = \dfrac{2\varepsilon}{L}$. Define

$$M_k := \left(\max_{\mathbf{p} \in \mathrm{cl}\left(\mathcal{N}^1_{\hat{\alpha}}(\mathbf{p}^*)\right)} \|\nabla \phi_k(\mathbf{p})\| \right) \sqrt{n_x - m_E}, \quad \forall k \in \{1, \cdots, m_E\},$$

$$K := \{k \in \{1, \cdots, m_E\} : M_k > 1\}.$$

1. If $\delta \geq 2r$, let $N = \prod_{k \in K} M_k$.

2. If $\dfrac{2r}{m-1} > \delta \geq \dfrac{2r}{m}$ for some $m \in \mathbb{N}$ with $m \leq n_x - m_E$ and $2 \leq m \leq 5$, then let

$$N = \left(\sum_{i=0}^{m-1} 2^i \binom{n_x - m_E}{i} + 2\left(n_x - m_E\right) \left\lceil \frac{m-3}{3} \right\rceil \right) \prod_{k \in K} M_k.$$

3. Otherwise, let

$$N = \left\lceil 2\left(\tau^*\right)^{\frac{1}{\beta^*}} \varepsilon^{\left(1 - \frac{1}{\beta^*}\right)} L^{-1} \right\rceil^{n_x - m_E - 1} \Bigg(\left\lceil 2\left(\tau^*\right)^{\frac{1}{\beta^*}} \varepsilon^{\left(1 - \frac{1}{\beta^*}\right)} L^{-1} \right\rceil + 2\left(n_x - m_E\right) \left\lceil \left(\tau^*\right)^{\frac{1}{\beta^*}} \varepsilon^{\left(1 - \frac{1}{\beta^*}\right)} L^{-1} \right\rceil \Bigg) \prod_{k \in K} M_k.$$

Then, N is an upper bound on the number of boxes of width δ required to cover $\hat{X}_5$.

Proof. Theorem 5.3.11 can be used to obtain an overestimate of the number of boxes of width δ required to cover the projection of $\hat{X}_5$, as defined by Lemma 5.3.13, on $\mathbf{p}$, i.e., $\left\{\mathbf{p} \in \mathcal{N}^1_{\hat{\alpha}}(\mathbf{p}^*) : L\|\mathbf{p} - \mathbf{p}^*\|_1 \leq 2\varepsilon\right\}$, by replacing n_x with $n_x - m_E$ in the expressions for N. This estimate can be extended to obtain a conservative estimate of the number of boxes of width δ required to cover $\hat{X}_5$ as follows.

Note that ϕ_k is Lipschitz continuous on $\mathrm{cl}\left(\mathcal{N}^1_{\hat{\alpha}}(\mathbf{p}^*)\right)$ with Lipschitz constant $\frac{M_k}{\sqrt{n_x - m_E}}$, $\forall k \in \{1, \cdots, m_E\}$. Consider any box B of width δ that is used to cover the projection of $\hat{X}_5$ on $\mathbf{p}$. We have

$$w\left(\overline{\phi_k}\left(B \cap \mathrm{cl}\left(\mathcal{N}^1_{\hat{\alpha}}(\mathbf{p}^*)\right)\right)\right) \leq M_k \delta, \quad \forall k \in \{1, \cdots, m_E\},$$

from the Lipschitz continuity of ϕ_k. Therefore, we can replace the box B using $\prod_{k \in K} M_k$ such boxes and translate them appropriately to cover the region

$$\left\{(\mathbf{z},\mathbf{p}) \in \mathcal{N}^1_{\alpha_{\mathbf{z}}}(\mathbf{z}^*) \times \left(B \cap \mathcal{N}^1_{\hat{\alpha}}(\mathbf{p}^*)\right) : L\|\mathbf{p}-\mathbf{p}^*\|_1 \leq 2\varepsilon,\ \mathbf{z} = \boldsymbol{\phi}(\mathbf{p})\right\}.$$

Since $\bigcup_B \left\{B \cap \mathcal{N}^1_{\hat{\alpha}}(\mathbf{p}^*)\right\}$ covers the projection of $\hat{X}_5$ on $\mathbf{p}$, the desired result follows by multiplying the estimate obtained from Theorem 5.3.11 (with n_x replaced by $n_x - m_E$) by $\prod_{k \in K} M_k$. □

The next result provides a natural extension of Lemma 5.3.2 to the case when the objective function is not differentiable at the minimizer $\mathbf{x}^*$ [237]. Note that a similar result was derived for the case of unconstrained optimization in [237, Section 2.3] under alternative assumptions.

Lemma 5.3.15. Consider Problem (P). Suppose $\mathbf{x}^*$ is nonisolated, f is locally Lipschitz continuous on X and directionally differentiable at $\mathbf{x}^*$, and $\exists \alpha > 0$ such that

$$L := \inf_{\left\{\mathbf{d}:\|\mathbf{d}\|_1 = 1,\ \exists t>0\,:\,(\mathbf{x}^*+t\mathbf{d}) \in \mathcal{N}^1_{\alpha}(\mathbf{x}^*)\cap\mathcal{F}(X)\right\}} f'(\mathbf{x}^*;\mathbf{d}) > 0.$$

Then, $\exists \hat{\alpha} \in (0,\alpha]$ such that the region $\mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) \cap X_5$ can be conservatively approximated by

$$\hat{X}_5 = \left\{\mathbf{x} \in \mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) : L\|\mathbf{x}-\mathbf{x}^*\|_1 \leq 2\varepsilon\right\}.$$

Proof. Let $\mathbf{x} = \mathbf{x}^* + t\mathbf{d} \in \mathcal{N}^1_{\alpha}(\mathbf{x}^*) \cap \mathcal{F}(X)$ with $\|\mathbf{d}\|_1 = 1$ and $t = \|\mathbf{x}-\mathbf{x}^*\|_1 > 0$. We have (see Theorem 3.1.2 in [204])

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}^* + t\mathbf{d}) \\ &= f(\mathbf{x}^*) + f'(\mathbf{x}^*;(\mathbf{x}-\mathbf{x}^*)) + o(\|\mathbf{x}-\mathbf{x}^*\|_1) \\ &= f(\mathbf{x}^*) + tf'(\mathbf{x}^*;\mathbf{d}) + o(t) \\ &\geq f(\mathbf{x}^*) + Lt + o(t), \end{aligned}$$

where Step 2 follows from the directional differentiability of f at $\mathbf{x}^*$. Consequently, there exists $\hat{\alpha} \in (0,\alpha]$ such that for all $\mathbf{x} = \mathbf{x}^* + t\mathbf{d} \in \mathcal{F}(X)$ with $\|\mathbf{d}\|_1 = 1$ and $t \in [0,\hat{\alpha})$:

$$f(\mathbf{x}) \geq f(\mathbf{x}^*) + Lt + o(t) \geq f(\mathbf{x}^*) + \frac{L}{2}t.$$

Therefore, $\forall \mathbf{x} \in \mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) \cap X_5$ we have $\mathbf{x} = \mathbf{x}^* + t\mathbf{d} \in \mathcal{F}(X)$ with $\|\mathbf{d}\|_1 = 1$ and $t = \|\mathbf{x} - \mathbf{x}^*\|_1 < \hat{\alpha}$, and

$$\varepsilon \geq f(\mathbf{x}) - f(\mathbf{x}^*) \geq \frac{L}{2}t \implies Lt = L\|\mathbf{x} - \mathbf{x}^*\|_1 \leq 2\varepsilon. \qquad \square$$

Remark 5.3.16. Theorem 5.3.11 can be extended to the case when the assumption that the function f is differentiable at $\mathbf{x}^*$ is relaxed by using Lemmata 5.2.4 and 5.3.15 and Corollary 2.1 in [237] (also see Theorem 5.3.11). Similar to the differentiable case, the dependence of N on ε disappears when the lower bounding scheme has first-order convergence on $\mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) \cap \mathcal{F}(X)$, i.e., $\beta^* = 1$. Additionally, Lemma 5.3.13 and Corollary 5.3.14 can also be extended to the case when f is not differentiable at $\mathbf{x}^*$ under suitable assumptions.

Thus far, we have established conditions under which first-order convergence of the lower bounding scheme at feasible points is sufficient to mitigate the cluster problem on X_5. In the remainder of this section, we will present conditions under which second-order convergence of the lower bounding scheme is sufficient to mitigate clustering on X_5. The first result in this regard provides a conservative estimate of the subset of X_5 around a nonisolated $\mathbf{x}^*$ under the assumption that the objective function grows quadratically (or faster) on the feasible region in some neighborhood of $\mathbf{x}^*$.

Lemma 5.3.17. Consider Problem (P), and suppose f is twice-differentiable at $\mathbf{x}^*$. Suppose $\exists \alpha > 0, \gamma > 0$ such that

$$\nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} + \frac{1}{2}\mathbf{d}^{\mathrm{T}}\nabla^2 f(\mathbf{x}^*)\mathbf{d} \geq \gamma \mathbf{d}^{\mathrm{T}}\mathbf{d}, \quad \forall \mathbf{d} \in \left\{\mathbf{d} : (\mathbf{x}^* + \mathbf{d}) \in \mathcal{N}^2_{\alpha}(\mathbf{x}^*) \cap \mathcal{F}(X)\right\}.$$

Then $\exists \hat{\alpha} \in (0, \alpha]$ such that the region $\mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*) \cap X_5$ can be conservatively approximated by

$$\hat{X}_5 = \left\{\mathbf{x} \in \mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*) : \gamma\|\mathbf{x} - \mathbf{x}^*\|^2 \leq 2\varepsilon\right\}.$$

Furthermore, $\mathbf{x}^*$ is the unique global minimizer for Problem (P) on $\mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*)$.

Proof. Let $\mathbf{x} = \mathbf{x}^* + \mathbf{d} \in \mathcal{N}^2_\alpha(\mathbf{x}^*) \cap \mathcal{F}(X)$. We have

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}^* + \mathbf{d}) \\ &= f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} + \frac{1}{2}\mathbf{d}^{\mathrm{T}}\nabla^2 f(\mathbf{x}^*)\mathbf{d} + o(\|\mathbf{d}\|^2) \\ &\geq f(\mathbf{x}^*) + \gamma\mathbf{d}^{\mathrm{T}}\mathbf{d} + o(\|\mathbf{d}\|^2). \end{aligned}$$

Consequently, there exists $\hat{\alpha} \in (0, \alpha]$ such that for all $\mathbf{x} = \mathbf{x}^* + \mathbf{d} \in \mathcal{F}(X)$ with $\|\mathbf{d}\| \in [0, \hat{\alpha})$:

$$f(\mathbf{x}) \geq f(\mathbf{x}^*) + \gamma\mathbf{d}^{\mathrm{T}}\mathbf{d} + o(\|\mathbf{d}\|^2) \geq f(\mathbf{x}^*) + \frac{\gamma}{2}\mathbf{d}^{\mathrm{T}}\mathbf{d}. \tag{5.1}$$

Therefore, $\forall \mathbf{x} \in \mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*) \cap X_5$ we have $\mathbf{x} = \mathbf{x}^* + \mathbf{d} \in \mathcal{F}(X)$ with $\|\mathbf{d}\| < \hat{\alpha}$, and

$$\varepsilon \geq f(\mathbf{x}) - f(\mathbf{x}^*) \geq \frac{\gamma}{2}\mathbf{d}^{\mathrm{T}}\mathbf{d} \implies \gamma\|\mathbf{d}\|^2 = \gamma\|\mathbf{x} - \mathbf{x}^*\|^2 \leq 2\varepsilon.$$

The conclusion that $\mathbf{x}^*$ is the unique global minimizer for Problem (P) on $\mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*)$ follows from Equation (5.1). □

Remark 5.3.18.

1. Lemma 5.3.17 is not applicable when $\nexists \alpha > 0$ and $\gamma > 0$, for example $X = (-2, 2) \times (-2, 2)$, $m_I = 2$, $m_E = 0$, $f(\mathbf{x}) = x_2$, $g_1(\mathbf{x}) = x_1^4 - x_2$, $g_2(\mathbf{x}) = x_2 - 1$, and $\mathbf{x}^* = (0, 0)$. In this case, for any $\alpha > 0$, there exist directions from $\mathbf{x}^*$ to feasible points in which f grows slower than quadratically near $\mathbf{x}^*$.

2. For the case of unconstrained global optimization, the assumption of Lemma 5.3.17 reduces to the assumption that $\nabla^2 f(\mathbf{x}^*)$ is positive definite, and γ can be taken to be equal to half the smallest eigenvalue of $\nabla^2 f(\mathbf{x}^*)$ (see Theorem 1 in [238]). When the minimum is constrained, γ may potentially be estimated as follows. The first possibility is to directly estimate γ using a quadratic underestimator of f on $\mathcal{N}^2_\alpha(\mathbf{x}^*) \cap \mathcal{F}(X)$. If such an underestimator cannot be constructed easily, γ may still be estimated relatively easily when additional assumptions are satisfied.

   Suppose $(\mathbf{x}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$ is a KKT point, where $\boldsymbol{\mu}^*$ and $\boldsymbol{\lambda}^*$ correspond to Lagrange multipliers for $\mathbf{g}$ and $\mathbf{h}$, respectively, at $\mathbf{x}^*$. Consider the restricted Lagrangian $L(\mathbf{x}; \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$, and suppose it is positive definite for all $\mathbf{x} \in \mathrm{cl}(\mathcal{N}^2_\alpha(\mathbf{x}^*) \cap \mathcal{F}(X))$ (cf. [13, Section 4.4] and Theorem 2.3.21). Then γ may be estimated from the eigenvalues of $\nabla^2 L(\mathbf{x};\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)$ on $\mathrm{cl}(\mathcal{N}^2_{\alpha}(\mathbf{x}^*)\cap\mathcal{F}(X))$. This is a consequence of the fact that $f(\mathbf{x}) \geq L(\mathbf{x};\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)$, $\forall \mathbf{x} \in \mathcal{F}(X)$, by weak duality, $f(\mathbf{x}^*) = L(\mathbf{x}^*;\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)$, and the stationarity condition $\nabla_{\mathbf{x}} L(\mathbf{x};\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*) = \mathbf{0}$. Otherwise, if $(\mathbf{x}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)$ is a KKT point and some convex combination of f and $L(\cdot;\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)$ grows quadratically or faster on $\mathcal{N}^2_{\alpha}(\mathbf{x}^*)\cap\mathcal{F}(X)$, then γ can be estimated using one of its quadratic underestimators on $\mathcal{N}^2_{\alpha}(\mathbf{x}^*)\cap\mathcal{F}(X)$.

3. The key assumption of Lemma 5.3.17, which assumes that f grows quadratically or faster on the feasible region in some neighborhood of $\mathbf{x}^*$, is a relaxation of the key assumption of Lemma 5.3.2, which assumes that f grows linearly on the feasible region in some neighborhood of $\mathbf{x}^*$. While it was shown in Theorem 5.3.11 that first-order convergence of the lower bounding scheme at feasible points may be sufficient to mitigate clustering on X_5 under the assumptions of Lemma 5.3.2, Theorem 5.3.21, which will be presented shortly, shows that second-order convergence of the lower bounding scheme at feasible points may be sufficient to mitigate clustering on X_5 under the assumptions of Lemma 5.3.17. Consequently, the assumptions of Lemmata 5.3.2 and 5.3.17 can be viewed as belonging to a hierarchy of conditions for certain convergence orders of the lower bounding scheme at feasible points being sufficient to mitigate clustering on X_5, with the condition for third-order convergence of the lower bounding scheme at feasible points to be sufficient to mitigate clustering on X_5 amounting to the third-order Taylor expansion of f growing faster than cubically on the feasible region in some neighborhood of $\mathbf{x}^*$, and so on.

4. Along the line of discussion in Remark 5.3.3, $\hat{\alpha}$ depends on the local behavior of f around $\mathbf{x}^*$, but is independent of ε. Consequently, for sufficiently small ε we can conservatively approximate the set $\mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*)\cap X_5$ by $\left\{\mathbf{x}\in X : \gamma\|\mathbf{x}-\mathbf{x}^*\|^2 \leq 2\varepsilon\right\}$. Additionally, if the objective function f is either an affine or a quadratic function of $\mathbf{x}$, then its second-order Taylor expansion around $\mathbf{x}^*$ equals f itself and we can choose $\hat{\alpha} = \alpha$. Furthermore, $\mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*)\cap X_5$ can be conservatively approximated by the set $\hat{X}_5 = \left\{\mathbf{x}\in X : \gamma\|\mathbf{x}-\mathbf{x}^*\|^2 \leq \varepsilon\right\}$.

5. Similar to Proposition 5.3.4, a less conservative estimate of $\mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*)\cap X_5$ can be obtained

as

$$\hat{X}_5 = \Big\{ \mathbf{x} \in \mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*) \; : \gamma \|\mathbf{x} - \mathbf{x}^*\|^2 \leq 2\varepsilon, \\ \nabla f(\mathbf{x}^*)^{\mathrm{T}} (\mathbf{x} - \mathbf{x}^*) + \frac{1}{2} (\mathbf{x} - \mathbf{x}^*)^{\mathrm{T}} \nabla^2 f(\mathbf{x}^*) (\mathbf{x} - \mathbf{x}^*) \geq \gamma \|\mathbf{x} - \mathbf{x}^*\|^2 \Big\}.$$

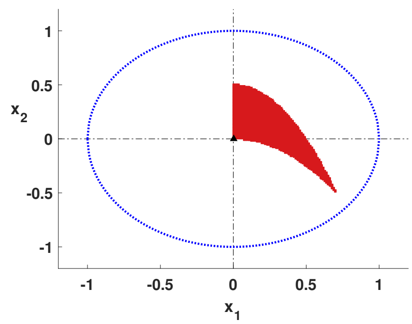
As an illustration of the application of Lemma 5.3.17, let us reconsider Example 5.2.9. Recall that $X = (0.4, 1.0) \times (0.5, 2.0)$, $m_I = 2$, $m_E = 1$, $f(\mathbf{x}) = -12x_1 - 7x_2 + x_2^2$, $g_1(\mathbf{x}) = x_1 - 0.9$, $g_2(\mathbf{x}) = 0.5 - x_1$, and $h(\mathbf{x}) = x_2 + 2x_1^4 - 2$ with $\mathbf{x}^* \approx (0.72, 1.47)$. Let $\varepsilon \leq 0.1$. We have $\mathcal{F}(X) = \{\mathbf{x} \in X : \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, h(\mathbf{x}) = 0\}$. Choose $\alpha = 0.1$, $\gamma = 2$, and $\hat{\alpha} = 0.1$ in Lemma 5.3.17. We have $X_5 = \left\{\mathbf{x} : x_2 = 2 - 2x_1^4, \; -12x_1 - 7x_2 + x_2^2 \leq f(\mathbf{x}^*) + \varepsilon\right\}$. From Lemma 5.3.17 and Remark 5.3.18, we have $\hat{X}_5 = \Big\{\mathbf{x} \in \mathcal{N}^2_{0.1}(\mathbf{x}^*) : \|\mathbf{x} - \mathbf{x}^*\|^2 \leq 0.5\varepsilon\Big\}$ (since f is quadratic). Note that an even better estimate of X_5 may be obtained using Lemma 5.3.23 by accounting for the fact that X_5 resides in a reduced-dimensional manifold.

The following examples illustrate two additional cases for which the assumptions of Lemma 5.3.17 hold.

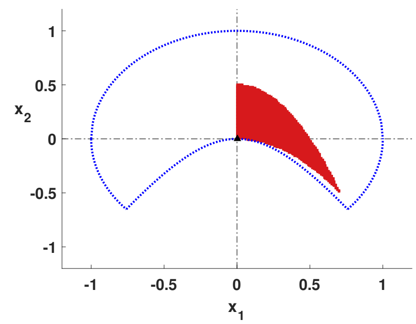
Example 5.3.19. Let $\varepsilon \leq 0.5$, $X = (-2, 2) \times (-2, 2)$, $m_I = 2$, and $m_E = 0$ with $f(\mathbf{x}) = x_2$, $g_1(\mathbf{x}) = x_1^2 - x_2$, $g_2(\mathbf{x}) = x_2 - 1$, and $\mathbf{x}^* = (0, 0)$. We have $\mathcal{F}(X) = \left\{\mathbf{x} : x_2 \geq x_1^2, \; x_2 \leq 1\right\}$. Choose $\alpha = 1$, $\gamma = 0.5$, and $\hat{\alpha} = 1$. From Remark 5.3.18, we have $X_5 = \left\{\mathbf{x} \in [-\sqrt{\varepsilon}, +\sqrt{\varepsilon}] \times [0, \varepsilon] : x_2 \geq x_1^2\right\} \subset \Big\{\mathbf{x} : \|\mathbf{x}\|^2 \leq 2\varepsilon\Big\} = \hat{X}_5$.

Example 5.3.20. Let $\varepsilon \leq 0.5$, $X = (-2, 2) \times (-2, 2)$, $m_I = 3$, and $m_E = 0$ with $f(\mathbf{x}) = 2x_1^2 + x_2$, $g_1(\mathbf{x}) = -x_1^2 - x_2$, $g_2(\mathbf{x}) = -x_1$, $g_3(\mathbf{x}) = x_1^2 + x_2^2 - 1$, and $\mathbf{x}^* = (0, 0)$. We have $\mathcal{F}(X) = \left\{\mathbf{x} : x_2 \geq -x_1^2, \; x_1 \geq 0, \; x_1^2 + x_2^2 \leq 1\right\}$ with $\alpha = 1$, $\gamma = 0.5$, $\hat{\alpha} = 1$, and $X_5 = \left\{\mathbf{x} : x_2 + 2x_1^2 \leq \varepsilon, x_2 \geq -x_1^2, x_1 \geq 0\right\} \subset \Big\{\mathbf{x} : \|\mathbf{x}\|^2 \leq 2\varepsilon\Big\} = \hat{X}_5$ (see Remark 5.3.18).

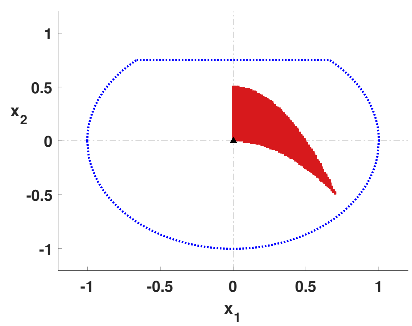
The overconservatism of the estimate $\hat{X}_5$ in the above two examples (with regards to its dependence on ε) is primarily due to the fact that the linear growth of the objective function in the direction of its gradient is not taken into account. This observation is formalized and taken advantage of in Lemma 5.3.25 to obtain a less conservative estimate. Figure 5-4 plots X_5 and $\hat{X}_5$, obtained using different estimation techniques, for $\varepsilon = 0.5$ and $\varepsilon = 0.1$ in Example 5.3.20. The benefit of using the estimate in Remark 5.3.18 over that of Lemma 5.3.17 is seen from Figures 5-4a and 5-4b, and the benefit of using the estimate from Lemma 5.3.25 (using $\rho_1 = 3$, $\rho_2 = 1.5$) over that of Lemma 5.3.17 is seen from Figures 5-4a



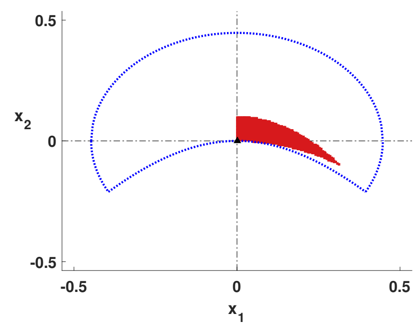
(a) X_5 & $\hat{X}_5$ from Lemma 5.3.17 for $\varepsilon = 0.5$



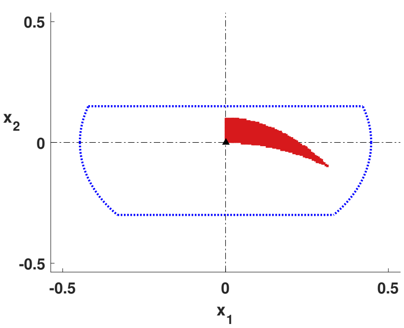
(b) X_5 & $\hat{X}_5$ from Remark 5.3.18 for $\varepsilon = 0.5$



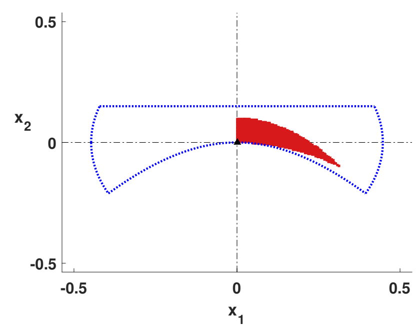
(c) X_5 & $\hat{X}_5$ from Lemma 5.3.25 for $\varepsilon = 0.5$



(d) X_5 & $\hat{X}_5$ from Remark 5.3.18 for $\varepsilon = 0.1$



(e) X_5 & $\hat{X}_5$ from Lemma 5.3.25 for $\varepsilon = 0.1$



(f) X_5 & $\hat{X}_5$ from Lemma 5.3.25 and Remark 5.3.18 for $\varepsilon = 0.1$

Figure 5-4: Plots of X_5 (solid regions) and $\hat{X}_5$ (area between the dotted lines) for Example 5.3.20. The filled-in triangles correspond to the minimizer $\mathbf{x}^*$, and the dash-dotted lines represent the axes translated to $\mathbf{x}^*$.

and 5-4c. It can be observed from Figure 5-4c that the constraint $-\rho_1\varepsilon \le \nabla f(\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}^*)$ in Lemma 5.3.25 is not active on the region $\left\{\mathbf{x} : \gamma\|\mathbf{x}-\mathbf{x}^*\|^2 \le \varepsilon\right\}$ for $\varepsilon = 0.5$. To illustrate the benefit of this constraint in Lemma 5.3.25, we consider $\varepsilon = 0.1$. Figures 5-4d and 5-4e demonstrate the advantages of using the estimates in Remark 5.3.18 and Lemma 5.3.25, respectively, over the estimate in Lemma 5.3.17, and Figure 5-4f combines the benefits of the estimates from Lemma 5.3.25 and Remark 5.3.18 by using the estimate

$$\begin{aligned}\hat{X}_5 = \Big\{\mathbf{x} \in \mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*) \ : \gamma\|\mathbf{x}-\mathbf{x}^*\|^2 \le 2\varepsilon,\ -\rho_1\varepsilon \le \nabla f(\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}^*) \le \rho_2\varepsilon, \\ \nabla f(\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}^*) + \frac{1}{2}(\mathbf{x}-\mathbf{x}^*)^{\mathrm{T}}\nabla^2 f(\mathbf{x}^*)(\mathbf{x}-\mathbf{x}^*) \ge \gamma\|\mathbf{x}-\mathbf{x}^*\|^2\Big\}.\end{aligned}$$

The following theorem follows from Lemma 3 in [238], and provides a conservative estimate of the number of boxes of width δ required to cover the estimate $\hat{X}_5$ from Lemma 5.3.17. Consequently, from Lemma 5.2.4 and the theorem below, we can get a conservative estimate of the number of boxes required to cover $\mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*) \cap X_5$ and estimate the extent of the cluster problem on that region.

Theorem 5.3.21. Consider Problem (P), and suppose the assumptions of Lemma 5.3.17 hold. Let $\delta = \left(\frac{\varepsilon}{\tau^*}\right)^{\frac{1}{\beta^*}}$ and $r = \sqrt{\frac{2\varepsilon}{\gamma}}$.

1. If $\delta \ge 2r$, let $N = 1$.
2. If $\frac{2r}{\sqrt{m-1}} > \delta \ge \frac{2r}{\sqrt{m}}$ for some $m \in \mathbb{N}$ with $m \le n_x$ and $2 \le m \le 18$, then let

$$N = \sum_{i=0}^{m-1} 2^i \binom{n_x}{i} + 2n_x \left\lceil \frac{m-9}{9} \right\rceil.$$

3. Otherwise, let

$$\begin{aligned}N = \left\lceil 2\,(\tau^*)^{\frac{1}{\beta^*}}\, \varepsilon^{\left(\frac{1}{2}-\frac{1}{\beta^*}\right)} \gamma^{-\frac{1}{2}} \right\rceil^{n_x-1} \Bigg(\left\lceil 2\,(\tau^*)^{\frac{1}{\beta^*}}\, \varepsilon^{\left(\frac{1}{2}-\frac{1}{\beta^*}\right)} \gamma^{-\frac{1}{2}} \right\rceil + \\ 2n_x \left\lceil (\sqrt{2}-1)\,(\tau^*)^{\frac{1}{\beta^*}}\, \varepsilon^{\left(\frac{1}{2}-\frac{1}{\beta^*}\right)} \gamma^{-\frac{1}{2}} \right\rceil \Bigg).\end{aligned}$$

Then, N is an upper bound on the number of boxes of width δ required to cover $\mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*) \cap X_5$.

Proof. From Lemma 5.3.17, we have that the set $\hat{X}_5 = \left\{\mathbf{x} \in \mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*) : \gamma\|\mathbf{x}-\mathbf{x}^*\|^2 \le 2\varepsilon\right\}$ provides a conservative estimate of $\mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*) \cap X_5$. The desired result follows from Lemma 3 in [238]. $\square$

For the case of unconstrained global optimization, Theorem 5.3.21 effectively reduces to Theorem 1 in [238] with γ equal to half the smallest eigenvalue of $\nabla^2 f(\mathbf{x}^*)$ $\big($note that there is a 'factor of two difference' from the analysis in [238] because we consider an appropriate $\hat{\alpha} \in (0, \alpha]\big)$.

Remark 5.3.22. Under the assumptions of Theorem 5.3.21, the dependence of N on ε disappears when the lower bounding scheme has second-order convergence on $\mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*) \cap \mathcal{F}(X)$. This is similar to the case of unconstrained global optimization where at least second-order convergent lower bounding schemes are required to eliminate the cluster problem.

Finally, we present two sets of additional assumptions over those of Lemma 5.3.17 under which less conservative estimates of the cluster problem on X_5 can be obtained. The first result in this regard, similar to Lemma 5.3.13, refines the analysis of Lemma 5.3.17 when Problem (P) contains equality constraints that can locally be eliminated using the implicit function theorem [192].

Lemma 5.3.23. Consider Problem (P) with $1 \leq m_E < n_x$. Suppose f is twice-differentiable at $\mathbf{x}^*$, and $\exists \alpha > 0$, $\gamma > 0$ such that $\mathbf{h}$ is continuously differentiable on $\mathcal{N}^2_{\alpha}(\mathbf{x}^*)$ and

$$\nabla f(\mathbf{x}^*)^{\mathrm{T}} \mathbf{d} + \frac{1}{2} \mathbf{d}^{\mathrm{T}} \nabla^2 f(\mathbf{x}^*) \mathbf{d} \geq \gamma \mathbf{d}^{\mathrm{T}} \mathbf{d}, \quad \forall \mathbf{d} \in \left\{ \mathbf{d} : (\mathbf{x}^* + \mathbf{d}) \in \mathcal{N}^2_{\alpha}(\mathbf{x}^*) \cap \mathcal{F}(X) \right\}.$$

Furthermore, suppose the variables $\mathbf{x}$ can be reordered and partitioned into dependent variables $\mathbf{z} \in \mathbb{R}^{m_E}$ and independent variables $\mathbf{p} \in \mathbb{R}^{n_x - m_E}$, with $\mathbf{x} \equiv (\mathbf{z}, \mathbf{p})$, such that $\nabla_{\mathbf{z}} \mathbf{h}((\mathbf{z}, \mathbf{p}))$ is nonsingular on $\mathcal{N}^2_{\alpha}((\mathbf{z}^*, \mathbf{p}^*))$, where $\mathbf{x}^* \equiv (\mathbf{z}^*, \mathbf{p}^*)$. Then, $\exists \alpha_{\mathbf{p}}, \alpha_{\mathbf{z}} \in (0, \alpha]$, a continuously differentiable function $\boldsymbol{\phi} : \mathcal{N}^2_{\alpha_{\mathbf{p}}}(\mathbf{p}^*) \to \mathcal{N}^2_{\alpha_{\mathbf{z}}}(\mathbf{z}^*)$, and $\hat{\alpha} \in (0, \alpha_{\mathbf{p}})$ such that the region $\big(\mathcal{N}^2_{\alpha_{\mathbf{z}}}(\mathbf{z}^*) \times \mathcal{N}^2_{\hat{\alpha}}(\mathbf{p}^*)\big) \cap X_5$ can be conservatively approximated by

$$\hat{X}_5 = \left\{ (\mathbf{z}, \mathbf{p}) \in \mathcal{N}^2_{\alpha_{\mathbf{z}}}(\mathbf{z}^*) \times \mathcal{N}^2_{\hat{\alpha}}(\mathbf{p}^*) : \mathbf{z} = \boldsymbol{\phi}(\mathbf{p}),\ \gamma \|\mathbf{p} - \mathbf{p}^*\|^2 \leq 2\varepsilon \right\}.$$

Proof. The result follows from the proof of Lemma 5.3.17 and the implicit function theorem [192, Chapter 9]. $\square$

Lemma 5.3.23 can be used to obtain a less conservative estimate of the number of boxes of width δ required to cover $\hat{X}_5$ as shown in the following corollary of Theorem 5.3.21. It provides an upper bound on the number of boxes of width δ required to cover X_5 that scales as $O\left(\varepsilon^{(n_x-m_E)\left(\frac{1}{2}-\frac{1}{\beta^*}\right)}\right)$ in contrast to the scaling $O\left(\varepsilon^{n_x\left(\frac{1}{2}-\frac{1}{\beta^*}\right)}\right)$ from Theorem 5.3.21.

Corollary 5.3.24. Suppose the assumptions of Lemma 5.3.23 hold. Let $\delta = \left(\frac{\varepsilon}{\tau^*}\right)^{\frac{1}{\beta^*}}$ and $r = \sqrt{\frac{2\varepsilon}{\gamma}}$. Define

$$M_k := \left(\max_{\mathbf{p}\in\mathrm{cl}\left(\mathcal{N}^2_{\hat{\alpha}}(\mathbf{p}^*)\right)} \|\nabla\phi_k(\mathbf{p})\|\right)\sqrt{n_x - m_E}, \quad \forall k \in \{1, \cdots, m_E\},$$
$$K := \{k \in \{1, \cdots, m_E\} : M_k > 1\}.$$

1. If $\delta \geq 2r$, let $N = \prod_{k\in K} M_k$.
2. If $\frac{2r}{\sqrt{m-1}} > \delta \geq \frac{2r}{\sqrt{m}}$ for some $m \in \mathbb{N}$ with $m \leq n_x - m_E$ and $2 \leq m \leq 18$, then let
$$N = \left(\sum_{i=0}^{m-1} 2^i \binom{n_x - m_E}{i} + 2\left(n_x - m_E\right)\left\lceil\frac{m-9}{9}\right\rceil\right)\prod_{k\in K} M_k.$$
3. Otherwise, let
$$N = \left\lceil 2\left(\tau^*\right)^{\frac{1}{\beta^*}}\varepsilon^{\left(\frac{1}{2}-\frac{1}{\beta^*}\right)}\gamma^{-\frac{1}{2}}\right\rceil^{n_x-m_E-1}\left(\left\lceil 2\left(\tau^*\right)^{\frac{1}{\beta^*}}\varepsilon^{\left(\frac{1}{2}-\frac{1}{\beta^*}\right)}\gamma^{-\frac{1}{2}}\right\rceil + 2\left(n_x - m_E\right)\left\lceil\left(\sqrt{2}-1\right)\left(\tau^*\right)^{\frac{1}{\beta^*}}\varepsilon^{\left(\frac{1}{2}-\frac{1}{\beta^*}\right)}\gamma^{-\frac{1}{2}}\right\rceil\right)\prod_{k\in K} M_k.$$

Then, N is an upper bound on the number of boxes of width δ required to cover $\hat{X}_5$.

Proof. The proof is similar to the proof of Corollary 5.3.14, and is therefore omitted. $\square$

The next result refines the analysis of Lemma 5.3.17 further, in part by accounting for the fact that f grows linearly around $\mathbf{x}^*$ in the direction of its gradient.

Lemma 5.3.25. Consider Problem (P), and suppose the assumptions of Lemma 5.3.17 hold. Then $\exists\hat{\alpha} \in (0, \alpha]$ and constants $\rho_1, \rho_2 \geq 0$ such that the region $\mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*) \cap X_5$ can be conservatively approximated by

$$\hat{X}_5 = \left\{\mathbf{x} \in \mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*) : \gamma\|\mathbf{x} - \mathbf{x}^*\|^2 \leq 2\varepsilon,\ -\rho_1\varepsilon \leq \nabla f(\mathbf{x}^*)^{\mathrm{T}}\left(\mathbf{x} - \mathbf{x}^*\right) \leq \rho_2\varepsilon\right\}.$$

Proof. The result trivially follows from Lemma 5.3.17 when $\nabla f(\mathbf{x}^*) = \mathbf{0}$.

Suppose $\nabla f(\mathbf{x}^*) \neq \mathbf{0}$. From Lemma 5.3.17, we have

$$\mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*) \cap X_5 \subset \left\{ \mathbf{x} \in \mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*) : \gamma \|\mathbf{x} - \mathbf{x}^*\|^2 \leq 2\varepsilon \right\}. \tag{5.2}$$

Suppose we represent each $\mathbf{x} \in \mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*) \cap \mathcal{F}(X)$ by $\mathbf{x} := \mathbf{x}^* + \beta_1 \nabla f(\mathbf{x}^*) + \beta_2 \mathbf{d}$, where $\beta_1, \beta_2 \in \mathbb{R}$ and $\mathbf{d} \perp \nabla f(\mathbf{x}^*)$ with $\|\mathbf{d}\| = 1$. Consider the case when $\beta_1 \geq 0$. We have

$$\begin{aligned}
f(\mathbf{x}) - f(\mathbf{x}^*) &= \nabla f(\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x} - \mathbf{x}^*) + \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^{\mathrm{T}} \nabla^2 f(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) + o\left(\|\mathbf{x} - \mathbf{x}^*\|^2\right) \\
&= \nabla f(\mathbf{x}^*)^{\mathrm{T}} \left(\beta_1 \nabla f(\mathbf{x}^*) + \beta_2 \mathbf{d}\right) + \\
&\qquad \frac{1}{2}(\beta_1 \nabla f(\mathbf{x}^*) + \beta_2 \mathbf{d})^{\mathrm{T}} \nabla^2 f(\mathbf{x}^*) \left(\beta_1 \nabla f(\mathbf{x}^*) + \beta_2 \mathbf{d}\right) + o\left(\beta_1^2 + \beta_2^2\right) \\
&= \beta_1 \|\nabla f(\mathbf{x}^*)\|^2 + \frac{1}{2}(\beta_1 \nabla f(\mathbf{x}^*) + \beta_2 \mathbf{d})^{\mathrm{T}} \nabla^2 f(\mathbf{x}^*) \left(\beta_1 \nabla f(\mathbf{x}^*) + \beta_2 \mathbf{d}\right) + \\
&\qquad o\left(\beta_1^2 + \beta_2^2\right).
\end{aligned}$$

Therefore, $\forall \mathbf{x} \in \mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*) \cap X_5$ with $\mathbf{x} = \mathbf{x}^* + \beta_1 \nabla f(\mathbf{x}^*) + \beta_2 \mathbf{d}$, $\beta_1 \geq 0, \beta_2 \in \mathbb{R}$ and $\mathbf{d} \perp \nabla f(\mathbf{x}^*)$ with $\|\mathbf{d}\| = 1$, we have

$$\begin{aligned}
&\beta_1 \|\nabla f(\mathbf{x}^*)\|^2 + \frac{1}{2}(\beta_1 \nabla f(\mathbf{x}^*) + \beta_2 \mathbf{d})^{\mathrm{T}} \nabla^2 f(\mathbf{x}^*) \left(\beta_1 \nabla f(\mathbf{x}^*) + \beta_2 \mathbf{d}\right) + o\left(\beta_1^2 + \beta_2^2\right) \leq \varepsilon \\
\Longrightarrow\ & \beta_1 \|\nabla f(\mathbf{x}^*)\|^2 \leq \varepsilon - \frac{1}{2}(\beta_1 \nabla f(\mathbf{x}^*) + \beta_2 \mathbf{d})^{\mathrm{T}} \nabla^2 f(\mathbf{x}^*) \left(\beta_1 \nabla f(\mathbf{x}^*) + \beta_2 \mathbf{d}\right) - o\left(\beta_1^2 + \beta_2^2\right) \\
\Longrightarrow\ & \beta_1 \|\nabla f(\mathbf{x}^*)\|^2 \leq \varepsilon - \frac{1}{2}(\beta_1 \nabla f(\mathbf{x}^*) + \beta_2 \mathbf{d})^{\mathrm{T}} \nabla^2 f(\mathbf{x}^*) \left(\beta_1 \nabla f(\mathbf{x}^*) + \beta_2 \mathbf{d}\right) + \\
&\qquad \frac{\gamma}{2}\left(\beta_1^2 \|\nabla f(\mathbf{x}^*)\|^2 + \beta_2^2\right),
\end{aligned} \tag{5.3}$$

where the last step uses the fact that $\hat{\alpha}$ is chosen such that (see Equation (5.1))

$$o\left(\beta_1^2 + \beta_2^2\right) \geq -\frac{\gamma}{2}\left(\beta_1^2 \|\nabla f(\mathbf{x}^*)\|^2 + \beta_2^2\right).$$

Note that $\beta_1 \leq \sqrt{\dfrac{2\varepsilon}{\gamma \|\nabla f(\mathbf{x}^*)\|^2}}$ and $|\beta_2| \leq \sqrt{\dfrac{2\varepsilon}{\gamma}}$ follow from Equation (5.2). The right hand side of Equation (5.3) is $O(\varepsilon)$ since $\beta_1 = \beta_2 = O(\sqrt{\varepsilon})$, thereby establishing the existence of $\rho_2 \geq 0$.

Next, suppose $\beta_1 \leq 0$. From the assumptions of Lemma 5.3.17, we have for each $\mathbf{x} \in \mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*) \cap X_5$ with $\mathbf{x} = \mathbf{x}^* + \beta_1 \nabla f(\mathbf{x}^*) + \beta_2 \mathbf{d}$, $\beta_1 \leq 0, \beta_2 \in \mathbb{R}$ and $\mathbf{d} \perp \nabla f(\mathbf{x}^*)$ with

$\|\mathbf{d}\| = 1$:

$$
\begin{aligned}
&\nabla f(\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}^*) + \frac{1}{2}(\mathbf{x}-\mathbf{x}^*)^{\mathrm{T}}\nabla^2 f(\mathbf{x}^*)(\mathbf{x}-\mathbf{x}^*) \geq \gamma\|\mathbf{x}-\mathbf{x}^*\|^2 \\
\implies &\nabla f(\mathbf{x}^*)^{\mathrm{T}}\left(\beta_1\nabla f(\mathbf{x}^*) + \beta_2\mathbf{d}\right) + \frac{1}{2}(\beta_1\nabla f(\mathbf{x}^*) + \beta_2\mathbf{d})^{\mathrm{T}}\nabla^2 f(\mathbf{x}^*)\left(\beta_1\nabla f(\mathbf{x}^*) + \beta_2\mathbf{d}\right) \\
&\geq \gamma\left(\beta_1^2\|\nabla f(\mathbf{x}^*)\|^2 + \beta_2^2\right) \\
\implies &\beta_1\|\nabla f(\mathbf{x}^*)\|^2 + \frac{1}{2}(\beta_1\nabla f(\mathbf{x}^*) + \beta_2\mathbf{d})^{\mathrm{T}}\nabla^2 f(\mathbf{x}^*)\left(\beta_1\nabla f(\mathbf{x}^*) + \beta_2\mathbf{d}\right) \\
&\geq \gamma\left(\beta_1^2\|\nabla f(\mathbf{x}^*)\|^2 + \beta_2^2\right) \\
\implies &\frac{1}{2}(\beta_1\nabla f(\mathbf{x}^*) + \beta_2\mathbf{d})^{\mathrm{T}}\nabla^2 f(\mathbf{x}^*)\left(\beta_1\nabla f(\mathbf{x}^*) + \beta_2\mathbf{d}\right) - \gamma\left(\beta_1^2\|\nabla f(\mathbf{x}^*)\|^2 + \beta_2^2\right) \\
&\geq -\beta_1\|\nabla f(\mathbf{x}^*)\|^2, \qquad (5.4)
\end{aligned}
$$

and $\beta_1 \geq -\sqrt{\dfrac{2\varepsilon}{\gamma\|\nabla f(\mathbf{x}^*)\|^2}}$, $|\beta_2| \leq \sqrt{\dfrac{2\varepsilon}{\gamma}}$ from Equation (5.2). The left hand side of Equation (5.4) is $O(\varepsilon)$ since $\beta_1 = \beta_2 = O(\sqrt{\varepsilon})$, thereby establishing the existence of $\rho_1 \geq 0$. $\square$

The previous lemma can be used to obtain a less conservative estimate of the number of boxes of width δ required to cover $\hat{X}_5$ when ε is sufficiently-small and the convergence order $\beta^* > 1$. This is presented in the following corollary of Theorem 5.3.21, which provides an upper bound on the number of boxes of width δ required to cover X_5 that scales as $O\left(\varepsilon^{(n_x-1)\left(\frac{1}{2}-\frac{1}{\beta^*}\right)}\right)$ in contrast to the scaling $O\left(\varepsilon^{n_x\left(\frac{1}{2}-\frac{1}{\beta^*}\right)}\right)$ from Theorem 5.3.21.

Corollary 5.3.26. Suppose the assumptions of Lemma 5.3.25 hold. Let $\delta = \left(\dfrac{\varepsilon}{\tau^*}\right)^{\frac{1}{\beta^*}}$ and $r = \sqrt{\dfrac{2\varepsilon}{\gamma}}$. Suppose $\beta^* > 1$, ε is sufficiently-small that $(\rho_1+\rho_2)\varepsilon \ll \delta$, and $\nabla f(\mathbf{x}^*) \neq \mathbf{0}$.

1. If $\delta \geq 2r$, let $N = 1$.

2. If $\dfrac{2r}{\sqrt{m-1}} > \delta \geq \dfrac{2r}{\sqrt{m}}$ for some $m \in \mathbb{N}$ with $m \leq n_x - 1$ and $2 \leq m \leq 18$, then let

$$
N = \sum_{i=0}^{m-1} 2^i \binom{n_x - 1}{i} + 2\,(n_x - 1)\left\lceil \frac{m-9}{9} \right\rceil.
$$

3. Otherwise, let

$$N = \left\lceil 2\,(\tau^*)^{\frac{1}{\beta^*}}\,\varepsilon^{\left(\frac{1}{2}-\frac{1}{\beta^*}\right)}\gamma^{-\frac{1}{2}} \right\rceil^{n_x-2} \Bigg(\left\lceil 2\,(\tau^*)^{\frac{1}{\beta^*}}\,\varepsilon^{\left(\frac{1}{2}-\frac{1}{\beta^*}\right)}\gamma^{-\frac{1}{2}} \right\rceil + 2\,(n_x-1) \left\lceil (\sqrt{2}-1)\,(\tau^*)^{\frac{1}{\beta^*}}\,\varepsilon^{\left(\frac{1}{2}-\frac{1}{\beta^*}\right)}\gamma^{-\frac{1}{2}} \right\rceil \Bigg).$$

Then, N is an upper bound on the number of boxes of width δ required to cover $\hat{X}_5$.

Proof. We have from Lemma 5.3.25 that the region $\mathcal{N}^2_{\hat{\alpha}}(\mathbf{x}^*) \cap X_5$ is conservatively estimated by a sphere with radius $= O(\sqrt{\varepsilon})$ truncated by the hyperplanes $\nabla f(\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}^*) \leq \rho_2\varepsilon$ and $\nabla f(\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}^*) \geq -\rho_1\varepsilon$. Therefore, when ε is chosen to be small enough that $(\rho_1+\rho_2)\varepsilon \ll \delta$, the desired result follows from Theorem 5.3.21 and the fact that any covering of the projection of $\hat{X}_5$ on to the subspace perpendicular to $\nabla f(\mathbf{x}^*)$ with boxes of width δ can be directly extended to cover $\hat{X}_5$ without using additional boxes. $\square$

Note that Corollary 5.3.26 can also be extended to the case when $0 < \beta^* \leq 1$, in which case the estimate N may additionally depend on the values of ρ_1 and ρ_2.

5.3.2 Estimates for the number of boxes required to cover $X_3 \backslash B_\delta$

This section assumes that Problem (P) has a finite number of global minimizers, and ε is small enough that X_3 is guaranteed to be contained in neighborhoods of constrained global minimizers under additional assumptions. An estimate for the number of boxes of certain widths required to cover some neighborhood of a constrained minimum $\mathbf{x}^*$ that contains the subset of X_3 around $\mathbf{x}^*$ is provided under suitable assumptions. An estimate for the number of boxes required to cover X_3 can be obtained by summing the above estimates over the set of constrained global minimizers. Throughout this section, we assume that $\mathbf{x}^*$ is a constrained global minimizer; otherwise $\exists \alpha > 0$ such that $\mathcal{N}^2_\alpha(\mathbf{x}^*) \cap X_3 = \emptyset$. Furthermore, we assume that $\mathbf{x}^*$ is at the center of a single box B_δ of width $\delta = \left(\frac{\varepsilon}{\tau^*}\right)^{\frac{1}{\beta^*}}$ placed while covering $\hat{X}_5$ (see Remark 5.3.28 for the reason for this assumption).

The first result in this section provides a conservative estimate of the subset of X_3 around a constrained minimizer $\mathbf{x}^*$ under the following assumption: the infeasible region in some neighborhood of $\mathbf{x}^*$ can be split into two subregions such that the objective function grows linearly in the first subregion and the measure of infeasibility grows linearly in the second subregion.

Lemma 5.3.27. Consider Problem (P). Suppose $\mathbf{x}^*$ is a constrained minimizer, and the functions f, g_j, $\forall j \in \mathcal{A}(\mathbf{x}^*)$, and h_k, $\forall k \in \{1, \cdots, m_E\}$, are locally Lipschitz continuous on X and directionally differentiable at $\mathbf{x}^*$. Furthermore, suppose $\exists \alpha > 0$ and a set $\mathcal{D}_0$ such that

$$L_f = \inf_{\mathbf{d} \in \mathcal{D}_0 \cap \mathcal{D}_I} f'(\mathbf{x}^*; \mathbf{d}) > 0,$$

$$L_I = \inf_{\mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0} \max \left\{ \max_{j \in \mathcal{A}(\mathbf{x}^*)} g_j'(\mathbf{x}^*; \mathbf{d}), \max_{k \in \{1, \cdots, m_E\}} \left| h_k'(\mathbf{x}^*; \mathbf{d}) \right| \right\} > 0,$$

where $\mathcal{D}_I$ is defined as

$$\mathcal{D}_I = \left\{ \mathbf{d} : \|\mathbf{d}\|_1 = 1, \, \exists t > 0 \, : \, (\mathbf{x}^* + t\mathbf{d}) \in \mathcal{N}_\alpha^1(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} \right\}.$$

Then, $\exists \hat{\alpha} \in (0, \alpha]$ such that the region

$$X_3^1 := \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap X_3 \cap \left\{ \mathbf{x} = (\mathbf{x}^* + t\mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_0 \cap \mathcal{D}_I, t > 0 \right\}$$

can be conservatively approximated as $\hat{X}_3^1 = \left\{ \mathbf{x} \in \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) : L_f \|\mathbf{x} - \mathbf{x}^*\|_1 \leq 2\varepsilon^o \right\}$, and the region

$$X_3^2 := \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap X_3 \cap \left\{ \mathbf{x} = (\mathbf{x}^* + t\mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0, t > 0 \right\}$$

can be conservatively approximated as $\hat{X}_3^2 = \left\{ \mathbf{x} \in \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) : L_I \|\mathbf{x} - \mathbf{x}^*\|_1 \leq 2\varepsilon^f \right\}$. Furthermore, suppose $\mathbf{x}^*$ is at the center of a box, B_δ, of width $\delta = \left(\frac{\varepsilon}{\tau^*} \right)^{\frac{1}{\beta^*}}$ placed while covering $\hat{X}_5$. Then, the region

$$X_3^2 \backslash B_\delta = \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap X_3 \cap \left\{ \mathbf{x} = (\mathbf{x}^* + t\mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0, t > 0 \right\} \Big\backslash B_\delta$$

is conservatively characterized by

$$\left\{ \mathbf{x} \in \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) : d \left(\begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix} (\mathbf{x}), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\} \right) \in \left(\frac{L_I}{4} \delta, \varepsilon^f \right] \right\}$$

whenever $L_I \delta < 4\varepsilon^f$.

Proof. Let $\mathbf{x} = \mathbf{x}^* + t\mathbf{d} \in \mathcal{N}_\alpha^1(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}}$ with $\|\mathbf{d}\|_1 = 1$, $\mathbf{d} \in \mathcal{D}_0$, and $t = \|\mathbf{x} - \mathbf{x}^*\|_1 > 0$. We have (see Theorem 3.1.2 in [204])

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}^* + t\mathbf{d}) \\ &= f(\mathbf{x}^*) + f'(\mathbf{x}^*; (\mathbf{x} - \mathbf{x}^*)) + o(\|\mathbf{x} - \mathbf{x}^*\|_1) \\ &= f(\mathbf{x}^*) + tf'(\mathbf{x}^*; \mathbf{d}) + o(t) \\ &\geq f(\mathbf{x}^*) + L_f t + o(t). \end{aligned}$$

Consequently, there exists $\hat{\alpha}_0 \in (0, \alpha]$ such that for all $\mathbf{x} = \mathbf{x}^* + t\mathbf{d} \in (\mathcal{F}(X))^{\mathrm{C}}$ with $\|\mathbf{d}\|_1 = 1$, $\mathbf{d} \in \mathcal{D}_0$ and $t \in [0, \hat{\alpha}_0)$:

$$f(\mathbf{x}) \geq f(\mathbf{x}^*) + L_f t + o(t) \geq f(\mathbf{x}^*) + \frac{L_f}{2} t.$$

Next, consider $\mathbf{x} = \mathbf{x}^* + t\mathbf{d} \in \mathcal{N}_\alpha^1(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}}$ with $\|\mathbf{d}\|_1 = 1$, $\mathbf{d} \notin \mathcal{D}_0$, and $t = \|\mathbf{x} - \mathbf{x}^*\|_1 > 0$. We have

$$\begin{aligned} &\max\left\{\max_{j \in \mathcal{A}(\mathbf{x}^*)} \{g_j(\mathbf{x})\}, \max_{k \in \{1,\cdots,m_E\}} \{|h_k(\mathbf{x})|\}\right\} \\ =&\max\left\{\max_{j \in \mathcal{A}(\mathbf{x}^*)} \{g_j(\mathbf{x}^* + t\mathbf{d})\}, \max_{k \in \{1,\cdots,m_E\}} \{|h_k(\mathbf{x}^* + t\mathbf{d})|\}\right\} \\ =&\max\left\{\max_{j \in \mathcal{A}(\mathbf{x}^*)} \left\{tg_j'(\mathbf{x}^*; \mathbf{d}) + o(t)\right\}, \max_{k \in \{1,\cdots,m_E\}} \left\{\left|th_k'(\mathbf{x}^*; \mathbf{d}) + o(t)\right|\right\}\right\}. \end{aligned}$$

Consequently, there exists $\hat{\alpha}_1 \in (0, \alpha]$ such that for all $\mathbf{x} = \mathbf{x}^* + t\mathbf{d} \in (\mathcal{F}(X))^{\mathrm{C}}$ with $\|\mathbf{d}\|_1 = 1$, $\mathbf{d} \notin \mathcal{D}_0$ and $t \in [0, \hat{\alpha}_1)$:

$$\begin{aligned} &d\left(\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}(\mathbf{x}), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\}\right) \\ \geq&\max\left\{\max_{j \in \mathcal{A}(\mathbf{x}^*)} \{g_j(\mathbf{x})\}, \max_{k \in \{1,\cdots,m_E\}} \{|h_k(\mathbf{x})|\}\right\} \\ =&\max\left\{\max_{j \in \mathcal{A}(\mathbf{x}^*)} \left\{tg_j'(\mathbf{x}^*; \mathbf{d}) + o(t)\right\}, \max_{k \in \{1,\cdots,m_E\}} \left\{\left|th_k'(\mathbf{x}^*; \mathbf{d}) + o(t)\right|\right\}\right\} \\ \geq&\frac{L_I}{2} t, \end{aligned}$$

where Step 1 follows from the fact that $\|\mathbf{z}\| \geq \|\mathbf{z}\|_\infty$, $\forall \mathbf{z} \in \mathbb{R}^{m_I} \times \mathbb{R}^{m_E}$.

Set $\hat{\alpha} = \min\{\hat{\alpha}_0, \hat{\alpha}_1\}$. Then

$$\forall \mathbf{x} \in X_3^1 := \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap X_3 \cap \left\{\mathbf{x} = (\mathbf{x}^* + t\mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_0 \cap \mathcal{D}_I, t > 0\right\},$$

we have $\mathbf{x} = \mathbf{x}^* + t\mathbf{d} \in (\mathcal{F}(X))^{\mathrm{C}}$ with $\|\mathbf{d}\|_1 = 1$, $\mathbf{d} \in \mathcal{D}_0$ and $t = \|\mathbf{x} - \mathbf{x}^*\|_1 < \hat{\alpha}$, and

$$\varepsilon^o \geq f(\mathbf{x}) - f(\mathbf{x}^*) \geq \frac{L_f}{2} t \implies L_f t = L_f \|\mathbf{x} - \mathbf{x}^*\|_1 \leq 2\varepsilon^o.$$

Furthermore,

$$\forall \mathbf{x} \in X_3^2 := \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap X_3 \cap \left\{\mathbf{x} = (\mathbf{x}^* + t\mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0, t > 0\right\},$$

we have $\mathbf{x} = \mathbf{x}^* + t\mathbf{d} \in (\mathcal{F}(X))^{\mathrm{C}}$ with $\|\mathbf{d}\|_1 = 1$, $\mathbf{d} \notin \mathcal{D}_0$ and $t = \|\mathbf{x} - \mathbf{x}^*\|_1 < \hat{\alpha}$, and

$$\varepsilon^f \geq d\left(\begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix}(\mathbf{x}), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\}\right) \geq \frac{L_I}{2} t \implies L_I t = L_I \|\mathbf{x} - \mathbf{x}^*\|_1 \leq 2\varepsilon^f.$$

Finally, for every

$$\mathbf{x} \in \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap X_3 \cap \left\{\mathbf{x} = (\mathbf{x}^* + t\mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0, t > 0\right\}$$

with $t \leq \dfrac{\delta}{2}$, we have $\mathbf{x} \in B_\delta$. Consequently, for each

$$\mathbf{x} \in \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap X_3 \cap \left\{\mathbf{x} = (\mathbf{x}^* + t\mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0, t > 0\right\} \Big\backslash B_\delta,$$

we have $t > \dfrac{\delta}{2}$ and therefore,

$$d\left(\begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix}(\mathbf{x}), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\}\right) \geq \frac{L_I}{2} t > \frac{L_I}{4} \delta.$$

The desired result follows when $L_I \delta < 4\varepsilon^f$; otherwise, if $L_I \delta \geq 4\varepsilon^f$, then

$$\mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap X_3 \cap \left\{\mathbf{x} = (\mathbf{x}^* + t\mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0, t > 0\right\} \subset B_\delta. \qquad \square$$

A conservative estimate of the number of boxes of certain widths required to cover

$\left(\mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) \cap X_3\right) \backslash B_\delta$ can be obtained by estimating the number of boxes of certain widths required to cover $\hat{X}^1_3$ and $\hat{X}^2_3 \backslash B_\delta$ (see Theorem 5.3.31). The following remark is in order.

Remark 5.3.28.

1. Lemma 5.3.27 does not hold when $\nexists \alpha > 0$, $\mathcal{D}_0$ such that both L_f and L_I are positive. Example 5.2.11 illustrates a case when no valid partition of $\mathcal{D}_I$ exists (since $[x^{\mathrm{L}}, 0)$, which is a subset of X_3, corresponds to $d = -1$ which has an empty intersection with every valid choice of $\mathcal{D}_0$, and $\nabla g_1(x^*) = 0$). Note that $\mathcal{D}_0$ may be chosen to be $\emptyset$, but it cannot be chosen to be $\mathcal{D}_I$ when the objective function is differentiable at $\mathbf{x}^*$. This is because when $\nabla f(\mathbf{x}^*) \neq \mathbf{0}$, the direction $-\nabla f(\mathbf{x}^*)$ leads to infeasible points around $\mathbf{x}^*$. One potential choice of $\mathcal{D}_0$ is

$$\mathcal{D}_0 = \Big\{ \mathbf{d} \, : \|\mathbf{d}\|_1 = 1, \, \exists t > 0 \, : \, (\mathbf{x}^* + t\mathbf{d}) \in \mathcal{N}^1_\alpha(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}}, \\ \max\left\{ \max_{j \in \mathcal{A}(\mathbf{x}^*)} \left\{ g'_j(\mathbf{x}^*; \mathbf{d}) \right\}, \max_{k \in \{1, \cdots, m_E\}} \left\{ \left| h'_k(\mathbf{x}^*; \mathbf{d}) \right| \right\} \right\} \leq \theta \Big\}$$

for some choice of $\theta > 0$, so long as $\inf\limits_{\mathbf{d} \in \mathcal{D}_0} f'(\mathbf{x}^*; \mathbf{d}) > 0$. Proposition 5.3.29 shows that the assumptions of Lemma 5.3.27 will not be satisfied when Problem (P) does not contain any active inequality constraints and the minimizer corresponds to a KKT point for Problem (P).

2. The inequality $L_I \delta < 4\varepsilon^f$ is equivalent to

$$L_I \delta = L_I \left(\frac{\varepsilon^f}{\tau^I} \right)^{\frac{1}{\beta^I}} < 4\varepsilon^f.$$

Since ε^f can be taken to be sufficiently-small, the above inequality holds only when $(\varepsilon^f)^{\frac{1}{\beta^I}} \leq \varepsilon^f \iff \beta^I \leq 1$, i.e., if $\beta^I > 1$, we can choose ε^f to be small-enough so that $L_I \delta \geq 4\varepsilon^f$. Note that if $L_I \delta \geq 4\varepsilon^f$, the region

$$\mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) \cap X_3 \cap \Big\{ \mathbf{x} = (\mathbf{x}^* + t\mathbf{d}) \in \mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0, t > 0 \Big\}$$

has already been covered while covering $\hat{X}_5$ since

$$\frac{L_I \delta}{4} \geq \varepsilon^f \geq \frac{L_I}{2} t \implies t \leq \frac{\delta}{2},$$

which implies $\mathbf{x} = \mathbf{x}^* + t\mathbf{d} \in B_\delta$.

The motivation for excluding the region B_δ from X_3 is as follows. Lemma 5.2.5 shows that if the measure of infeasibility, as determined by the distance function d, is strictly greater than ε^f at each point in the domain of a node, the node can be fathomed by a box of width δ. However, if $\mathbf{x}^*$ is a constrained minimizer, we will have points in X_3 which are arbitrarily close to $\mathbf{x}^*$ and have a measure of infeasibility that is arbitrarily close to 0. Such points will then have to be fathomed by boxes of width much smaller than δ (and arbitrarily close to 0). To avoid this issue, such points are assumed to be eliminated when X_5 is covered by boxes of width δ.

3. $\hat{\alpha}$ depends on the local behavior of f, g_j, $\forall j \in \mathcal{A}(\mathbf{x}^*)$, and h_k, $\forall k \in \{1, \cdots, m_E\}$, around $\mathbf{x}^*$, but is independent of ε. Consequently, for sufficiently small ε we have $\hat{X}_3^1 = \{\mathbf{x} \in X : L_f \|\mathbf{x} - \mathbf{x}^*\|_1 \leq 2\varepsilon^o\}$ and $\hat{X}_3^2 = \left\{\mathbf{x} \in X : L_I \|\mathbf{x} - \mathbf{x}^*\|_1 \leq 2\varepsilon^f\right\}$. Additionally, if f and g_j, $\forall j \in \mathcal{A}(\mathbf{x}^*)$, are convex on $\mathcal{N}_\alpha^1(\mathbf{x}^*)$ and h_k, $\forall k \in \{1, \cdots, m_E\}$, are affine on $\mathcal{N}_\alpha^1(\mathbf{x}^*)$, we can choose $\hat{\alpha} = \alpha$. Furthermore,

$$X_3^1 := \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap X_3 \cap \left\{\mathbf{x} = (\mathbf{x}^* + t\mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_0 \cap \mathcal{D}_I, t > 0\right\}$$

can be conservatively approximated as $\hat{X}_3^1 = \{\mathbf{x} \in X : L_f \|\mathbf{x} - \mathbf{x}^*\|_1 \leq \varepsilon^o\}$,

$$X_3^2 := \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap X_3 \cap \left\{\mathbf{x} = (\mathbf{x}^* + t\mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0, t > 0\right\}$$

can be conservatively approximated as $\hat{X}_3^2 = \left\{\mathbf{x} \in X : L_I \|\mathbf{x} - \mathbf{x}^*\|_1 \leq \varepsilon^f\right\}$, and the region

$$\mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap X_3 \cap \left\{\mathbf{x} = (\mathbf{x}^* + t\mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0, t > 0\right\} \Big\backslash B_\delta$$

is conservatively characterized by

$$\left\{\mathbf{x} \in \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) : d\left(\begin{bmatrix}\mathbf{g} \\ \mathbf{h}\end{bmatrix}(\mathbf{x}), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\}\right) \in \left(\frac{L_I}{2}\delta, \varepsilon^f\right]\right\}$$

whenever $L_I \delta < 2\varepsilon^f$.

4. Similar to Proposition 5.3.4, the following less conservative estimates of X_3^1 and X_3^2 can

be obtained:

$$\hat{X}_3^1 = \left\{\mathbf{x} \in \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \,:\, L_f\|\mathbf{x}-\mathbf{x}^*\|_1 \leq 2\varepsilon^o,\; f'(\mathbf{x}^*;\mathbf{x}-\mathbf{x}^*) \geq L_f\|\mathbf{x}-\mathbf{x}^*\|_1\right\},$$
$$\hat{X}_3^2 = \Bigg\{\mathbf{x} \in \mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \,:\, L_I\|\mathbf{x}-\mathbf{x}^*\|_1 \leq 2\varepsilon^f,$$
$$\max\left\{\max_{j\in\mathcal{A}(\mathbf{x}^*)} g_j'(\mathbf{x}^*;\mathbf{x}-\mathbf{x}^*),\; \max_{k\in\{1,\cdots,m_E\}} \left|h_k'(\mathbf{x}^*;\mathbf{x}-\mathbf{x}^*)\right|\right\} \geq L_I\|\mathbf{x}-\mathbf{x}^*\|_1\Bigg\}.$$

As an illustration of the application of Lemma 5.3.27, let us reconsider Example 5.2.8. Recall that $X = (2.2, 2.5) \times (2.9, 3.3)$, $m_I = 3$, $m_E = 0$, $f(\mathbf{x}) = -x_1 - x_2$, $g_1(\mathbf{x}) = x_2 - 2x_1^4 + 8x_1^3 - 8x_1^2 - 2$, $g_2(\mathbf{x}) = x_2 - 4x_1^4 + 32x_1^3 - 88x_1^2 + 96x_1 - 36$, and $g_3(\mathbf{x}) = 3 - x_2$ with $\mathbf{x}^* \approx (2.33, 3.18)$. Let $\varepsilon^o \leq 0.03$ and $\varepsilon^f \leq 0.05$. We have $\mathcal{F}(X) = \{\mathbf{x} \in X : \mathbf{g}(\mathbf{x}) \leq \mathbf{0}\}$, $\nabla f(\mathbf{x}^*) = (-1,-1)$, $\nabla g_1(\mathbf{x}^*) \approx (-8.164, 1)$, and $\nabla g_2(\mathbf{x}^*) \approx (4.700, 1)$. Choose $\alpha = +\infty$. $\mathcal{D}_I = \left\{\mathbf{d} : \|\mathbf{d}\|_1 = 1,\, \exists t > 0 \,:\, (\mathbf{x}^* + t\mathbf{d}) \in (\mathcal{F}(X))^{\mathrm{C}}\right\}$. Choose the set of unit-norm directions $\mathcal{D}_0 = \left\{\mathbf{d} : \|\mathbf{d}\|_1 = 1,\, \nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} \geq 0.298\right\}$ and $\hat{\alpha} = +\infty$ in Lemma 5.3.27. From Lemma 5.3.27 and Remark 5.3.28, we have $L_f = 0.298$ and $L_I = 1$ with the estimates $\hat{X}_3^1 = \{\mathbf{x} : 0.298\|\mathbf{x}-\mathbf{x}^*\|_1 \leq \varepsilon^o\}$ (since f is convex), and $\hat{X}_3^2 = \left\{\mathbf{x} : \|\mathbf{x}-\mathbf{x}^*\|_1 \leq 2\varepsilon^f\right\}$. Figure 5-5 illustrates the set $\mathcal{D}_0$, and plots the sets X_3^1 and X_3^2 along with their estimates $\hat{X}_3^1$ and $\hat{X}_3^2$ for $\varepsilon^o = 0.03$ and $\varepsilon^f = 0.05$.

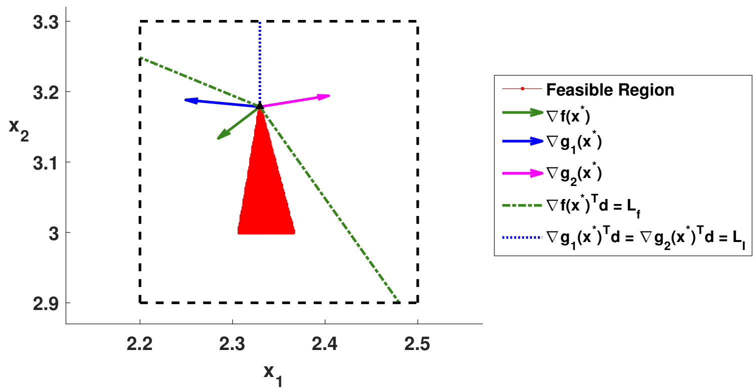
The next result provides conditions under which the assumptions of Lemma 5.3.27 will not be satisfied. In particular, it is shown that the assumptions of Lemma 5.3.27 will not be satisfied if Problem (P) is purely equality-constrained and all the functions in Problem (P) are differentiable at a nonisolated constrained minimizer $\mathbf{x}^*$.

Proposition 5.3.29. Consider Problem (P) with $m_E \geq 1$. Suppose $\mathbf{x}^*$ is a nonisolated constrained minimizer, f is differentiable at $\mathbf{x}^*$, functions h_k, $k = 1, \cdots, m_E$, are differentiable at $\mathbf{x}^*$, and $\mathcal{A}(\mathbf{x}^*) = \emptyset$. Furthermore, suppose there exist multipliers $\boldsymbol{\lambda}^* \in \mathbb{R}^{m_E}$ corresponding to the equality constraints such that $(\mathbf{x}^*, \mathbf{0}, \boldsymbol{\lambda}^*)$ is a KKT point. Then $\nexists\alpha > 0$, $\mathcal{D}_0$ such that the assumptions of Lemma 5.3.27 are satisfied.

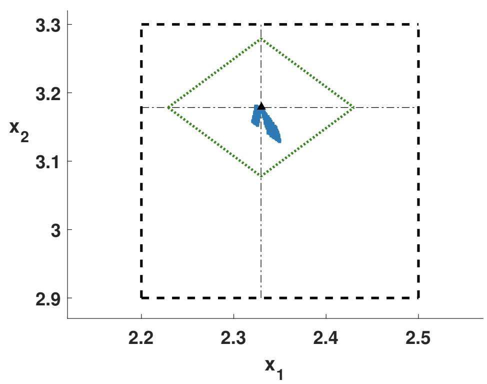
Proof. Since $(\mathbf{x}^*, \mathbf{0}, \boldsymbol{\lambda}^*)$ is a KKT point, we have

$$\nabla f(\mathbf{x}^*) + \sum_{k=1}^{m_E} \lambda_k^* \nabla h_k(\mathbf{x}^*) = \mathbf{0}.$$

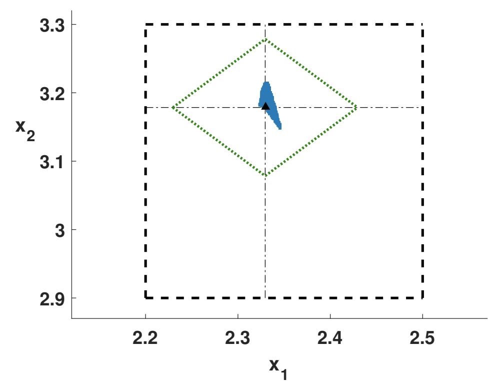
From the assumption that $\mathbf{x}^*$ is a nonisolated feasible point, we have that the set of unit-



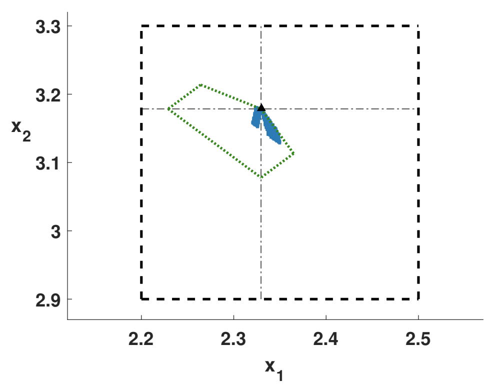
(a) Illustration of the sets $\mathcal{D}_0$ and $\mathcal{D}_I \backslash \mathcal{D}_0$



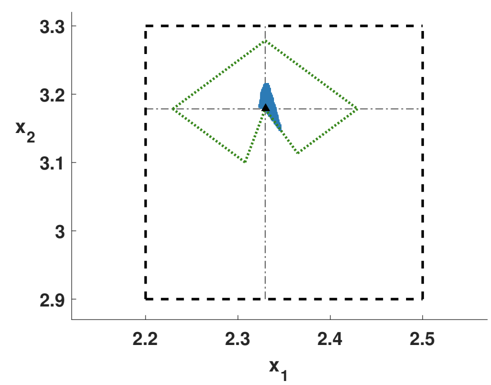
(b) X_3^1 and estimate $\hat{X}_3^1$ from Lemma 5.3.27



(c) X_3^2 and estimate $\hat{X}_3^2$ from Lemma 5.3.27



(d) X_3^1 and estimate $\hat{X}_3^1$ from Remark 5.3.28



(e) X_3^2 and estimate $\hat{X}_3^2$ from Remark 5.3.28

Figure 5-5: Illustration of the sets $\mathcal{D}_0$ and $\mathcal{D}_I \backslash \mathcal{D}_0$, the sets X_3^1 and X_3^2, and their estimates $\hat{X}_3^1$ and $\hat{X}_3^2$ for Example 5.2.8. The dashed lines represent the set X, and the filled-in triangles represent the minimum $\mathbf{x}^*$. (Top Plot) The solid region represents the feasible region and the solid vectors represent the gradients of the objective and the constraints. The set of directions between the dot-dashed lines (the part in which the feasible region resides) defines the set $\mathcal{D}_0$, and the remaining directions define the set $\mathcal{D}_I \backslash \mathcal{D}_0$. The dotted line represents the direction in $\mathcal{D}_I \backslash \mathcal{D}_0$ in which both constraints grow equally quickly in a first-order sense. (Other Plots) The solid regions represent the set X_3^1 or X_3^2, the area between the dotted lines represent the estimate $\hat{X}_3^1$ or $\hat{X}_3^2$, and the dash-dotted lines represent the axes translated to $\mathbf{x}^*$. All plots use $\varepsilon^o = 0.03$ and $\varepsilon^f = 0.05$.

norm directions $\{\mathbf{d} : \|\mathbf{d}\|_1 = 1, \mathbf{d} \in T(\mathbf{x}^*)\}$ is nonempty. Additionally, we have from the proof of Proposition 5.3.7 that for each $\mathbf{d} \in T(\mathbf{x}^*)$ with $\|\mathbf{d}\|_1 = 1$, $\nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} = 0$ and $\nabla h_k(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} = 0, \forall k \in \{1, \cdots, m_E\}$.

Assume, by way of contradiction that $\exists \alpha > 0$ and a set $\mathcal{D}_0$ satisfying the assumptions of Lemma 5.3.27. Consequently, $\exists L_f, L_I > 0$ such that

$$L_f = \inf_{\mathbf{d} \in \mathcal{D}_0 \cap \mathcal{D}_I} \nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d}$$

and

$$L_I = \inf_{\mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0} \max_{k \in \{1, \cdots, m_E\}} \left| \nabla h_k(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} \right|.$$

Since $\exists \mathbf{d} \in T(\mathbf{x}^*)$ with $\|\mathbf{d}\|_1 = 1$ such that $\nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} = 0$ and $\nabla h_k(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} = 0, \forall k \in \{1, \cdots, m_E\}$, we have that the set

$$S := \left\{ \mathbf{d} \in \mathbb{R}^{n_x} : \|\mathbf{d}\|_1 = 1,\ |\nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d}| < L_f,\ |\nabla h_k(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d}| < L_I, \forall k \in \{1, \cdots, m_E\} \right\}$$

is nonempty. All that remains to reach a contradiction is to show that $\exists \bar{\mathbf{d}} \in S \cap \mathcal{D}_I$.

From the above arguments, we have the existence of $\bar{\mathbf{d}} \in S$, $\bar{k} \in \{1, \cdots, m_E\}$ such that $|\nabla h_{\bar{k}}(\mathbf{x}^*)^{\mathrm{T}}\bar{\mathbf{d}}| \in (0, L_I)$, since the assumption $L_I > 0$ implies all of the equality constraint gradients $\nabla h_k(\mathbf{x}^*)$, $k \in \{1, \cdots, m_E\}$, cannot simultaneously be $\mathbf{0}$. Since $\nabla h_{\bar{k}}(\mathbf{x}^*)^{\mathrm{T}}\bar{\mathbf{d}} \neq 0$, we have $\bar{\mathbf{d}} \notin T(\mathbf{x}^*)$ (this follows from the arguments made in the proof of Proposition 5.3.7). Consequently, $\exists t \in (0, \alpha)$ such that $(\mathbf{x}^* + t\bar{\mathbf{d}}) \in \mathcal{N}^1_\alpha(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} \implies \bar{\mathbf{d}} \in \mathcal{D}_I$. This implies that either $\bar{\mathbf{d}} \in \mathcal{D}_0$, or $\bar{\mathbf{d}} \in \mathcal{D}_I \backslash \mathcal{D}_0$, which contradicts the definition of L_f or L_I since $\nabla f(\mathbf{x}^*)^{\mathrm{T}}\bar{\mathbf{d}} < L_f$ and $|\nabla h_k(\mathbf{x}^*)^{\mathrm{T}}\bar{\mathbf{d}}| < L_I, \forall k \in \{1, \cdots, m_E\}$. $\square$

The above result can be extended to the case when there exist active inequality constraints if all such constraints are strongly active at $\mathbf{x}^*$ (see [13, Section 4.4]) and there exists $\mathbf{d} \in T(\mathbf{x}^*)$ such that $\nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} = 0$.

Next, we revisit two equality-constrained examples from Section 5.3.1 for which the assumptions of Lemma 5.3.27 hold, and which do not satisfy individual assumptions of Proposition 5.3.29. Consider Example 5.3.8, and recall that $X = (-2, 2) \times (-2, 2)$, $m_I = 1$, and $m_E = 1$ with $f(\mathbf{x}) = x_1 + 10x_2^2$, $g_1(\mathbf{x}) = x_1 - 1$, $h(\mathbf{x}) = x_1 - |x_2|$, and $\mathbf{x}^* = (0, 0)$. Let $\varepsilon^o, \varepsilon^f \leq 0.25$. We have $\mathcal{F}(X) = \{\mathbf{x} \in X : x_1 = |x_2|, x_1 \leq 1\}$, $\nabla f(\mathbf{x}^*) = (1, 0)$, and $h'(\mathbf{x}^*; \mathbf{d}) = d_1 - |d_2|$. Choose $\alpha = +\infty$. We have $\mathcal{D}_I = \left\{ \mathbf{d} : \|\mathbf{d}\|_1 = 1,\ \exists t > 0 \ :\ (\mathbf{x}^* + t\mathbf{d}) \in (\mathcal{F}(X))^{\mathrm{C}} \right\}$.

Choose $\mathcal{D}_0 = \left\{\mathbf{d} : \|\mathbf{d}\|_1 = 1,\ \nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} \geq 0.25\right\}$ and $\hat{\alpha} = +\infty$ in Lemma 5.3.27. From Lemma 5.3.27 and Remark 5.3.28, we have $L_f = 0.25$ and $L_I = 0.5$ with the estimates $\hat{X}_3^1 = \{\mathbf{x} : 0.25\|\mathbf{x} - \mathbf{x}^*\|_1 \leq \varepsilon^o\}$ (since f is convex), and $\hat{X}_3^2 = \left\{\mathbf{x} : 0.5\|\mathbf{x} - \mathbf{x}^*\|_1 \leq 2\varepsilon^f\right\}$.

Consider Example 5.3.9, and recall that $X = (-2, 2) \times (-2, 2)$, $m_I = 4$, and $m_E = 1$ with $f(\mathbf{x}) = x_1 + x_2$, $g_1(\mathbf{x}) = -x_1$, $g_2(\mathbf{x}) = -x_2$, $g_3(\mathbf{x}) = x_1 - 1$, $g_4(\mathbf{x}) = x_2 - 1$, $h(\mathbf{x}) = x_2 - x_1^3$, and $\mathbf{x}^* = (0, 0)$. Let $\varepsilon^o, \varepsilon^f \leq \frac{1}{3}$. $\mathcal{F}(X) = \left\{\mathbf{x} \in [0, 1]^2 : x_2 = x_1^3\right\}$, $\nabla f(\mathbf{x}^*) = (1, 1)$, $\nabla g_1(\mathbf{x}^*) = (-1, 0)$, $\nabla g_2(\mathbf{x}^*) = (0, -1)$, and $\nabla h(\mathbf{x}^*) = (0, 1)$. Choose $\alpha = +\infty$. $\mathcal{D}_I = \left\{\mathbf{d} : \|\mathbf{d}\|_1 = 1, \exists t > 0 : (\mathbf{x}^* + t\mathbf{d}) \in (\mathcal{F}(X))^{\mathrm{C}}\right\}$. Choose the set of unit-norm directions $\mathcal{D}_0 = \left\{\mathbf{d} : \|\mathbf{d}\|_1 = 1,\ \nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} \geq \frac{1}{3}\right\}$ and $\hat{\alpha} = +\infty$ in Lemma 5.3.27. From Lemma 5.3.27 and Remark 5.3.28, we have $L_f = \frac{1}{3}$ and $L_I = \frac{1}{3}$ with the estimates $\hat{X}_3^1 = \{\mathbf{x} : \|\mathbf{x} - \mathbf{x}^*\|_1 \leq 3\varepsilon^o\}$ (since f is convex), and $\hat{X}_3^2 = \left\{\mathbf{x} : \|\mathbf{x} - \mathbf{x}^*\|_1 \leq 3\varepsilon^f\right\}$ (since g_1 and g_2 are convex).

The next example illustrates a simple one-dimensional case which satisfies the assumptions of Lemma 5.3.27 with $\mathcal{D}_0 = \emptyset$.

Example 5.3.30. Let $\varepsilon^f \leq 0.5$, $X = (-2, 2)$, $m_I = 2$, and $m_E = 0$ with $f(x) = x^3$, $g_1(x) = x - 1$, $g_2(x) = -x$, and $x^* = 0$. We have $\mathcal{F}(X) = [0, 1]$, $\nabla f(x^*) = 0$, $\nabla g_2(x^*) = -1$, and $X_3 = [-\varepsilon^f, 0)$. Choose $\alpha = +\infty$. We have $\mathcal{D}_I = \{-1\}$. Choose $\mathcal{D}_0 = \emptyset$ and $\hat{\alpha} = +\infty$ in Lemma 5.3.27. From Lemma 5.3.27 and Remark 5.3.28, we have $L_I = 1$ and $\hat{X}_3^2 = [-\varepsilon^f, +\varepsilon^f]$ (since g_2 is convex).

The following result follows from Corollary 2.1 in [237] (also see the proof of Theorem 5.3.11). It provides a conservative estimate of the number of boxes of certain widths required to cover $\hat{X}_3^1$ and $\hat{X}_3^2 \backslash B_\delta$ from Lemma 5.3.27. Therefore, from Lemmata 5.2.5 and 5.2.6 and the result below, we can get an upper bound on the worst-case number of boxes required to cover $\mathcal{N}_{\hat{\alpha}}^1(\mathbf{x}^*) \cap X_3$ and estimate the extent of the cluster problem on that region.

Theorem 5.3.31. Suppose the assumptions of Lemma 5.3.27 hold. Let $\delta = \delta_f = \left(\frac{\varepsilon}{\tau^*}\right)^{\frac{1}{\beta^*}} = \left(\frac{\varepsilon^o}{\tau^f}\right)^{\frac{1}{\beta^f}} = \left(\frac{\varepsilon^f}{\tau^I}\right)^{\frac{1}{\beta^I}}$, $\delta_I = \left(\frac{L_I \delta}{4\tau^I}\right)^{\frac{1}{\beta^I}} = \left(\frac{L_I}{4\tau^I}\right)^{\frac{1}{\beta^I}} \left(\frac{\varepsilon^f}{\tau^I}\right)^{\frac{1}{(\beta^I)^2}}$, $r_I = \frac{2\varepsilon^f}{L_I}$, $r_f = \frac{2\varepsilon^o}{L_f}$.

1. If $\delta_I \geq 2r_I$, let $N_I = 1$.

2. If $\frac{2r_I}{\bar{m}_I - 1} > \delta_I \geq \frac{2r_I}{\bar{m}_I}$ for some $\bar{m}_I \in \mathbb{N}$ with $\bar{m}_I \leq n_x$ and $2 \leq \bar{m}_I \leq 5$, then let

$$N_I = \sum_{i=0}^{\bar{m}_I - 1} 2^i \binom{n_x}{i} + 2n_x \left\lceil \frac{\bar{m}_I - 3}{3} \right\rceil.$$

3. Otherwise, let

$$N_I = \left\lceil 2B_I(\varepsilon^f; \beta^I, L_I, \tau_I) \right\rceil^{n_x - 1} \left(\left\lceil 2B_I(\varepsilon^f; \beta^I, L_I, \tau_I) \right\rceil + 2n_x \left\lceil B_I(\varepsilon^f; \beta^I, L_I, \tau_I) \right\rceil \right),$$

   where

$$B_I(\varepsilon^f; \beta^I, L_I, \tau_I) := 4^{\frac{1}{\beta^I}} \left(\tau^I\right)^{\left(\frac{1}{\beta^I} + \frac{1}{(\beta^I)^2}\right)} \left(\varepsilon^f\right)^{\left(1 - \frac{1}{(\beta^I)^2}\right)} L_I^{-\left(1 + \frac{1}{\beta^I}\right)}.$$

4. If $\delta_f \geq 2r_f$, let $N_f = 1$.

5. If $\frac{2r_f}{m_f - 1} > \delta_f \geq \frac{2r_f}{m_f}$ for some $m_f \in \mathbb{N}$ with $m_f \leq n_x$ and $2 \leq m_f \leq 5$, then let

$$N_f = \sum_{i=0}^{m_f - 1} 2^i \binom{n_x}{i} + 2n_x \left\lceil \frac{m_f - 3}{3} \right\rceil.$$

6. Otherwise, let

$$N_f = \left\lceil 2\left(\tau^f\right)^{\frac{1}{\beta^f}} (\varepsilon^o)^{\left(1 - \frac{1}{\beta^f}\right)} L_f^{-1} \right\rceil^{n_x - 1} \Bigg(\left\lceil 2\left(\tau^f\right)^{\frac{1}{\beta^f}} (\varepsilon^o)^{\left(1 - \frac{1}{\beta^f}\right)} L_f^{-1} \right\rceil + 2n_x \left\lceil \left(\tau^f\right)^{\frac{1}{\beta^f}} (\varepsilon^o)^{\left(1 - \frac{1}{\beta^f}\right)} L_f^{-1} \right\rceil \Bigg).$$

Then, N_I is an upper bound on the number of boxes of width δ_I required to cover $\hat{X}_3^2 \backslash B_\delta$, and N_f is an upper bound on the number of boxes of width δ_f required to cover $\hat{X}_3^1$.

Proof. The result on N_f follows from Lemmata 5.2.6 and 5.3.27 and Corollary 2.1 in [237] (also see the proof of Theorem 5.3.11). To deduce the result on N_I, note that we cover $\hat{X}_3^2 \backslash B_\delta$ with boxes of width $\delta_I = \left(\frac{L_I \delta}{4\tau^I}\right)^{\frac{1}{\beta^I}}$ since, from Lemma 5.3.27, we have

$$\hat{X}_3^2 \backslash B_\delta \subset \left\{ \mathbf{x} \in \mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) : d\left(\begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix} (\mathbf{x}), \mathbb{R}^{m_I}_- \times \{\mathbf{0}\} \right) \in \left(\frac{L_I}{4}\delta, \varepsilon^f \right] \right\}$$

and, from Lemma 5.2.5, we have that a box B_{δ_I} of width δ_I with each $\mathbf{x} \in B_{\delta_I}$ satisfying $d\left(\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}(\mathbf{x}), \mathbb{R}_{-}^{m_I} \times \{\mathbf{0}\}\right) > \frac{L_I}{4}\delta$ can be fathomed by infeasibility. The desired result then follows from Corollary 2.1 in [237]. □

Remark 5.3.32. Under the assumptions of Lemma 5.3.27, the dependence of N_I on ε^f disappears when the lower bounding scheme has first-order convergence on $\mathcal{N}^1_{\hat{\alpha}}(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}}$, i.e., $\beta^I = 1$, and the dependence of N_f on ε^o disappears when the scheme $(f^{\mathrm{cv}}_Z)_{Z \in \mathbb{I}X}$ has first-order convergence on X, i.e., $\beta^f = 1$. Therefore, the cluster problem on X_3 can be eliminated even using first-order convergent schemes with sufficiently small prefactors. Note that the dependence of N_f and N_I on the prefactors τ^f and τ^I, respectively, can be detailed in a manner similar to Table 1 in [238].

The following results illustrate one set of assumptions under which second-order convergence of the lower bounding scheme at infeasible points is sufficient to eliminate the cluster problem on $X_3 \backslash B_\delta$. First, we provide a conservative estimate of the subset of X_3 around a constrained minimizer $\mathbf{x}^*$ under the following assumption: the infeasible region in some neighborhood of $\mathbf{x}^*$ can be split into two subregions such that the objective function grows quadratically (or faster) in the first subregion and the measure of infeasibility grows quadratically (or faster) in the second subregion. Note that better estimates of X_3 may be derived either under the (stronger) assumption that the objective function grows linearly in the directions $\mathcal{D}_0 \cap \mathcal{D}_I$, or under the (stronger) assumption that the measure of infeasibility grows linearly in the directions $\mathcal{D}_I \backslash \mathcal{D}_0$.

Lemma 5.3.33. Consider Problem (P). Suppose $\mathbf{x}^*$ is a constrained minimizer, functions f, g_j, $\forall j \in \mathcal{A}(\mathbf{x}^*)$, and h_k, $\forall k \in \{1, \cdots, m_E\}$, are twice-differentiable at $\mathbf{x}^*$, and $\exists \alpha > 0, \gamma_1 > 0, \gamma_2 > 0$ and a set $\mathcal{D}_0$ such that

$$\nabla f(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} + \frac{1}{2}\mathbf{d}^{\mathrm{T}}\nabla^2 f(\mathbf{x}^*)\mathbf{d} \geq \gamma_1 \mathbf{d}^{\mathrm{T}}\mathbf{d}, \quad \forall \mathbf{d} \in \mathcal{D}_0 \cap \mathcal{D}_I,$$

$$\max\left\{\max_{j \in \mathcal{A}(\mathbf{x}^*)}\left\{\nabla g_j(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} + \frac{1}{2}\mathbf{d}^{\mathrm{T}}\nabla^2 g_j(\mathbf{x}^*)\mathbf{d}\right\}, \max_{k \in \{1,\cdots,m_E\}}\left\{\left|\nabla h_k(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} + \frac{1}{2}\mathbf{d}^{\mathrm{T}}\nabla^2 h_k(\mathbf{x}^*)\mathbf{d}\right|\right\}\right\}$$
$$\geq \gamma_2 \mathbf{d}^{\mathrm{T}}\mathbf{d}, \quad \forall \mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0,$$

where $\mathcal{D}_I$ is defined as $\mathcal{D}_I = \left\{\mathbf{d} : (\mathbf{x}^* + \mathbf{d}) \in \mathcal{N}^2_\alpha(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}}\right\}$. Then, $\exists \hat{\alpha} \in (0, \alpha]$ such that the region

$$X_3^1 := \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap X_3 \cap \left\{ \mathbf{x} = (\mathbf{x}^* + \mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_0 \cap \mathcal{D}_I \right\}$$

can be conservatively approximated as $\hat{X}_3^1 = \left\{ \mathbf{x} \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) : \gamma_1 \|\mathbf{x} - \mathbf{x}^*\|^2 \leq 2\varepsilon^o \right\}$, and the region

$$X_3^2 := \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap X_3 \cap \left\{ \mathbf{x} = (\mathbf{x}^* + \mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0 \right\}$$

can be conservatively approximated as $\hat{X}_3^2 = \left\{ \mathbf{x} \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) : \gamma_2 \|\mathbf{x} - \mathbf{x}^*\|^2 \leq 2\varepsilon^f \right\}$. Furthermore, suppose $\mathbf{x}^*$ is at the center of a box, B_δ, of width $\delta = \left(\frac{\varepsilon}{\tau^*} \right)^{\frac{1}{\beta^*}}$ placed while covering $\hat{X}_5$. Then, the region

$$X_3^2 \backslash B_\delta = \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap X_3 \cap \left\{ \mathbf{x} = (\mathbf{x}^* + \mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0 \right\} \Big\backslash B_\delta$$

is conservatively characterized by

$$\left\{ \mathbf{x} \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) : d \left(\begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix} (\mathbf{x}), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\} \right) \in \left(\frac{\gamma_2}{8} \delta^2, \varepsilon^f \right] \right\},$$

whenever $\gamma_2 \delta^2 < 8\varepsilon^f$.

Proof. From Lemma 5.3.17, we have the existence of $\hat{\alpha}_0 > 0$ such that

$$\mathcal{N}_{\hat{\alpha}_0}^2(\mathbf{x}^*) \cap X_3 \cap \left\{ \mathbf{x} = (\mathbf{x}^* + \mathbf{d}) \in \mathcal{N}_{\hat{\alpha}_0}^2(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_0 \cap \mathcal{D}_I \right\}$$

can be conservatively approximated as $\left\{ \mathbf{x} \in \mathcal{N}_{\hat{\alpha}_0}^2(\mathbf{x}^*) : \gamma_1 \|\mathbf{x} - \mathbf{x}^*\|^2 \leq 2\varepsilon^o \right\}$.

Consider $\mathbf{x} = \mathbf{x}^* + \mathbf{d} \in \mathcal{N}_{\alpha}^2(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}}$ with $\mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0$. We have

$$\begin{aligned}
&\max \left\{ \max_{j \in \mathcal{A}(\mathbf{x}^*)} \{g_j(\mathbf{x})\}, \max_{k \in \{1, \cdots, m_E\}} \{|h_k(\mathbf{x})|\} \right\} \\
&= \max \left\{ \max_{j \in \mathcal{A}(\mathbf{x}^*)} \{g_j(\mathbf{x}^* + \mathbf{d})\}, \max_{k \in \{1, \cdots, m_E\}} \{|h_k(\mathbf{x}^* + \mathbf{d})|\} \right\} \\
&= \max \left\{ \max_{j \in \mathcal{A}(\mathbf{x}^*)} \left\{ \nabla g_j(\mathbf{x}^*)^{\mathrm{T}} \mathbf{d} + \frac{1}{2} \mathbf{d}^{\mathrm{T}} \nabla^2 g_j(\mathbf{x}^*) \mathbf{d} + o(\|\mathbf{d}\|^2) \right\}, \right. \\
&\qquad \left. \max_{k \in \{1, \cdots, m_E\}} \left\{ \left| \nabla h_k(\mathbf{x}^*)^{\mathrm{T}} \mathbf{d} + \frac{1}{2} \mathbf{d}^{\mathrm{T}} \nabla^2 h_k(\mathbf{x}^*) \mathbf{d} + o(\|\mathbf{d}\|^2) \right| \right\} \right\}.
\end{aligned}$$

Consequently, there exists $\hat{\alpha}_1 \in (0, \alpha]$ such that for all $\mathbf{x} = \mathbf{x}^* + \mathbf{d} \in (\mathcal{F}(X))^{\mathrm{C}}$ with $\|\mathbf{d}\| \in [0, \hat{\alpha}_1)$, $\mathbf{d} \notin \mathcal{D}_0$:

$$
\begin{aligned}
d\left(\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}(\mathbf{x}), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\}\right) &\geq \max\left\{\max_{j\in\mathcal{A}(\mathbf{x}^*)}\{g_j(\mathbf{x})\}, \max_{k\in\{1,\cdots,m_E\}}\{|h_k(\mathbf{x})|\}\right\} \\
&= \max\left\{\max_{j\in\mathcal{A}(\mathbf{x}^*)}\left\{\nabla g_j(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} + \frac{1}{2}\mathbf{d}^{\mathrm{T}}\nabla^2 g_j(\mathbf{x}^*)\mathbf{d} + o(\|\mathbf{d}\|^2)\right\},\right. \\
&\qquad \left.\max_{k\in\{1,\cdots,m_E\}}\left\{\left|\nabla h_k(\mathbf{x}^*)^{\mathrm{T}}\mathbf{d} + \frac{1}{2}\mathbf{d}^{\mathrm{T}}\nabla^2 h_k(\mathbf{x}^*)\mathbf{d} + o(\|\mathbf{d}\|^2)\right|\right\}\right\} \\
&\geq \frac{\gamma_2}{2}\|\mathbf{d}\|^2,
\end{aligned}
$$

where Step 1 follows from the fact that $\|\mathbf{z}\| \geq \|\mathbf{z}\|_\infty$, $\forall \mathbf{z} \in \mathbb{R}^{m_I} \times \mathbb{R}^{m_E}$.

Choose $\hat{\alpha} = \min\{\hat{\alpha}_0, \hat{\alpha}_1\}$. The region

$$
X_3^1 := \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap X_3 \cap \left\{\mathbf{x} = (\mathbf{x}^* + \mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_0 \cap \mathcal{D}_I\right\}
$$

can be conservatively approximated as $\hat{X}_3^1 = \left\{\mathbf{x} \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) : \gamma_1\|\mathbf{x} - \mathbf{x}^*\|^2 \leq 2\varepsilon^o\right\}$, and

$$
\forall \mathbf{x} \in X_3^2 := \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap X_3 \cap \left\{\mathbf{x} = (\mathbf{x}^* + \mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0\right\},
$$

we have $\mathbf{x} = \mathbf{x}^* + \mathbf{d} \in (\mathcal{F}(X))^{\mathrm{C}}$ with $\mathbf{d} \notin \mathcal{D}_0$, $\|\mathbf{d}\| < \hat{\alpha}$, and

$$
\varepsilon^f \geq d\left(\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}(\mathbf{x}), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\}\right) \geq \frac{\gamma_2}{2}\|\mathbf{x} - \mathbf{x}^*\|^2 \implies \gamma_2\|\mathbf{x} - \mathbf{x}^*\|^2 \leq 2\varepsilon^f.
$$

Finally, for every $\mathbf{x} \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap X_3 \cap \left\{\mathbf{x} = (\mathbf{x}^* + \mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0\right\}$ with $\|\mathbf{d}\| \leq \frac{\delta}{2}$, we have $\mathbf{x} \in B_\delta$. Consequently, for each

$$
\mathbf{x} \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap X_3 \cap \left\{\mathbf{x} = (\mathbf{x}^* + \mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0\right\} \Big\backslash B_\delta,
$$

we have $\|\mathbf{d}\| > \frac{\delta}{2}$ and therefore,

$$
d\left(\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}(\mathbf{x}), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\}\right) > \frac{\gamma_2}{8}\delta^2.
$$

The desired result follows when $\gamma_2\delta^2 < 8\varepsilon^f$; otherwise, if $\gamma_2\delta^2 \geq 8\varepsilon^f$, then

$$\mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap X_3 \cap \left\{ \mathbf{x} = (\mathbf{x}^* + \mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0 \right\} \subset B_\delta. \qquad \square$$

A conservative estimate of the number of boxes of certain widths required to cover $\left(\mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap X_3\right) \backslash B_\delta$ can be obtained by estimating the number of boxes of certain widths required to cover $\hat{X}_3^1$ and $\hat{X}_3^2 \backslash B_\delta$ (see Theorem 5.3.35). The following remark is in order.

Remark 5.3.34.

1. Lemma 5.3.33 does not hold when $\nexists \alpha, \gamma_1, \gamma_2 > 0$, and $\mathcal{D}_0$, for example $X = (0,2) \times (0,2)$, $m_I = 0$, $m_E = 2$, $f(\mathbf{x}) = -x_1$, $h_1(\mathbf{x}) = x_2 - (1-x_1)^3$, $h_2(\mathbf{x}) = -x_2 - (1-x_1)^3$, and $\mathbf{x}^* = (1,0)$ (see [13, Example 4.3.5]). Note that $\mathcal{D}_0$ may be chosen to be $\emptyset$, but it cannot be chosen to be $\mathcal{D}_I$ (see Remark 5.3.28 for an explanation).

2. The inequality $\gamma_2\delta^2 < 8\varepsilon^f$ is equivalent to

$$\gamma_2\delta^2 = \gamma_2 \left(\frac{\varepsilon^f}{\tau^I} \right)^{\frac{2}{\beta^I}} < 8\varepsilon^f.$$

   Since ε^f can be taken to be sufficiently-small, the above inequality holds only when $(\varepsilon^f)^{\frac{2}{\beta^I}} \leq \varepsilon^f \iff \beta^I \leq 2$, i.e., if $\beta^I > 2$, we can choose ε^f to be small-enough so that $\gamma_2\delta^2 \geq 8\varepsilon^f$. Note that if $\gamma_2\delta^2 \geq 8\varepsilon^f$, the region

$$\mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap X_3 \cap \left\{ \mathbf{x} = (\mathbf{x}^* + \mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0 \right\}$$

   has already been covered while covering $\hat{X}_5$ since

$$\frac{\gamma_2\delta^2}{8} \geq \varepsilon^f \geq \frac{\gamma_2 \|\mathbf{d}\|^2}{2} \implies \|\mathbf{d}\| \leq \frac{\delta}{2},$$

   which implies $\mathbf{x} = \mathbf{x}^* + \mathbf{d} \in B_\delta$.

3. $\hat{\alpha}$ depends on the local behavior of f, g_j, $\forall j \in \mathcal{A}(\mathbf{x}^*)$, and h_k, $\forall k \in \{1, \cdots, m_E\}$, around $\mathbf{x}^*$, but is independent of ε. Consequently, for sufficiently small ε we have $\hat{X}_3^1 = \left\{ \mathbf{x} \in X : \gamma_1 \|\mathbf{x} - \mathbf{x}^*\|^2 \leq 2\varepsilon^o \right\}$ and $\hat{X}_3^2 = \left\{ \mathbf{x} \in X : \gamma_2 \|\mathbf{x} - \mathbf{x}^*\|^2 \leq 2\varepsilon^f \right\}$. Additionally, if the objective function and the active constraints are all either affine or quadratic functions of $\mathbf{x}$, then their second-order Taylor expansions around $\mathbf{x}^*$ equal themselves

and we can choose $\hat{\alpha} = \alpha$. Furthermore,

$$X_3^1 := \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap X_3 \cap \left\{ \mathbf{x} = (\mathbf{x}^* + \mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_0 \cap \mathcal{D}_I \right\}$$

can be conservatively approximated as $\hat{X}_3^1 = \left\{ \mathbf{x} \in X : \gamma_1 \|\mathbf{x} - \mathbf{x}^*\|^2 \leq \varepsilon^o \right\}$, the region

$$X_3^2 := \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap X_3 \cap \left\{ \mathbf{x} = (\mathbf{x}^* + \mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0 \right\}$$

can be conservatively approximated as $\hat{X}_3^2 = \left\{ \mathbf{x} \in X : \gamma_2 \|\mathbf{x} - \mathbf{x}^*\|^2 \leq \varepsilon^f \right\}$, and the region

$$\mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap X_3 \cap \left\{ \mathbf{x} = (\mathbf{x}^* + \mathbf{d}) \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}} : \mathbf{d} \in \mathcal{D}_I \backslash \mathcal{D}_0 \right\} \Big\backslash B_\delta$$

is conservatively characterized by

$$\left\{ \mathbf{x} \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) : d\left(\begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix} (\mathbf{x}), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\} \right) \in \left(\frac{\gamma_2}{4}\delta^2, \varepsilon^f \right] \right\}$$

whenever $\gamma_2 \delta^2 \geq 4\varepsilon^f$.

4. Similar to Proposition 5.3.4, the following less conservative estimates of X_3^1 and X_3^2 can be obtained:

$$\begin{aligned}
\hat{X}_3^1 = \Big\{ & \mathbf{x} \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \ : \ \gamma_1 \|\mathbf{x} - \mathbf{x}^*\|^2 \leq 2\varepsilon^o, \\
& \nabla f(\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x} - \mathbf{x}^*) + \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^{\mathrm{T}} \nabla^2 f(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) \geq \gamma_1 \|\mathbf{x} - \mathbf{x}^*\|^2 \Big\}, \\
\hat{X}_3^2 = \Big\{ & \mathbf{x} \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \ : \ \gamma_2 \|\mathbf{x} - \mathbf{x}^*\|^2 \leq 2\varepsilon^f, \\
& \max\Big\{ \max_{j \in \mathcal{A}(\mathbf{x}^*)} \left\{ \nabla g_j(\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x} - \mathbf{x}^*) + \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^{\mathrm{T}} \nabla^2 g_j(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) \right\}, \\
& \max_{k \in \{1, \cdots, m_E\}} \left\{ \left| \nabla h_k(\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x} - \mathbf{x}^*) + \frac{1}{2}(\mathbf{x} - \mathbf{x}^*)^{\mathrm{T}} \nabla^2 h_k(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) \right| \right\} \Big\} \geq \gamma_2 \|\mathbf{x} - \mathbf{x}^*\|^2 \Big\}.
\end{aligned}$$

To illustrate the application of Lemma 5.3.33, let us reconsider Example 5.2.11 with $\varepsilon^o, \varepsilon^f \leq 1$. Recall that $X = (-2, 2)$, $m_I = 3$, $m_E = 0$, $f(x) = x$, $g_1(x) = x^2$, $g_2(x) = x - 1$, and $g_3(x) = -1 - x$ with $x^* = 0$. We have $\mathcal{F}(X) = \{0\}$ and $X_3 = [-\sqrt{\varepsilon^f}, 0) \cup \left(0, \min\{\varepsilon^o, \sqrt{\varepsilon^f}\}\right]$. Choose $\alpha = 1$. We have $\mathcal{D}_I = (-1, 1)\backslash\{0\}$. Choose the set of directions $\mathcal{D}_0 = \{d \in \mathcal{D}_I : d > 0\}$, $\gamma_1 = 1$, $\gamma_2 = 1$, and $\hat{\alpha} = 1$ in Lemma 5.3.33. From Lemma 5.3.33

and Remark 5.3.34, we have $\hat{X}_3^1 = \left\{x : x^2 \leq \varepsilon^o\right\}$ and $\hat{X}_3^2 = \left\{x : x^2 \leq \varepsilon^f\right\}$ (since f is linear and g_1 is quadratic). In fact, for this example, we can get a better estimate of X_3^1 by taking into account the fact that f grows linearly on $\mathcal{D}_0 \cap \mathcal{D}_I$.

Next, we revisit a couple of examples from Section 5.3.1 for which the assumptions of Lemma 5.3.33 hold. First, consider Example 5.3.19 with $\varepsilon^o \leq 0.6$, $\varepsilon^f \leq 0.5$, and recall that $X = (-2,2) \times (-2,2)$, $m_I = 2$, and $m_E = 0$ with $f(\mathbf{x}) = x_2$, $g_1(\mathbf{x}) = x_1^2 - x_2$, $g_2(\mathbf{x}) = x_2 - 1$, and $\mathbf{x}^* = (0,0)$. We have $\mathcal{F}(X) = \left\{\mathbf{x} : x_2 \geq x_1^2,\ x_2 \leq 1\right\}$ and $X_3 = \left\{\mathbf{x} : x_1^2 - \varepsilon^f \leq x_2 < x_1^2, x_2 \leq \varepsilon^o\right\}$. Choose $\alpha = 1$. We have $\mathcal{D}_I = \left\{\mathbf{d} \in \mathcal{N}_1^2(\mathbf{0}) : d_2 < d_1^2\right\}$. Choose $\mathcal{D}_0 = \left\{\mathbf{d} \in \mathcal{D}_I : d_2 \geq 0.5 d_1^2\right\}$, $\gamma_1 = 0.3$, $\gamma_2 = 0.25$ and $\hat{\alpha} = 1$ in Lemma 5.3.33. From Lemma 5.3.33 and Remark 5.3.34, we have $\hat{X}_3^1 = \left\{\mathbf{x} \in \mathcal{N}_1^2(\mathbf{x}^*) : 0.3\|\mathbf{x}\|^2 \leq \varepsilon^o\right\}$ and $\hat{X}_3^2 = \left\{\mathbf{x} \in \mathcal{N}_1^2(\mathbf{x}^*) : \|\mathbf{x}\|^2 \leq 4\varepsilon^f\right\}$ (since f is linear and g_1 is quadratic).

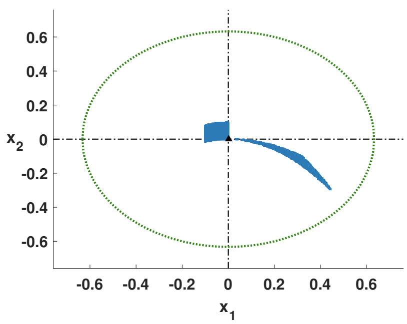
Finally, consider Example 5.3.20 with $\varepsilon^o, \varepsilon^f \leq 0.1$, and recall that $X = (-2,2) \times (-2,2)$, $m_I = 3$, and $m_E = 0$ with $f(\mathbf{x}) = 2x_1^2 + x_2$, $g_1(\mathbf{x}) = -x_1^2 - x_2$, $g_2(\mathbf{x}) = -x_1$, $g_3(\mathbf{x}) = x_1^2 + x_2^2 - 1$, and $\mathbf{x}^* = (0,0)$. We have $\mathcal{F}(X) = \left\{\mathbf{x} : x_2 \geq -x_1^2,\ x_1 \geq 0,\ x_1^2 + x_2^2 \leq 1\right\}$ and

$$X_3 = \Bigg\{\mathbf{x} \in X \ : \sqrt{\left(\max\{0, -x_1^2 - x_2\}\right)^2 + \left(\max\{0, -x_1\}\right)^2 + \left(\max\{0, x_1^2 + x_2^2 - 1\}\right)^2} \in (0, \varepsilon^f], \\ 2x_1^2 + x_2 \leq \varepsilon^o\Bigg\}.$$

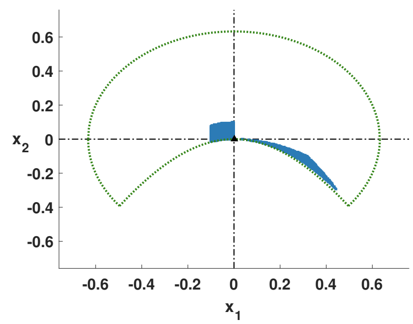
Choose $\alpha = \frac{2}{3}$. We have $\mathcal{D}_I = \left\{\mathbf{d} \in \mathcal{N}_{\frac{2}{3}}^2(\mathbf{0}) : (\mathbf{x}^* + \mathbf{d}) \in (\mathcal{F}(X))^{\mathrm{C}}\right\}$. Choose $\mathcal{D}_0 = \left\{\mathbf{d} \in \mathcal{D}_I : d_2 \geq -1.5 d_1^2\right\}$, $\gamma_1 = 0.25$, $\gamma_2 = 0.25$ and $\hat{\alpha} = \frac{2}{3}$ in Lemma 5.3.33. We have from Lemma 5.3.33 and Remark 5.3.34 that $\hat{X}_3^1 = \left\{\mathbf{x} : \|\mathbf{x}\|^2 \leq 4\varepsilon^o\right\}$ and $\hat{X}_3^2 = \left\{\mathbf{x} : \|\mathbf{x}\|^2 \leq 4\varepsilon^f\right\}$ (since f and g_2 are quadratic, and g_1 is linear). Figure 5-6 plots the sets X_3^1 and X_3^2 along with their estimates $\hat{X}_3^1$ and $\hat{X}_3^2$ for $\varepsilon^o = \varepsilon^f = 0.1$. The benefit of using the estimates in Remark 5.3.34 over that of Lemma 5.3.33 is seen from Figure 5-6.

The following result follows from Lemma 3 in [238]. It provides a conservative estimate of the number of boxes of certain widths required to cover $\hat{X}_3^1$ and $\hat{X}_3^2 \backslash B_\delta$ from Lemma 5.3.33. Therefore, from Lemmata 5.2.5 and 5.2.6 and the result below, we can get an upper bound on the worst-case number of boxes required to cover $\mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap X_3$ and estimate the extent of the cluster problem on that region.

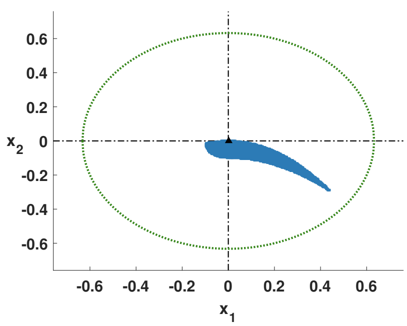
Theorem 5.3.35. Suppose the assumptions of Lemma 5.3.33 hold. Let $\delta = \left(\frac{\varepsilon}{\tau^*}\right)^{\frac{1}{\beta^*}} = \delta_f = \left(\frac{\varepsilon^o}{\tau^f}\right)^{\frac{1}{\beta^f}} = \left(\frac{\varepsilon^f}{\tau^I}\right)^{\frac{1}{\beta^I}}$, $\delta_I = \left(\frac{\gamma_2 \delta^2}{8\tau^I}\right)^{\frac{1}{\beta^I}} = \left(\frac{\gamma_2}{8\tau^I}\right)^{\frac{1}{\beta^I}} \left(\frac{\varepsilon^f}{\tau^I}\right)^{\frac{2}{(\beta^I)^2}}$, $r_I = \sqrt{\frac{2\varepsilon^f}{\gamma_2}}$, $r_f = \sqrt{\frac{2\varepsilon^o}{\gamma_1}}$.



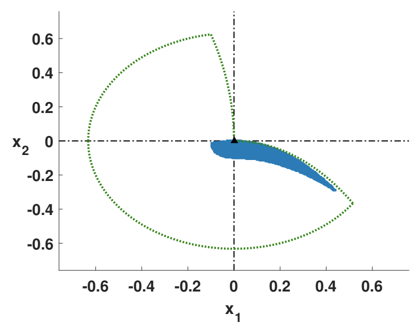
(a) X_3^1 and estimate $\hat{X}_3^1$ from Lemma 5.3.33



(b) X_3^1 and estimate $\hat{X}_3^1$ from Remark 5.3.34



(c) X_3^2 and estimate $\hat{X}_3^2$ from Lemma 5.3.33



(d) X_3^2 and estimate $\hat{X}_3^2$ from Remark 5.3.34

Figure 5-6: Plots of X_3^1 and X_3^2 (solid regions) and their estimates $\hat{X}_3^1$ and $\hat{X}_3^2$ (area between the dotted lines) for Example 5.3.20. The filled-in triangles correspond to the minimizer $\mathbf{x}^*$, and the dash-dotted lines represent the axes translated to $\mathbf{x}^*$. All plots use $\varepsilon^o, \varepsilon^f = 0.1$.

1. If $\delta_I \geq 2r_I$, let $N_I = 1$.

2. If $\dfrac{2r_I}{\sqrt{\bar{m}_I - 1}} > \delta_I \geq \dfrac{2r_I}{\sqrt{\bar{m}_I}}$ for some $\bar{m}_I \in \mathbb{N}$ with $\bar{m}_I \leq n_x$ and $2 \leq \bar{m}_I \leq 18$, then let

$$N_I = \sum_{i=0}^{\bar{m}_I - 1} 2^i \binom{n_x}{i} + 2n_x \left\lceil \frac{\bar{m}_I - 9}{9} \right\rceil.$$

3. Otherwise, let

$$N_I = \left\lceil 2B_I(\varepsilon^f; \beta^I, \gamma_2, \tau_I) \right\rceil^{n_x - 1} \left(\left\lceil 2B_I(\varepsilon^f; \beta^I, \gamma_2, \tau_I) \right\rceil + 2n_x \left\lceil (\sqrt{2} - 1) B_I(\varepsilon^f; \beta^I, \gamma_2, \tau_I) \right\rceil \right),$$

where

$$B_I(\varepsilon^f; \beta^I, \gamma_2, \tau_I) := 8^{\frac{1}{\beta^I}} \left(\tau^I\right)^{\left(\frac{1}{\beta^I} + \frac{2}{(\beta^I)^2}\right)} \left(\varepsilon^f\right)^{\left(\frac{1}{2} - \frac{2}{(\beta^I)^2}\right)} \gamma_2^{-\left(\frac{1}{2} + \frac{1}{\beta^I}\right)}.$$

4. If $\delta_f \geq 2r_f$, let $N_f = 1$.

5. If $\dfrac{2r_f}{\sqrt{m_f - 1}} > \delta_f \geq \dfrac{2r_f}{\sqrt{m_f}}$ for some $m_f \in \mathbb{N}$ with $m_f \leq n_x$ and $2 \leq m_f \leq 18$, then let

$$N_f = \sum_{i=0}^{m_f - 1} 2^i \binom{n_x}{i} + 2n_x \left\lceil \frac{m_f - 9}{9} \right\rceil.$$

6. Otherwise, let

$$N_f = \left\lceil 2\left(\tau^f\right)^{\frac{1}{\beta^f}} (\varepsilon^o)^{\left(\frac{1}{2} - \frac{1}{\beta^f}\right)} \gamma_1^{-\frac{1}{2}} \right\rceil^{n_x - 1} \left(\left\lceil 2\left(\tau^f\right)^{\frac{1}{\beta^f}} (\varepsilon^o)^{\left(\frac{1}{2} - \frac{1}{\beta^f}\right)} \gamma_1^{-\frac{1}{2}} \right\rceil + \right.$$
$$\left. 2n_x \left\lceil (\sqrt{2} - 1)\left(\tau^f\right)^{\frac{1}{\beta^f}} (\varepsilon^o)^{\left(\frac{1}{2} - \frac{1}{\beta^f}\right)} \gamma_1^{-\frac{1}{2}} \right\rceil \right).$$

Then, N_I is an upper bound on the number of boxes of width δ_I required to cover $\hat{X}_3^2 \backslash B_\delta$, and N_f is an upper bound on the number of boxes of width δ_f required to cover $\hat{X}_3^1$.

Proof. The result on N_f follows from Lemmata 5.2.6 and 5.3.33, and Lemma 3 in [238]. To deduce the result on N_I, note that we cover $\hat{X}_3^2 \backslash B_\delta$ with boxes of width $\delta_I = \left(\dfrac{\gamma_2 \delta^2}{8\tau^I}\right)^{\frac{1}{\beta^I}}$

since, from Lemma 5.3.33, we have

$$\hat{X}_3^2 \backslash B_\delta \subset \left\{ \mathbf{x} \in \mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) : d\left(\begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix} (\mathbf{x}), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\} \right) \in \left(\frac{\gamma_2}{8}\delta^2, \varepsilon^f \right] \right\}$$

and, from Lemma 5.2.5, we have that a box B_{δ_I} of width δ_I with each $\mathbf{x} \in B_{\delta_I}$ satisfying $d\left(\begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix} (\mathbf{x}), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\} \right) > \frac{\gamma_2}{8}\delta^2$ can be fathomed by infeasibility. The desired result then follows from Lemma 3 in [238]. □

Remark 5.3.36.

1. Under the assumptions of Lemma 5.3.33, the dependence of N_I on ε^f disappears when the lower bounding scheme has second-order convergence on $\mathcal{N}_{\hat{\alpha}}^2(\mathbf{x}^*) \cap (\mathcal{F}(X))^{\mathrm{C}}$, i.e., $\beta^I = 2$, and the dependence of N_f on ε^o disappears when the scheme $(f_Z^{\mathrm{cv}})_{Z \in \mathbb{I}X}$ has second-order convergence on X, i.e., $\beta^f = 2$. Therefore, the cluster problem on X_3 can be eliminated using second-order convergent schemes with sufficiently small prefactors.

2. The dependence of N_I on ε^f for $\beta^I = 1$, i.e., $N_I \propto \left(\varepsilon^f\right)^{-1.5 n_x}$, scales worse than the corresponding dependence of N on ε for $\beta^* = 1$ when second-order convergence on X_5 is required to mitigate clustering, i.e., $N \propto \varepsilon^{-0.5 n_x}$ (see Theorem 5.3.21). Note, however, that this worse scaling may be an artifact of the conservative requirement that all of $\hat{X}_3^2 \backslash B_\delta$ has to be covered using boxes of size δ_I instead of simply requiring that the subset of $\hat{X}_3^2$ that is not fathomed by value dominance (the rest of $\hat{X}_3^2$, including B_δ, would have already been accounted for while covering $\hat{X}_5$ and $\hat{X}_3^1$) be covered using boxes of appropriate size.

3. Similar to Lemma 5.3.25, less conservative estimates (with respect to the dependence on ε^o and ε^f) may be obtained for X_3^1 and X_3^2 by taking into account the fact that the objective function and the measure of infeasibility grow linearly in certain directions.

Remark 5.3.37. The main assumptions of Lemmata 5.3.2 and 5.3.27, which assume that the objective function and the measure of infeasibility grow linearly on certain regions in some neighborhood of $\mathbf{x}^*$, are similar to the linear growth condition in [104], and the main assumptions of Lemmata 5.3.17 and 5.3.33, which assume that the objective function and the measure of infeasibility grow quadratically on certain regions in some neighborhood of $\mathbf{x}^*$, are similar to the quadratic growth condition in [43, 104]. Furthermore, the assumptions of Lemmata 5.3.2, 5.3.17, 5.3.27, and 5.3.33 may be weakened based on the linear and quadratic growth conditions in [43, 104] to account for cases in which $\mathbf{x}^*$ is not a strict local minimum.

5.4 Conclusion

This chapter provides an analysis of the cluster problem for constrained problems. The analysis indicates different scaling of the number of boxes required to cover regions close to a global minimizer based on the convergence order and corresponding prefactor of the lower bounding scheme on nearly-optimal and nearly-feasible regions in the vicinity of the global minimizer.

It is shown that lower bounding schemes with first-order convergence may eliminate the cluster problem at a constrained minimizer if: i. the objective function grows linearly in directions leading to feasible points in some neighborhood of the minimizer, ii. either the objective function, or a measure of constraint violation grows linearly in directions leading to infeasible points in some neighborhood of the minimizer, and iii. the corresponding convergence order prefactors are sufficiently-small. This is shown to be possible because nodes containing nearly-optimal and nearly-feasible points may be fathomed relatively easily, by value dominance or by infeasibility, even using first-order convergent lower bounding schemes when the objective function or the measure of constraint violation grows linearly in directions around the minimizer. The above result is in contrast to the case of unconstrained minimization where at least second-order convergence is required to eliminate the cluster problem at a point of differentiability of the objective function. When the objective function is twice-differentiable at an unconstrained minimizer, this is a consequence of the fact that the objective function grows quadratically or slower around the minimizer.

It is also shown that at least second-order convergence is required to mitigate the cluster problem at a nonisolated constrained minimizer that satisfies certain regularity conditions when the problem is purely equality-constrained. Conditions under which second-order convergence of bounding schemes is sufficient to mitigate clustering are presented. This analysis reduces to previous analyses for unconstrained problems under suitable assumptions.

Chapter 6

Convergence-order analysis of branch-and-bound algorithms for constrained problems

The performance of branch-and-bound algorithms for deterministic global optimization is strongly dependent on the ability to construct tight and rapidly convergent schemes of lower bounds. One metric of the efficiency of a branch-and-bound algorithm is the convergence order of its bounding scheme. This chapter develops a notion of convergence order for lower bounding schemes for constrained problems, and defines (and analyzes) the convergence order of convex relaxation-based and Lagrangian dual-based lower bounding schemes (see Chapter 5 for the motivation behind the analysis in this chapter). The material in this chapter has been published as the article [109].

6.1 Introduction

Global optimization has found widespread applications in various areas of engineering and the sciences [80]. Deterministic global optimization algorithms attempt to determine an approximate optimal solution within a specified tolerance and terminate with a certificate of its optimality in finite time [101]. While efficient algorithms are known for classes of convex optimization problems [24], no such algorithms are currently known for most classes of nonconvex problems. Deterministic global optimization algorithms for nonconvex problems usually involve the concept of partitioning the domain of ('branching on') the decision variables [101]. The performance of branch-and-bound algorithms for deterministic global optimization is strongly dependent on the ability to construct tight and rapidly convergent schemes of relaxations of nonconvex functions.

Since the worst-case running time of all known branch-and-bound algorithms is exponential in the dimension of the variables partitioned, it may be advantageous to utilize 'reduced-space' algorithms which only require branching on a subset of the variables (as opposed to 'full-space' branch-and-bound algorithms which may branch on all of the variables, see Section 2.3.2.3 of Chapter 2) to guarantee convergence. Despite the potential advantages of reduced-space algorithms for nonconvex problems [20, 69, 76, 237], such methods have not been widely adopted in the literature and in commercial software. One potential reason is that most widely-applicable reduced-space branch-and-bound algorithms often do not seem to exhibit favorable convergence rates compared to their full-space counterparts. The convergence properties of reduced-space branch-and-bound algorithms have not been thoroughly investigated, although some progress has been made in this direction [70, 237]. The reader is directed to the work of Epperly and Pistikopoulos [76] for a survey of reduced-space branch-and-bound algorithms.

One metric of the efficiency of a deterministic branch-and-bound algorithm is the order of convergence of its bounding scheme, which, for the case of unconstrained optimization, compares the rate of convergence of an estimated range of a function to its true range [172]. Recently, Bompadre and coworkers [38, 39] developed the notions of Hausdorff and pointwise convergence orders of bounding schemes and established sharp rules for the propagation of convergence orders of bounding schemes constructed using McCormick [154], Taylor [184], and McCormick-Taylor [197] models. In addition, they showed that if a function is twice continuously differentiable, the scheme of relaxations corresponding to its envelopes is at least second-order pointwise convergent which, in turn, implies Hausdorff convergence of at least second-order (see Theorem 2.3.38 and Lemma 2.3.36). Najman and Mitsos [174] used the framework developed in [38, 39] to establish sharp rules for the propagation of convergence orders of multivariate McCormick relaxations [227]. Khan and coworkers [124] developed a continuously differentiable variant of McCormick relaxations [154, 227], and established second-order pointwise convergence of schemes of the differentiable McCormick relaxations for twice continuously differentiable functions. Also note the definition of rate of convergence of bounding schemes for geometric branch-and-bound methods proposed by Schöbel and Scholz [203], and the proof of second-order Hausdorff convergence of centered forms in [127, 205]. Establishing that a scheme of relaxations is at least second-order Hausdorff convergent is important from many viewpoints, notably in mitigating the so-called cluster effect in unconstrained global optimization [68, 238]. Chapter 5 analyzed the cluster problem for constrained global optimization where it was shown that, under certain conditions, first-order convergence of the lower bounding scheme may be sufficient to avoid the cluster problem at constrained minima (see Theorems 5.3.11 and 5.3.31). However, an analysis of convergence order for constrained problems is lacking.

In this chapter, we investigate the convergence orders of some full-space and reduced-space deterministic branch-and-bound algorithms by extending the convergence analysis of Bompadre and coworkers to constrained problems. This chapter develops a notion of convergence order for lower bounding schemes for constrained problems, and defines the convergence order of convex relaxation-based and Lagrangian dual-based lower bounding schemes. It is shown that full-space convex relaxation-based lower bounding schemes can achieve first-order convergence under mild assumptions. Furthermore, such schemes can achieve second-order convergence at KKT points, at Slater points, and at infeasible points when second-order pointwise convergent schemes of relaxations are used. Lagrangian dual-based full-space lower bounding schemes are shown to have at least as high a convergence order as convex relaxation-based full-space lower bounding schemes. Additionally, it is shown that Lagrangian dual-based full-space lower bounding schemes achieve first-order convergence even when the dual problem is not solved to optimality. The convergence order of some widely-applicable reduced-space lower bounding schemes is also analyzed, and it is shown that such schemes can achieve first-order convergence under suitable assumptions. Furthermore, such schemes can achieve second-order convergence at KKT points, at unconstrained points in the reduced-space, and at infeasible points under suitable assumptions when the problem exhibits a specific separable structure. The importance of constraint propagation techniques in boosting the convergence order of reduced-space lower bounding schemes (and helping mitigate clustering in the process) for problems which do not possess such a structure is demonstrated. Throughout this chapter, we tacitly assume that a branch-and-bound algorithm utilizes efficient heuristics for finding feasible points which determine a global optimal solution early on in the branch-and-bound tree (if one exists).

This chapter is organized as follows. Section 6.2 formulates the problem of interest, and provides some background definitions. Section 6.3 develops the notion of convergence order of a lower bounding scheme, and Section 6.4 provides some results on the convergence orders of commonly-used full-space lower bounding schemes. Section 6.5 lists some widely-applicable reduced-space lower bounding schemes in the literature, provides some results on their convergence orders, and highlights the importance of constraint propagation in reduced-space branch-and-bound algorithms. Finally, Section 6.6 lists the conclusions and some avenues for future work.

6.2 Problem formulation and background

In this chapter, we consider the formulation

$$
\begin{aligned}
\min_{\mathbf{x},\mathbf{y}} \ & f(\mathbf{x},\mathbf{y}) && \text{(P)}\\
\text{s.t.} \ & \mathbf{g}(\mathbf{x},\mathbf{y}) \leq \mathbf{0},\\
& \mathbf{h}(\mathbf{x},\mathbf{y}) = \mathbf{0},\\
& \mathbf{x} \in X, \ \ \mathbf{y} \in Y,
\end{aligned}
$$

where $X \subset \mathbb{R}^{n_x}$ and $Y \subset \mathbb{R}^{n_y}$ are nonempty convex sets, $f : X \times Y \to \mathbb{R}$ and $\mathbf{g} : X \times Y \to \mathbb{R}^{m_I}$ are partially convex with respect to $\mathbf{x}$, i.e., $f(\cdot, \mathbf{y})$ and $\mathbf{g}(\cdot, \mathbf{y})$ are convex on X for each $\mathbf{y} \in Y$, and $\mathbf{h} : X \times Y \to \mathbb{R}^{m_E}$ is affine with respect to $\mathbf{x}$, i.e., $\mathbf{h}(\cdot, \mathbf{y})$ is affine on X for each $\mathbf{y} \in Y$. The following assumption will be made throughout this chapter.

Assumption 6.2.1. The sets X and Y are compact, and the functions f, $\mathbf{g}$, and $\mathbf{h}$ are continuous on $X \times Y$.

When the dimension n_y of the Y-space corresponding to the nonconvexities in the functions in Problem (P) is significantly smaller than the dimension n_x of the X-space, it may be computationally advantageous to partition only the Y-space during the course of a branch-and-bound algorithm (assuming, of course, that the reduced-space algorithm is guaranteed to converge). However, the convergence rate of a reduced-space branch-and-bound algorithm may be different compared to a similar full-space algorithm, which makes it difficult to judge *a priori* whether using a reduced-space branch-and-bound approach would be advantageous. Before we analyze the convergence orders of some full-space and reduced-space lower bounding schemes in the literature, we need to define formally the notion of convergence order for constrained problems. For this purpose, we assume that the reader is familiar with the notation and the background definitions introduced in Chapter 2 (in particular, we will use Definitions 2.2.4, 2.2.5, 2.3.5, 2.3.23, 2.3.24, 2.3.25, 2.3.27, 2.3.28, 2.3.29, 2.3.31, and 2.3.34, Lemmata 2.2.2, 2.3.6, 2.3.30, and 2.3.35, and Corollary 2.3.7). We restate the definition of convex and concave relaxations from Chapter 2 because it is central to the analysis in this chapter.

Definition 2.3.28. [Convex and Concave Relaxations] Given a convex set $Z \subset \mathbb{R}^n$ and a function $f : Z \to \mathbb{R}$, a convex function $f^{\text{cv}}_Z : Z \to \mathbb{R}$ is called a convex relaxation of f on Z if $f^{\text{cv}}_Z(\mathbf{z}) \leq f(\mathbf{z})$, $\forall \mathbf{z} \in Z$. Similarly, a concave function $f^{\text{cc}}_Z : Z \to \mathbb{R}$ is called a concave relaxation of f on Z if $f^{\text{cc}}_Z(\mathbf{z}) \geq f(\mathbf{z})$, $\forall \mathbf{z} \in Z$.

Remark 6.2.2. Although convex and concave relaxations of classes of functions can be constructed on general convex sets, the typical application requires construction of relaxations on bounded intervals. Therefore, we will implicitly assume that the sets X and Y are intervals and that relaxations are constructed on intervals in subsequent sections. The assumption that X and Y are intervals is not restrictive since general convex constraints defining X and Y that are available in factorable form can be equivalently reformulated to appear as part of the constraints $\mathbf{g}$ and $\mathbf{h}$. The proposed definitions of convergence order in the next section will be based on schemes of relaxations constructed on intervals. Note that similar notions of convergence order can be developed for schemes of relaxations constructed, for instance, on simplices.

6.3 Definitions of convergence order

This section reviews the definitions of convergence orders of (reduced-space) schemes of relaxations [38, 39] and defines the convergence order of a (reduced-space) lower bounding scheme (cf. the related full-space definitions in Section 2.3.2.1.2 of Chapter 2). It is also shown that the convergence order of a convergent scheme of relaxations at a point is governed by the tiny intervals around that point. We begin with the following definition, adapted from [38, Definition 6], that defines schemes of relaxations in a reduced-space.

Definition 6.3.1. [Schemes of Convex and Concave Relaxations] Let $V \subset \mathbb{R}^{n_v}$ and $W \subset \mathbb{R}^{n_w}$ be nonempty convex sets, and let $f : V \times W \to \mathbb{R}$. Suppose, for every $Z \in \mathbb{I}W$, we can construct functions $f^{\mathrm{cv}}_{V\times Z} : V \times Z \to \mathbb{R}$ and $f^{\mathrm{cc}}_{V\times Z} : V \times Z \to \mathbb{R}$ that are convex and concave relaxations, respectively, of f on $V \times Z$. The sets of functions $(f^{\mathrm{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ and $(f^{\mathrm{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ define schemes of convex and concave relaxations of f in W, respectively, and the set of pairs of functions $(f^{\mathrm{cv}}_{V\times Z}, f^{\mathrm{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ defines a scheme of relaxations of f in W. The schemes of relaxations are said to be continuous when $f^{\mathrm{cv}}_{V\times Z}$ and $f^{\mathrm{cc}}_{V\times Z}$ are continuous on $V \times Z$ for each $Z \in \mathbb{I}W$.

Bompadre and coworkers [38, 39] define Hausdorff convergence of inclusion functions. Note that an inclusion function can be associated with schemes of relaxations in a natural way (see [38, Definition 7]).

Definition 6.3.2. [Hausdorff Convergence Order of an Inclusion Function] Let $V \in \mathbb{IR}^{n_v}$ and $W \subset \mathbb{R}^{n_w}$ be nonempty sets, $h : V \times W \to \mathbb{R}$ be a continuous function, and H be an inclusion function of h on $\mathbb{I}(V \times W)$.

The inclusion function H is said to have Hausdorff convergence of order $\beta > 0$ at a point $\mathbf{w} \in W$ if for each bounded $Q \subset W$ with $\mathbf{w} \in Q$, there exists $\tau \geq 0$ such that

$$d_H(\overline{h}(V \times Z), H(V \times Z)) \leq \tau w(Z)^{\beta}, \quad \forall Z \in \mathbb{I}Q \text{ with } \mathbf{w} \in Z.$$

Moreover, H is said to have Hausdorff convergence of order $\beta > 0$ on W if it has Hausdorff convergence of order (at least) β at each $\mathbf{w} \in W$, with the constant τ independent of $\mathbf{w}$.

In the context of (constrained) global optimization, the following definition of convergence of schemes of convex and concave relaxations is more pertinent.

Definition 6.3.3. [Convergence Order of Schemes of Convex and Concave Relaxations] Let $V \subset \mathbb{R}^{n_v}$, $W \subset \mathbb{R}^{n_w}$ be nonempty convex sets, and $f : V \times W \to \mathbb{R}$ be a continuous function. Let $(f^{\mathrm{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ and $(f^{\mathrm{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ respectively denote schemes of convex and concave relaxations of f in W.

The scheme of convex relaxations $(f^{\mathrm{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ is said to have convergence of order $\beta > 0$ at $\mathbf{w} \in W$ if for each bounded $Q \subset W$ with $\mathbf{w} \in Q$, there exists $\tau^{\mathrm{cv}} \geq 0$ such that

$$\inf_{(\mathbf{v},\mathbf{z})\in V\times Z} f(\mathbf{v},\mathbf{z}) - \inf_{(\mathbf{v},\mathbf{z})\in V\times Z} f^{\mathrm{cv}}_{V\times Z}(\mathbf{v},\mathbf{z}) \leq \tau^{\mathrm{cv}} w(Z)^{\beta}, \quad \forall Z \in \mathbb{I}Q \text{ with } \mathbf{w} \in Z.$$

Similarly, the scheme of concave relaxations $(f^{\mathrm{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ is said to have convergence of order $\beta > 0$ at $\mathbf{w} \in W$ if for each bounded $Q \subset W$ with $\mathbf{w} \in Q$, there exists $\tau^{\mathrm{cc}} \geq 0$ such that

$$\sup_{(\mathbf{v},\mathbf{z})\in V\times Z} f^{\mathrm{cc}}_{V\times Z}(\mathbf{v},\mathbf{z}) - \sup_{(\mathbf{v},\mathbf{z})\in V\times Z} f(\mathbf{v},\mathbf{z}) \leq \tau^{\mathrm{cc}} w(Z)^{\beta}, \quad \forall Z \in \mathbb{I}Q \text{ with } \mathbf{w} \in Z.$$

The scheme of relaxations $(f^{\mathrm{cv}}_{V\times Z}, f^{\mathrm{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ is said to have (Hausdorff) convergence of order $\beta > 0$ at $\mathbf{w} \in W$ if the corresponding schemes of convex and concave relaxations have convergence of orders (at least) β at $\mathbf{w}$. The schemes $(f^{\mathrm{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ and $(f^{\mathrm{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ are said to have convergence of order $\beta > 0$ on W if they have convergence of order (at least) β at each $\mathbf{w} \in W$, with constants τ^{cv} and τ^{cc} independent of $\mathbf{w}$.

Definition 6.3.4. [Pointwise Convergence Order of Schemes of Convex and Concave Relaxations] Let $V \subset \mathbb{R}^{n_v}$, $W \subset \mathbb{R}^{n_w}$ be nonempty convex sets, and $f : V \times W \to \mathbb{R}$ be a continuous function. Let $(f^{\mathrm{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ and $(f^{\mathrm{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ respectively denote schemes of convex and concave relaxations of f in W. The scheme of convex relaxations $(f^{\mathrm{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ is said to have pointwise convergence of order $\gamma > 0$ at $\mathbf{w} \in W$ if for each bounded $Q \subset W$ with $\mathbf{w} \in Q$, there exists $\tau^{\mathrm{cv}} \geq 0$ such that

$$\sup_{(\mathbf{v},\mathbf{z})\in V\times Z} \left| f(\mathbf{v},\mathbf{z}) - f^{\mathrm{cv}}_{V\times Z}(\mathbf{v},\mathbf{z}) \right| \leq \tau^{\mathrm{cv}} w(Z)^{\gamma}, \quad \forall Z \in \mathbb{I}Q \text{ with } \mathbf{w} \in Z.$$

Similarly, the scheme of concave relaxations $(f^{\mathrm{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ is said to have pointwise convergence of order $\gamma > 0$ at $\mathbf{w} \in W$ if for each bounded $Q \subset W$ with $\mathbf{w} \in Q$, there exists $\tau^{\mathrm{cc}} \geq 0$ such that

$$\sup_{(\mathbf{v},\mathbf{z})\in V\times Z} \left| f^{\mathrm{cc}}_{V\times Z}(\mathbf{v},\mathbf{z}) - f(\mathbf{v},\mathbf{z}) \right| \leq \tau^{\mathrm{cc}} w(Z)^{\gamma}, \quad \forall Z \in \mathbb{I}Q \text{ with } \mathbf{w} \in Z.$$

The scheme of relaxations $(f^{\mathrm{cv}}_{V\times Z}, f^{\mathrm{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ is said to have pointwise convergence of order $\gamma > 0$ at $\mathbf{w} \in W$ if the corresponding schemes of convex and concave relaxations have pointwise convergence of orders (at least) γ at $\mathbf{w}$. Furthermore, the schemes of relaxations are said to have pointwise convergence of order $\gamma > 0$ on W if they have pointwise convergence of order at least γ at each $\mathbf{w} \in W$, with constants τ^{cv} and τ^{cc} independent of $\mathbf{w}$.

Note that we simply say that a scheme of relaxations, $(f^{\mathrm{cv}}_{V\times Z}, f^{\mathrm{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$, of a function

$f : V \times W \to \mathbb{R}$ in W has (pointwise) convergence order of $\gamma > 0$ if it has (pointwise) convergence of order γ on W.

Remark 6.3.5. Definitions 6.3.2, 6.3.3, and 6.3.4 are based on a modification (see [123, Definition 9.2.35]) of the definitions of convergence order proposed in [38, 39], which incorporates the set Q. Note that the use of the set Q is necessary when the schemes of relaxations are constructed on unbounded sets, but may be omitted (set to W) when the schemes of relaxations are constructed over bounded sets (which is the typical application). Henceforth, the use of Q shall be omitted for brevity since we are only interested in compact sets V and W (see Assumption 6.2.1).

Remark 6.3.6. The pointwise convergence order of a convergent scheme of convex and concave relaxations on W is governed by the strength of the relaxations over small intervals in W. This observation is made precise in Lemma 6.3.8. Also note that the pointwise convergence order of schemes of either convex, or concave relaxations (as per Definition 6.3.4) can be arbitrarily high for nonlinear functions in contrast to the pointwise convergence order of schemes of convex *and* concave relaxations (see Theorem 2 in [38]). For instance, consider the function $f : [0,1] \times [0,1] =: V \times W \to \mathbb{R}$ with $f(v, w) = v^2 - \sqrt{w}$ and a corresponding scheme of convex relaxations $(f^{\text{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ defined by $f^{\text{cv}}_{V\times Z}(v, z) = v^2 - \sqrt{w}$ on $[w^{\text{L}}, w^{\text{U}}] \subset [0, 1]$. The scheme of convex relaxations $(f^{\text{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ has arbitrarily high pointwise convergence order on W.

Remark 6.3.7. Unlike the pointwise convergence order of a scheme of relaxations, the convergence order of a scheme of convex and concave relaxations can be arbitrarily high for any function. For instance, consider the scheme of constant relaxations of the function $f : [0, 1]\times [0, 1] =: V \times W \to \mathbb{R}$ with $f(v, w) = w - \sqrt{v}$ defined by $f^{\text{cv}}_{V\times Z}(v, z) = w^{\text{L}} - 1$, $f^{\text{cc}}_{V\times Z}(v, z) = w^{\text{U}}$ on $[w^{\text{L}}, w^{\text{U}}] \subset [0, 1]$. The scheme of constant relaxations $(f^{\text{cv}}_{V\times Z}, f^{\text{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ has arbitrarily high convergence order on W, but is not pointwise convergent of any order on W.

Lemma 6.3.8. Let $V \subset \mathbb{R}^{n_v}$, $W \subset \mathbb{R}^{n_w}$ be nonempty compact convex sets and $f : V \times W \to \mathbb{R}$. Let $(f^{\text{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ denote a scheme of convex relaxations of f in W with pointwise convergence order $\gamma^{\text{cv}} > 0$ and corresponding prefactor $\tau^{\text{cv}} \geq 0$ (on W). If there exist constants $\gamma \geq \gamma^{\text{cv}}$, $\tau \geq 0$, and $\delta > 0$ such that for every $Z \in \mathbb{I}W$ with $w(Z) \leq \delta$,

$$\sup_{(\mathbf{v},\mathbf{z})\in V\times Z} \left| f(\mathbf{v}, \mathbf{z}) - f^{\text{cv}}_{V\times Z}(\mathbf{v}, \mathbf{z}) \right| \leq \tau w(Z)^{\gamma},$$

then $(f^{\mathrm{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ converges pointwise with order γ to f on W.

Proof. Since $(f^{\mathrm{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ converges pointwise with order γ^{cv} to f on W which is compact, there exists $M \geq 0$ such that

$$\sup_{(\mathbf{v},\mathbf{z})\in V\times Z} \left| f(\mathbf{v},\mathbf{z}) - f^{\mathrm{cv}}_{V\times Z}(\mathbf{v},\mathbf{z}) \right| \leq \tau^{\mathrm{cv}} w(Z)^{\gamma^{\mathrm{cv}}} \leq M, \quad \forall Z \in \mathbb{I}W.$$

The desired result then follows from the fact that for every $Z \in \mathbb{I}W$,

$$\sup_{(\mathbf{v},\mathbf{z})\in V\times Z} \left| f(\mathbf{v},\mathbf{z}) - f^{\mathrm{cv}}_{V\times Z}(\mathbf{v},\mathbf{z}) \right| \leq \left(\tau + \frac{M}{\delta^{\gamma}} \right) w(Z)^{\gamma}. \qquad \square$$

Results similar to Lemma 6.3.8 are applicable to other notions of convergence order presented in this chapter and will be used freely. Note that if the constant δ in Lemma 6.3.8 is relatively small, then the bound on the prefactor obtained can be relatively large making the result weak on intervals with $w(Z) \gg \delta$.

The next result shows that for schemes of relaxations, the notion of pointwise convergence is stronger than the notion of convergence in Definition 6.3.3 (also see [38, Theorem 1]).

Lemma 6.3.9. Let $V \subset \mathbb{R}^{n_v}$, $W \subset \mathbb{R}^{n_w}$ be nonempty compact convex sets, and $(f^{\mathrm{cv}}_{V\times Z})_{Z\in\mathbb{I}W}$ and $(f^{\mathrm{cc}}_{V\times Z})_{Z\in\mathbb{I}W}$ respectively denote schemes of convex and concave relaxations of a bounded function $f : V \times W \to \mathbb{R}$ in W. If either scheme has pointwise convergence of order $\gamma > 0$, it has convergence of order $\beta \geq \gamma$.

Proof. The proof is similar to that of Lemma 2.3.36, and is therefore omitted. $\square$

The following lemma establishes mild sufficient conditions under which the scheme of envelopes of a function is first-order pointwise convergent.

Lemma 6.3.10. Let $W \subset \mathbb{R}^{n_w}$ be a nonempty compact convex set and $f : W \to \mathbb{R}$ be Lipschitz continuous on W. Let $(f^{\mathrm{cv,env}}_{Z}, f^{\mathrm{cc,env}}_{Z})_{Z\in\mathbb{I}W}$ denote the scheme of envelopes of f in W. Then the scheme $(f^{\mathrm{cv,env}}_{Z}, f^{\mathrm{cc,env}}_{Z})_{Z\in\mathbb{I}W}$ is at least first-order pointwise convergent on W.

Proof. We wish to show that there exists $\tau \geq 0$ such that for every $Z \in \mathbb{I}W$,

$$\sup_{\mathbf{z}\in Z}\left|f(\mathbf{z}) - f_Z^{\text{cv,env}}(\mathbf{z})\right| \leq \tau w(Z),$$
$$\sup_{\mathbf{z}\in Z}\left|f(\mathbf{z}) - f_Z^{\text{cc,env}}(\mathbf{z})\right| \leq \tau w(Z).$$

Consider the scheme of relaxations $(f_Z^{\text{cv}}, f_Z^{\text{cc}})_{Z\in\mathbb{I}W}$ defined by

$$f_Z^{\text{cv}}(\mathbf{z}) = \min_{\mathbf{w}\in Z} f(\mathbf{w}), \;\; f_Z^{\text{cc}}(\mathbf{z}) = \max_{\mathbf{w}\in Z} f(\mathbf{w}), \quad \forall Z \in \mathbb{I}W.$$

From the fact that f_Z^{cv} and f_Z^{cc} are convex and concave relaxations of f in Z and the assumption that f is Lipschitz continuous, we have that $(f_Z^{\text{cv}}, f_Z^{\text{cc}})_{Z\in\mathbb{I}W}$ is at least first-order pointwise convergent on W. The desired result then follows from the definition of $(f_Z^{\text{cv,env}}, f_Z^{\text{cc,env}})_{Z\in\mathbb{I}W}$. $\square$

Remark 6.3.11. Locally Lipschitz continuous functions are Lipschitz continuous on compact subsets of their domains (see Definition 2.2.4). Therefore, the assumption that the functions f, $\mathbf{g}$, and $\mathbf{h}$ in Problem (P) are Lipschitz continuous on $X \times Y$ is not particularly strong when Assumption 6.2.1 is made.

The definitions provided thus far facilitate a theoretical analysis of the (reduced-space) convergence order of a scheme of relaxations to a corresponding scalar function, or, in the context of global optimization, provide a way to analyze theoretically the (reduced-space) convergence order of a (lower) bounding scheme for an unconstrained problem. The subsequent definitions seek to extend naturally the analysis of convergence order to constrained problems.

Definition 6.3.12. [Convergence Order of a Lower Bounding Scheme] Consider Problem (P) (satisfying Assumption 6.2.1). For any $Z \in \mathbb{I}Y$, let

$$\mathcal{F}(Z) = \{(\mathbf{x}, \mathbf{y}) \in X \times Z : \mathbf{g}(\mathbf{x}, \mathbf{y}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x}, \mathbf{y}) = \mathbf{0}\}$$

denote the feasible set of Problem (P) with $\mathbf{y}$ restricted to Z.

Consider a scheme of lower bounding problems $(\mathcal{L}(Z))_{Z\in\mathbb{I}Y}$ for Problem (P). We associate with the scheme $(\mathcal{L}(Z))_{Z\in\mathbb{I}Y}$ a scheme of pairs $(\mathcal{O}(Z), \mathcal{I}_C(Z))_{Z\in\mathbb{I}Y}$, where $(\mathcal{O}(Z))_{Z\in\mathbb{I}Y}$

is a scheme of lower bounds on the scheme of problems $\left(\min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}(Z)} f(\mathbf{x},\mathbf{y})\right)_{Z\in\mathbb{I}Y}$ and the scheme $(\mathcal{I}_C(Z))_{Z\in\mathbb{I}Y}$ is a scheme of subsets of $\mathbb{R}^{m_I+m_E}$ that indicate the feasibility of the lower bounding scheme $(\mathcal{L}(Z))_{Z\in\mathbb{I}Y}$. The schemes $(\mathcal{O}(Z))_{Z\in\mathbb{I}Y}$ and $(\mathcal{I}_C(Z))_{Z\in\mathbb{I}Y}$ (are required to) satisfy

$$\begin{aligned}
&\mathcal{O}(Z) \leq \min_{(\mathbf{x},\mathbf{z})\in\mathcal{F}(Z)} f(\mathbf{x},\mathbf{z}), \quad \forall Z \in \mathbb{I}Y,\\
&d(\mathcal{I}_C(Z), \mathbb{R}^{m_I}_{-} \times \{\mathbf{0}\}) \leq d\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(X\times Z), \mathbb{R}^{m_I}_{-} \times \{\mathbf{0}\}\right), \quad \forall Z \in \mathbb{I}Y,\\
&\mathcal{O}(Z) = +\infty \iff d(\mathcal{I}_C(Z), \mathbb{R}^{m_I}_{-} \times \{\mathbf{0}\}) > 0, \quad \forall Z \in \mathbb{I}Y,
\end{aligned}$$

where $\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(X\times Z)$ denotes the image of $X \times Z$ under the vector-valued function $\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}$. The scheme of lower bounding problems $(\mathcal{L}(Z))_{Z\in\mathbb{I}Y}$ is said to have convergence of order $\beta > 0$ at

1. a feasible point $\mathbf{y} \in Y$ if there exists $\tau \geq 0$ such that for every $Z \in \mathbb{I}Y$ with $\mathbf{y} \in Z$,

$$\min_{(\mathbf{x},\mathbf{z})\in\mathcal{F}(Z)} f(\mathbf{x},\mathbf{z}) \; - \mathcal{O}(Z) \leq \tau w(Z)^{\beta}.$$

2. an infeasible point $\mathbf{y} \in Y$ if there exists $\bar{\tau} \geq 0$ such that for every $Z \in \mathbb{I}Y$ with $\mathbf{y} \in Z$,

$$d\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(X\times Z), \mathbb{R}^{m_I}_{-} \times \{\mathbf{0}\}\right) - d\left(\mathcal{I}_C(Z), \mathbb{R}^{m_I}_{-} \times \{\mathbf{0}\}\right) \leq \bar{\tau} w(Z)^{\beta}.$$

The scheme of lower bounding problems is said to have convergence of order $\beta > 0$ on Y if it has convergence of order (at least) β at each $\mathbf{y} \in Y$, with constants τ and $\bar{\tau}$ independent of $\mathbf{y}$.

Remark 6.3.13. Definition 6.3.12 is motivated by the requirements of a lower bounding scheme to fathom feasible and infeasible regions in a branch-and-bound procedure [101] (also see Definition 5.2.3). The first condition requires that the sequence of lower bounds converges rapidly to the corresponding sequence of minimum objective values on nested sequences of intervals converging to a feasible point of Problem (P). On nested sequences of intervals converging to an infeasible point of Problem (P), the second condition requires that the sequence of lower bounding problems rapidly detect the (eventual) infeasibility of the corresponding sequences of intervals for Problem (P). In simple terms, the first condition can be used to require that feasible points with 'good objective values' are fathomed rather easily, while the second condition can be used to require that infeasible points that are 'close to the feasible region', as determined by the distance function d, are fathomed with relatively less effort (see Section 5.3 of Chapter 5). Note that Definition 6.3.12 reduces to the definition of convergence order for unconstrained minimization in [238, Definition 1] when n_x, m_I, and m_E are all set to zero.

Definition 6.3.12 can be readily applied to analyze the convergence order of a convex relaxation-based lower bounding scheme as follows.

Suppose, for each $Z \in \mathbb{I}Y$, we associate a convex set $X(Z) \subset \mathbb{R}^{n_x}$ such that $X \supset X(Z) \supset \mathcal{F}_X(Z)$, where $\mathcal{F}_X(Z) := \{\mathbf{x} \in X : \exists \mathbf{y} \in Z \text{ s.t. } \mathbf{g}(\mathbf{x},\mathbf{y}) \leq \mathbf{0}, \mathbf{h}(\mathbf{x},\mathbf{y}) = \mathbf{0}\}$ denotes the projection of $\mathcal{F}(Z)$ on X. The set $X(Z)$ could, for instance, correspond to an interval subset of X that is obtained using bounds tightening techniques [19] when $\mathbf{y}$ is restricted to Z (the motivation for considering the set $X(Z)$ in the definition of convergence order below will become clear in Section 6.5). Note that the restriction $X(Z) \supset \mathcal{F}_X(Z)$ can be relaxed when optimality-based bounds tightening techniques are employed. Also note that unless otherwise specified, we simply use $X(Z) = X$, $\forall Z \in \mathbb{I}Y$.

By an abuse of Definition 6.3.1, let $(f^{\text{cv}}_{X(Z)\times Z})_{Z\in\mathbb{I}Y}$ and $(\mathbf{g}^{\text{cv}}_{X(Z)\times Z})_{Z\in\mathbb{I}Y}$ denote continuous schemes of convex relaxations of f and $\mathbf{g}$, respectively, in Y, and let $(\mathbf{h}^{\text{cv}}_{X(Z)\times Z}, \mathbf{h}^{\text{cc}}_{X(Z)\times Z})_{Z\in\mathbb{I}Y}$ denote a continuous scheme of relaxations of $\mathbf{h}$ in Y. For any $Z \in \mathbb{I}Y$, let

$$\mathcal{F}^{\text{cv}}(Z) = \left\{(\mathbf{x},\mathbf{y}) \in X(Z) \times Z : \mathbf{g}^{\text{cv}}_{X(Z)\times Z}(\mathbf{x},\mathbf{y}) \leq \mathbf{0}, \mathbf{h}^{\text{cv}}_{X(Z)\times Z}(\mathbf{x},\mathbf{y}) \leq \mathbf{0}, \mathbf{h}^{\text{cc}}_{X(Z)\times Z}(\mathbf{x},\mathbf{y}) \geq \mathbf{0}\right\}$$

denote the feasible set of the convex relaxation-based lower bounding scheme. The lower bounding scheme $(\mathcal{L}(Z))_{Z\in\mathbb{I}Y}$ with

$$\begin{aligned}
(\mathcal{O}(Z))_{Z\in\mathbb{I}Y} &:= \left(\min_{(\mathbf{x},\mathbf{z})\in\mathcal{F}^{\text{cv}}(Z)} f^{\text{cv}}_{X(Z)\times Z}(\mathbf{x},\mathbf{z})\right)_{Z\in\mathbb{I}Y}, \\
(\mathcal{I}_C(Z))_{Z\in\mathbb{I}Y} &:= \Big(\Big\{(\mathbf{v},\mathbf{w}) \in \mathbb{R}^{m_I} \times \mathbb{R}^{m_E} : \mathbf{v} = \mathbf{g}^{\text{cv}}_{X(Z)\times Z}(\mathbf{x},\mathbf{z}), \mathbf{h}^{\text{cv}}_{X(Z)\times Z}(\mathbf{x},\mathbf{z}) \leq \mathbf{w} \leq \mathbf{h}^{\text{cc}}_{X(Z)\times Z}(\mathbf{x},\mathbf{z}) \\
&\qquad \text{for some } (\mathbf{x},\mathbf{z}) \in X(Z) \times Z\Big\}\Big)_{Z\in\mathbb{I}Y}
\end{aligned} \tag{6.1}$$

is said to have convergence of order $\beta > 0$ at

1. a feasible point $\mathbf{y} \in Y$ if there exists $\tau \geq 0$ such that for every $Z \in \mathbb{I}Y$ with $\mathbf{y} \in Z$,

$$\min_{(\mathbf{x},\mathbf{z}) \in \mathcal{F}(Z)} f(\mathbf{x}, \mathbf{z}) - \min_{(\mathbf{x},\mathbf{z}) \in \mathcal{F}^{\text{cv}}(Z)} f^{\text{cv}}_{X(Z)\times Z}(\mathbf{x}, \mathbf{z}) \leq \tau w(Z)^{\beta}.$$

2. an infeasible point $\mathbf{y} \in Y$ if there exists $\bar{\tau} \geq 0$ such that for every $Z \in \mathbb{I}Y$ with $\mathbf{y} \in Z$,

$$d\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(X(Z) \times Z), \mathbb{R}^{m_I}_{-} \times \{\mathbf{0}\}\right) - d\left(\mathcal{I}_C(Z), \mathbb{R}^{m_I}_{-} \times \{\mathbf{0}\}\right) \leq \bar{\tau} w(Z)^{\beta},$$

   where $\mathcal{I}_C(Z)$ is defined by Equation (6.1).

Definition 6.3.12 can also be used to analyze the convergence orders of alternative lower bounding schemes such as those based on Lagrangian duality (see Section 6.4.2).

6.4 Full-space branch-and-bound algorithms

In this section, we present some results on the convergence order of lower bounding schemes for Problem (P) when both the X and Y sets may be partitioned during the course of the branch-and-bound algorithm (we consider schemes of relaxations in $X \times Y$ instead of schemes of relaxations in Y as was considered in Section 6.3). This section is divided into two parts. First, we look at the convergence order of lower bounding schemes which utilize convex and concave relaxations (see, for instance, [4, 124, 154, 225, 227] for techniques to construct relaxations) of the objective and the constraints in their construction. Next, the convergence order of duality-based lower bounding schemes (see, for instance, [69]) is investigated.

6.4.1 Convex relaxation-based branch-and-bound

This section derives bounds on the convergence order of convex relaxation-based lower bounding schemes by making assumptions on the convergence orders of the schemes of relaxations used by the lower bounding schemes. The reader is directed to [38], [39], [174], and [124] for details on how to construct schemes of (convex) relaxations that have the requisite convergence orders.

The following result establishes a lower bound on the convergence order of the lower bounding scheme at infeasible points based on the convergence orders of schemes of convex relaxations of the inequality constraints and schemes of relaxations of the equality constraints. Note that this is the primary result that is used to derive a lower bound on the convergence order of such relaxation-based lower bounding schemes at infeasible points.

Lemma 6.4.1. Let $(g^{\text{cv}}_{j,Z})_{Z\in\mathbb{I}(X\times Y)}$, $j = 1,\cdots, m_I$, denote continuous schemes of convex relaxations of $g_1,\cdots,g_{m_I}$, respectively, in $X\times Y$ with pointwise convergence orders $\gamma^{\text{cv}}_{g,1} > 0,\cdots,\gamma^{\text{cv}}_{g,m_I} > 0$ and corresponding constants $\tau^{\text{cv}}_{g,1},\cdots,\tau^{\text{cv}}_{g,m_I}$, and $(h^{\text{cv}}_{k,Z}, h^{\text{cc}}_{k,Z})_{Z\in\mathbb{I}(X\times Y)}$, $k = 1,\cdots, m_E$, denote continuous schemes of relaxations of $h_1,\cdots,h_{m_E}$, respectively, in $X\times Y$ with pointwise convergence orders $\gamma_{h,1} > 0,\cdots,\gamma_{h,m_E} > 0$ and corresponding constants $\tau_{h,1},\cdots,\tau_{h,m_E}$. Then, there exists $\bar{\tau}\geq 0$ such that for every $Z\in\mathbb{I}(X\times Y)$

$$d\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(Z), \mathbb{R}^{m_I}_{-}\times\{\mathbf{0}\}\right) - d\left(\mathcal{I}_C(Z), \mathbb{R}^{m_I}_{-}\times\{\mathbf{0}\}\right) \leq \bar{\tau}w(Z)^{\beta},$$

where $\mathcal{I}_C(Z)$ is defined as

$$\mathcal{I}_C(Z) := \left\{(\mathbf{v},\mathbf{w})\in\mathbb{R}^{m_I}\times\mathbb{R}^{m_E} : \mathbf{v} = \mathbf{g}^{\text{cv}}_Z(\mathbf{x},\mathbf{y}), \mathbf{h}^{\text{cv}}_Z(\mathbf{x},\mathbf{y}) \leq \mathbf{w} \leq \mathbf{h}^{\text{cc}}_Z(\mathbf{x},\mathbf{y}) \text{ for some } (\mathbf{x},\mathbf{y})\in Z\right\},$$

and β is defined as

$$\beta := \min\left\{\min_{j\in\{1,\cdots,m_I\}}\gamma^{\text{cv}}_{g,j}, \min_{k\in\{1,\cdots,m_E\}}\gamma_{h,k}\right\}.$$

Proof. Suppose $Z\in\mathbb{I}(X\times Y)$. Then for each $j\in\{1,\cdots,m_I\}$, $k\in\{1,\cdots,m_E\}$, we have from Definition 6.3.4 that

$$\begin{aligned}
\max_{(\mathbf{x},\mathbf{y})\in Z}\left|g_j(\mathbf{x},\mathbf{y}) - g^{\text{cv}}_{j,Z}(\mathbf{x},\mathbf{y})\right| &\leq \tau^{\text{cv}}_{g,j}w(Z)^{\gamma^{\text{cv}}_{g,j}},\\
\max_{(\mathbf{x},\mathbf{y})\in Z}\left|h_k(\mathbf{x},\mathbf{y}) - h^{\text{cv}}_{k,Z}(\mathbf{x},\mathbf{y})\right| &\leq \tau_{h,k}w(Z)^{\gamma_{h,k}},\\
\max_{(\mathbf{x},\mathbf{y})\in Z}\left|h_k(\mathbf{x},\mathbf{y}) - h^{\text{cc}}_{k,Z}(\mathbf{x},\mathbf{y})\right| &\leq \tau_{h,k}w(Z)^{\gamma_{h,k}},
\end{aligned}$$

since $(g^{\text{cv}}_{j,Z})_{Z\in\mathbb{I}(X\times Y)}$ and $(h^{\text{cv}}_{k,Z}, h^{\text{cc}}_{k,Z})_{Z\in\mathbb{I}(X\times Y)}$ converge pointwise to g_j and h_k, respectively, on $X\times Y$ with orders $\gamma^{\text{cv}}_{g,j}$ and $\gamma_{h,k}$. Let $(\mathbf{x}^{\text{cv}}_Z,\mathbf{y}^{\text{cv}}_Z)\in Z$ and $(\mathbf{v}^{\text{cv}}_Z,\mathbf{w}^{\text{cv}}_Z)\in\mathcal{I}_C(Z)$ such that $\mathbf{v}^{\text{cv}}_Z = \mathbf{g}^{\text{cv}}_Z(\mathbf{x}^{\text{cv}}_Z,\mathbf{y}^{\text{cv}}_Z), \mathbf{h}^{\text{cv}}_Z(\mathbf{x}^{\text{cv}}_Z,\mathbf{y}^{\text{cv}}_Z) \leq \mathbf{w}^{\text{cv}}_Z \leq \mathbf{h}^{\text{cc}}_Z(\mathbf{x}^{\text{cv}}_Z,\mathbf{y}^{\text{cv}}_Z)$, and $d(\{(\mathbf{v}^{\text{cv}}_Z,\mathbf{w}^{\text{cv}}_Z)\}, \mathbb{R}^{m_I}_{-}\times\{\mathbf{0}\}) =$

$d\left(\mathcal{I}_C(Z), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\}\right)$. The existence of $(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}})$ and $(\mathbf{v}_Z^{\mathrm{cv}}, \mathbf{w}_Z^{\mathrm{cv}})$ follows from the continuity of $\mathbf{g}_Z^{\mathrm{cv}}$, $\mathbf{h}_Z^{\mathrm{cv}}$, and $\mathbf{h}_Z^{\mathrm{cc}}$ and the compactness of Z. We have

$$
\begin{aligned}
& d\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(Z), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\}\right) - d\left(\mathcal{I}_C(Z), \mathbb{R}_-^{m_I} \times \{\mathbf{0}\}\right)\\
&\leq d\left(\left\{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}})\right\}, \mathbb{R}_-^{m_I} \times \{\mathbf{0}\}\right) - d\left(\{(\mathbf{v}_Z^{\mathrm{cv}}, \mathbf{w}_Z^{\mathrm{cv}})\}, \mathbb{R}_-^{m_I} \times \{\mathbf{0}\}\right)\\
&\leq d\left(\left\{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}}) - (\mathbf{v}_Z^{\mathrm{cv}}, \mathbf{w}_Z^{\mathrm{cv}})\right\}, \mathbb{R}_-^{m_I} \times \{\mathbf{0}\}\right)\\
&\leq \|\mathbf{g}(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}}) - \mathbf{v}_Z^{\mathrm{cv}}\| + \|\mathbf{h}(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}}) - \mathbf{w}_Z^{\mathrm{cv}}\|\\
&\leq \|\mathbf{g}(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}}) - \mathbf{g}_Z^{\mathrm{cv}}(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}})\| +\\
&\qquad \max\left\{\|\mathbf{h}(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}}) - \mathbf{h}_Z^{\mathrm{cv}}(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}})\|, \|\mathbf{h}(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}}) - \mathbf{h}_Z^{\mathrm{cc}}(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}})\|\right\}\\
&\leq \max_{(\mathbf{x},\mathbf{y})\in Z}\|\mathbf{g}(\mathbf{x}, \mathbf{y}) - \mathbf{g}_Z^{\mathrm{cv}}(\mathbf{x}, \mathbf{y})\| +\\
&\qquad \max\left\{\max_{(\mathbf{x},\mathbf{y})\in Z}\|\mathbf{h}(\mathbf{x}, \mathbf{y}) - \mathbf{h}_Z^{\mathrm{cv}}(\mathbf{x}, \mathbf{y})\|, \max_{(\mathbf{x},\mathbf{y})\in Z}\|\mathbf{h}(\mathbf{x}, \mathbf{y}) - \mathbf{h}_Z^{\mathrm{cc}}(\mathbf{x}, \mathbf{y})\|\right\}\\
&\leq \sum_{j=1}^{m_I}\max_{(\mathbf{x},\mathbf{y})\in Z}\left|g_j(\mathbf{x}, \mathbf{y}) - g_{j,Z}^{\mathrm{cv}}(\mathbf{x}, \mathbf{y})\right| +\\
&\qquad \max\left\{\sum_{k=1}^{m_E}\max_{(\mathbf{x},\mathbf{y})\in Z}\left|h_k(\mathbf{x}, \mathbf{y}) - h_{k,Z}^{\mathrm{cv}}(\mathbf{x}, \mathbf{y})\right|, \sum_{k=1}^{m_E}\max_{(\mathbf{x},\mathbf{y})\in Z}\left|h_k(\mathbf{x}, \mathbf{y}) - h_{k,Z}^{\mathrm{cc}}(\mathbf{x}, \mathbf{y})\right|\right\}\\
&\leq \sum_{j=1}^{m_I}\tau_{g,j}^{\mathrm{cv}} w(Z)^{\gamma_{g,j}^{\mathrm{cv}}} + \sum_{k=1}^{m_E}\tau_{h,k} w(Z)^{\gamma_{h,k}}\\
&\leq \left(\sum_{j=1}^{m_I}\tau_{g,j}^{\mathrm{cv}} w(X\times Y)^{\gamma_{g,j}^{\mathrm{cv}}-\beta} + \sum_{k=1}^{m_E}\tau_{h,k} w(X\times Y)^{\gamma_{h,k}-\beta}\right) w(Z)^{\beta},
\end{aligned}
$$

where Corollary 2.3.7 is used to derive Step 2, Step 3 follows from the triangle inequality, and Lemma 2.2.2 is used to derive Step 6. The desired result follows by choosing $\bar{\tau}$ as

$$\bar{\tau} = \left(\sum_{j=1}^{m_I}\tau_{g,j}^{\mathrm{cv}} w(X\times Y)^{\gamma_{g,j}^{\mathrm{cv}}-\beta} + \sum_{k=1}^{m_E}\tau_{h,k} w(X\times Y)^{\gamma_{h,k}-\beta}\right). \qquad \square$$

Note that the conclusions of Lemma 6.4.1 hold even when the schemes of convex relaxations $(g_{j,Z}^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$, $\forall j \in \{1, \cdots, m_I\}$, and $(h_{k,Z}^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$, $\forall k \in \{1, \cdots, m_E\}$, are merely lower semicontinuous, and the schemes of concave relaxations $(h_{k,Z}^{\mathrm{cc}})_{Z\in\mathbb{I}(X\times Y)}$,

$\forall k \in \{1, \cdots, m_E\}$, are merely upper semicontinuous.

Remark 6.4.2. The analysis in Lemma 6.4.1 can be refined under the following assumptions. Let $(g^{\mathrm{cv}}_{j,Z})_{Z\in\mathbb{I}(X\times Y)}$, $j = 1, \cdots, m_I$, denote schemes of convex relaxations of $g_1, \cdots, g_{m_I}$, respectively, in $X \times Y$ with convergence orders $\beta^{\mathrm{cv}}_{g,1} > 0, \cdots, \beta^{\mathrm{cv}}_{g,m_I} > 0$ and corresponding constants $\tau^{\mathrm{cv}}_{g,1}, \cdots, \tau^{\mathrm{cv}}_{g,m_I}$, and let $(h^{\mathrm{cv}}_{k,Z}, h^{\mathrm{cc}}_{k,Z})_{Z\in\mathbb{I}(X\times Y)}$, $k = 1, \cdots, m_E$, denote schemes of relaxations of $h_1, \cdots, h_{m_E}$, respectively, in $X \times Y$ with convergence orders $\beta_{h,1} > 0, \cdots, \beta_{h,m_E} > 0$ and corresponding constants $\tau_{h,1}, \cdots, \tau_{h,m_E}$. Suppose for each interval $Z \in \mathbb{I}(X \times Y)$, there exists $(\mathbf{x}_Z, \mathbf{y}_Z) \in Z$ such that:

$$d\left(\{(\mathbf{x}_Z, \mathbf{y}_Z)\}, \mathbb{R}^{m_I}_- \times \{\mathbf{0}\}\right) = d\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(Z), \mathbb{R}^{m_I}_- \times \{\mathbf{0}\}\right),$$

$$(\mathbf{x}_Z, \mathbf{y}_Z) \in \underset{(\mathbf{x},\mathbf{y})\in Z}{\arg\min}\, g_j(\mathbf{x}, \mathbf{y}), \quad \forall j \in \{1, \cdots, m_I\}, \text{ and}$$

$$\text{either } (\mathbf{x}_Z, \mathbf{y}_Z) \in \underset{(\mathbf{x},\mathbf{y})\in Z}{\arg\min}\, h_k(\mathbf{x}, \mathbf{y}), \text{ or } (\mathbf{x}_Z, \mathbf{y}_Z) \in \underset{(\mathbf{x},\mathbf{y})\in Z}{\arg\max}\, h_k(\mathbf{x}, \mathbf{y}), \quad \forall k \in \{1, \cdots, m_E\}.$$

Then, there exists $\bar{\tau} \geq 0$ such that for every $Z \in \mathbb{I}(X \times Y)$

$$d\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(Z), \mathbb{R}^{m_I}_- \times \{\mathbf{0}\}\right) - d\left(\mathcal{I}_C(Z), \mathbb{R}^{m_I}_- \times \{\mathbf{0}\}\right) \leq \bar{\tau} w(Z)^{\beta},$$

where β is defined as $\beta := \min\left\{\min\limits_{j\in\{1,\cdots,m_I\}} \beta^{\mathrm{cv}}_{g,j}, \min\limits_{k\in\{1,\cdots,m_E\}} \beta_{h,k}\right\}$. Note that the above assumptions are trivially satisfied when Problem (P) only has one inequality constraint (cf. Example 6.4.3).

The following example demonstrates the importance of a sufficiently high convergence order at nearly-feasible points (also see [108, Example 4]).

Example 6.4.3. Let $X = [0, 0], Y = [-1, 1]$, $m_I = 1$, and $m_E = 0$ with $f(x, y) = y$ and $g(x, y) = -y$. For any $[0, 0] \times [y^{\mathrm{L}}, y^{\mathrm{U}}] =: Z \in \mathbb{I}(X \times Y)$, let $f^{\mathrm{cv}}_Z(x, y) = y$, $g^{\mathrm{cv}}_Z(x, y) = -y^{\mathrm{U}} - (y^{\mathrm{U}} - y^{\mathrm{L}})^{\alpha}$ for some constant $\alpha > 0$. Note that $(f^{\mathrm{cv}}_Z)_{Z\in\mathbb{I}(X\times Y)}$ has arbitrarily high pointwise convergence order and arbitrarily high convergence order on $X \times Y$, whereas $(g^{\mathrm{cv}}_Z)_{Z\in\mathbb{I}(X\times Y)}$ has $\min\{\alpha, 1\}$-order pointwise convergence and α-order convergence on $X \times Y$.

Pick $\delta \in (0, 1)$ and let $\varepsilon \in (0, \delta)$. Let $y^{\mathrm{L}} = -\delta - \varepsilon$, $y^{\mathrm{U}} = -\delta + \varepsilon$. The width of Z is $w(Z) = 2\varepsilon$. We have $\overline{g}(Z) = [\delta - \varepsilon, \delta + \varepsilon]$, which yields $d(\overline{g}(Z), \mathbb{R}^{m_I}_-) = \delta - \varepsilon$ (this confirms that g is infeasible at each $(x, y) \in Z$). Furthermore, $\overline{g}^{\text{cv}}_Z(Z) = [\delta - \varepsilon - (2\varepsilon)^\alpha, \delta - \varepsilon - (2\varepsilon)^\alpha]$, which yields $d(\overline{g}^{\text{cv}}_Z(Z), \mathbb{R}^{m_I}_-) = \max\{0, \delta - \varepsilon - (2\varepsilon)^\alpha\}$. Therefore, for ε sufficiently small, the lower bounding problem detects the infeasibility of Z and we have $d(\overline{g}(Z), \mathbb{R}^{m_I}_-) - d(\overline{g}^{\text{cv}}_Z(Z), \mathbb{R}^{m_I}_-) = (2\varepsilon)^\alpha$, which implies that convergence of the lower bounding scheme at the infeasible point $(0, -\delta)$ is at most of order α.

For $\alpha = 1$, the maximum value of ε for which the interval Z can be fathomed by infeasibility by the lower bounding scheme is $\varepsilon = \frac{\delta}{3}$, whereas for $\alpha = 0.5$, the maximum value of ε for which the interval Z can be fathomed by infeasibility by the lower bounding scheme is $\varepsilon = \left(\frac{-\sqrt{2}+\sqrt{2}\sqrt{1+2\delta}}{2}\right)^2$, which is $O(\delta^2)$ for $\delta \ll 1$.

Therefore, a lower bounding scheme with a low convergence order at infeasible points may result in a large number of partitions on nearly-feasible points before they are fathomed, thereby resulting in the cluster problem.

The next result shows that under mild assumptions on the objective, the constraints, and the schemes of relaxations, first-order convergence to a global minimum is guaranteed.

Theorem 6.4.4. Consider Problem (P). Suppose f, g_j, $j = 1, \cdots, m_I$, and h_k, $k = 1, \cdots, m_E$, are Lipschitz continuous on $X \times Y$ with Lipschitz constants M_f, $M_{g,1}, \cdots, M_{g,m_I}$, $M_{h,1}, \cdots, M_{h,m_E}$, respectively. Let $(f^{\text{cv}}_Z)_{Z\in\mathbb{I}(X\times Y)}$, $(g^{\text{cv}}_{j,Z})_{Z\in\mathbb{I}(X\times Y)}$, $j = 1, \cdots, m_I$, denote continuous schemes of convex relaxations of f, $g_1, \cdots, g_{m_I}$, respectively, in $X \times Y$ with pointwise convergence orders $\gamma^{\text{cv}}_f \geq 1$, $\gamma^{\text{cv}}_{g,1} \geq 1, \cdots, \gamma^{\text{cv}}_{g,m_I} \geq 1$ and corresponding constants τ^{cv}_f, $\tau^{\text{cv}}_{g,1}, \cdots, \tau^{\text{cv}}_{g,m_I}$. Let $(h^{\text{cv}}_{k,Z}, h^{\text{cc}}_{k,Z})_{Z\in\mathbb{I}(X\times Y)}$, $k = 1, \cdots, m_E$, denote continuous schemes of relaxations of $h_1, \cdots, h_{m_E}$, respectively, in $X \times Y$ with pointwise convergence orders $\gamma_{h,1} \geq 1, \cdots, \gamma_{h,m_E} \geq 1$ and corresponding constants $\tau_{h,1}, \cdots, \tau_{h,m_E}$. The scheme of lower bounding problems $(\mathcal{L}(Z))_{Z\in\mathbb{I}(X\times Y)}$ with

$$(\mathcal{O}(Z))_{Z\in\mathbb{I}(X\times Y)} := \left(\min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}^{\text{cv}}(Z)} f^{\text{cv}}_Z(\mathbf{x}, \mathbf{y})\right)_{Z\in\mathbb{I}(X\times Y)},$$

$$(\mathcal{I}_C(Z))_{Z\in\mathbb{I}(X\times Y)} := \Big(\big\{(\mathbf{v}, \mathbf{w}) \in \mathbb{R}^{m_I} \times \mathbb{R}^{m_E} : \mathbf{v} = \mathbf{g}^{\text{cv}}_Z(\mathbf{x}, \mathbf{y}), \mathbf{h}^{\text{cv}}_Z(\mathbf{x}, \mathbf{y}) \leq \mathbf{w} \leq \mathbf{h}^{\text{cc}}_Z(\mathbf{x}, \mathbf{y}) \text{ for some } (\mathbf{x}, \mathbf{y}) \in Z\big\}\Big)_{Z\in\mathbb{I}(X\times Y)}$$

is at least first-order convergent on $X \times Y$.

Proof. Lemma 6.4.1 establishes first-order convergence at infeasible points $(\mathbf{x}, \mathbf{y}) \in X \times Y$

with the prefactor $\bar{\tau}$ independent of $(\mathbf{x}, \mathbf{y})$; therefore, it suffices to prove first-order convergence at feasible points $(\mathbf{x}, \mathbf{y}) \in X \times Y$ with a prefactor independent of $(\mathbf{x}, \mathbf{y})$.

In order to do so, suppose $\mathcal{F}(X \times Y) \neq \emptyset$ and consider $Z \in \mathbb{I}(X \times Y)$ such that $Z \cap \mathcal{F}(X \times Y) \neq \emptyset$. Let $\mathcal{F}^{\text{cv}}(Z) := \{(\mathbf{x}, \mathbf{y}) \in Z : \mathbf{g}^{\text{cv}}_Z(\mathbf{x}, \mathbf{y}) \leq \mathbf{0}, \mathbf{h}^{\text{cv}}_Z(\mathbf{x}, \mathbf{y}) \leq \mathbf{0}, \mathbf{h}^{\text{cc}}_Z(\mathbf{x}, \mathbf{y}) \geq \mathbf{0}\}$ denote the feasible set of the convex relaxation-based lower bounding scheme. Then

$$
\begin{aligned}
&\min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}(Z)} f(\mathbf{x}, \mathbf{y}) \; - \min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}^{\text{cv}}(Z)} f^{\text{cv}}_Z(\mathbf{x}, \mathbf{y}) \\
&= \left(\min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}(Z)} f(\mathbf{x}, \mathbf{y}) - \min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}^{\text{cv}}(Z)} f(\mathbf{x}, \mathbf{y}) \right) + \left(\min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}^{\text{cv}}(Z)} f(\mathbf{x}, \mathbf{y}) - \min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}^{\text{cv}}(Z)} f^{\text{cv}}_Z(\mathbf{x}, \mathbf{y}) \right) \\
&\leq \left(\min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}(Z)} f(\mathbf{x}, \mathbf{y}) \; - \min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}^{\text{cv}}(Z)} f(\mathbf{x}, \mathbf{y}) \right) \; + \max_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}^{\text{cv}}(Z)} \left| f(\mathbf{x}, \mathbf{y}) - f^{\text{cv}}_Z(\mathbf{x}, \mathbf{y}) \right|, \qquad (6.2)
\end{aligned}
$$

where the above inequality follows from Lemma 2.3.35. The second term in Equation (6.2) can be bounded from above as

$$
\max_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}^{\text{cv}}(Z)} \left| f(\mathbf{x}, \mathbf{y}) - f^{\text{cv}}_Z(\mathbf{x}, \mathbf{y}) \right| \leq \tau^{\text{cv}}_f w(Z)^{\gamma^{\text{cv}}_f},
$$

since $(f^{\text{cv}}_Z)_{Z \in \mathbb{I}(X \times Y)}$ converges pointwise to f on $X \times Y$ with order $\gamma^{\text{cv}}_f \geq 1$.

Let $(\mathbf{x}^*_Z, \mathbf{y}^*_Z) \in \underset{(\mathbf{x},\mathbf{y}) \in \mathcal{F}(Z)}{\arg\min} f(\mathbf{x}, \mathbf{y})$ and $(\mathbf{x}^{\text{cv}}_Z, \mathbf{y}^{\text{cv}}_Z) \in \underset{(\mathbf{x},\mathbf{y}) \in \mathcal{F}^{\text{cv}}(Z)}{\arg\min} f(\mathbf{x}, \mathbf{y})$. The first term in Equation (6.2) can be bounded from above as

$$
\begin{aligned}
\left(\min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}(Z)} f(\mathbf{x}, \mathbf{y}) \; - \min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}^{\text{cv}}(Z)} f(\mathbf{x}, \mathbf{y}) \right) \; &= f(\mathbf{x}^*_Z, \mathbf{y}^*_Z) - f(\mathbf{x}^{\text{cv}}_Z, \mathbf{y}^{\text{cv}}_Z) \\
&\leq M_f \sqrt{n_x + n_y}\, w(Z),
\end{aligned}
$$

where the last step follows from the Lipschitz continuity of f on $X \times Y$ and Lemma 2.2.2.

Plugging in the above bounds in Equation (6.2), we get

$$
\min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}(Z)} f(\mathbf{x}, \mathbf{y}) \; - \min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}^{\text{cv}}(Z)} f^{\text{cv}}_Z(\mathbf{x}, \mathbf{y}) \leq \left(M_f \sqrt{n_x + n_y} + \tau^{\text{cv}}_f w(X \times Y)^{\gamma^{\text{cv}}_f - 1} \right) w(Z),
$$

which establishes first-order convergence of $(\mathcal{L}(Z))_{Z \in \mathbb{I}(X \times Y)}$ at feasible points $(\mathbf{x}, \mathbf{y}) \in X \times Y$ with the prefactor independent of $(\mathbf{x}, \mathbf{y})$. $\square$

The following examples show that the convergence order of the lower bounding scheme may be as low as one under the assumptions of Theorem 6.4.4.

Example 6.4.5. Let $X = [-1, 1], Y = [-1, 1]$, $m_I = 1$, and $m_E = 0$ with $f(x, y) = 2x + 2y$

and $g(x,y) = -x-y$. For any $[x^{\mathrm{L}}, x^{\mathrm{U}}] \times [y^{\mathrm{L}}, y^{\mathrm{U}}] =: Z \in \mathbb{I}(X \times Y)$, let $f_Z^{\mathrm{cv}}(x,y) = 2x+2y$ and $g_Z^{\mathrm{cv}}(x,y) = -x^{\mathrm{U}} - y^{\mathrm{U}}$. The scheme $(f_Z^{\mathrm{cv}})_{Z \in \mathbb{I}(X \times Y)}$ has arbitrarily high pointwise convergence order on $X \times Y$ and the scheme $(g_Z^{\mathrm{cv}})_{Z \in \mathbb{I}(X \times Y)}$ has first-order pointwise convergence on $X \times Y$. Note that $(g_Z^{\mathrm{cv}})_{Z \in \mathbb{I}(X \times Y)}$ has arbitrarily high convergence order on $X \times Y$.

Let $x^{\mathrm{L}} = y^{\mathrm{L}} = -\varepsilon$, $x^{\mathrm{U}} = y^{\mathrm{U}} = \varepsilon$ with $0 < \varepsilon \leq 1$. The width of Z is $w(Z) = 2\varepsilon$. The optimal objective value of Problem (P) on Z is 0, while the optimal objective of the lower bounding problem on Z is -4ε. Convergence at the point $(0,0)$ is, therefore, at most first-order.

Example 6.4.6. Let $X = [-1,1], Y = [-1,1]$, $m_I = 1$, and $m_E = 0$ with $f(x,y) = 2x+2y$ and $g(x,y) = -x-y$. For any $[x^{\mathrm{L}}, x^{\mathrm{U}}] \times [y^{\mathrm{L}}, y^{\mathrm{U}}] =: Z \in \mathbb{I}(X \times Y)$, let $f_Z^{\mathrm{cv}}(x,y) = 2x^{\mathrm{L}} + 2y^{\mathrm{L}}$ and $g_Z^{\mathrm{cv}}(x,y) = -x-y$. The scheme $(f_Z^{\mathrm{cv}})_{Z \in \mathbb{I}(X \times Y)}$ has first-order pointwise convergence on $X \times Y$ and the scheme $(g_Z^{\mathrm{cv}})_{Z \in \mathbb{I}(X \times Y)}$ has arbitrarily high pointwise convergence order on $X \times Y$. Note that $(f_Z^{\mathrm{cv}})_{Z \in \mathbb{I}(X \times Y)}$ has arbitrarily high convergence order on $X \times Y$.

Let $x^{\mathrm{L}} = y^{\mathrm{L}} = -\varepsilon$, $x^{\mathrm{U}} = y^{\mathrm{U}} = \varepsilon$ with $0 < \varepsilon \leq 1$. The width of Z is $w(Z) = 2\varepsilon$. The optimal objective value of Problem (P) on Z is 0, while the optimal objective of the lower bounding problem on Z is -4ε. Convergence at the point $(0,0)$ is, therefore, at most first-order.

Example 6.4.7. Let $X = [0,0], Y = [-1,1]$, $m_I = 1$, and $m_E = 0$ with $f(x,y) = y$ and $g(x,y) = \min\{-0.5y, -y\}$. For any $[0,0] \times [y^{\mathrm{L}}, y^{\mathrm{U}}] =: Z \in \mathbb{I}(X \times Y)$ with $y^{\mathrm{L}} < 0 < y^{\mathrm{U}}$, let

$$f_Z^{\mathrm{cv}}(x,y) = y, \quad g_Z^{\mathrm{cv}}(x,y) = -\frac{y^{\mathrm{U}} - 0.5y^{\mathrm{L}}}{y^{\mathrm{U}} - y^{\mathrm{L}}} y + \frac{0.5y^{\mathrm{L}} y^{\mathrm{U}}}{y^{\mathrm{U}} - y^{\mathrm{L}}}.$$

Note that g_Z^{cv} corresponds to the convex envelope of g on Z. The scheme $(f_Z^{\mathrm{cv}})_{Z \in \mathbb{I}(X \times Y)}$ has arbitrarily high pointwise convergence order on $X \times Y$ and the scheme $(g_Z^{\mathrm{cv}})_{Z \in \mathbb{I}(X \times Y)}$ has first-order pointwise convergence on $X \times Y$. Note that $(g_Z^{\mathrm{cv}})_{Z \in \mathbb{I}(X \times Y)}$ has arbitrarily high convergence order on $X \times Y$.

Let $y^{\mathrm{L}} = -\varepsilon$, $y^{\mathrm{U}} = \varepsilon$ with $0 < \varepsilon \leq 1$. The width of Z is $w(Z) = 2\varepsilon$. The optimal objective value of Problem (P) on Z is 0, while the optimal objective of the lower bounding problem on Z is $-\frac{\varepsilon}{3}$. Convergence at the point $(0,0)$ is, therefore, at most first-order.

Example 6.4.8. Let $X = [0,0], Y = [-1,1]$, $m_I = 0$, and $m_E = 1$ with $f(x,y) = y$ and

$h(x,y) = \min\{-0.5y, -y\}$. For any $[0,0] \times [y^{\mathrm{L}}, y^{\mathrm{U}}] =: Z \in \mathbb{I}(X \times Y)$ with $y^{\mathrm{L}} < 0 < y^{\mathrm{U}}$, let

$$f_Z^{\mathrm{cv}}(x,y) = y, \quad h_Z^{\mathrm{cv}}(x,y) = -\frac{y^{\mathrm{U}} - 0.5y^{\mathrm{L}}}{y^{\mathrm{U}} - y^{\mathrm{L}}} y + \frac{0.5 y^{\mathrm{L}} y^{\mathrm{U}}}{y^{\mathrm{U}} - y^{\mathrm{L}}}, \quad h_Z^{\mathrm{cc}}(x,y) = \min\{-0.5y, -y\}.$$

Note that h_Z^{cv} and h_Z^{cc} correspond to the convex and concave envelopes of h on Z, respectively. The scheme $(f_Z^{\mathrm{cv}})_{Z \in \mathbb{I}(X \times Y)}$ has arbitrarily high pointwise convergence order on $X \times Y$ and the scheme $(h_Z^{\mathrm{cv}}, h_Z^{\mathrm{cc}})_{Z \in \mathbb{I}(X \times Y)}$ has first-order pointwise convergence on $X \times Y$. Note that $(h_Z^{\mathrm{cv}}, h_Z^{\mathrm{cc}})_{Z \in \mathbb{I}(X \times Y)}$ has arbitrarily high convergence order on $X \times Y$.

Let $y^{\mathrm{L}} = -\varepsilon$, $y^{\mathrm{U}} = \varepsilon$ with $0 < \varepsilon \leq 1$. The width of Z is $w(Z) = 2\varepsilon$. The optimal objective value of Problem (P) on Z is 0, while the optimal objective of the lower bounding problem on Z is $-\frac{\varepsilon}{3}$. Convergence at the point $(0,0)$ is, therefore, at most first-order.

Despite the fact that the schemes of relaxations used in Examples 6.4.7 and 6.4.8 correspond to the envelopes of the functions involved (unlike those used in Examples 6.4.5 and 6.4.6), we only have first-order convergence at the global minimizer $(0,0)$. However, the reader can verify that first-order convergent lower bounding schemes may be sufficient to mitigate the cluster problem in Examples 6.4.7 and 6.4.8, whereas at least second-order convergent lower bounding schemes are required to mitigate the cluster problem in Examples 6.4.5 and 6.4.6 (see Chapter 5). Furthermore, Examples 6.4.5 to 6.4.8 illustrate that high convergence orders of schemes of relaxations of the objective and constraints do not guarantee a high convergence order of the lower bounding scheme (cf. Remark 6.4.2) at constrained minima (which may be required to mitigate clustering). This is because a high convergence order of a scheme of relaxations of the objective function may only place a restriction on the gap between the minimum value of the relaxation and the minimum value of the objective function without taking the feasible region into account; this restriction may not be sufficient in a constrained setting because the gap between the minimum value of the relaxed problem and the minimum value of the original problem may be relatively large when their respective feasible regions are taken into consideration (see Example 6.4.6 for an extreme case). Similarly, a high convergence order of a scheme of relaxations of the constraints may not exclude infeasible regions of the search space in which the objective function value is less than the optimal (constrained) objective value (Example 6.4.5 provides an extreme case), potentially leading to relatively large underestimation gaps.

The following result proves second-order convergence at certain points in $X \times Y$.

Theorem 6.4.9. Consider Problem (P). Suppose f is Lipschitz continuous on $X \times Y$ with Lipschitz constant M_f. Let $(f_Z^{\mathrm{cv}})_{Z \in \mathbb{I}(X \times Y)}$ denote a continuous scheme of convex relaxations of f with pointwise convergence order $\gamma_f^{\mathrm{cv}} \geq 2$ and corresponding constant τ_f^{cv}.

Suppose there exists a feasible point $(\mathbf{x}^{\mathrm{f}}, \mathbf{y}^{\mathrm{f}}) \in \mathcal{F}(X \times Y)$, continuous schemes of convex relaxations $(g_{j,Z}^{\mathrm{cv}})_{Z \in \mathbb{I}(X \times Y)}$, $j = 1, \cdots, m_I$, of $g_1, \cdots, g_{m_I}$, respectively, in $X \times Y$, continuous schemes of relaxations $(h_{k,Z}^{\mathrm{cv}}, h_{k,Z}^{\mathrm{cc}})_{Z \in \mathbb{I}(X \times Y)}$, $k = 1, \cdots, m_E$, of $h_1, \cdots, h_{m_E}$, respectively, in $X \times Y$, and a constant $\delta > 0$ such that for each $Z \in \mathbb{I}(X \times Y)$ with $(\mathbf{x}^{\mathrm{f}}, \mathbf{y}^{\mathrm{f}}) \in Z$ and $w(Z) \leq \delta$, we have

$$d\left(\mathcal{F}(Z), \underset{(\mathbf{x},\mathbf{y}) \in \mathcal{F}^{\mathrm{cv}}(Z)}{\arg\min} f(\mathbf{x}, \mathbf{y})\right) \leq \hat{\tau} w(Z)^{\gamma} \tag{6.3}$$

for constants $\gamma \geq 2$ and $\hat{\tau} \geq 0$. Then, the lower bounding scheme $(\mathcal{L}(Z))_{Z \in \mathbb{I}(X \times Y)}$ with

$$(\mathcal{O}(Z))_{Z \in \mathbb{I}(X \times Y)} := \left(\min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}^{\mathrm{cv}}(Z)} f_Z^{\mathrm{cv}}(\mathbf{x}, \mathbf{y})\right)_{Z \in \mathbb{I}(X \times Y)},$$

$$\begin{aligned}(\mathcal{I}_C(Z))_{Z \in \mathbb{I}(X \times Y)} := \big(\{(\mathbf{v}, \mathbf{w}) \in \mathbb{R}^{m_I} \times \mathbb{R}^{m_E} : \mathbf{v} = \mathbf{g}_Z^{\mathrm{cv}}(\mathbf{x}, \mathbf{y}), \mathbf{h}_Z^{\mathrm{cv}}(\mathbf{x}, \mathbf{y}) \leq \mathbf{w} \leq \mathbf{h}_Z^{\mathrm{cc}}(\mathbf{x}, \mathbf{y}) \\ \text{for some } (\mathbf{x}, \mathbf{y}) \in Z\}\big)_{Z \in \mathbb{I}(X \times Y)}\end{aligned}$$

is at least $\min\{\gamma_f^{\mathrm{cv}}, \gamma\}$-order convergent at $(\mathbf{x}^{\mathrm{f}}, \mathbf{y}^{\mathrm{f}})$.

Proof. Suppose $Z \in \mathbb{I}(X \times Y)$ such that $(\mathbf{x}^{\mathrm{f}}, \mathbf{y}^{\mathrm{f}}) \in Z$ and $w(Z) \leq \delta$. From the proof of Theorem 6.4.4, we have

$$\begin{aligned}&\min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}(Z)} f(\mathbf{x}, \mathbf{y}) - \min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}^{\mathrm{cv}}(Z)} f_Z^{\mathrm{cv}}(\mathbf{x}, \mathbf{y}) \\ &\leq \left(\min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}(Z)} f(\mathbf{x}, \mathbf{y}) - \min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}^{\mathrm{cv}}(Z)} f(\mathbf{x}, \mathbf{y})\right) + \left(\min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}^{\mathrm{cv}}(Z)} f(\mathbf{x}, \mathbf{y}) - \min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}^{\mathrm{cv}}(Z)} f_Z^{\mathrm{cv}}(\mathbf{x}, \mathbf{y})\right) \\ &\leq \left(\min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}(Z)} f(\mathbf{x}, \mathbf{y}) - \min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}^{\mathrm{cv}}(Z)} f(\mathbf{x}, \mathbf{y})\right) + \tau_f^{\mathrm{cv}} w(Z)^{\gamma_f^{\mathrm{cv}}}. \end{aligned} \tag{6.4}$$

Consider $(\mathbf{x}_Z^*, \mathbf{y}_Z^*) \in \underset{(\mathbf{x},\mathbf{y}) \in \mathcal{F}(Z)}{\arg\min} f(\mathbf{x}, \mathbf{y})$. Choose a feasible point $(\hat{\mathbf{x}}_Z, \hat{\mathbf{y}}_Z) \in \mathcal{F}(Z)$ and $(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}}) \in \underset{(\mathbf{x},\mathbf{y}) \in \mathcal{F}^{\mathrm{cv}}(Z)}{\arg\min} f(\mathbf{x}, \mathbf{y})$ such that $d\left(\{(\hat{\mathbf{x}}_Z, \hat{\mathbf{y}}_Z)\}, \{(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}})\}\right) \leq \hat{\tau} w(Z)^{\gamma}$ (note that $(\hat{\mathbf{x}}_Z, \hat{\mathbf{y}}_Z)$ and $(\mathbf{x}_Z^{\mathrm{cv}}, \mathbf{y}_Z^{\mathrm{cv}})$ exist by assumption). The first term in Equation (6.4) can be bounded from above as

$$\begin{aligned}
\min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}(Z)} f(\mathbf{x},\mathbf{y}) \ - \min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}^{\text{cv}}(Z)} f(\mathbf{x},\mathbf{y}) &= f(\mathbf{x}^*_Z,\mathbf{y}^*_Z) - f(\mathbf{x}^{\text{cv}}_Z,\mathbf{y}^{\text{cv}}_Z) \\
&\le f(\hat{\mathbf{x}}_Z,\hat{\mathbf{y}}_Z) - f(\mathbf{x}^{\text{cv}}_Z,\mathbf{y}^{\text{cv}}_Z) \\
&\le M_f\, d\left(\{(\hat{\mathbf{x}}_Z,\hat{\mathbf{y}}_Z)\},\{(\mathbf{x}^{\text{cv}}_Z,\mathbf{y}^{\text{cv}}_Z)\}\right) \\
&\le M_f \hat{\tau} w(Z)^{\gamma},
\end{aligned}$$

where Step 3 above follows from the Lipschitz continuity of f. Therefore, from Equation (6.4),

$$\begin{aligned}
&\min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}(Z)} f(\mathbf{x},\mathbf{y}) \ - \min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}^{\text{cv}}(Z)} f^{\text{cv}}_Z(\mathbf{x},\mathbf{y}) \\
\le &\left(M_f\hat{\tau}w(X\times Y)^{\gamma-\min\{\gamma^{\text{cv}}_f,\gamma\}} + \tau^{\text{cv}}_f w(X\times Y)^{\gamma^{\text{cv}}_f-\min\{\gamma^{\text{cv}}_f,\gamma\}}\right) w(Z)^{\min\{\gamma^{\text{cv}}_f,\gamma\}}.
\end{aligned}$$

The desired result follows by analogy to Lemma 6.3.8 by noting that the lower bounding scheme $(\mathcal{L}(Z))_{Z\in\mathbb{I}(X\times Y)}$ has convergence of order at least one at $(\mathbf{x}^{\text{f}},\mathbf{y}^{\text{f}})$ from Theorem 6.4.4.

□

The key assumption of Theorem 6.4.9, Equation (6.3), is rather unwieldy since verifying it involves the solution of the optimization problem $\min\limits_{(\mathbf{x},\mathbf{y})\in\mathcal{F}^{\text{cv}}(Z)} f(\mathbf{x},\mathbf{y})$ for each $Z\in\mathbb{I}(X\times Y)$ with $(\mathbf{x}^{\text{f}},\mathbf{y}^{\text{f}})\in Z$ and $w(Z)\le\delta$. The following more restrictive (but potentially more easily verifiable) condition implies Equation (6.3):

$$\begin{aligned}
\exists\delta>0,\hat{\tau}\ge 0,\gamma\ge 2:\ \ d_H(\mathcal{F}(Z),\mathcal{F}^{\text{cv}}(Z)) \le \hat{\tau}w(Z)^{\gamma}, \quad &\forall Z\in\mathbb{I}(X\times Y) \text{ with} \\
&(\mathbf{x}^{\text{f}},\mathbf{y}^{\text{f}})\in Z \text{ and } w(Z)\le\delta.
\end{aligned}$$

The following example shows that the convergence order may be as low as two under the assumptions of Theorem 6.4.9.

Example 6.4.10. Let $X=[-1,1], Y=[-1,1]$, $m_I=1$, and $m_E=0$ with $f(x,y)=-xy$ and $g(x,y)=x+y-1$. For any $[x^{\text{L}},x^{\text{U}}]\times[y^{\text{L}},y^{\text{U}}] =: Z\in\mathbb{I}(X\times Y)$, let

$$\begin{aligned}
f^{\text{cv}}_Z(x,y) &= \max\{-x^{\text{U}}y+(-x)y^{\text{L}}-(-x^{\text{U}})y^{\text{L}}, -x^{\text{L}}y+(-x)y^{\text{U}}-(-x^{\text{L}})y^{\text{U}}\}, \\
g^{\text{cv}}_Z(x,y) &= x+y-1.
\end{aligned}$$

The scheme $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$, which corresponds to the scheme of convex envelopes of f on $X \times Y$, has (at least) second-order pointwise convergence on $X \times Y$ (see [38, Theorem 10]) and the scheme $(g_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has arbitrarily high pointwise convergence order on $X \times Y$. Note that $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has arbitrarily high convergence order on $X \times Y$.

Let $x^{\mathrm{L}} = y^{\mathrm{L}} = 0.5-\varepsilon$, $x^{\mathrm{U}} = y^{\mathrm{U}} = 0.5+\varepsilon$ with $0 < \varepsilon \leq 0.5$. The width of Z is $w(Z) = 2\varepsilon$. The optimal objective value of Problem (P) on Z is -0.25, while $f_Z^{\mathrm{cv}}(0.5, 0.5) = -0.25 - \varepsilon^2$ and $g_Z^{\mathrm{cv}}(0.5, 0.5) = 0$. Convergence at the point $(0.5, 0.5)$ is, therefore, at most second-order.

Note that the use of feasibility-based bounds tightening techniques is ineffective in boosting the convergence order for the above example. This is in contrast to the similar Example 6.5.19 where the use of constraint propagation techniques improves the convergence order of reduced-space branch-and-bound algorithms (also see Examples 6.5.21 and 6.5.22 in Section 6.5.2).

Remark 6.4.11. Theorem 6.4.9 requires, at a minimum, second-order pointwise convergence of the scheme of convex relaxations $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$, which cannot be achieved in general by relaxations constructed purely using interval arithmetic [172]. Consequently, lower bounding schemes constructed using interval arithmetic can, in general, only be expected to possess first-order convergence (see Theorem 6.4.4). When the functions f, $\mathbf{g}$, and $\mathbf{h}$ are twice continuously differentiable, references [191] and [224] imply that polyhedral outer-approximation schemes of second-order pointwise convergent schemes of relaxations, that are employed by most state-of-the-art software for nonconvex problems (P) [19, 162, 225], also produce second-order pointwise convergent schemes of relaxations.

The following corollary of Theorem 6.4.9 shows that second-order convergence is guaranteed at points $(\mathbf{x}, \mathbf{y}) \in X \times Y$ such that $\mathbf{g}(\mathbf{x}, \mathbf{y}) < \mathbf{0}$, assuming Problem (P) contains no equality constraints (note the weaker assumption on the pointwise convergence order of the scheme $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$, and the slight abuse of notation in the description of $\mathcal{I}_C(Z)$ where we simply discard the components corresponding to $\mathbf{h}$ since $m_E = 0$). A consequence of the corollary is that second-order convergence to unconstrained minima is guaranteed.

Corollary 6.4.12. Consider Problem (P) with $m_E = 0$. Suppose f is Lipschitz continuous on $X \times Y$. Let $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ denote a continuous scheme of convex relaxations of f in $X \times Y$ with pointwise convergence order $\gamma_f^{\mathrm{cv}} \geq 1$, and convergence order $\beta_f^{\mathrm{cv}} \geq 2$ with corresponding constant τ_f^{cv}. Furthermore, let $(g_{j,Z}^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$, $j = 1,\cdots,m_I$, denote continuous schemes of convex relaxations of $g_1,\cdots,g_{m_I}$, respectively, in $X\times Y$ with pointwise convergence orders $\gamma_{g,1}^{\mathrm{cv}} > 0,\cdots,\gamma_{g,m_I}^{\mathrm{cv}} > 0$ and corresponding constants $\tau_{g,1}^{\mathrm{cv}},\cdots,\tau_{g,m_I}^{\mathrm{cv}}$.

Suppose $(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}}) \in X\times Y$ is such that $\mathbf{g}(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}}) < \mathbf{0}$ (i.e., $(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})$ is a Slater point, cf. Definition 2.3.20). Then, the scheme of lower bounding problems $(\mathcal{L}(Z))_{Z\in\mathbb{I}(X\times Y)}$ with

$$(\mathcal{O}(Z))_{Z\in\mathbb{I}(X\times Y)} := \left(\min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}^{\mathrm{cv}}(Z)} f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y})\right)_{Z\in\mathbb{I}(X\times Y)},$$
$$(\mathcal{I}_C(Z))_{Z\in\mathbb{I}(X\times Y)} := (\overline{\mathbf{g}}_Z^{\mathrm{cv}}(Z))_{Z\in\mathbb{I}(X\times Y)}$$

is at least β_f^{cv}-order convergent at $(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})$.

Proof. Since we are interested in the convergence order at the feasible point $(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})$, it suffices to show that the assumptions of Theorem 6.4.9 hold.

Let $g_j(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}}) = -\varepsilon_j < 0$, $j = 1,\cdots,m_I$. Since g_j is continuous for each $j \in \{1,\cdots,m_I\}$ by virtue of Assumption 6.2.1, there exists $\delta_j > 0$, $\forall j \in \{1,\cdots,m_I\}$, such that $\left\|(\mathbf{x},\mathbf{y}) - (\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})\right\|_\infty < \delta_j$ implies $\left|g_j(\mathbf{x},\mathbf{y}) - g_j(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})\right| < \frac{\varepsilon_j}{2}$ (see Lemma 2.2.2).

Define $\delta := \min\limits_{j\in\{1,\cdots,m_I\}} \delta_j$, and note that $\delta > 0$. Consider $Z \in \mathbb{I}(X\times Y)$ such that $(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}}) \in Z$ and $w(Z) \le \delta$. For each $(\mathbf{x},\mathbf{y}) \in Z$ and $j \in \{1,\cdots,m_I\}$ we have that $\left|g_j(\mathbf{x},\mathbf{y}) - g_j(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})\right| < \frac{\varepsilon_j}{2}$. Consequently, $\forall j \in \{1,\cdots,m_I\}$, $g_j(\mathbf{x},\mathbf{y}) < -\frac{\varepsilon_j}{2}$, $\forall(\mathbf{x},\mathbf{y}) \in Z$. Since $\mathbf{g}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) \le \mathbf{g}(\mathbf{x},\mathbf{y}) < -\frac{\varepsilon_j}{2}$, $\forall(\mathbf{x},\mathbf{y}) \in Z$, we have $\mathbf{g}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) < -\frac{\varepsilon_j}{2}$, $\forall(\mathbf{x},\mathbf{y}) \in Z$, i.e., every point in Z is feasible for Problem (P) and the lower bounding problem $\mathcal{L}(Z)$.

Therefore, $\delta := \min\limits_{j\in\{1,\cdots,m_I\}} \delta_j$, any $(\hat{\mathbf{x}}_Z,\hat{\mathbf{y}}_Z) \in \operatorname*{arg\,min}\limits_{(\mathbf{x},\mathbf{y})\in\mathcal{F}^{\mathrm{cv}}(Z)} f(\mathbf{x},\mathbf{y})$, $\gamma = \beta_f^{\mathrm{cv}} + 1$, and $\hat{\tau} = 0$ satisfies the (necessary) assumptions of Theorem 6.4.9 which yield an upper bound on the first term in Equation (6.4). The second term in Equation (6.4) can be bounded from above as

$$\begin{aligned}\min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}^{\mathrm{cv}}(Z)} f(\mathbf{x},\mathbf{y}) \;-\; \min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}^{\mathrm{cv}}(Z)} f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) &= \min_{(\mathbf{x},\mathbf{y})\in Z} f(\mathbf{x},\mathbf{y}) \;-\; \min_{(\mathbf{x},\mathbf{y})\in Z} f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) \\ &\le \tau_f^{\mathrm{cv}} w(Z)^{\beta_f^{\mathrm{cv}}}\end{aligned}$$

since f_Z^{cv} converges with order β_f^{cv} to f on $X\times Y$, and $\mathcal{F}^{\mathrm{cv}}(Z) = Z$. Substituting the above bounds in Equation (6.4), we obtain

$$\min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}(Z)} f(\mathbf{x},\mathbf{y}) \ - \min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}^{\mathrm{cv}}(Z)} f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) \le \tau_f^{\mathrm{cv}} w(Z)^{\beta_f^{\mathrm{cv}}}.$$

The desired result follows by analogy to Lemma 6.3.8 by noting that the lower bounding scheme $(\mathcal{L}(Z))_{Z\in\mathbb{I}(X\times Y)}$ is at least first-order convergent at $(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})$ from Theorem 6.4.4. $\square$

Note that the bound on the prefactor obtained from Corollary 6.4.12 for convergence at points where a constraint is 'nearly active' can be relatively large (also see the comment after Lemma 6.3.8).

Remark 6.4.13. Corollary 6.4.12 does not apply to Problem (P) with active constraints; however, Theorem 6.4.9 can be used to demonstrate second-order convergence when Problem (P) contains active convex constraints (note that this includes affine equality constraints) if the schemes of relaxations of the active constraints are the (convex) functions themselves and the scheme of convex relaxations of the objective function is second-order pointwise convergent. Examples 6.4.15 and 6.4.16 illustrate cases where the above modification of Corollary 6.4.12 does not apply when the schemes of relaxations of active convex constraints are not the constraints themselves (note that if the schemes of relaxations of active convex constraints used are the constraints themselves, then the convergence orders of the lower bounding schemes in these examples would be arbitrarily high at their minimizers), thereby highlighting the importance of convexity detection in boosting the convergence order.

The following example shows that the convergence order of the lower bounding scheme is dictated by the convergence order, β_f^{cv}, of the scheme $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ when the assumptions of Corollary 6.4.12 are satisfied.

Example 6.4.14. Let $X=[0,0], Y=[0,1]$, $m_I=1$, and $m_E=0$ with $f(x,y)=y^4-y^2$ and $g(x,y)=1-2y$. For any $[0,0]\times[y^{\mathrm{L}},y^{\mathrm{U}}]=: Z\in\mathbb{I}(X\times Y)$, let $f_Z^{\mathrm{cv}}(x,y)=y^4-(y^{\mathrm{L}}+y^{\mathrm{U}})y+y^{\mathrm{L}}y^{\mathrm{U}}$, $g_Z^{\mathrm{cv}}(x,y)=1-2y$. The scheme $(f_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has second-order pointwise convergence and second-order convergence on $X\times Y$, while the scheme $(g_Z^{\mathrm{cv}})_{Z\in\mathbb{I}(X\times Y)}$ has arbitrarily high pointwise convergence order on $X\times Y$.

Let $y^{\mathrm{L}}=\frac{1}{\sqrt{2}}-\varepsilon$, $y^{\mathrm{U}}=\frac{1}{\sqrt{2}}+\varepsilon$ with $0<\varepsilon\le 0.25$. The width of Z is $w(Z)=2\varepsilon$. The optimal objective value of Problem (P) on Z is -0.25, while the optimal objective of the lower bounding problem on Z is $-0.25-\varepsilon^2$. Convergence at the point $\left(0, \frac{1}{\sqrt{2}}\right)$ is, therefore, at most second-order.

Example 6.4.15. Let $X = [-3, 3], Y = [-3, 3]$, $m_I = 1$, and $m_E = 0$ with $f(x, y) = x + y$ and $g(x, y) = x^2 + y^2 - 8$. For any $[x^{\mathrm{L}}, x^{\mathrm{U}}] \times [y^{\mathrm{L}}, y^{\mathrm{U}}] =: Z \in \mathbb{I}(X \times Y)$, let $f^{\mathrm{cv}}_Z(x, y) = x + y$, $g^{\mathrm{cv}}_Z(x, y) = x^2 + y^2 - 8 - (w(Z))^2$. The scheme $(f^{\mathrm{cv}}_Z)_{Z\in\mathbb{I}(X\times Y)}$ has arbitrarily high pointwise convergence order on $X \times Y$, while the scheme $(g^{\mathrm{cv}}_Z)_{Z\in\mathbb{I}(X\times Y)}$ has second-order pointwise convergence on $X \times Y$.

Let $x^{\mathrm{L}} = y^{\mathrm{L}} = -2 - \varepsilon$, $x^{\mathrm{U}} = y^{\mathrm{U}} = -2 + \varepsilon$ with $0 < \varepsilon \leq 1$. The width of Z is $w(Z) = 2\varepsilon$. The optimal objective value of Problem (P) on Z is -4, while the optimal objective of the lower bounding problem on Z is $-\sqrt{16 + 8\varepsilon^2} = -4 - \varepsilon^2 + O(\varepsilon^4)$ for $\varepsilon \ll 1$. Convergence at the point $(-2, -2)$ is, therefore, at most second-order.

Example 6.4.16. Let $X = [0, 1], Y = [0, 1]$, $m_I = 1$, and $m_E = 0$ with $f(x, y) = -x - y$ and $g(x, y) = x^2 + 2xy + y^2 - 1$. For any $[x^{\mathrm{L}}, x^{\mathrm{U}}] \times [y^{\mathrm{L}}, y^{\mathrm{U}}] =: Z \in \mathbb{I}(X \times Y)$, let

$$f^{\mathrm{cv}}_Z(x, y) = -x-y, \quad g^{\mathrm{cv}}_Z(x, y) = x^2+2\max\left\{x^{\mathrm{L}}y + y^{\mathrm{L}}x - x^{\mathrm{L}}y^{\mathrm{L}}, x^{\mathrm{U}}y + y^{\mathrm{U}}x - x^{\mathrm{U}}y^{\mathrm{U}}\right\}+y^2-1.$$

The scheme $(f^{\mathrm{cv}}_Z)_{Z\in\mathbb{I}(X\times Y)}$ has arbitrarily high pointwise convergence order on $X\times Y$, while the scheme $(g^{\mathrm{cv}}_Z)_{Z\in\mathbb{I}(X\times Y)}$ has second-order pointwise convergence on $X \times Y$.

Let $x^{\mathrm{L}} = y^{\mathrm{L}} = 0.5 - \varepsilon$, $x^{\mathrm{U}} = y^{\mathrm{U}} = 0.5 + \varepsilon$ with $0 < \varepsilon \leq 0.5$. The width of Z is $w(Z) = 2\varepsilon$. The optimal objective value of Problem (P) on Z is -1, while the point $\left(0.5 - 0.25\varepsilon^2, 0.5 + 0.5\varepsilon^2\right)$ is feasible for the lower bounding problem on Z with objective value $-1 - 0.25\varepsilon^2$. Convergence at the point $(0.5, 0.5)$ is, therefore, at most second-order.

The next result provides a slight generalization of Corollary 6.4.12 by showing that under the assumptions of Corollary 6.4.12, the lower bounding scheme $(\mathcal{L}(Z))_{Z\in\mathbb{I}(X\times Y)}$ in fact exhibits (at least) second-order convergence on a neighborhood of $(\mathbf{x}^{\mathrm{S}}, \mathbf{y}^{\mathrm{S}})$ (this result is motivated by the assumptions on the convergence order of the lower bounding scheme in the analysis of the cluster problem in Chapter 5).

Corollary 6.4.17. Consider Problem (P) with $m_E = 0$. Suppose f is Lipschitz continuous on $X \times Y$. Let $(f^{\mathrm{cv}}_Z)_{Z\in\mathbb{I}(X\times Y)}$ denote a continuous scheme of convex relaxations of f in $X \times Y$ with pointwise convergence order $\gamma^{\mathrm{cv}}_f \geq 1$, and convergence order $\beta^{\mathrm{cv}}_f \geq 1$ with corresponding constant τ^{cv}_f. Furthermore, let $(g^{\mathrm{cv}}_{j,Z})_{Z\in\mathbb{I}(X\times Y)}$, $j = 1, \cdots, m_I$, denote continuous schemes of convex relaxations of $g_1, \cdots, g_{m_I}$, respectively, in $X \times Y$ with pointwise convergence orders $\gamma^{\text{cv}}_{g,1} > 0, \cdots, \gamma^{\text{cv}}_{g,m_I} > 0$ and corresponding constants $\tau^{\text{cv}}_{g,1}, \cdots, \tau^{\text{cv}}_{g,m_I}$.

Suppose $(\mathbf{x}^{\text{S}}, \mathbf{y}^{\text{S}}) \in X \times Y$ such that $\mathbf{g}(\mathbf{x}^{\text{S}}, \mathbf{y}^{\text{S}}) < \mathbf{0}$ (i.e., $(\mathbf{x}^{\text{S}}, \mathbf{y}^{\text{S}})$ is a Slater point). Then, $\exists \delta > 0$ such that the scheme of lower bounding problems $(\mathcal{L}(Z))_{Z \in \mathbb{I}(X \times Y)}$ with

$$
\begin{aligned}
(\mathcal{O}(Z))_{Z \in \mathbb{I}(X \times Y)} &:= \left(\min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}^{\text{cv}}(Z)} f^{\text{cv}}_Z(\mathbf{x}, \mathbf{y}) \right)_{Z \in \mathbb{I}(X \times Y)}, \\
(\mathcal{I}_C(Z))_{Z \in \mathbb{I}(X \times Y)} &:= (\overline{\mathbf{g}}^{\text{cv}}_Z(Z))_{Z \in \mathbb{I}(X \times Y)}
\end{aligned}
$$

is at least β^{cv}_f-order convergent on $\left\{ (\mathbf{x}, \mathbf{y}) : \left\| (\mathbf{x}, \mathbf{y}) - (\mathbf{x}^{\text{S}}, \mathbf{y}^{\text{S}}) \right\|_\infty < \delta \right\}$.

Proof. Let $g_j(\mathbf{x}^{\text{S}}, \mathbf{y}^{\text{S}}) = -\varepsilon_j < 0$, $j = 1, \cdots, m_I$. Since g_j is continuous for each $j \in \{1, \cdots, m_I\}$, there exists $\delta_j > 0$, $\forall j \in \{1, \cdots, m_I\}$, such that $\left\| (\mathbf{x}, \mathbf{y}) - (\mathbf{x}^{\text{S}}, \mathbf{y}^{\text{S}}) \right\|_\infty < \delta_j$ implies $\left| g_j(\mathbf{x}, \mathbf{y}) - g_j(\mathbf{x}^{\text{S}}, \mathbf{y}^{\text{S}}) \right| < \varepsilon_j$ (see Lemma 2.2.2). Define $\bar{\delta} := \min\limits_{j \in \{1, \cdots, m_I\}} \delta_j$, note that $\bar{\delta} > 0$, and let $\delta := \frac{\bar{\delta}}{2}$.

Consider $Z \in \mathbb{I}(X \times Y)$ with $Z \cap \left\{ (\mathbf{x}, \mathbf{y}) : \left\| (\mathbf{x}, \mathbf{y}) - (\mathbf{x}^{\text{S}}, \mathbf{y}^{\text{S}}) \right\|_\infty < \delta \right\} \neq \emptyset$ and $w(Z) \leq \delta$. Similar to the proof of Corollary 6.4.12, it can be shown that

$$
\min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}(Z)} f(\mathbf{x}, \mathbf{y}) \; - \min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}^{\text{cv}}(Z)} f^{\text{cv}}_Z(\mathbf{x}, \mathbf{y}) \leq \tau^{\text{cv}}_f w(Z)^{\beta^{\text{cv}}_f}.
$$

The desired result follows by analogy to Lemma 6.3.8 by noting from Theorem 6.4.4 that the lower bounding scheme $(\mathcal{L}(Z))_{Z \in \mathbb{I}(X \times Y)}$ has at least first-order convergence on the feasible set $\left\{ (\mathbf{x}, \mathbf{y}) : \left\| (\mathbf{x}, \mathbf{y}) - (\mathbf{x}^{\text{S}}, \mathbf{y}^{\text{S}}) \right\|_\infty < \delta \right\}$. $\square$

While it may appear that the neighborhood of a Slater point on which second-order convergence of the lower bounding scheme is guaranteed by Corollary 6.4.17 can be unnecessarily small, Example 6.4.18 shows that a stronger result cannot be deduced without additional assumptions.

A natural question is whether second-order convergence is guaranteed on $X \times Y$ when second-order pointwise convergent schemes of (convex) relaxations of the objective function f, the inequality constraint functions $g_1, \cdots, g_{m_I}$, and the equality constraint functions $h_1, \cdots, h_{m_E}$ are used by the lower bounding scheme. The following example shows that even when schemes of (convex) envelopes are used to underestimate smooth functions f, $\mathbf{g}$, and $\mathbf{h}$, at most first-order convergence can be guaranteed at certain points in $X \times Y$.

Example 6.4.18. Let $X = [0, 0], Y = [-1, 1]$, $m_I = 1$, and $m_E = 0$ with $f(x, y) = -y$ and $g(x, y) = y^3$. For any $[0, 0] \times [-\varepsilon, \varepsilon] =: Z \in \mathbb{I}(X \times Y)$ with $\varepsilon > 0$, let

$$f_Z^{\text{cv}}(x, y) = -y, \quad g_Z^{\text{cv}}(x, y) = \begin{cases} -0.25\varepsilon^3 + 0.75\varepsilon^2 y, & \text{if } y < 0.5\varepsilon \\ y^3, & \text{if } y \geq 0.5\varepsilon \end{cases}.$$

Note that the scheme $(f_Z^{\text{cv}})_{Z \in \mathbb{I}(X \times Y)}$ has arbitrarily high pointwise convergence order on $X \times Y$ and the scheme $(g_Z^{\text{cv}})_{Z \in \mathbb{I}(X \times Y)}$, which is the scheme of convex envelopes of g on $X \times Y$ [142], has (at least) second-order pointwise convergence on $X \times Y$. Also note that $(g_Z^{\text{cv}})_{Z \in \mathbb{I}(X \times Y)}$ has arbitrarily high convergence order on $X \times Y$.

The width of Z is $w(Z) = 2\varepsilon$. The optimal objective value of Problem (P) on Z is 0, while the optimal objective of the lower bounding problem on Z is $-\frac{\varepsilon}{3}$. Convergence at the point $(0, 0)$ is, therefore, at most first-order.

Despite the fact that we only have first-order convergence at the global minimizer in Example 6.4.18, the reader can verify that the natural interval extension-based lower bounding scheme along with the interval bisection branching rule and lowest lower bound node selection rule is sufficient to mitigate the cluster problem for this case [108].

The following result establishes second-order convergence of a convex relaxation-based lower bounding scheme at a feasible point $(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}}) \in X \times Y$ when second-order pointwise convergent schemes of relaxations are used and the dual lower bounding scheme (see Section 6.4.2) is second-order convergent at $(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}})$. This result will be used to prove second-order convergence of such convex relaxation-based lower bounding schemes at KKT points (see Definition 2.3.18) in Corollary 6.4.28.

Theorem 6.4.19. Consider Problem (P), and let $(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}}) \in X \times Y$ be a feasible point. Suppose the dual lower bounding scheme has convergence of order $\beta_d > 0$ at $(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}})$ with a corresponding scheme of bounded dual variables $\left(\left(\boldsymbol{\mu}_Z^{(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}})}, \boldsymbol{\lambda}_Z^{(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}})}\right)\right)_{Z \in \mathbb{I}(X \times Y)}$ (not necessarily optimal, but which yield β_d-order convergence at $(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}})$) with $\left(\boldsymbol{\mu}_Z^{(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}})}, \boldsymbol{\lambda}_Z^{(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}})}\right) \in \mathbb{R}_+^{m_I} \times \mathbb{R}^{m_E}$, $\left\|\boldsymbol{\mu}_Z^{(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}})}\right\|_\infty \leq \bar{\mu}$ and $\left\|\boldsymbol{\lambda}_Z^{(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}})}\right\|_\infty \leq \bar{\lambda}$, $\forall Z \in \mathbb{I}(X \times Y)$, for some constants $\bar{\mu}, \bar{\lambda} \geq 0$ (see Section 6.4.2). Let $(f_Z^{\text{cv}})_{Z \in \mathbb{I}(X \times Y)}$, $(g_{j,Z}^{\text{cv}})_{Z \in \mathbb{I}(X \times Y)}$, $j = 1, \cdots, m_I$, denote continuous schemes of convex relaxations of f, $g_1, \cdots, g_{m_I}$, respectively, in $X \times Y$ with pointwise convergence orders $\gamma_f^{\text{cv}} \geq 1$, $\gamma_{g,1}^{\text{cv}} \geq 1, \cdots, \gamma_{g,m_I}^{\text{cv}} \geq 1$ and corresponding constants τ_f^{cv}, $\tau_{g,1}^{\text{cv}}, \cdots, \tau_{g,m_I}^{\text{cv}}$. Let $(h_{k,Z}^{\text{cv}}, h_{k,Z}^{\text{cc}})_{Z \in \mathbb{I}(X \times Y)}$, $k = 1, \cdots, m_E$, denote continuous schemes of relaxations of $h_1, \cdots, h_{m_E}$, respectively, in $X \times Y$ with pointwise convergence orders $\gamma_{h,1} \geq 1, \cdots, \gamma_{h,m_E} \geq 1$ and corresponding constants $\tau_{h,1}, \cdots, \tau_{h,m_E}$. Then, the scheme of lower bounding problems $(\mathcal{L}(Z))_{Z\in\mathbb{I}(X\times Y)}$ with

$$(\mathcal{O}(Z))_{Z\in\mathbb{I}(X\times Y)} := \left(\min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}^{\text{cv}}(Z)} f^{\text{cv}}_Z(\mathbf{x},\mathbf{y})\right)_{Z\in\mathbb{I}(X\times Y)},$$

$$(\mathcal{I}_C(Z))_{Z\in\mathbb{I}(X\times Y)} := \big(\big\{(\mathbf{v},\mathbf{w}) \in \mathbb{R}^{m_I} \times \mathbb{R}^{m_E} : \mathbf{v} = \mathbf{g}^{\text{cv}}_Z(\mathbf{x},\mathbf{y}), \mathbf{h}^{\text{cv}}_Z(\mathbf{x},\mathbf{y}) \leq \mathbf{w} \leq \mathbf{h}^{\text{cc}}_Z(\mathbf{x},\mathbf{y}) \text{ for some } (\mathbf{x},\mathbf{y}) \in Z\big\}\big)_{Z\in\mathbb{I}(X\times Y)}$$

is at least $\min\left\{\min\left\{\gamma^{\text{cv}}_f, \min_{j\in\{1,\cdots,m_I\}} \gamma^{\text{cv}}_{g,j}, \min_{k\in\{1,\cdots,m_E\}} \gamma_{h,k}\right\}, \beta_d\right\}$-order convergent at $(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}})$.

Proof. Let $\beta := \min\left\{\min\left\{\gamma^{\text{cv}}_f, \min_{j\in\{1,\cdots,m_I\}} \gamma^{\text{cv}}_{g,j}, \min_{k\in\{1,\cdots,m_E\}} \gamma_{h,k}\right\}, \beta_d\right\}$, and let $\boldsymbol{\mu}_Z := \boldsymbol{\mu}^{(\mathbf{x}^{\text{f}},\mathbf{y}^{\text{f}})}_Z$, $\boldsymbol{\lambda}_Z := \boldsymbol{\lambda}^{(\mathbf{x}^{\text{f}},\mathbf{y}^{\text{f}})}_Z$, $\forall Z \in \mathbb{I}(X \times Y)$, denote the scheme of dual variables corresponding to the dual lower bounding scheme (we omit the dependence of the dual variables on $(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}})$ for ease of exposition). Since we are concerned about the convergence order at the feasible point $(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}})$, it suffices to show the existence of $\tau \geq 0$ such that for every $Z \in \mathbb{I}(X \times Y)$ with $(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}}) \in Z$,

$$\min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}(Z)} f(\mathbf{x},\mathbf{y}) \; - \min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}^{\text{cv}}(Z)} f^{\text{cv}}_Z(\mathbf{x},\mathbf{y}) \leq \tau w(Z)^{\beta}.$$

Consider $Z \in \mathbb{I}(X \times Y)$ with $(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}}) \in Z$. By virtue of the assumption that the dual lower bounding scheme, with the dual variables fixed to $((\boldsymbol{\mu}_Z, \boldsymbol{\lambda}_Z))_{Z\in\mathbb{I}(X\times Y)}$, has convergence of order β_d at $(\mathbf{x}^{\text{f}}, \mathbf{y}^{\text{f}})$, we have

$$\min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}(Z)} f(\mathbf{x},\mathbf{y}) - \min_{(\mathbf{x},\mathbf{y})\in Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\text{T}}_Z \mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}^{\text{T}}_Z \mathbf{h}(\mathbf{x},\mathbf{y})\right] \leq \tau_d w(Z)^{\beta_d}. \tag{6.5}$$

Choose $\boldsymbol{\lambda}_{Z,+}, \boldsymbol{\lambda}_{Z,-} \in \mathbb{R}^{m_E}_+$ such that $\boldsymbol{\lambda}_Z = \boldsymbol{\lambda}_{Z,+} - \boldsymbol{\lambda}_{Z,-}$, $\|\boldsymbol{\lambda}_{Z,+}\|_\infty \leq \bar{\lambda}$, and $\|\boldsymbol{\lambda}_{Z,-}\|_\infty \leq \bar{\lambda}$.

Let $\bar{\gamma} := \min\left\{\gamma^{\text{cv}}_f, \min_{j\in\{1,\cdots,m_I\}} \gamma^{\text{cv}}_{g,j}, \min_{k\in\{1,\cdots,m_E\}} \gamma_{h,k}\right\}$. We have

$$\begin{aligned}
&\min_{(\mathbf{x},\mathbf{y})\in Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\text{T}}_Z \mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}^{\text{T}}_Z \mathbf{h}(\mathbf{x},\mathbf{y})\right] - \min_{\mathcal{F}^{\text{cv}}(Z)} f^{\text{cv}}_Z(\mathbf{x},\mathbf{y}) \\
\leq &\min_{(\mathbf{x},\mathbf{y})\in Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\text{T}}_Z \mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}^{\text{T}}_Z \mathbf{h}(\mathbf{x},\mathbf{y})\right] - \\
&\qquad \sup_{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}_1\geq\mathbf{0},\boldsymbol{\lambda}_2\leq\mathbf{0}} \; \min_{(\mathbf{x},\mathbf{y})\in Z} \left[f^{\text{cv}}_Z(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\text{T}} \mathbf{g}^{\text{cv}}_Z(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}^{\text{T}}_1 \mathbf{h}^{\text{cv}}_Z(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}^{\text{T}}_2 \mathbf{h}^{\text{cc}}_Z(\mathbf{x},\mathbf{y})\right]
\end{aligned}$$

$$
\begin{aligned}
&\leq \min_{(\mathbf{x},\mathbf{y})\in Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}_Z^{\mathrm{T}} \mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_{Z,+}^{\mathrm{T}} \mathbf{h}(\mathbf{x},\mathbf{y}) - \boldsymbol{\lambda}_{Z,-}^{\mathrm{T}} \mathbf{h}(\mathbf{x},\mathbf{y}) \right] - \\
&\qquad\qquad \min_{(\mathbf{x},\mathbf{y})\in Z} \left[f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}_Z^{\mathrm{T}} \mathbf{g}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_{Z,+}^{\mathrm{T}} \mathbf{h}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) - \boldsymbol{\lambda}_{Z,-}^{\mathrm{T}} \mathbf{h}_Z^{\mathrm{cc}}(\mathbf{x},\mathbf{y}) \right] \\
&\leq \max_{(\mathbf{x},\mathbf{y})\in Z} \Big[\left(f(\mathbf{x},\mathbf{y}) - f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) \right) + \boldsymbol{\mu}_Z^{\mathrm{T}} \left(\mathbf{g}(\mathbf{x},\mathbf{y}) - \mathbf{g}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) \right) + \\
&\qquad\qquad \boldsymbol{\lambda}_{Z,+}^{\mathrm{T}} \left(\mathbf{h}(\mathbf{x},\mathbf{y}) - \mathbf{h}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) \right) + \boldsymbol{\lambda}_{Z,-}^{\mathrm{T}} \left(\mathbf{h}_Z^{\mathrm{cc}}(\mathbf{x},\mathbf{y}) - \mathbf{h}(\mathbf{x},\mathbf{y}) \right) \Big] \\
&\leq \max_{(\mathbf{x},\mathbf{y})\in Z} \left(f(\mathbf{x},\mathbf{y}) - f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) \right) + \max_{(\mathbf{x},\mathbf{y})\in Z} \boldsymbol{\mu}_Z^{\mathrm{T}} \left(\mathbf{g}(\mathbf{x},\mathbf{y}) - \mathbf{g}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) \right) + \\
&\qquad\qquad \max_{(\mathbf{x},\mathbf{y})\in Z} \boldsymbol{\lambda}_{Z,+}^{\mathrm{T}} \left(\mathbf{h}(\mathbf{x},\mathbf{y}) - \mathbf{h}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) \right) + \max_{(\mathbf{x},\mathbf{y})\in Z} \boldsymbol{\lambda}_{Z,-}^{\mathrm{T}} \left(\mathbf{h}_Z^{\mathrm{cc}}(\mathbf{x},\mathbf{y}) - \mathbf{h}(\mathbf{x},\mathbf{y}) \right) \\
&\leq \tau_f^{\mathrm{cv}} w(Z)^{\gamma_f^{\mathrm{cv}}} + \sum_{j=1}^{m_I} \bar{\mu} \tau_{g,j}^{\mathrm{cv}} w(Z)^{\gamma_{g,j}^{\mathrm{cv}}} + 2 \sum_{k=1}^{m_E} \bar{\lambda} \tau_{h,k} w(Z)^{\gamma_{h,k}} \\
&\leq \left(\tau_f^{\mathrm{cv}} w(X \times Y)^{\gamma_f^{\mathrm{cv}} - \bar{\gamma}} + \sum_{j=1}^{m_I} \bar{\mu} \tau_{g,j}^{\mathrm{cv}} w(X \times Y)^{\gamma_{g,j}^{\mathrm{cv}} - \bar{\gamma}} + 2 \sum_{k=1}^{m_E} \bar{\lambda} \tau_{h,k} w(X \times Y)^{\gamma_{h,k} - \bar{\gamma}} \right) w(Z)^{\bar{\gamma}},
\end{aligned}
\tag{6.6}
$$

where Step 1 follows from weak duality and Step 3 follows from Lemma 2.3.35.

Therefore, from Equations (6.5) and (6.6), we have

$$
\begin{aligned}
& \min_{(\mathbf{x},\mathbf{y})\in \mathcal{F}(Z)} f(\mathbf{x},\mathbf{y}) \;-\; \min_{(\mathbf{x},\mathbf{y})\in \mathcal{F}^{\mathrm{cv}}(Z)} f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) \\
=& \left(\min_{(\mathbf{x},\mathbf{y})\in \mathcal{F}(Z)} f(\mathbf{x},\mathbf{y}) \;-\; \min_{(\mathbf{x},\mathbf{y})\in Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}_Z^{\mathrm{T}} \mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_Z^{\mathrm{T}} \mathbf{h}(\mathbf{x},\mathbf{y}) \right] \right) + \\
& \quad \left(\min_{(\mathbf{x},\mathbf{y})\in Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}_Z^{\mathrm{T}} \mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_Z^{\mathrm{T}} \mathbf{h}(\mathbf{x},\mathbf{y}) \right] - \min_{(\mathbf{x},\mathbf{y})\in \mathcal{F}^{\mathrm{cv}}(Z)} f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) \right) \\
\leq& \tau w(Z)^{\beta},
\end{aligned}
$$

where the prefactor τ is defined as

$$
\tau := \left(\tau_f^{\mathrm{cv}} w(X \times Y)^{\gamma_f^{\mathrm{cv}} - \beta} + \sum_{j=1}^{m_I} \bar{\mu} \tau_{g,j}^{\mathrm{cv}} w(X \times Y)^{\gamma_{g,j}^{\mathrm{cv}} - \beta} + 2 \sum_{k=1}^{m_E} \bar{\lambda} \tau_{h,k} w(X \times Y)^{\gamma_{h,k} - \beta} + \tau_d w(X \times Y)^{\beta_d - \beta} \right).
$$

□

6.4.2 Duality-based branch-and-bound

In this section, we investigate the convergence order of a Lagrangian dual-based lower bounding scheme. Before we define the convergence order of the scheme, the Lagrangian dual problem is introduced (cf. Section 2.3.2.1.3) and some of its properties are outlined.

The dual problem of Problem (P) that is obtained by dualizing all of the constraints $\mathbf{g}(\mathbf{x},\mathbf{y}) \leq \mathbf{0}$ and $\mathbf{h}(\mathbf{x},\mathbf{y}) = \mathbf{0}$ is given by

$$\begin{aligned} &\sup_{\boldsymbol{\mu},\boldsymbol{\lambda}} \; q(\boldsymbol{\mu},\boldsymbol{\lambda}) \\ &\text{ s.t. } \boldsymbol{\mu} \in \mathbb{R}^{m_I}_+, \boldsymbol{\lambda} \in \mathbb{R}^{m_E}, \end{aligned} \tag{D}$$

where $q : \mathbb{R}^{m_I}_+ \times \mathbb{R}^{m_E} \to \mathbb{R}$, defined by

$$q(\boldsymbol{\mu},\boldsymbol{\lambda}) := \min_{(\mathbf{x},\mathbf{y})\in X\times Y} f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y}), \quad \forall (\boldsymbol{\mu},\boldsymbol{\lambda}) \in \mathbb{R}^{m_I}_+ \times \mathbb{R}^{m_E},$$

is the Lagrangian dual function. Let $\min$ (P) and $\sup$ (D) respectively denote the optimal objective values of Problem (P) and Problem (D). From weak duality, we know that $\min$ (P) $\geq \sup$ (D), which validates the use of Problem (D) as a lower bounding problem.

The following result shows that the lower bounds obtained by solving the Lagrangian dual Problem (D) are stronger than those obtained by solving any convex relaxation-based lower bounding problem.

Lemma 6.4.20. Consider Problem (P), and suppose $Z \in \mathbb{I}(X \times Y)$. Let f^{cv}_Z and $\mathbf{g}^{\mathrm{cv}}_Z$ denote convex relaxations of f and $\mathbf{g}$, respectively, on Z, and let $\mathbf{h}^{\mathrm{cv}}_Z$ and $\mathbf{h}^{\mathrm{cc}}_Z$ denote convex and concave relaxations, respectively, of $\mathbf{h}$ on Z. Furthermore, assume that strong duality holds for the convex relaxation-based lower bounding problem $\min\limits_{(\mathbf{x},\mathbf{y})\in\mathcal{F}^{\mathrm{cv}}(Z)} f^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y})$. Then the lower bound obtained by solving the Lagrangian dual problem is at least as strong as that obtained by solving the above convex relaxation-based lower bounding problem, i.e.,

$$\sup_{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}} \min_{(\mathbf{x},\mathbf{y})\in Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y})\right] - \min_{\mathcal{F}^{\mathrm{cv}}(Z)} f^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y}) \geq 0.$$

Proof. Since strong duality holds for the convex relaxation-based lower bounding problem, the difference between the lower bounds can be rewritten as

$$\begin{aligned} &\sup_{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}} \min_{(\mathbf{x},\mathbf{y})\in Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y})\right] - \\ &\qquad \sup_{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}_1\geq\mathbf{0},\boldsymbol{\lambda}_2\leq\mathbf{0}} \min_{(\mathbf{x},\mathbf{y})\in Z} \left[f^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_1^{\mathrm{T}}\mathbf{h}^{\mathrm{cv}}_Z(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_2^{\mathrm{T}}\mathbf{h}^{\mathrm{cc}}_Z(\mathbf{x},\mathbf{y})\right] \end{aligned}$$

$$
\begin{aligned}
&= \sup_{\boldsymbol{\mu}\geq \mathbf{0},\boldsymbol{\lambda}_1\geq\mathbf{0},\boldsymbol{\lambda}_2\leq\mathbf{0}} \min_{(\mathbf{x},\mathbf{y})\in Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_1^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_2^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y})\right] - \\
&\qquad \sup_{\boldsymbol{\mu}\geq \mathbf{0},\boldsymbol{\lambda}_1\geq\mathbf{0},\boldsymbol{\lambda}_2\leq\mathbf{0}} \min_{(\mathbf{x},\mathbf{y})\in Z} \left[f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_1^{\mathrm{T}}\mathbf{h}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_2^{\mathrm{T}}\mathbf{h}_Z^{\mathrm{cc}}(\mathbf{x},\mathbf{y})\right] \\
&\geq 0,
\end{aligned}
$$

where the last step follows from the fact that $\forall(\mathbf{x},\mathbf{y})\in Z, \boldsymbol{\mu}\geq\mathbf{0}, \boldsymbol{\lambda}_1\geq\mathbf{0}, \boldsymbol{\lambda}_2\leq\mathbf{0}$,

$$
\begin{aligned}
&f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_1^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_2^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y}) \\
\geq\ & f_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_1^{\mathrm{T}}\mathbf{h}_Z^{\mathrm{cv}}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}_2^{\mathrm{T}}\mathbf{h}_Z^{\mathrm{cc}}(\mathbf{x},\mathbf{y}).
\end{aligned}
$$

□

The following result due to Dür [70] establishes the condition under which the dual lower bounding problem detects infeasibility.

Lemma 6.4.21. Consider Problem (P) (satisfying Assumption 6.2.1). We have

$$
\sup(\mathrm{D}) = +\infty \iff \mathrm{conv}\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}(X\times Y)}\right) \cap \left(\mathbb{R}_-^{m_I}\times\{\mathbf{0}\}\right) = \emptyset.
$$

Proof. The result follows, in part, by replacing $\mathbf{h}(\mathbf{x},\mathbf{y}) = \mathbf{0}$ with $\mathbf{h}(\mathbf{x},\mathbf{y}) \leq \mathbf{0}$ and $-\mathbf{h}(\mathbf{x},\mathbf{y}) \leq \mathbf{0}$ and using Theorem 2 in [70]. □

Definition 6.3.12 can be applied to analyze the convergence order of the above duality-based lower bounding scheme as follows.

The scheme of dual lower bounding problems $(\mathcal{L}(Z))_{Z\in\mathbb{I}(X\times Y)}$ with

$$
\begin{aligned}
(\mathcal{O}(Z))_{Z\in\mathbb{I}(X\times Y)} &:= \left(\sup_{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}} \min_{(\mathbf{x},\mathbf{y})\in Z}\left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y})\right]\right)_{Z\in\mathbb{I}(X\times Y)}, \\
(\mathcal{I}_C(Z))_{Z\in\mathbb{I}(X\times Y)} &:= \left(\mathrm{conv}\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}(Z)}\right)\right)_{Z\in\mathbb{I}(X\times Y)}
\end{aligned}
$$

is thus said to have convergence of order $\beta > 0$ at

1. a feasible point $(\mathbf{x},\mathbf{y}) \in X\times Y$ if there exists $\tau \geq 0$ such that for every $Z \in \mathbb{I}(X\times Y)$

with $(\mathbf{x}, \mathbf{y}) \in Z$,

$$\min_{(\mathbf{v},\mathbf{w}) \in \mathcal{F}(Z)} f(\mathbf{v}, \mathbf{w}) - \sup_{\boldsymbol{\mu} \geq \mathbf{0}, \boldsymbol{\lambda}} \min_{(\mathbf{v},\mathbf{w}) \in Z} \left[f(\mathbf{v}, \mathbf{w}) + \boldsymbol{\mu}^{\mathrm{T}} \mathbf{g}(\mathbf{v}, \mathbf{w}) + \boldsymbol{\lambda}^{\mathrm{T}} \mathbf{h}(\mathbf{v}, \mathbf{w}) \right] \leq \tau w(Z)^{\beta}.$$

2. an infeasible point $(\mathbf{x}, \mathbf{y}) \in X \times Y$ if there exists $\bar{\tau} \geq 0$ such that for every $Z \in \mathbb{I}(X \times Y)$ with $(\mathbf{x}, \mathbf{y}) \in Z$,

$$d\left(\overline{\begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix}}(Z), \mathbb{R}_{-}^{m_I} \times \{\mathbf{0}\}\right) - d\left(\operatorname{conv}\left(\overline{\begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix}}(Z)\right), \mathbb{R}_{-}^{m_I} \times \{\mathbf{0}\}\right) \leq \bar{\tau} w(Z)^{\beta}.$$

We associate with the dual lower bounding scheme, $(\mathcal{L}(Z))_{Z \in \mathbb{I}(X \times Y)}$, at a feasible point $(\mathbf{x}, \mathbf{y})$, a scheme of dual variables $((\boldsymbol{\mu}_Z^{(\mathbf{x},\mathbf{y})}, \boldsymbol{\lambda}_Z^{(\mathbf{x},\mathbf{y})}))_{Z \in \mathbb{I}(X \times Y)}$ corresponding to the solution of the scheme of dual lower bounding problems $(\mathcal{O}(Z))_{Z \in \mathbb{I}(X \times Y)}$ with $(\mathbf{x}, \mathbf{y}) \in Z$ (note that sup (D) may not be attained, in which case we assume that dual variables that yield a dual function value arbitrarily close to the supremum are available). Using Lemma 6.4.21, we next show that if the convex relaxation-based lower bounding problem corresponding to Problem (P) that is obtained by replacing the functions in Problem (P) with their envelopes is infeasible, then sup (D) $= +\infty$.

Lemma 6.4.22. Let $(g_{j,Z}^{\mathrm{cv}})_{Z \in \mathbb{I}(X \times Y)}$, $j = 1, \cdots, m_I$, denote (any) schemes of convex relaxations of $g_1, \cdots, g_{m_I}$ in $X \times Y$ and $(h_{k,Z}^{\mathrm{cv}}, h_{k,Z}^{\mathrm{cc}})_{Z \in \mathbb{I}(X \times Y)}$, $k = 1, \cdots, m_E$, denote (any) schemes of relaxations of $h_1, \cdots, h_{m_E}$ in $X \times Y$. Then for each $Z \in \mathbb{I}(X \times Y)$, we have

$$d\left(\overline{\begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix}}(Z), \mathbb{R}_{-}^{m_I} \times \{\mathbf{0}\}\right) \geq d\left(\operatorname{conv}\left(\overline{\begin{bmatrix} \mathbf{g} \\ \mathbf{h} \end{bmatrix}}(Z)\right), \mathbb{R}_{-}^{m_I} \times \{\mathbf{0}\}\right) \geq d\left(\mathcal{I}_C(Z), \mathbb{R}_{-}^{m_I} \times \{\mathbf{0}\}\right),$$

where $\mathcal{I}_C(Z)$ is defined as

$$\mathcal{I}_C(Z) := \left\{ (\mathbf{v}, \mathbf{w}) \in \mathbb{R}^{m_I} \times \mathbb{R}^{m_E} : \mathbf{v} = \mathbf{g}_Z^{\mathrm{cv}}(\mathbf{x}, \mathbf{y}), \mathbf{h}_Z^{\mathrm{cv}}(\mathbf{x}, \mathbf{y}) \leq \mathbf{w} \leq \mathbf{h}_Z^{\mathrm{cc}}(\mathbf{x}, \mathbf{y}) \text{ for some } (\mathbf{x}, \mathbf{y}) \in Z \right\}.$$

Proof. The first inequality trivially holds. To prove the second inequality, we first notice that

$$d\left(\mathcal{I}_C(Z), \mathbb{R}_{-}^{m_I} \times \{\mathbf{0}\}\right) = d\left(\bar{\mathcal{I}}_C(Z), \mathbb{R}_{-}^{m_I} \times \{\mathbf{0}\}\right),$$

where $\bar{\mathcal{I}_C}(Z)$ is defined as

$$\bar{\mathcal{I}_C}(Z) := \{(\mathbf{v}, \mathbf{w}) \in \mathbb{R}^{m_I} \times \mathbb{R}^{m_E} : \mathbf{v} \geq \mathbf{g}^{\mathrm{cv}}_Z(\mathbf{x}, \mathbf{y}), \mathbf{h}^{\mathrm{cv}}_Z(\mathbf{x}, \mathbf{y}) \leq \mathbf{w} \leq \mathbf{h}^{\mathrm{cc}}_Z(\mathbf{x}, \mathbf{y}) \text{ for some } (\mathbf{x}, \mathbf{y}) \in Z\}.$$

Note that $\bar{\mathcal{I}_C}(Z)$ is a convex set since it can be represented as the direct sum of two convex sets.

Since $\mathrm{conv}\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}(Z)}\right)$ is the smallest convex set that encloses $\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}(Z)}$, the desired result follows. $\square$

Theorem 6.4.23. Consider Problem (P). Suppose strong duality holds for the scheme of convex relaxation-based lower bounding problems for Problem (P) obtained by using the schemes of (convex) envelopes of f, $\mathbf{g}$, and $\mathbf{h}$. If the assumptions of Theorem 6.4.4 hold for the functions f, $\mathbf{g}$, $\mathbf{h}$, and the schemes of (convex) envelopes of f, $\mathbf{g}$, and $\mathbf{h}$, the dual lower bounding scheme is (at least) first-order convergent on $X \times Y$. Furthermore, if the assumptions of Theorem 6.4.9 hold for the schemes of (convex) envelopes of f, $\mathbf{g}$, and $\mathbf{h}$ and $(\mathbf{x}^{\mathrm{f}}, \mathbf{y}^{\mathrm{f}}) \in X \times Y$, the dual lower bounding scheme is (at least) second-order convergent at $(\mathbf{x}^{\mathrm{f}}, \mathbf{y}^{\mathrm{f}})$.

Proof. From Lemma 6.4.22, we have that the convergence order of the dual lower bounding scheme at an infeasible point $(\mathbf{x}, \mathbf{y}) \in X \times Y$ is at least as high as the convergence order at $(\mathbf{x}, \mathbf{y})$ of the convex relaxation-based lower bounding scheme obtained by using the schemes of (convex) envelopes of f, $\mathbf{g}$, and $\mathbf{h}$. Lemma 6.4.20 implies that the lower bounds obtained using the dual lower bounding scheme are at least as tight as the lower bounds obtained using the schemes of (convex) envelopes of f, $\mathbf{g}$, and $\mathbf{h}$. The desired result follows from Definition 6.3.12. $\square$

Note that the conclusions of Theorem 6.4.23 hold even if the schemes of relaxations of f, $\mathbf{g}$, and $\mathbf{h}$ do not correspond to their envelopes, so long as the (remaining) assumptions of Theorem 6.4.23 are satisfied.

Remark 6.4.24. The assumption of strong duality is in fact not required to show first-order convergence of the dual lower bounding scheme when all functions in Problem (P) are Lipschitz continuous. For this case, the proof of first-order convergence at infeasible points follows from Lemmata 6.3.10, 6.4.1, and 6.4.22, and the proof of first-order convergence at feasible points follows from Proposition 1 in [70].

Theorem 6.4.23 makes no assumptions on the boundedness of schemes of dual variables. This is reflected in the application of the dual lower bounding scheme to Example 6.4.18 where the optimal scheme of dual variables can be unbounded (note, however, that first-order convergence of the dual lower bounding scheme at the global minimizer of Example 6.4.18 can be achieved using bounded schemes of dual variables when the dual problem is not solved to optimality). Furthermore, Example 6.4.10 shows that the convergence order of the dual lower bounding scheme can be as low as two at $(\mathbf{x}^{\mathrm{f}}, \mathbf{y}^{\mathrm{f}})$ when the assumptions of Theorem 6.4.9 are satisfied for the schemes of (convex) envelopes of f, $\mathbf{g}$, and $\mathbf{h}$ (see Lemma 6.5.12). The following result shows that in the absence of equality constraints, the dual lower bounding scheme has arbitrarily high convergence order at unconstrained points.

Proposition 6.4.25. Consider Problem (P) with $m_E = 0$. Suppose f and g_j, $\forall j \in \{1, \cdots, m_I\}$, are Lipschitz continuous on $X \times Y$. Furthermore, suppose $(\mathbf{x}^{\mathrm{S}}, \mathbf{y}^{\mathrm{S}}) \in X \times Y$ such that $\mathbf{g}(\mathbf{x}^{\mathrm{S}}, \mathbf{y}^{\mathrm{S}}) < \mathbf{0}$ (i.e., $(\mathbf{x}^{\mathrm{S}}, \mathbf{y}^{\mathrm{S}})$ is a Slater point). The dual lower bounding scheme has arbitrarily high convergence order at $(\mathbf{x}^{\mathrm{S}}, \mathbf{y}^{\mathrm{S}})$.

Proof. The arguments below are closely related to the proof of Corollary 6.4.12.

Since we wish to prove that the dual lower bounding scheme has arbitrarily high convergence order at the feasible point $(\mathbf{x}^{\mathrm{S}}, \mathbf{y}^{\mathrm{S}})$, it suffices to show that for each $\beta > 0$, there exists $\tau \geq 0$, $\delta > 0$ such that for every $Z \in \mathbb{I}(X \times Y)$ with $(\mathbf{x}^{\mathrm{S}}, \mathbf{y}^{\mathrm{S}}) \in Z$ and $w(Z) \leq \delta$,

$$\min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}(Z)} f(\mathbf{x}, \mathbf{y}) - \sup_{\boldsymbol{\mu} \geq \mathbf{0}} \min_{(\mathbf{x},\mathbf{y}) \in Z} \left[f(\mathbf{x}, \mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}} \mathbf{g}(\mathbf{x}, \mathbf{y}) \right] \leq \tau w(Z)^{\beta},$$

and the desired result follows by analogy to Lemma 6.3.8 by observing that the dual lower bounding scheme is at least first-order convergent at $(\mathbf{x}^{\mathrm{S}}, \mathbf{y}^{\mathrm{S}})$.

Let $g_j(\mathbf{x}^{\mathrm{S}}, \mathbf{y}^{\mathrm{S}}) = -\varepsilon_j < 0$, $\forall j \in \{1, \cdots, m_I\}$. Since g_j is continuous for each $j \in \{1, \cdots, m_I\}$, there exists $\delta_j > 0$, $\forall j \in \{1, \cdots, m_I\}$, such that $\left\| (\mathbf{x}, \mathbf{y}) - (\mathbf{x}^{\mathrm{S}}, \mathbf{y}^{\mathrm{S}}) \right\|_{\infty} < \delta_j$ implies $\left| g_j(\mathbf{x}, \mathbf{y}) - g_j(\mathbf{x}^{\mathrm{S}}, \mathbf{y}^{\mathrm{S}}) \right| < \frac{\varepsilon_j}{2}$ (see Lemma 2.2.2).

Define $\delta := \min\limits_{j \in \{1, \cdots, m_I\}} \delta_j$, and note that $\delta > 0$. Consider $Z \in \mathbb{I}(X \times Y)$ such that $(\mathbf{x}^{\mathrm{S}}, \mathbf{y}^{\mathrm{S}}) \in Z$ and $w(Z) \leq \delta$. For each $(\mathbf{x}, \mathbf{y}) \in Z$ and $j \in \{1, \cdots, m_I\}$ we have that $\left| g_j(\mathbf{x}, \mathbf{y}) - g_j(\mathbf{x}^{\mathrm{S}}, \mathbf{y}^{\mathrm{S}}) \right| < \frac{\varepsilon_j}{2}$. Therefore, for each $j \in \{1, \cdots, m_I\}$, $g_j(\mathbf{x}, \mathbf{y}) < -\frac{\varepsilon_j}{2} <$

0, $\forall(\mathbf{x},\mathbf{y}) \in Z$. Consequently,

$$\sup_{\boldsymbol{\mu}\geq\mathbf{0}} \min_{(\mathbf{x},\mathbf{y})\in Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y})\right] \geq \min_{(\mathbf{x},\mathbf{y})\in Z} f(\mathbf{x},\mathbf{y})$$
$$= \min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}(Z)} f(\mathbf{x},\mathbf{y})$$

since Problem (P) is effectively unconstrained over the small intervals Z around $(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})$, which implies $\tau = 0$ and $\delta = \min\limits_{j\in\{1,\cdots,m_I\}} \delta_j$ satisfy the requirements. $\square$

Remark 6.4.26. Proposition 6.4.25 as stated does not apply to Problem (P) with active constraints; however, it can be modified to demonstrate second-order convergence when Problem (P) contains active convex constraints (note that this includes affine equality constraints) if f is twice continuously differentiable, and strong duality holds for the scheme of relaxations of Problem (P) in which only the active (convex) constraints are included and f is replaced by its scheme of convex envelopes (see Remark 6.4.13). Proposition 6.4.25 can also be extended to demonstrate arbitrarily high convergence order of the dual lower bounding scheme on a neighborhood of $(\mathbf{x}^{\mathrm{S}},\mathbf{y}^{\mathrm{S}})$ in a manner similar to Corollary 6.4.17.

The next result shows that the dual lower bounding scheme is second-order convergent at KKT points (see Definition 2.3.18) when the functions f, $\mathbf{g}$, and $\mathbf{h}$ in Problem (P) are twice continuously differentiable.

Theorem 6.4.27. Consider Problem (P). Suppose $\mathrm{int}(X \times Y)$ is nonempty, and f, $\mathbf{g}$, and $\mathbf{h}$ are twice continuously differentiable on $\mathrm{int}(X \times Y)$. Furthermore, suppose there exists $(\mathbf{x}^*,\mathbf{y}^*) \in \mathrm{int}(X \times Y)$, $\boldsymbol{\mu}^* \in \mathbb{R}^{m_I}_+$, and $\boldsymbol{\lambda}^* \in \mathbb{R}^{m_E}$ such that $(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)$ is a KKT point for Problem (P). The dual lower bounding scheme is at least second-order convergent at $(\mathbf{x}^*,\mathbf{y}^*)$.

Proof. Let $L(\mathbf{x},\mathbf{y},\boldsymbol{\mu},\boldsymbol{\lambda}) := f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y})$ denote the Lagrangian of Problem (P). Since we are concerned about the convergence order at the feasible point $(\mathbf{x}^*,\mathbf{y}^*)$, it suffices to show the existence of $\tau \geq 0$ such that for every $Z \in \mathbb{I}(X \times Y)$ with $(\mathbf{x}^*,\mathbf{y}^*) \in Z$,

$$\min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}(Z)} f(\mathbf{x},\mathbf{y}) - \sup_{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}} \min_{(\mathbf{x},\mathbf{y})\in Z} L(\mathbf{x},\mathbf{y},\boldsymbol{\mu},\boldsymbol{\lambda}) \leq \tau w(Z)^2.$$

We have

$$\begin{aligned}
\sup_{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}} \min_{(\mathbf{x},\mathbf{y})\in Z} L(\mathbf{x},\mathbf{y},\boldsymbol{\mu},\boldsymbol{\lambda}) &\geq \min_{(\mathbf{x},\mathbf{y})\in Z} L(\mathbf{x},\mathbf{y},\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*) \\
&= \min_{(\mathbf{x},\mathbf{y})\in Z} \Big[L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*) + \nabla_{\mathbf{x}} L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}^*) \\
&\qquad\qquad + \nabla_{\mathbf{y}} L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)^{\mathrm{T}}(\mathbf{y}-\mathbf{y}^*) - O(w(Z)^2)\Big] \\
&= \min_{(\mathbf{x},\mathbf{y})\in Z} \big[f(\mathbf{x}^*,\mathbf{y}^*) - O(w(Z)^2)\big] \\
&\geq f(\mathbf{x}^*,\mathbf{y}^*) - O(w(Z)^2).
\end{aligned}$$

Note that Step 3 above uses $L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*) = f(\mathbf{x}^*,\mathbf{y}^*)$, $\nabla_{\mathbf{x}} L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*) = \mathbf{0}$, and $\nabla_{\mathbf{y}} L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*) = \mathbf{0}$ by virtue of the assumption that $(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)$ is a KKT point for Problem (P). Therefore,

$$\min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}(Z)} f(\mathbf{x},\mathbf{y}) - \sup_{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}} \min_{(\mathbf{x},\mathbf{y})\in Z} L(\mathbf{x},\mathbf{y},\boldsymbol{\mu},\boldsymbol{\lambda}) \leq O(w(Z)^2),$$

which establishes the existence of τ for all $Z \in \mathbb{I}(X \times Y)$ with $(\mathbf{x}^*,\mathbf{y}^*) \in Z$ by analogy to Lemma 6.3.8 since the dual lower bounding scheme is at least first-order convergent at $(\mathbf{x}^*,\mathbf{y}^*)$. $\square$

A corollary of Theorems 6.4.19 and 6.4.27 is that second-order convergence at KKT points is guaranteed for convex relaxation-based lower bounding schemes in which second-order pointwise convergent schemes of relaxations are used.

Corollary 6.4.28. Consider Problem (P). Suppose $\mathrm{int}(X \times Y)$ is nonempty and f, $\mathbf{g}$, and $\mathbf{h}$ are twice continuously differentiable on $\mathrm{int}(X \times Y)$. Furthermore, suppose there exists $(\mathbf{x}^*,\mathbf{y}^*) \in \mathrm{int}(X \times Y)$, $\boldsymbol{\mu}^* \in \mathbb{R}^{m_I}_+$, and $\boldsymbol{\lambda}^* \in \mathbb{R}^{m_E}$ such that $(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)$ is a KKT point for Problem (P). Let $(f^{\mathrm{cv}}_Z)_{Z\in\mathbb{I}(X\times Y)}$, $(g^{\mathrm{cv}}_{j,Z})_{Z\in\mathbb{I}(X\times Y)}$, $j = 1,\cdots,m_I$, denote continuous schemes of convex relaxations of f, $g_1,\cdots,g_{m_I}$, respectively, in $X \times Y$ with pointwise convergence orders $\gamma^{\mathrm{cv}}_f \geq 2$, $\gamma^{\mathrm{cv}}_{g,1} \geq 2,\cdots,\gamma^{\mathrm{cv}}_{g,m_I} \geq 2$ and corresponding constants τ^{cv}_f, $\tau^{\mathrm{cv}}_{g,1},\cdots,\tau^{\mathrm{cv}}_{g,m_I}$. Let $(h^{\mathrm{cv}}_{k,Z}, h^{\mathrm{cc}}_{k,Z})_{Z\in\mathbb{I}(X\times Y)}$, $k = 1,\cdots,m_E$, denote continuous schemes of relaxations of $h_1,\cdots,h_{m_E}$, respectively, in $X \times Y$ with pointwise convergence orders $\gamma_{h,1} \geq 2,\cdots,\gamma_{h,m_E} \geq 2$ and corresponding constants $\tau_{h,1},\cdots,\tau_{h,m_E}$. Then, the resulting scheme of convex relaxation-based lower bounding problems for Problem (P) is at least second-order convergent at $(\mathbf{x}^*, \mathbf{y}^*)$.

Proof. The result holds as a consequence of Theorems 6.4.19 and 6.4.27, by using $\boldsymbol{\mu}_Z = \boldsymbol{\mu}^*$, $\boldsymbol{\lambda}_Z = \boldsymbol{\lambda}^*$, $\bar{\mu} = \|\boldsymbol{\mu}^*\|_\infty$, $\bar{\lambda} = \|\boldsymbol{\lambda}^*\|_\infty$ in Theorem 6.4.19. $\square$

The following example shows that the convergence order may be as low as two under the assumptions of Theorem 6.4.27.

Example 6.4.29. Let $X = [-2, 2], Y = [0, 3]$, $m_I = 1$, and $m_E = 1$ with $f(x, y) = x + y$, $g(x, y) = -y^2 + y + 2$, and $h(x, y) = x$. Consider intervals $[0, 0] \times [2 - \varepsilon, 2 + \varepsilon] =: Z \in \mathbb{I}(X \times Y)$ with $0 < \varepsilon \leq 1$. Note that $w(Z) = 2\varepsilon$, and that $\left(0, 2, \frac{1}{3}, -1\right)$ is a KKT point for Problem (P). The optimal objective value of Problem (P) on Z is 2, while the optimal objective value of the Lagrangian dual-based lower bounding problem on Z can be derived as

$$
\begin{aligned}
\mathcal{O}(Z) &= \sup_{\mu \geq 0, \lambda} \min_{(x,y) \in Z} x + y + \mu\left(-y^2 + y + 2\right) + \lambda x \\
&= \sup_{\mu \geq 0} \min\left\{(2 - \varepsilon) + \mu\left(2 + (2 - \varepsilon) - (2 - \varepsilon)^2\right), (2 + \varepsilon) + \mu\left(2 + (2 + \varepsilon) - (2 + \varepsilon)^2\right)\right\} \\
&= \sup_{\mu \geq 0} \min\left\{(2 - \varepsilon) + \mu\left(3\varepsilon - \varepsilon^2\right), (2 + \varepsilon) + \mu\left(-3\varepsilon - \varepsilon^2\right)\right\} \\
&= (2 - \varepsilon) + \frac{1}{3}\left(3\varepsilon - \varepsilon^2\right) \\
&= 2 - \frac{\varepsilon^2}{3},
\end{aligned}
$$

where Step 2 follows from the fact that the minimum of a concave function on an interval is attained at one of its endpoints, and the value of μ in Step 4 is obtained by equating the two arguments of the inner min function in Step 3. Convergence of the dual lower bounding scheme at the point $(0, 2)$ is, therefore, at most second-order.

Finally, we show that the dual lower bounding scheme is (at least) first-order convergent even when the dual problem is not solved to optimality.

Theorem 6.4.30. Consider Problem (P). Suppose f, g_j, $j = 1, \cdots, m_I$, and h_k, $k = 1, \cdots, m_E$, are Lipschitz continuous on $X \times Y$ with Lipschitz constants M_f, $M_{g,1}, \cdots, M_{g,m_I}$, $M_{h,1}, \cdots, M_{h,m_E}$, respectively. Furthermore, suppose the dual lower bounding scheme involves at most $n_d \geq 1$ iterations of an algorithm applied to the dual at each node of the branch-and-bound tree. In addition, suppose the branch-and-bound algorithm uses first-order (Hausdorff) convergent schemes of constant relaxations $\left(g^{\mathrm{L}}_{j,Z}, g^{\mathrm{U}}_{j,Z}\right)_{Z \in \mathbb{I}(X \times Y)}$,

$j = 1, \cdots, m_I$, on $X \times Y$ to overestimate $\left(\overline{g}_j(Z)\right)_{Z \in \mathbb{I}(X \times Y)}$ and first-order (Hausdorff) convergent schemes of constant relaxations $\left(h^{\mathrm{L}}_{k,Z}, h^{\mathrm{U}}_{k,Z}\right)_{Z \in \mathbb{I}(X \times Y)}$, $k = 1, \cdots, m_E$, on $X \times Y$ to overestimate $\left(\overline{h}_k(Z)\right)_{Z \in \mathbb{I}(X \times Y)}$ (such schemes of constant relaxations can be obtained, for example, using interval arithmetic [172]), sets $\mu_j = 0$ at each iteration of the algorithm applied to the dual on Z if $g^{\mathrm{U}}_{j,Z} < 0$ (i.e., when inequality constraint j is determined to be inactive on Z by $g^{\mathrm{U}}_{j,Z}$), and determines the dual lower bounding problem on Z to be infeasible either when $g^{\mathrm{L}}_{j,Z} > 0$ for any $j \in \{1, \cdots, m_I\}$ (i.e., when inequality constraint j is determined to be unsatisfiable on Z by $g^{\mathrm{L}}_{j,Z}$), or when $0 \notin \left[h^{\mathrm{L}}_{k,Z}, h^{\mathrm{U}}_{k,Z}\right]$ for any $k \in \{1, \cdots, m_E\}$ (i.e., when equality constraint k is determined to be unsatisfiable on Z by $(h^{\mathrm{L}}_{k,Z}, h^{\mathrm{U}}_{k,Z})$). Assume that the absolute values of the schemes of dual variables generated by the dual lower bounding scheme are bounded from above by M_∞. Then the dual lower bounding scheme is at least first-order convergent on $X \times Y$.

Proof. From the assumption that the schemes $\left(g^{\mathrm{L}}_{j,Z}, g^{\mathrm{U}}_{j,Z}\right)_{Z \in \mathbb{I}(X \times Y)}$, $j = 1, \cdots, m_I$, and $\left(h^{\mathrm{L}}_{k,Z}, h^{\mathrm{U}}_{k,Z}\right)_{Z \in \mathbb{I}(X \times Y)}$, $k = 1, \cdots, m_E$, are first-order convergent on $X \times Y$, the determination of infeasibility of the dual lower bounding problem on Z if $g^{\mathrm{L}}_{j,Z} > 0$ for any $j \in \{1, \cdots, m_I\}$, or if $0 \notin \left[h^{\mathrm{L}}_{k,Z}, h^{\mathrm{U}}_{k,Z}\right]$ for any $k \in \{1, \cdots, m_E\}$, Proposition 1 in [39], and Lemma 6.4.1, we conclude that the dual lower bounding scheme has at least first-order convergence at infeasible points (although the dual lower bounding scheme detects infeasibility of infeasible points in $X \times Y$ at least as quickly as any convex relaxation-based lower bounding scheme (see Lemma 6.4.22), we assume that the schemes $\left(g^{\mathrm{L}}_{j,Z}, g^{\mathrm{U}}_{j,Z}\right)_{Z \in \mathbb{I}(X \times Y)}$, $j = 1, \cdots, m_I$, and $\left(h^{\mathrm{L}}_{k,Z}, h^{\mathrm{U}}_{k,Z}\right)_{Z \in \mathbb{I}(X \times Y)}$, $k = 1, \cdots, m_E$, are available to detect infeasibility since we are only allowed to use at most n_d iterations of an algorithm applied to the dual).

Next, suppose $\mathcal{F}(X \times Y) \neq \emptyset$ and $Z \in \mathbb{I}(X \times Y)$ with $Z \cap \mathcal{F}(X \times Y) \neq \emptyset$. Let J_Z denote the set of inequality constraints that are potentially active at some point in Z as determined by $\left(g^{\mathrm{L}}_{j,Z}, g^{\mathrm{U}}_{j,Z}\right)$, i.e., $J_Z := \left\{j \in \{1, \cdots, m_I\} : g^{\mathrm{U}}_{j,Z} \geq 0\right\}$. Let $(\bar{\boldsymbol{\mu}}_Z, \bar{\boldsymbol{\lambda}}_Z) \in \mathbb{R}^{m_I}_{+} \times \mathbb{R}^{m_E}$ denote the pair of dual variables corresponding to the dual lower bound on Z after at most n_d iterations of an algorithm applied to the dual with $\bar{\mu}_{j,Z} = 0$, $\forall j \in \{1, \cdots, m_I\} \backslash J_Z$, and let $(\bar{\mathbf{x}}_Z, \bar{\mathbf{y}}_Z) \in \underset{(\mathbf{x},\mathbf{y}) \in Z}{\arg\min}\, f(\mathbf{x}, \mathbf{y}) + \bar{\boldsymbol{\mu}}^{\mathrm{T}}_Z \mathbf{g}(\mathbf{x}, \mathbf{y}) + \bar{\boldsymbol{\lambda}}^{\mathrm{T}}_Z \mathbf{h}(\mathbf{x}, \mathbf{y})$. Note that the condition $\bar{\mu}_{j,Z} = 0$, $\forall j \in \{1, \cdots, m_I\} \backslash J_Z$ can be guaranteed by a suitable initialization of the dual variables and by suitably modifying the dual variables generated by the algorithm applied to the dual (this modification of the dual lower bounding scheme is once again necessitated by the assumption that at most n_d iterations of an algorithm applied to the dual are used). For each $j \in J_Z$, we have

$$g^{\mathrm{U}}_{j,Z} - \max_{(\mathbf{x},\mathbf{y})\in Z} g_j(\mathbf{x},\mathbf{y}) \leq \tau_j w(Z)$$

for some constant $\tau_j \geq 0$ by virtue of the fact that the scheme of constant concave relaxations $\left(g^{\mathrm{U}}_{j,Z}\right)_{Z\in\mathbb{I}(X\times Y)}$ has first-order convergence on $X \times Y$. Since $g^{\mathrm{U}}_{j,Z} \geq 0$, $\forall j \in J_Z$, and g_j is Lipschitz continuous on $X \times Y$, this implies

$$g_j(\mathbf{x},\mathbf{y}) \geq -\left(\tau_j + M_{g,j}\sqrt{n_x+n_y}\right) w(Z), \quad \forall (\mathbf{x},\mathbf{y}) \in Z,\ \forall j \in J_Z.$$

Let $(\mathbf{x}^*_Z, \mathbf{y}^*_Z) \in \underset{(\mathbf{x},\mathbf{y})\in\mathcal{F}(Z)}{\arg\min} f(\mathbf{x},\mathbf{y})$. We have

$$\begin{aligned}
&\min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}(Z)} f(\mathbf{x},\mathbf{y}) - \min_{(\mathbf{x},\mathbf{y})\in Z}\left[f(\mathbf{x},\mathbf{y}) + \bar{\boldsymbol{\mu}}^{\mathrm{T}}_Z\mathbf{g}(\mathbf{x},\mathbf{y}) + \bar{\boldsymbol{\lambda}}^{\mathrm{T}}_Z\mathbf{h}(\mathbf{x},\mathbf{y})\right]\\
&= f(\mathbf{x}^*_Z,\mathbf{y}^*_Z) - \left[f(\bar{\mathbf{x}}_Z,\bar{\mathbf{y}}_Z) + \bar{\boldsymbol{\mu}}^{\mathrm{T}}_Z\mathbf{g}(\bar{\mathbf{x}}_Z,\bar{\mathbf{y}}_Z) + \bar{\boldsymbol{\lambda}}^{\mathrm{T}}_Z\mathbf{h}(\bar{\mathbf{x}}_Z,\bar{\mathbf{y}}_Z)\right]\\
&= \left(f(\mathbf{x}^*_Z,\mathbf{y}^*_Z) - f(\bar{\mathbf{x}}_Z,\bar{\mathbf{y}}_Z)\right) - \sum_{j\in J_Z}\bar{\mu}_{j,Z}\, g_j(\bar{\mathbf{x}}_Z,\bar{\mathbf{y}}_Z) + \bar{\boldsymbol{\lambda}}^{\mathrm{T}}_Z\left(\mathbf{h}(\mathbf{x}^*_Z,\mathbf{y}^*_Z) - \mathbf{h}(\bar{\mathbf{x}}_Z,\bar{\mathbf{y}}_Z)\right)\\
&\leq M_f\|(\mathbf{x}^*_Z,\mathbf{y}^*_Z) - (\bar{\mathbf{x}}_Z,\bar{\mathbf{y}}_Z)\| + \sum_{j\in J_Z}\bar{\mu}_{j,Z}\left(\tau_j + M_{g,j}\sqrt{n_x+n_y}\right)w(Z) +\\
&\qquad \sum_{k=1}^{m_E}|\bar{\lambda}_{k,Z}|M_{h,k}\|(\mathbf{x}^*_Z,\mathbf{y}^*_Z) - (\bar{\mathbf{x}}_Z,\bar{\mathbf{y}}_Z)\|\\
&\leq \left(M_f\sqrt{n_x+n_y} + \sum_{j\in J_Z}M_\infty\left(\tau_j + M_{g,j}\sqrt{n_x+n_y}\right) + \sum_{k=1}^{m_E}M_\infty M_{h,k}\sqrt{n_x+n_y}\right)w(Z)\\
&\leq \left(M_f\sqrt{n_x+n_y} + \sum_{j=1}^{m_I}M_\infty\left(\tau_j + M_{g,j}\sqrt{n_x+n_y}\right) + \sum_{k=1}^{m_E}M_\infty M_{h,k}\sqrt{n_x+n_y}\right)w(Z),
\end{aligned}$$

which establishes the desired result. □

6.5 Reduced-space branch-and-bound algorithms

In this section, we present some results on the convergence orders of some widely-applicable reduced-space lower bounding schemes in the literature [69, 76] for Problem (P) when only the set Y may be partitioned during the course of the algorithm. This section is divided into two parts. First, we consider a convex relaxation-based reduced-space lower bounding scheme for a subclass of Problem (P) [76] and investigate its convergence order. Next, we look at the convergence order of a duality-based reduced-space lower bounding scheme [69, Section 3.3] for Problem (P).

Algorithm 2.2 in Chapter 2 outlined a generic reduced-space branch-and-bound algorithm for Problem (P). The reader is directed to references [76] and [69] for two widely-applicable instances of Algorithm 2.2, and for examples of their application. In the remainder of this section, we investigate the convergence orders of the reduced-space lower bounding schemes described in [76] and [69].

6.5.1 Convex relaxation-based branch-and-bound for problems with special structure

Epperly and Pistikopoulos [76] proposed a reduced-space branch-and-bound algorithm for Problem (P) when $m_E = 0$ (note that this condition can be relaxed as detailed below), and the functions f and g_j, $\forall j \in \{1, \cdots, m_I\}$, in Problem (P) are each of the form

$$w(\mathbf{x}, \mathbf{y}) = w^A(\mathbf{x}) + \sum_{i \in Q} w_i^B(\mathbf{x}) w_i^C(\mathbf{y}) + w^D(\mathbf{y}), \tag{W}$$

where Q is a finite set of indices, and the functions $w^A, \mathbf{w}^B, \mathbf{w}^C$, and w^D satisfy:

1. w^A and $\mathbf{w}^B$ are convex on X.

2. $\mathbf{w}^C$ and w^D are continuous on Y.

3. Strongly consistent convex and concave relaxations are available for $\mathbf{w}^C$ and w^D on Y.

4. $\mathbf{w}^B$ and $\mathbf{w}^C$ have continuous tight bounds.

5. For each $i \in Q$, at least one of the following two conditions must hold:

   a. $w_i^B(\mathbf{x}) = \mathbf{c}_i^{\mathrm{T}} \mathbf{x}$ for some constant $\mathbf{c}_i \in \mathbb{R}^{n_x}$,

   b. $w_i^C(\mathbf{y}) \geq 0$ for all $\mathbf{y} \in Y$.

Epperly and Pistikopoulos [76] state that equality constraints can be equivalently reformulated using pairs of inequalities; however, the above assumptions restrict the functional forms of the equality constraints h_k, $k = 1, \cdots, m_E$, to

$$h_k(\mathbf{x}, \mathbf{y}) = \sum_{i \in Q} \left(\mathbf{c}_i^{\mathrm{T}} \mathbf{x}\right) w_i^C(\mathbf{y}) + w^D(\mathbf{y}). \tag{W$_{\text{eq}}$}$$

Suppose for each $Z \in \mathbb{I}Y$, we associate an interval $X(Z)$ such that $\square X \supset X(Z) \supset \mathcal{F}_X(Z)$, where $\square X$ denotes the interval hull of X (note that we make the implicit assumption (see Remark 6.2.2) that X is an interval in this section). Assumption 3 can be restated as follows: there exists a continuous scheme, $(w_{i,Z}^{C,\text{cv}}, w_{i,Z}^{C,\text{cc}})_{Z \in \mathbb{I}Y}$, of relaxations of w_i^C, $i \in Q$, in Y with pointwise convergence order $\gamma_i^C > 0$, and there exists a continuous scheme of convex relaxations, $(w_Z^{D,\text{cv}})_{Z \in \mathbb{I}Y}$, of w^D in Y with pointwise convergence order $\gamma^{D,\text{cv}} > 0$. Assumption 4 can be replaced by the following: there exist schemes of constant relaxations $(w_{i,Z}^{B,\text{L}}, w_{i,Z}^{B,\text{U}})_{Z \in \mathbb{I}\square X}$ and $(w_{i,Z}^{C,\text{L}}, w_{i,Z}^{C,\text{U}})_{Z \in \mathbb{I}Y}$, $i \in Q$, of w_i^B and w_i^C in X and Y, respectively, with (Hausdorff) convergence orders $\beta_i^{B,\text{c}} > 0$ and $\beta_i^{C,\text{c}} > 0$. In addition, we assume that the range order of w_i^C, $\forall i \in Q$, is greater than zero on Y (cf. Lemma 6.5.1).

Under the above assumptions, Epperly and Pistikopoulos [76] show that underestimating each function $w(\mathbf{x}, \mathbf{y})$ of the form (W) using the scheme $(w_{X(Z)\times Z}^{\text{cv}})_{Z \in \mathbb{I}Y}$ of convex relaxations defined by

$$w_{X(Z)\times Z}^{\text{cv}}(\mathbf{x}, \mathbf{y}) = w^A(\mathbf{x}) + \sum_{i \in Q} w_{i,X(Z)\times Z}^{BC,\text{cv}}(\mathbf{x}, \mathbf{y}) + w_Z^{D,\text{cv}}(\mathbf{y}), \tag{W$^{\text{cv}}$}$$

where, for each $i \in Q$, the scheme of convex relaxations $(w_{i,X(Z)\times Z}^{BC,\text{cv}})_{Z \in \mathbb{I}Y}$ is obtained using McCormick's product rule [154] as

$$w_{i,X(Z)\times Z}^{BC,\text{cv}}(\mathbf{x}, \mathbf{y}) = \begin{cases} \max\begin{Bmatrix} w_{i,X(Z)}^{B,\text{U}} w_{i,Z}^{C,\text{cv}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\text{U}} - w_{i,X(Z)}^{B,\text{U}} w_{i,Z}^{C,\text{U}}, \\ w_{i,X(Z)}^{B,\text{L}} w_{i,Z}^{C,\text{cv}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\text{L}} - w_{i,X(Z)}^{B,\text{L}} w_{i,Z}^{C,\text{L}} \end{Bmatrix}, & \text{if } w_{i,X(Z)}^{B,\text{L}} \geq 0 \\ \max\begin{Bmatrix} w_{i,X(Z)}^{B,\text{U}} w_{i,Z}^{C,\text{cc}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\text{U}} - w_{i,X(Z)}^{B,\text{U}} w_{i,Z}^{C,\text{U}}, \\ w_{i,X(Z)}^{B,\text{L}} w_{i,Z}^{C,\text{cc}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\text{L}} - w_{i,X(Z)}^{B,\text{L}} w_{i,Z}^{C,\text{L}} \end{Bmatrix}, & \text{if } w_{i,X(Z)}^{B,\text{U}} < 0 \\ \max\begin{Bmatrix} w_{i,X(Z)}^{B,\text{U}} w_{i,Z}^{C,\text{cv}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\text{U}} - w_{i,X(Z)}^{B,\text{U}} w_{i,Z}^{C,\text{U}}, \\ w_{i,X(Z)}^{B,\text{L}} w_{i,Z}^{C,\text{cc}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\text{L}} - w_{i,X(Z)}^{B,\text{L}} w_{i,Z}^{C,\text{L}} \end{Bmatrix}, & \text{otherwise} \end{cases},$$

yields a convergent reduced-space lower bounding scheme with any accumulation point of the sequence of lower bounding solutions solving Problem (P) when the subdivision process is exhaustive on Y and the selection procedure is bound improving.

Before we investigate the convergence order of the reduced-space lower bounding scheme in [76], we look at the propagation of the convergence orders of the relaxation schemes $(w_{i,Z}^{B,\mathrm{L}}, w_{i,Z}^{B,\mathrm{U}})_{Z\in\mathbb{I}\square X}$, $(w_{i,Z}^{C,\mathrm{cv}}, w_{i,Z}^{C,\mathrm{cc}})_{Z\in\mathbb{I}Y}$, $(w_{i,Z}^{C,\mathrm{L}}, w_{i,Z}^{C,\mathrm{U}})_{Z\in\mathbb{I}Y}$, $\forall i \in Q$, and $(w_{Z}^{D,\mathrm{cv}})_{Z\in\mathbb{I}Y}$ to the convergence order of the reduced-space scheme of convex relaxations $(w_{X(Z)\times Z}^{\mathrm{cv}})_{Z\in\mathbb{I}Y}$. Note that unless otherwise specified, we simply use $X(Z) = \square X (= X)$, $\forall Z \in \mathbb{I}Y$. The following result provides sufficient conditions for the scheme of convex relaxations defined by (W^{cv}) to have pointwise convergence of a given order on Y.

Lemma 6.5.1. Let $X \subset \mathbb{R}^{n_x}$, $Y \subset \mathbb{R}^{n_y}$ be nonempty compact convex sets and $f : X \times Y \to \mathbb{R}$ be a function of the form (W) such that

$$f : X \times Y \ni (\mathbf{x}, \mathbf{y}) \longmapsto w^A(\mathbf{x}) + \sum_{i\in Q} w_i^B(\mathbf{x}) w_i^C(\mathbf{y}) + w^D(\mathbf{y}).$$

Assume that w^A, w_i^B, $\forall i \in Q$, and w^D are continuous, and for each $i \in Q$, w_i^C has range of order $\alpha_i^C \geq 1$ on Y with corresponding constant $\tau_i^{C,\mathrm{r}}$. Let $(w_{i,Z}^{C,\mathrm{cv}}, w_{i,Z}^{C,\mathrm{cc}})_{Z\in\mathbb{I}Y}$ and $(w_{Z}^{D,\mathrm{cv}})_{Z\in\mathbb{I}Y}$ respectively denote continuous schemes of relaxations of w_i^C, $i \in Q$, and w^D in Y with pointwise convergence orders $\gamma_i^C \geq 1$ and $\gamma^{D,\mathrm{cv}} \geq 1$ and corresponding constants τ_i^C and $\tau^{D,\mathrm{cv}}$. Let $(w_{i,Z}^{B,\mathrm{L}}, w_{i,Z}^{B,\mathrm{U}})_{Z\in\mathbb{I}\square X}$ and $(w_{i,Z}^{C,\mathrm{L}}, w_{i,Z}^{C,\mathrm{U}})_{Z\in\mathbb{I}Y}$ respectively denote schemes of constant relaxations of w_i^B in $\square X$ and w_i^C in Y, $\forall i \in Q$, with (Hausdorff) convergence orders $\beta_i^{B,\mathrm{c}} > 0$ and $\beta_i^{C,\mathrm{c}} \geq 1$ and corresponding constants $\tau_i^{B,\mathrm{c}}$ and $\tau_i^{C,\mathrm{c}}$. Then the continuous scheme of convex relaxations $(f_{X(Z)\times Z}^{\mathrm{cv}})_{Z\in\mathbb{I}Y}$ of the form (W^{cv}) defined by

$$f_{X(Z)\times Z}^{\mathrm{cv}}(\mathbf{x}, \mathbf{y}) := w^A(\mathbf{x}) + \sum_{i\in Q} w_{i,X(Z)\times Z}^{BC,\mathrm{cv}}(\mathbf{x}, \mathbf{y}) + w_Z^{D,\mathrm{cv}}(\mathbf{y}), \quad \forall (\mathbf{x}, \mathbf{y}) \in X(Z) \times Z,$$

has pointwise convergence of order at least $\min\left\{\min_{i\in Q}\left\{\ \min\left\{\alpha_i^C, \beta_i^{C,\mathrm{c}}, \gamma_i^C\right\}\right\}, \gamma^{D,\mathrm{cv}}\right\}$ on Y.

Proof. From Equation (W^{cv}), we have for each $(\mathbf{x},\mathbf{y}) \in X(Z)\times Z$:

$$
\begin{aligned}
&f(\mathbf{x},\mathbf{y}) - f^{\mathrm{cv}}_{X(Z)\times Z}(\mathbf{x},\mathbf{y}) \\
&= \left(w^A(\mathbf{x}) + \sum_{i\in Q} w_i^B(\mathbf{x})w_i^C(\mathbf{y}) + w^D(\mathbf{y})\right) - \left(w^A(\mathbf{x}) + \sum_{i\in Q} w^{BC,\mathrm{cv}}_{i,X(Z)\times Z}(\mathbf{x},\mathbf{y}) + w_Z^{D,\mathrm{cv}}(\mathbf{y})\right) \\
&= \sum_{i\in Q}\left(w_i^B(\mathbf{x})w_i^C(\mathbf{y}) - w^{BC,\mathrm{cv}}_{i,X(Z)\times Z}(\mathbf{x},\mathbf{y})\right) + \left(w^D(\mathbf{y}) - w_Z^{D,\mathrm{cv}}(\mathbf{y})\right).
\end{aligned}
$$

Depending on whether $w^{B,\mathrm{L}}_{i,X(Z)} \geq 0$, $w^{B,\mathrm{U}}_{i,X(Z)} < 0$, or $0 \in \left(w^{B,\mathrm{L}}_{i,X(Z)}, w^{B,\mathrm{U}}_{i,X(Z)}\right]$ for each $i \in Q$, we have that $\left(w_i^B(\mathbf{x})w_i^C(\mathbf{y}) - w^{BC,\mathrm{cv}}_{i,X(Z)\times Z}(\mathbf{x},\mathbf{y})\right)$ is bounded from above either by

$$
\left[w_i^B(\mathbf{x})w_i^C(\mathbf{y}) - \left(w^{B,\mathrm{U}}_{i,X(Z)}w^{C,\mathrm{cv}}_{i,Z}(\mathbf{y}) + w_i^B(\mathbf{x})w^{C,\mathrm{U}}_{i,Z} - w^{B,\mathrm{U}}_{i,X(Z)}w^{C,\mathrm{U}}_{i,Z}\right)\right],
$$

or by

$$
\left[w_i^B(\mathbf{x})w_i^C(\mathbf{y}) - \left(w^{B,\mathrm{U}}_{i,X(Z)}w^{C,\mathrm{cc}}_{i,Z}(\mathbf{y}) + w_i^B(\mathbf{x})w^{C,\mathrm{U}}_{i,Z} - w^{B,\mathrm{U}}_{i,X(Z)}w^{C,\mathrm{U}}_{i,Z}\right)\right]
$$

for each $(\mathbf{x},\mathbf{y}) \in X(Z)\times Z$. Consequently, it is sufficient to show the existence of constants $\tau_1, \tau_2 \geq 0$ such that

$$
\begin{aligned}
\max_{(\mathbf{x},\mathbf{y})\in X(Z)\times Z} \Bigg| &\left(w^A(\mathbf{x}) + \sum_{i\in Q} w_i^B(\mathbf{x})w_i^C(\mathbf{y}) + w^D(\mathbf{y})\right) - \\
&\left(w^A(\mathbf{x}) + \sum_{i\in Q}\left(w^{B,\mathrm{U}}_{i,X(Z)}w^{C,\mathrm{cv}}_{i,Z}(\mathbf{y}) + w_i^B(\mathbf{x})w^{C,\mathrm{U}}_{i,Z} - w^{B,\mathrm{U}}_{i,X(Z)}w^{C,\mathrm{U}}_{i,Z}\right) + w_Z^{D,\mathrm{cv}}(\mathbf{y})\right)\Bigg| \leq \tau_1 w(Z)^\gamma
\end{aligned}
$$

and

$$
\begin{aligned}
\max_{(\mathbf{x},\mathbf{y})\in X(Z)\times Z} \Bigg| &\left(w^A(\mathbf{x}) + \sum_{i\in Q} w_i^B(\mathbf{x})w_i^C(\mathbf{y}) + w^D(\mathbf{y})\right) - \\
&\left(w^A(\mathbf{x}) + \sum_{i\in Q}\left(w^{B,\mathrm{U}}_{i,X(Z)}w^{C,\mathrm{cc}}_{i,Z}(\mathbf{y}) + w_i^B(\mathbf{x})w^{C,\mathrm{U}}_{i,Z} - w^{B,\mathrm{U}}_{i,X(Z)}w^{C,\mathrm{U}}_{i,Z}\right) + w_Z^{D,\mathrm{cv}}(\mathbf{y})\right)\Bigg| \leq \tau_2 w(Z)^\gamma,
\end{aligned}
$$

where $\gamma := \min\left\{\min_{i\in Q}\left\{\min\left\{\alpha_i^C, \beta_i^{C,\mathrm{c}}, \gamma_i^C\right\}\right\}, \gamma^{D,\mathrm{cv}}\right\}$, to prove that $(f^{\mathrm{cv}}_{X(Z)\times Z})_{Z\in\mathbb{I}Y}$ converges pointwise to f with order γ on Y. The ensuing arguments prove the existence of τ_1; the existence of τ_2 can be proven analogously.

We have $\forall(\mathbf{x},\mathbf{y}) \in X(Z) \times Z$:

$$\begin{aligned}
&\Bigg(\bigg(w^A(\mathbf{x}) + \sum_{i\in Q} w_i^B(\mathbf{x}) w_i^C(\mathbf{y}) + w^D(\mathbf{y})\bigg) - \\
&\qquad \bigg(w^A(\mathbf{x}) + \sum_{i\in Q}\Big(w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{cv}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\mathrm{U}} - w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{U}}\Big) + w_Z^{D,\mathrm{cv}}(\mathbf{y})\bigg)\Bigg) \\
&= \left(\sum_{i\in Q} w_i^B(\mathbf{x}) w_i^C(\mathbf{y}) - \Big(w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{cv}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\mathrm{U}} - w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{U}}\Big)\right) + \\
&\qquad \Big(w^D(\mathbf{y}) - w_Z^{D,\mathrm{cv}}(\mathbf{y})\Big). \qquad (6.7)
\end{aligned}$$

Note that $\left|w_i^C(\mathbf{y}) - w_{i,Z}^{C,\mathrm{U}}\right|$ can be bounded from above as

$$\begin{aligned}
&\left|w_i^C(\mathbf{y}) - w_{i,Z}^{C,\mathrm{U}}\right| \\
=&\left|\Big(w_i^C(\mathbf{y}) - \max_{\mathbf{y}\in Z}\ w_i^C(\mathbf{y})\Big) + \Big(\max_{\mathbf{y}\in Z}\ w_i^C(\mathbf{y}) - w_{i,Z}^{C,\mathrm{U}}\Big)\right| \\
\leq &\left|w_i^C(\mathbf{y}) - \max_{\mathbf{y}\in Z}\ w_i^C(\mathbf{y})\right| + \left|\max_{\mathbf{y}\in Z}\ w_i^C(\mathbf{y}) - w_{i,Z}^{C,\mathrm{U}}\right| \\
\leq &\ \Big(\tau_i^{C,\mathrm{r}} w(Z)^{\alpha_i^C - \min\left\{\alpha_i^C,\beta_i^{C,\mathrm{c}}\right\}} + \tau_i^{C,\mathrm{c}} w(Z)^{\beta_i^{C,\mathrm{c}} - \min\left\{\alpha_i^C,\beta_i^{C,\mathrm{c}}\right\}}\Big) w(Z)^{\min\left\{\alpha_i^C,\beta_i^{C,\mathrm{c}}\right\}} \\
\leq &\, M_i^C w(Z)^{\beta_i^{C,\mathrm{r}}}, \quad \forall \mathbf{y} \in Z,
\end{aligned}$$

with $M_i^C := \tau_i^{C,\mathrm{r}} w(Y)^{\alpha_i^C - \beta_i^{C,\mathrm{r}}} + \tau_i^{C,\mathrm{c}} w(Y)^{\beta_i^{C,\mathrm{c}} - \beta_i^{C,\mathrm{r}}}$ and $\beta_i^{C,\mathrm{r}} := \min\left\{\alpha_i^C, \beta_i^{C,\mathrm{c}}\right\}$.

The first term in Equation (6.7) can be bounded as

$$\begin{aligned}
&\sum_{i\in Q}\Big(w_i^B(\mathbf{x}) w_i^C(\mathbf{y}) - \Big(w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{cv}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\mathrm{U}} - w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{U}}\Big)\Big) \\
=&\sum_{i\in Q}\Big[\Big(w_i^B(\mathbf{x}) - w_{i,X(Z)}^{B,\mathrm{U}}\Big)\Big(w_i^C(\mathbf{y}) - w_{i,Z}^{C,\mathrm{U}}\Big) + w_{i,X(Z)}^{B,\mathrm{U}}\Big(w_i^C(\mathbf{y}) - w_{i,Z}^{C,\mathrm{cv}}(\mathbf{y})\Big)\Big] \\
\leq&\sum_{i\in Q}\left|w_i^B(\mathbf{x}) - w_{i,X(Z)}^{B,\mathrm{U}}\right|\left|w_i^C(\mathbf{y}) - w_{i,Z}^{C,\mathrm{U}}\right| + \left|w_{i,X(Z)}^{B,\mathrm{U}}\Big(w_i^C(\mathbf{y}) - w_{i,Z}^{C,\mathrm{cv}}(\mathbf{y})\Big)\right| \\
\leq&\sum_{i\in Q} M_i^{BC} w(Z)^{\gamma_i^{BC}} \\
\leq& M^{BC} w(Z)^{\gamma^{BC}}, \quad \forall(\mathbf{x},\mathbf{y}) \in X(Z) \times Z, \qquad (6.8)
\end{aligned}$$

where the constants M^{BC}, γ^{BC}, and M_i^{BC}, γ_i^{BC}, $\forall i \in Q$, can be computed as

$$M^{BC} := \sum_{i \in Q} M_i^{BC} w(Y)^{\gamma_i^{BC} - \gamma^{BC}}, \quad \gamma^{BC} := \min_{i \in Q} \gamma_i^{BC}, \quad \gamma_i^{BC} := \min\left\{\beta_i^{C,\mathrm{r}}, \gamma_i^{C}\right\},$$
$$M_i^{BC} := \left[M_i^{B,1} M_i^{C} w(Y)^{\beta_i^{C,\mathrm{r}} - \gamma_i^{BC}} + M_i^{B,2} \tau_i^{C} w(Y)^{\gamma_i^{C} - \gamma_i^{BC}}\right],$$
$$M_i^{B,1} := \max_{\mathbf{x} \in X} w_i^B(\mathbf{x}) - \min_{\mathbf{x} \in X} w_i^B(\mathbf{x}) + \tau_i^{B,\mathrm{c}} w(X)^{\beta_i^{B,\mathrm{c}}}, \quad M_i^{B,2} := \max_{\mathbf{x} \in X} w_i^B(\mathbf{x}) + \tau_i^{B,\mathrm{c}} w(X)^{\beta_i^{B,\mathrm{c}}}.$$

The second term in Equation (6.7) is simply bounded as

$$w^D(\mathbf{y}) - w_Z^{D,\mathrm{cv}}(\mathbf{y}) \leq \tau^{D,\mathrm{cv}} w(Z)^{\gamma^{D,\mathrm{cv}}}, \quad \forall \mathbf{y} \in Z. \tag{6.9}$$

From Equations (6.8) and (6.9), we have

$$\begin{aligned}
&\max_{(\mathbf{x},\mathbf{y}) \in X(Z) \times Z} \left| \left(w^A(\mathbf{x}) + \sum_{i \in Q} w_i^B(\mathbf{x}) w_i^C(\mathbf{y}) + w^D(\mathbf{y}) \right) - \right. \\
&\qquad \left. \left(w^A(\mathbf{x}) + \sum_{i \in Q} \left(w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{cv}}(\mathbf{y}) + w_i^B(\mathbf{x}) w_{i,Z}^{C,\mathrm{U}} - w_{i,X(Z)}^{B,\mathrm{U}} w_{i,Z}^{C,\mathrm{U}} \right) + w_Z^{D,\mathrm{cv}}(\mathbf{y}) \right) \right| \\
&\leq \left(M^{BC} w(Y)^{\gamma^{BC} - \gamma} + \tau^{D,\mathrm{cv}} w(Y)^{\gamma^{D,\mathrm{cv}} - \gamma} \right) w(Z)^{\gamma},
\end{aligned}$$

which proves the existence of τ_1. □

The following remark is in order.

Remark 6.5.2.

1. Suppose w_i^C is Lipschitz continuous on Y for each $i \in Q$. We then have $\alpha_i^C \geq 1$, $\forall i \in Q$. If $\gamma_i^C \geq 1$ and $\beta_i^{C,\mathrm{c}} \geq 1$, $\forall i \in Q$, and $\gamma^{D,\mathrm{cv}} \geq 1$, we have from Lemma 6.5.1 that $(f_{X(Z)\times Z}^{\mathrm{cv}})_{Z \in \mathbb{I}Y}$ has at least first-order convergence on Y.

2. Let $X = [1,2]$, $Y = [-1,1]$, and $f(x,y) = xy$. For any $[-\varepsilon, \varepsilon] =: Z \in \mathbb{I}Y$ with $\varepsilon > 0$, consider the scheme of convex relaxations $(f_{X(Z)\times Z}^{\mathrm{cv}})_{Z \in \mathbb{I}Y}$ of f in Y with

$$f_{X(Z)\times Z}^{\mathrm{cv}}(x,y) = \max\{y - \varepsilon x + \varepsilon, 2y + \varepsilon x - 2\varepsilon\}.$$

The above scheme corresponds to the tightest possible scheme of convex relaxations in the reduced-space, but has at most first-order pointwise convergence on Y. This is in contrast to Theorem 10 in [38] where the scheme of convex envelopes of any twice continuously differentiable function was shown to have pointwise convergence order of at least two on $X \times Y$. Note that if $Q = \emptyset$, the pointwise convergence order of the scheme of convex relaxations $(f^{\text{cv}}_{X(Z)\times Z})_{Z\in\mathbb{I}Y}$ is dictated by the pointwise convergence order of the scheme $(w^{D,\text{cv}}_{Z})_{Z\in\mathbb{I}Y}$, and second-order pointwise convergence of $(f^{\text{cv}}_{X(Z)\times Z})_{Z\in\mathbb{I}Y}$ can be achieved by using the scheme of convex envelopes of w^D if w^D is twice continuously differentiable. Also note that Theorem 2 in [38], which states that the pointwise convergence order of a scheme of relaxations of a nonlinear twice continuously differentiable function can be at most two on $X \times Y$, naturally holds over Y as well.

The following result establishes a lower bound on the convergence order of the reduced-space lower bounding scheme proposed in [76] at infeasible points.

Lemma 6.5.3. Consider Problem (P), and suppose functions g_j, $j = 1, \cdots, m_I$, are each of the form (W) and functions h_k, $k = 1, \cdots, m_E$, are each of the form (W_{eq}). Let $(g^{\text{cv}}_{j,X(Z)\times Z})_{Z\in\mathbb{I}Y}$, $j = 1, \cdots, m_I$, denote continuous schemes of convex relaxations of $g_1, \cdots, g_{m_I}$, respectively, in Y with pointwise convergence orders $\gamma^{\text{cv}}_{g,1} > 0, \cdots, \gamma^{\text{cv}}_{g,m_I} > 0$ and corresponding constants $\tau^{\text{cv}}_{g,1}, \cdots, \tau^{\text{cv}}_{g,m_I}$, and let $(h^{\text{cv}}_{k,X(Z)\times Z}, h^{\text{cc}}_{k,X(Z)\times Z})_{Z\in\mathbb{I}Y}$, $k = 1, \cdots, m_E$, denote continuous schemes of relaxations of $h_1, \cdots, h_{m_E}$, respectively, in Y with pointwise convergence orders $\gamma_{h,1} > 0, \cdots, \gamma_{h,m_E} > 0$ and corresponding constants $\tau_{h,1}, \cdots, \tau_{h,m_E}$. Then, there exists $\bar{\tau} \geq 0$ such that for every $Z \in \mathbb{I}Y$

$$d\left(\overline{\begin{bmatrix}\mathbf{g}\\ \mathbf{h}\end{bmatrix}}(X(Z)\times Z), \mathbb{R}^{m_I}_{-} \times \{\mathbf{0}\}\right) - d\left(\mathcal{I}_C(Z), \mathbb{R}^{m_I}_{-} \times \{\mathbf{0}\}\right) \leq \bar{\tau} w(Z)^{\beta},$$

where $\mathcal{I}_C(Z)$ is defined as

$$\mathcal{I}_C(Z) := \Big\{(\mathbf{v}, \mathbf{w}) \in \mathbb{R}^{m_I} \times \mathbb{R}^{m_E} : \mathbf{v} = \mathbf{g}^{\text{cv}}_{X(Z)\times Z}(\mathbf{x}, \mathbf{y}), \mathbf{h}^{\text{cv}}_{X(Z)\times Z}(\mathbf{x}, \mathbf{y}) \leq \mathbf{w} \leq \mathbf{h}^{\text{cc}}_{X(Z)\times Z}(\mathbf{x}, \mathbf{y}) \\ \text{for some } (\mathbf{x}, \mathbf{y}) \in X(Z) \times Z\Big\}$$

and β is defined as

$$\beta := \min\left\{\min_{j\in\{1,\cdots,m_I\}} \gamma^{\text{cv}}_{g,j}, \min_{k\in\{1,\cdots,m_E\}} \gamma_{h,k}\right\}.$$

Proof. The proof is similar to that of Lemma 6.4.1 and is therefore omitted. ☐

Definition 6.5.4. [Feasible Point in the Reduced-Space] Consider Problem (P). A point $\mathbf{y} \in Y$ is said to be feasible (in the reduced-space) if there exists $\mathbf{x} \in X$ such that $(\mathbf{x}, \mathbf{y})$ is feasible for Problem (P).

The following result establishes first-order convergence of the reduced-space lower bounding scheme proposed in [76] at a feasible point in the reduced-space when first-order pointwise convergent schemes of relaxations are used and the reduced-space dual lower bounding scheme (see Section 6.5.2) is first-order convergent.

Theorem 6.5.5. Consider Problem (P). Suppose the functions f and g_j, $j = 1, \cdots, m_I$, are each of the form (W) and functions h_k, $k = 1, \cdots, m_E$, are each of the form ($\mathrm{W_{eq}}$). Let $\mathbf{y}^{\mathrm{f}} \in Y$ be a feasible point in the reduced-space for Problem (P). Suppose the reduced-space dual lower bounding scheme (see Section 6.5.2) has convergence of order β_d at $\mathbf{y}^{\mathrm{f}}$ and a corresponding scheme of dual variables $\left(\left(\boldsymbol{\mu}_Z^{\mathbf{y}^{\mathrm{f}}}, \boldsymbol{\lambda}_Z^{\mathbf{y}^{\mathrm{f}}}\right)\right)_{Z \in \mathbb{I}Y}$ (not necessarily optimal, but which yield β_d-order convergence at $\mathbf{y}^{\mathrm{f}}$) with $\left(\boldsymbol{\mu}_Z^{\mathbf{y}^{\mathrm{f}}}, \boldsymbol{\lambda}_Z^{\mathbf{y}^{\mathrm{f}}}\right) \in \mathbb{R}_+^{m_I} \times \mathbb{R}^{m_E}$, $\left\|\boldsymbol{\mu}_Z^{\mathbf{y}^{\mathrm{f}}}\right\|_\infty \leq \bar{\mu}$ and $\left\|\boldsymbol{\lambda}_Z^{\mathbf{y}^{\mathrm{f}}}\right\|_\infty \leq \bar{\lambda}$, $\forall Z \in \mathbb{I}Y$, for some constants $\bar{\mu}, \bar{\lambda} \geq 0$. Let $(f^{\mathrm{cv}}_{X(Z)\times Z})_{Z \in \mathbb{I}Y}$, $(g^{\mathrm{cv}}_{j,X(Z)\times Z})_{Z \in \mathbb{I}Y}$, $j = 1, \cdots, m_I$, denote continuous schemes of convex relaxations of f, $g_1, \cdots, g_{m_I}$, respectively, in Y with pointwise convergence orders $\gamma_f^{\mathrm{cv}} \geq 1$, $\gamma_{g,1}^{\mathrm{cv}} \geq 1, \cdots, \gamma_{g,m_I}^{\mathrm{cv}} \geq 1$ and corresponding constants τ_f^{cv}, $\tau_{g,1}^{\mathrm{cv}}, \cdots, \tau_{g,m_I}^{\mathrm{cv}}$. Let $(h^{\mathrm{cv}}_{k,X(Z)\times Z}, h^{\mathrm{cc}}_{k,X(Z)\times Z})_{Z \in \mathbb{I}Y}$, $k = 1, \cdots, m_E$, denote continuous schemes of relaxations of $h_1, \cdots, h_{m_E}$, respectively, in Y with pointwise convergence orders $\gamma_{h,1} \geq 1, \cdots, \gamma_{h,m_E} \geq 1$ and corresponding constants $\tau_{h,1}, \cdots, \tau_{h,m_E}$. Then, the scheme of lower bounding problems $(\mathcal{L}(Z))_{Z \in \mathbb{I}Y}$ with

$$(\mathcal{O}(Z))_{Z \in \mathbb{I}Y} := \left(\min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}^{\mathrm{cv}}(Z)} f^{\mathrm{cv}}_{X(Z)\times Z}(\mathbf{x}, \mathbf{y})\right)_{Z \in \mathbb{I}Y},$$

$$(\mathcal{I}_C(Z))_{Z \in \mathbb{I}Y} := \left(\left\{(\mathbf{v}, \mathbf{w}) \in \mathbb{R}^{m_I} \times \mathbb{R}^{m_E} : \mathbf{v} = \mathbf{g}^{\mathrm{cv}}_{X(Z)\times Z}(\mathbf{x}, \mathbf{y}), \mathbf{h}^{\mathrm{cv}}_{X(Z)\times Z}(\mathbf{x}, \mathbf{y}) \leq \mathbf{w} \leq \mathbf{h}^{\mathrm{cc}}_{X(Z)\times Z}(\mathbf{x}, \mathbf{y}) \right.\right.$$
$$\left.\left. \text{for some } (\mathbf{x}, \mathbf{z}) \in X(Z) \times Z\right\}\right)_{Z \in \mathbb{I}Y}$$

is at least $\min\left\{\min\left\{\gamma_f^{\mathrm{cv}}, \min\limits_{j \in \{1,\cdots,m_I\}} \gamma_{g,j}^{\mathrm{cv}}, \min\limits_{k \in \{1,\cdots,m_E\}} \gamma_{h,k}\right\}, \beta_d\right\}$-order convergent at $\mathbf{y}^{\mathrm{f}}$.

Proof. The proof is similar to that of Theorem 6.4.19 and is therefore omitted. □

Definition 6.5.6. [Unconstrained Point in the Reduced-Space] Consider Problem (P) with $m_E = 0$. A point $\mathbf{y} \in Y$ is said to be unconstrained (in the reduced-space) if there exists $\delta > 0$ such that $\forall \mathbf{z} \in Y$ with $\|\mathbf{z} - \mathbf{y}\| < \delta$, we have $\mathbf{g}(\mathbf{x}, \mathbf{z}) < \mathbf{0}$, $\forall \mathbf{x} \in X$.

The next result establishes first-order convergence of the reduced-space lower bounding scheme proposed in [76] at unconstrained points in the reduced-space when a first-order convergent scheme of relaxations of the objective is used by the (convergent) lower bounding scheme.

Proposition 6.5.7. Consider Problem (P) with $m_E = 0$. Suppose the functions f and g_j, $j = 1, \cdots, m_I$, are each of the form (W). Let $(f^{\mathrm{cv}}_{X(Z)\times Z})_{Z\in\mathbb{I}Y}$ denote a continuous scheme of convex relaxations of f in Y with convergence order $\beta^{\mathrm{cv}}_f > 0$ and corresponding constant τ^{cv}_f, $(g^{\mathrm{cv}}_{j,X(Z)\times Z})_{Z\in\mathbb{I}Y}$, $j = 1, \cdots, m_I$, denote continuous schemes of convex relaxations of $g_1, \cdots, g_{m_I}$, respectively, in Y with pointwise convergence orders $\gamma^{\mathrm{cv}}_{g,1} > 0, \cdots, \gamma^{\mathrm{cv}}_{g,m_I} > 0$ and corresponding constants $\tau^{\mathrm{cv}}_{g,1}, \cdots, \tau^{\mathrm{cv}}_{g,m_I}$.

Suppose $\mathbf{y}^{\mathrm{S}} \in Y$ is an unconstrained point in the reduced-space, and the scheme of lower bounding problems $(\mathcal{L}(Z))_{Z\in\mathbb{I}Y}$ with

$$(\mathcal{O}(Z))_{Z\in\mathbb{I}Y} := \left(\min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}^{\mathrm{cv}}(Z)} f^{\mathrm{cv}}_{X(Z)\times Z}(\mathbf{x},\mathbf{y})\right)_{Z\in\mathbb{I}Y},$$
$$(\mathcal{I}_C(Z))_{Z\in\mathbb{I}Y} := \left(\overline{\mathbf{g}}^{\mathrm{cv}}_{X(Z)\times Z}(X(Z)\times Z)\right)_{Z\in\mathbb{I}Y}$$

has convergence of order $\beta \in (0, \beta^{\mathrm{cv}}_f]$ at $\mathbf{y}^{\mathrm{S}}$. Then the scheme of lower bounding problems $(\mathcal{L}(Z))_{Z\in\mathbb{I}Y}$ is at least β^{cv}_f-order convergent at $\mathbf{y}^{\mathrm{S}}$.

Proof. The proof is similar to the proof of Corollary 6.4.12.

Since $\mathbf{y}^{\mathrm{S}}$ is an unconstrained point in the reduced-space and g_j is continuous for each $j \in \{1, \cdots, m_I\}$ by virtue of Assumption 6.2.1, $\exists\delta > 0$ such that $\forall \mathbf{z} \in Y$ with $\left\|\mathbf{z} - \mathbf{y}^{\mathrm{S}}\right\|_\infty \leq \delta$ (see Lemma 2.2.2), we have $\mathbf{g}(\mathbf{x},\mathbf{z}) < \mathbf{0}$, $\forall \mathbf{x} \in X$.

Consider $Z \in \mathbb{I}Y$ with $\mathbf{y}^{\mathrm{S}} \in Z$ and $w(Z) \leq \delta$. We have $\overline{\mathbf{g}}(X(Z)\times Z) \subset \mathbb{R}^{m_I}_-$ and $\overline{\mathbf{g}}^{\mathrm{cv}}_{X(Z)\times Z}(X(Z)\times Z) \subset \mathbb{R}^{m_I}_-$. Consequently,

$$\begin{aligned}
&\min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}(Z)} f(\mathbf{x},\mathbf{y}) \; - \min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}^{\mathrm{cv}}(Z)} f^{\mathrm{cv}}_{X(Z)\times Z}(\mathbf{x},\mathbf{y}) \\
=&\min_{(\mathbf{x},\mathbf{y})\in X(Z)\times Z} f(\mathbf{x},\mathbf{y}) \; - \min_{(\mathbf{x},\mathbf{y})\in X(Z)\times Z} f^{\mathrm{cv}}_{X(Z)\times Z}(\mathbf{x},\mathbf{y}) \\
\leq&\tau^{\mathrm{cv}}_f w(Z)^{\beta^{\mathrm{cv}}_f}.
\end{aligned}$$

The desired result follows by analogy to Lemma 6.3.8 based on the assumption that the reduced-space lower bounding scheme $(\mathcal{L}(Z))_{Z\in\mathbb{I}Y}$ is at least β-order convergent at $\mathbf{y}^{\mathrm{S}}$. $\square$

Note that Proposition 6.5.7 can be generalized in a manner similar to Corollary 6.4.17 to show that the above lower bounding scheme has β_f^{cv}-order convergence on a neighborhood of $\mathbf{y}^{\text{S}}$.

The following example shows that the convergence order of the reduced-space lower bounding scheme is dictated by the convergence order, β_f^{cv}, of the scheme $(f^{\text{cv}}_{X(Z)\times Z})_{Z\in\mathbb{I}Y}$ under the assumptions of Proposition 6.5.7.

Example 6.5.8. Let $X = [-1, 0.1], Y = [-1, 1]$, $m_I = 1$, and $m_E = 0$ with $f(x, y) = x^2+y^2$ and $g(x, y) = x + y - 0.5$. For any $[y^{\text{L}}, y^{\text{U}}] =: Z \in \mathbb{I}Y$, let

$$f^{\text{cv}}_{X(Z)\times Z}(x, y) = \begin{cases} x^2 - (y^{\text{U}} - y^{\text{L}})^3, & \text{if } 0 \in [y^{\text{L}}, y^{\text{U}}] \\ x^2 + \min\left\{(y^{\text{L}})^2, (y^{\text{U}})^2\right\} - (y^{\text{U}} - y^{\text{L}})^3, & \text{otherwise} \end{cases},$$

$$g^{\text{cv}}_{X(Z)\times Z}(x, y) = x + y - 0.5.$$

The scheme $(f^{\text{cv}}_{X(Z)\times Z})_{Z\in\mathbb{I}Y}$ has first-order pointwise convergence on Y and third-order convergence on Y, while the scheme $(g^{\text{cv}}_{X(Z)\times Z})_{Z\in\mathbb{I}Y}$ has arbitrarily high pointwise convergence order on Y.

Let $y^{\text{L}} = -\varepsilon$, $y^{\text{U}} = \varepsilon$ with $0 < \varepsilon \leq 0.1$. The width of Z is $w(Z) = 2\varepsilon$. The optimal objective value of Problem (P) on Z is 0, while the optimal objective of the lower bounding problem on Z is $-8\varepsilon^3$. Convergence at the point $y = 0$ is, therefore, at most third-order.

It is natural to wonder at this stage whether the reduced-space lower bounding scheme in [76] has 'similar convergence properties' to the full-space lower bounding scheme that was analyzed in Section 6.4.1. Example 6.5.19 presents a case where the reduced-space lower bounding scheme in [76] only has first-order convergence at a constrained minimizer that is a KKT point (cf. Example 6.4.10, Theorem 6.4.9 and Corollary 6.4.28). The following example shows that the reduced-space lower bounding scheme in [76] may have a convergence order as low as one even for unconstrained problems with smooth objective functions.

Example 6.5.9. Consider the following instance of Problem (P):

$$\begin{aligned} &\min_{x,y}\ 2x^2 + x^2y - xy^2 + (y - 0.5)^2 \\ &\text{s.t. } x \in [-1, 1], y \in [0, 1]. \end{aligned}$$

The global minimum, (x^*, y^*), of the above 'unconstrained problem' is $x^* = \frac{2\sqrt{21}}{3} - 3$, $y^* = \frac{\sqrt{21}}{3} - 1$ with optimal objective value $\nu^* = 2(x^*)^2 + (x^*)^2 y^* - x^*(y^*)^2 + (y^* - 0.5)^2$.

Consider $[y^* - \varepsilon, y^* + \varepsilon] =: Z \in \mathbb{I}Y$ with $\varepsilon \in (0, 0.25]$. The reduced-space lower bounding scheme in [76] yields

$$\begin{aligned}
\mathcal{O}(Z) = \min_{x,y,w_1,w_2} \; & 2x^2 + w_1 + w_2 + (y - 0.5)^2 \\
\text{s.t.} \quad & w_1 \geq x^2(y^* - \varepsilon), \\
& w_1 \geq y + x^2(y^* + \varepsilon) - (y^* + \varepsilon), \\
& w_2 \geq y^2 - x(y^* + \varepsilon)^2 - (y^* + \varepsilon)^2, \\
& w_2 \geq (y^*)^2 - 2y^* y - \varepsilon^2 - x(y^* - \varepsilon)^2 + (y^* - \varepsilon)^2, \\
& x \in [-1, 1], y \in [y^* - \varepsilon, y^* + \varepsilon].
\end{aligned}$$

Note that the point $(x_Z^{\mathrm{f}}, y_Z^{\mathrm{f}}, w_{1,Z}^{\mathrm{f}}, w_{2,Z}^{\mathrm{f}}) = (x^*, y^*, (x^*)^2(y^* - \varepsilon), -(y^*)^2 - \varepsilon^2 - x^*(y^* - \varepsilon)^2 + (y^* - \varepsilon)^2)$ is feasible for the lower bounding scheme with objective value $2(x^*)^2 + w_{1,Z}^{\mathrm{f}} + w_{2,Z}^{\mathrm{f}} + (y^* - 0.5)^2 = \nu^* + 2x^* y^* \varepsilon - (x^*)^2 \varepsilon - x^* \varepsilon^2 - 2y^* \varepsilon$. Therefore, we have

$$\begin{aligned}
\min_{(x,y) \in \mathcal{F}(Z)} f(x,y) \; - \min_{(x,y) \in \mathcal{F}^{\mathrm{cv}}(Z)} f_{X(Z) \times Z}^{\mathrm{cv}}(x,y) &\geq (x^*)^2 \varepsilon + x^* \varepsilon^2 + 2y^* \varepsilon - 2x^* y^* \varepsilon \\
&= \left(0.5(x^*)^2 + 0.5x^* \varepsilon + y^* - x^* y^*\right) w(Z) \\
&= 0.5 \left(1 + \varepsilon x^*\right) w(Z) \\
&\geq 0.5 w(Z),
\end{aligned}$$

which establishes that the reduced-space lower bounding scheme in [76] has at most first-order convergence at (the reduced-space minimizer) y^*.

Remark 6.5.10. Example 6.5.9 provides an instance of Problem (P) for which the minimum is unconstrained but the reduced-space lower bounding scheme in [76] is only first-order convergent at the reduced-space minimizer. Therefore, the lower bounding scheme in [76] could face severe clustering for this example [68, 238]. Note that this is in contrast to the full-space lower bounding schemes in Section 6.4 which can achieve at least second-order convergence at the above minimizer and thereby mitigate clustering.

The presence of the terms $x^2 y$ and $-xy^2$ in the objective function in Example 6.5.9 plays a crucial role in limiting the convergence order of the reduced-space lower bounding scheme in [76] (see Remark 6.5.2). Additionally, the analysis in Example 6.5.9 implies that the scheme of relaxations of its objective function has at most first-order Hausdorff convergence on Y. Theorem 6.5.23 in Section 6.5.2 implies that the reduced-space lower bounding scheme in [76] has second-order convergence at KKT points when all of the functions in Problem (P) are twice continuously differentiable and separable in $\mathbf{x}$ and $\mathbf{y}$.

6.5.2 Duality-based branch-and-bound

Dür and Horst [69, Section 3.3] outlined a reduced-space branch-and-bound algorithm in which they used Lagrangian duality to obtain lower bounds (also see [20, 70]). Dür and Horst [69] prove that when a constraint qualification holds for the reduced-space convex relaxation-based lower bounding scheme with each function in Problem (P) replaced by its (convex) envelope on $X \times Z$ (for each $Z \in \mathbb{I}Y$), the subdivision process is exhaustive on Y, and the selection procedure is bound improving, then any accumulation point of the sequence of reduced-space dual lower bounding solutions solves Problem (P).

The reduced-space Lagrangian dual lower bounding problem is in essence the same as its full-space counterpart Problem (D), with the major difference being that we only branch on the Y-space in the reduced-space dual lower bounding scheme to converge. We associate with the reduced-space dual lower bounding scheme, $(\mathcal{L}(Z))_{Z\in\mathbb{I}Y}$, at a feasible point in the reduced-space $\mathbf{y}$, a scheme of dual variables $((\boldsymbol{\mu}_Z^{\mathbf{y}}, \boldsymbol{\lambda}_Z^{\mathbf{y}}))_{Z\in\mathbb{I}Y}$ corresponding to the solution of the scheme of dual lower bounding problems $(\mathcal{O}(Z))_{Z\in\mathbb{I}Y}$ with $\mathbf{y} \in Z$. Dür and Horst [69, Section 4] also outlined classes of problems for which the reduced-space dual lower bounding problem can be solved to optimality. The following result, analogous to Theorem 6.4.23, holds.

Theorem 6.5.11. Consider Problem (P). Suppose strong duality holds for the reduced-space convex relaxation-based lower bounding scheme for Problem (P) obtained by using the schemes of (convex) envelopes of f, $\mathbf{g}$, and $\mathbf{h}$. Then, the reduced-space dual lower bounding scheme has at least as high a convergence order as the reduced-space convex relaxation-based lower bounding scheme obtained by using schemes of (convex) envelopes.

Proof. The proof is similar to that of Theorem 6.4.23 and is therefore omitted. $\square$

The following result from [69] states that when the constraints in Problem (P) are affine on $X \times Y$, the lower bounding scheme corresponding to schemes of (convex) envelopes provides the same scheme of lower bounds as that obtained using the dual lower bounding scheme.

Lemma 6.5.12. Consider Problem (P), and suppose the constraints in Problem (P) are affine in $\mathbf{x}$ and $\mathbf{y}$, i.e., $\mathbf{g} : X \times Y \ni (\mathbf{x}, \mathbf{y}) \mapsto \mathbf{A}_g\mathbf{x} + \mathbf{B}_g\mathbf{y} - \mathbf{c}_g$ and $\mathbf{h} : X \times Y \ni (\mathbf{x}, \mathbf{y}) \mapsto \mathbf{A}_h\mathbf{x} + \mathbf{B}_h\mathbf{y} - \mathbf{c}_h$. In addition, suppose Problem (P) is feasible and strong duality holds for Problem (P) for $\mathbf{y}$ restricted to any feasible point in Y. Then the lower bound obtained by solving the dual problem on $Z \in \mathbb{I}Y$ is the same as the lower bound obtained by solving the relaxation of the original problem on Z with the objective function f replaced by its convex envelope on $X \times Z$.

Proof. See Proposition 2.1 in [69]. □

Lemma 6.5.3 (in conjunction with Lemma 6.4.22) guarantees that the reduced-space dual lower bounding scheme has at least first-order convergence at infeasible points for the subclass of Problem (P) for which the algorithm of Epperly and Pistikopoulos is applicable when the functions w_i^C, $\forall i \in Q$, and w^D corresponding to each of the constraints are Lipschitz continuous. The following result shows that first-order convergence at infeasible points is guaranteed for a more general class of problems in the form of Problem (P) even when constraint propagation techniques are not used.

Lemma 6.5.13. Let $X \subset \mathbb{R}^{n_x}$, $Y \subset \mathbb{R}^{n_y}$ be nonempty compact convex sets, $f : X \times Y \to \mathbb{R}$ be Lipschitz continuous on $X \times Y$ with Lipschitz constant M_f. Suppose f is partially convex with respect to $\mathbf{x}$, i.e., $f(\cdot, \mathbf{y})$ is convex on X for each $\mathbf{y} \in Y$. For any $Z \in \mathbb{I}Y$, let $f^{\text{cv,env}}_{X\times Z} : X \times Z \to \mathbb{R}$ denote the convex envelope of f on $X \times Z$. Assume that for each $\bar{\mathbf{x}} \in X$, there exists a subgradient $\mathbf{s}(\mathbf{y}; \bar{\mathbf{x}}) \in \partial_{\mathbf{x}} f(\mathbf{x}, \mathbf{y})|_{\mathbf{x}=\bar{\mathbf{x}}}$ such that each $s_i(\mathbf{y}; \bar{\mathbf{x}})$, $i = 1, \cdots, n_x$, is Lipschitz continuous on Y with Lipschitz constant M_s. Then, the reduced-space scheme of convex envelopes $\left(f^{\text{cv,env}}_{X\times Z}\right)_{Z\in\mathbb{I}Y}$ has pointwise convergence of order at least one on Y.

Proof. We wish to prove the existence of a constant $\tau \geq 0$ such that

$$\max_{(\mathbf{x},\mathbf{y})\in X\times Z} \left| f(\mathbf{x}, \mathbf{y}) - f^{\text{cv,env}}_{X\times Z}(\mathbf{x}, \mathbf{y}) \right| \leq \tau w(Z), \quad \forall Z \in \mathbb{I}Y.$$

Note that the existence of the maximum in the above expression follows from the (Lipschitz) continuity of f, Lemma 2.3.30, and the compactness of $X \times Y$. Consider $Z \in \mathbb{I}Y$, and let

$(\mathbf{x}_Z^*, \mathbf{y}_Z^*) \in \underset{(\mathbf{x},\mathbf{y})\in X\times Z}{\arg\max} \left|f(\mathbf{x},\mathbf{y}) - f_{X\times Z}^{\text{cv,env}}(\mathbf{x},\mathbf{y})\right|$. We have

$$\begin{aligned}\max_{(\mathbf{x},\mathbf{y})\in X\times Z}\left|f(\mathbf{x},\mathbf{y}) - f_{X\times Z}^{\text{cv,env}}(\mathbf{x},\mathbf{y})\right| &= f(\mathbf{x}_Z^*,\mathbf{y}_Z^*) - f_{X\times Z}^{\text{cv,env}}(\mathbf{x}_Z^*,\mathbf{y}_Z^*) \\ &= \max_{\mathbf{y}\in Z}\left|f(\mathbf{x}_Z^*,\mathbf{y}) - f_{X\times Z}^{\text{cv,env}}(\mathbf{x}_Z^*,\mathbf{y})\right|. \qquad (6.10)\end{aligned}$$

Since $f(\cdot,\mathbf{y})$ is convex on X for each $\mathbf{y}\in Y$, we have

$$\begin{aligned}f(\mathbf{x},\mathbf{y}) &\geq f(\mathbf{x}_Z^*,\mathbf{y}) + \mathbf{s}(\mathbf{y};\mathbf{x}_Z^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}_Z^*) \\ &= f(\mathbf{x}_Z^*,\mathbf{y}) + w_Z(\mathbf{x},\mathbf{y}) \\ &\geq f_Z^{\text{cv,env}}(\mathbf{x}_Z^*,\mathbf{y}) + w_{X\times Z}^{\text{cv}}(\mathbf{x},\mathbf{y}), \quad \forall(\mathbf{x},\mathbf{y})\in X\times Z,\end{aligned}$$

where $\mathbf{s}(\mathbf{y};\mathbf{x}_Z^*) \in \partial_{\mathbf{x}} f(\mathbf{x},\mathbf{y})|_{\mathbf{x}=\mathbf{x}_Z^*}$ is a subgradient of $f(\cdot,\mathbf{y})$ at $\mathbf{x}_Z^*$ such that $s_i(\mathbf{y};\mathbf{x}_Z^*)$, $\forall i\in\{1,\cdots,n_x\}$, is Lipschitz continuous on Z with Lipschitz constant M_s, $f_Z^{\text{cv,env}}(\mathbf{x}_Z^*,\cdot)$ denotes the convex envelope of $f(\mathbf{x}_Z^*,\cdot)$ on Z, $w_Z(\mathbf{x},\mathbf{y}) := \mathbf{s}(\mathbf{y};\mathbf{x}_Z^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}_Z^*)$ is a function of the form (W), and $w_{X\times Z}^{\text{cv}}$ is a convex relaxation of w_Z on $X\times Z$ of the form (W^{cv}) with first-order (pointwise) convergent schemes of estimators of $\mathbf{s}(\mathbf{y};\mathbf{x}_Z^*)$ used in its construction.

Since f is Lipschitz continuous on $X\times Y$ and $f_Z^{\text{cv,env}}(\mathbf{x}_Z^*,\cdot)$ is the convex envelope of $f(\mathbf{x}_Z^*,\cdot)$ on Z, we have from Lemma 6.3.10 that

$$\max_{\mathbf{y}\in Z}\left|f(\mathbf{x}_Z^*,\mathbf{y}) - f_Z^{\text{cv,env}}(\mathbf{x}_Z^*,\mathbf{y})\right| \leq M_f w(Z).$$

Using Lemma 6.5.1 with $w_i^B(\mathbf{x}) = (x_i - x_{i,Z}^*)$, $w_i^C(\mathbf{y}) = s_i(\mathbf{y};\mathbf{x}_Z^*)$, $w_{i,X}^{B,\mathrm{L}} = \min_{\mathbf{x}\in X}(x_i - x_{i,Z}^*)$, $w_{i,X}^{B,\mathrm{U}} = \max_{\mathbf{x}\in X}(x_i - x_{i,Z}^*)$, $w_{i,Z}^{C,\text{cv}}(\mathbf{y}) = w_{i,Z}^{C,\mathrm{L}} = \min_{\mathbf{y}\in Z} s_i(\mathbf{y};\mathbf{x}_Z^*)$, and $w_{i,Z}^{C,\text{cc}}(\mathbf{y}) = w_{i,Z}^{C,\mathrm{U}} = \max_{\mathbf{y}\in Z} s_i(\mathbf{y};\mathbf{x}_Z^*)$, we can show the existence of a constant $\bar{\tau}\geq 0$ such that

$$\max_{(\mathbf{x},\mathbf{y})\in X\times Z}\left|w_Z(\mathbf{x},\mathbf{y}) - w_{X\times Z}^{\text{cv}}(\mathbf{x},\mathbf{y})\right| \leq \bar{\tau} w(Z).$$

From the above two inequalities, we have

$$\max_{(\mathbf{x},\mathbf{y})\in X\times Z}\left|\left(f(\mathbf{x}_Z^*,\mathbf{y}) + w_Z(\mathbf{x},\mathbf{y})\right) - \left(f_Z^{\text{cv,env}}(\mathbf{x}_Z^*,\mathbf{y}) + w_{X\times Z}^{\text{cv}}(\mathbf{x},\mathbf{y})\right)\right| \leq (M_f + \bar{\tau})\, w(Z).$$

Using $w_Z(\mathbf{x}_Z^*, \mathbf{y}) = 0$, we obtain

$$\max_{\mathbf{y}\in Z}\left|f(\mathbf{x}_Z^*, \mathbf{y}) - \left(f_Z^{\text{cv,env}}(\mathbf{x}_Z^*, \mathbf{y}) + w_{X\times Z}^{\text{cv}}(\mathbf{x}_Z^*, \mathbf{y})\right)\right| \leq (M_f + \bar{\tau})\, w(Z).$$

Since the convex envelope of f on $X \times Z$, $f_{X\times Z}^{\text{cv,env}}$, is, by definition, tighter than the convex relaxation
$f_Z^{\text{cv,env}}(\mathbf{x}_Z^*, \cdot) + w_{X\times Z}^{\text{cv}}$ at $\mathbf{x}_Z^*$, we have from Equation (6.10) that

$$\max_{\mathbf{y}\in Z}\left|f(\mathbf{x}_Z^*, \mathbf{y}) - f_{X\times Z}^{\text{cv,env}}(\mathbf{x}_Z^*, \mathbf{y})\right| \leq (M_f + \bar{\tau})\, w(Z),$$

which proves the existence of τ. □

Note that the assumptions of Lemma 6.5.13 are readily satisfied if f is a Lipschitz continuous function of the form (W) that is composed of continuous functions w^A, w_i^B, $\forall i \in Q$, and w^D and Lipschitz continuous functions w_i^C, $\forall i \in Q$. An instance for which the assumptions of Lemma 6.5.13 are not satisfied is $f(x, y) = |y||x + y + 1|$ with $X = [-1, 1]$ and $Y = [-1, 1]$. The following examples provide instances for which the assumptions of Lemma 6.5.13 are satisfied, but where the functions involved are not in the form (W).

Example 6.5.14. Let $X = [-1, 1]$, $Y = [-1, 1]$, and $f(x, y) = \exp(xy)$. We have $M_f = \sqrt{2}\exp(1)$, $s(y; x) = y\exp(xy)$, and $M_s = 2\exp(1)$ satisfying the assumptions of Lemma 6.5.13.

Example 6.5.15. Let $X = [-1, 1]$, $Y = [-1, 1]$, and $f(x, y) = -|y|\sqrt{x + y + 3}$. We have $M_f = 4$, $s(y; x) = -\frac{|y|}{2\sqrt{x+y+3}}$, and $M_s = 1$ satisfying the assumptions of Lemma 6.5.13.

The next result shows that the reduced-space dual lower bounding scheme has arbitrarily high convergence order at unconstrained points in the reduced-space.

Proposition 6.5.16. Consider Problem (P) with $m_E = 0$. Suppose $\mathbf{y}^{\text{S}} \in Y$ is an unconstrained point in the reduced-space. Furthermore, suppose the reduced-space dual lower bounding scheme has convergence of order $\beta > 0$ at $\mathbf{y}^{\text{S}}$. Then the reduced-space dual lower bounding scheme has arbitrarily high convergence order at $\mathbf{y}^{\text{S}}$.

Proof. The proof is closely related to the proof of Proposition 6.4.25.

Since $\mathbf{y}^{\mathrm{S}}$ is an unconstrained point in the reduced-space and g_j is continuous for each $j \in \{1, \cdots, m_I\}$ by virtue of Assumption 6.2.1, there exists $\delta > 0$ such that $\forall \mathbf{z} \in Y$ satisfying $\left\|\mathbf{z} - \mathbf{y}^{\mathrm{S}}\right\|_{\infty} \leq \delta$ (see Lemma 2.2.2), we have $\mathbf{g}(\mathbf{x}, \mathbf{z}) < 0$, $\forall \mathbf{x} \in X$.

Consider $Z \in \mathbb{I}Y$ with $w(Z) \leq \delta$. Since $\overline{\mathbf{g}}(X(Z) \times Z) \subset \mathbb{R}^{m_I}_{-}$, Problem (P) can be reformulated as

$$\min_{(\mathbf{x},\mathbf{y}) \in \mathcal{F}(Z)} f(\mathbf{x}, \mathbf{y}) = \min_{(\mathbf{x},\mathbf{y}) \in X(Z) \times Z} f(\mathbf{x}, \mathbf{y}).$$

The dual lower bound can be bounded from below as

$$\sup_{\boldsymbol{\mu} \geq \mathbf{0}} \min_{(\mathbf{x},\mathbf{y}) \in X(Z) \times Z} \left[f(\mathbf{x}, \mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}} \mathbf{g}(\mathbf{x}, \mathbf{y}) \right] \geq \min_{(\mathbf{x},\mathbf{y}) \in X(Z) \times Z} f(\mathbf{x}, \mathbf{y}).$$

The desired result follows by analogy to Lemma 6.3.8 and the assumption that the dual lower bounding scheme is at least β-order convergent at $\mathbf{y}^{\mathrm{S}}$. $\square$

The following result establishes first-order convergence of the reduced-space dual lower bounding scheme even in the absence of constraint propagation.

Theorem 6.5.17. Consider Problem (P). Suppose f, g_j, $j = 1, \cdots, m_I$, and h_k, $k = 1, \cdots, m_E$, are Lipschitz continuous on $X \times Y$ with Lipschitz constants $M_f, M_{g,1}, \cdots, M_{g,m_I}$, $M_{h,1}, \cdots, M_{h,m_E}$, respectively, and assume that the assumptions of Lemma 6.5.13 hold for $\mathbf{g}$ and $\mathbf{h}$. Assume, in addition, that Problem (P) is feasible, and that strong duality holds for Problem (P) for $\mathbf{y}$ restricted to any feasible point in Y. Furthermore, suppose the set of optimal dual variables for Problem (P) restricted to any feasible $\mathbf{y} \in Y$ is bounded (with the bound independent of $\mathbf{y}$). Then the reduced-space dual lower bounding scheme is at least first-order convergent on Y.

Proof. Lemmata 6.4.22, 6.5.3, and 6.5.13 imply that the dual lower bounding scheme is at least first-order convergent at any infeasible point $\mathbf{y} \in Y$ with the prefactor independent of $\mathbf{y}$ (note that the conclusion of Lemma 6.5.3 does not depend on the schemes of relaxations of the constraints being in the form (W^{cv})).

Define $F(\mathbf{y}, \boldsymbol{\mu}, \boldsymbol{\lambda}) := \min_{\mathbf{x} \in X} f(\mathbf{x}, \mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}} \mathbf{g}(\mathbf{x}, \mathbf{y}) + \boldsymbol{\lambda}^{\mathrm{T}} \mathbf{h}(\mathbf{x}, \mathbf{y})$. We first show that $F(\cdot, \boldsymbol{\mu}, \boldsymbol{\lambda})$ is Lipschitz continuous on Y for any $(\boldsymbol{\mu}, \boldsymbol{\lambda}) \in \mathbb{R}^{m_I}_{+} \times \mathbb{R}^{m_E}$. Consider $(\boldsymbol{\mu}, \boldsymbol{\lambda}) \in \mathbb{R}^{m_I}_{+} \times \mathbb{R}^{m_E}$

and $\mathbf{y}_1, \mathbf{y}_2 \in Y$. We have

$$\begin{aligned}
&|F(\mathbf{y}_1, \boldsymbol{\mu}, \boldsymbol{\lambda}) - F(\mathbf{y}_2, \boldsymbol{\mu}, \boldsymbol{\lambda})| \\
=&\left|\left(\min_{\mathbf{x}\in X} f(\mathbf{x}, \mathbf{y}_1) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x}, \mathbf{y}_1) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{x}, \mathbf{y}_1)\right) - \left(\min_{\mathbf{x}\in X} f(\mathbf{x}, \mathbf{y}_2) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x}, \mathbf{y}_2) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{x}, \mathbf{y}_2)\right)\right| \\
\leq& \max_{\mathbf{x}\in X} \left|(f(\mathbf{x}, \mathbf{y}_1) - f(\mathbf{x}, \mathbf{y}_2)) + \boldsymbol{\mu}^{\mathrm{T}}\left(\mathbf{g}(\mathbf{x}, \mathbf{y}_1) - \mathbf{g}(\mathbf{x}, \mathbf{y}_2)\right) + \boldsymbol{\lambda}^{\mathrm{T}}\left(\mathbf{h}(\mathbf{x}, \mathbf{y}_1) - \mathbf{h}(\mathbf{x}, \mathbf{y}_2)\right)\right| \\
\leq& \max_{\mathbf{x}\in X} |f(\mathbf{x}, \mathbf{y}_1) - f(\mathbf{x}, \mathbf{y}_2)| + \max_{\mathbf{x}\in X} \left|\boldsymbol{\mu}^{\mathrm{T}}(\mathbf{g}(\mathbf{x}, \mathbf{y}_1) - \mathbf{g}(\mathbf{x}, \mathbf{y}_2))\right| + \max_{\mathbf{x}\in X} \left|\boldsymbol{\lambda}^{\mathrm{T}}(\mathbf{h}(\mathbf{x}, \mathbf{y}_1) - \mathbf{h}(\mathbf{x}, \mathbf{y}_2))\right| \\
\leq& \left(M_f + \sum_{j=1}^{m_I}|\mu_j|M_{g,j} + \sum_{k=1}^{m_E}|\lambda_k|M_{h,k}\right)\|\mathbf{y}_1 - \mathbf{y}_2\|,
\end{aligned}$$

where Step 2 follows from Lemma 2.3.35, and Step 4 follows from the Lipschitz continuity of the functions involved.

Suppose $\mathcal{F}(Y) \neq \emptyset$ and $Z \in \mathbb{I}Y$ such that $Z \cap \mathcal{F}(Y) \neq \emptyset$. Since strong duality holds for Problem (P) with $\mathbf{y}$ restricted to any feasible point in Y, Problem (P) can be equivalently expressed on Z as

$$\min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}(Z)} f(\mathbf{x}, \mathbf{y}) = \min_{\mathbf{y}\in Z} \sup_{(\boldsymbol{\mu},\boldsymbol{\lambda})\in\mathbb{R}_+^{m_I}\times\mathbb{R}^{m_E}} F(\mathbf{y}, \boldsymbol{\mu}, \boldsymbol{\lambda}).$$

By strong duality and $Z \cap \mathcal{F}(Y) \neq \emptyset$, there exists a minimizer $(\mathbf{y}_Z^*, \boldsymbol{\mu}_Z^*, \boldsymbol{\lambda}_Z^*)$ of the above 'dual form' of Problem (P) when $\mathbf{y}$ is restricted to Z. We have

$$\begin{aligned}
&\left|\min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}(Z)} f(\mathbf{x}, \mathbf{y}) \;-\; \sup_{(\boldsymbol{\mu},\boldsymbol{\lambda})\in\mathbb{R}_+^{m_I}\times\mathbb{R}^{m_E}} \min_{(\mathbf{x},\mathbf{y})\in X\times Z} \left[f(\mathbf{x}, \mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x}, \mathbf{y}) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{x}, \mathbf{y})\right]\right| \\
=& \left|F(\mathbf{y}_Z^*, \boldsymbol{\mu}_Z^*, \boldsymbol{\lambda}_Z^*) - \sup_{(\boldsymbol{\mu},\boldsymbol{\lambda})\in\mathbb{R}_+^{m_I}\times\mathbb{R}^{m_E}} \min_{\mathbf{y}\in Z} F(\mathbf{y}, \boldsymbol{\mu}, \boldsymbol{\lambda})\right| \\
\leq& \left|F(\mathbf{y}_Z^*, \boldsymbol{\mu}_Z^*, \boldsymbol{\lambda}_Z^*) - \min_{\mathbf{y}\in Z} F(\mathbf{y}, \boldsymbol{\mu}_Z^*, \boldsymbol{\lambda}_Z^*)\right| \\
=& |F(\mathbf{y}_Z^*, \boldsymbol{\mu}_Z^*, \boldsymbol{\lambda}_Z^*) - F(\bar{\mathbf{y}}_Z, \boldsymbol{\mu}_Z^*, \boldsymbol{\lambda}_Z^*)| \\
\leq& \left(M_f + \sum_{j=1}^{m_I}\left|\mu_{j,Z}^*\right|M_{g,j} + \sum_{k=1}^{m_E}\left|\lambda_{k,Z}^*\right|M_{h,k}\right)\|\mathbf{y}_Z^* - \bar{\mathbf{y}}_Z\| \\
\leq& \left(M_f + \sum_{j=1}^{m_I} M_\infty M_{g,j} + \sum_{k=1}^{m_E} M_\infty M_{h,k}\right)\sqrt{n_y}w(Z),
\end{aligned}$$

where $\bar{\mathbf{y}}_Z \in \arg\min_{\mathbf{y}\in Z} F(\mathbf{y}, \boldsymbol{\mu}_Z^*, \boldsymbol{\lambda}_Z^*)$, $M_\infty := \sup_{\mathbf{y}\in Y} \max\{\|\boldsymbol{\mu}^*(\mathbf{y})\|_\infty, \|\boldsymbol{\lambda}^*(\mathbf{y})\|_\infty\}$ is an upper bound on the norm of pairs of optimal dual variables $(\boldsymbol{\mu}^*(\mathbf{y}), \boldsymbol{\lambda}^*(\mathbf{y})) \in \underset{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}}{\arg\max}\ F(\mathbf{y}, \boldsymbol{\mu}, \boldsymbol{\lambda})$, and the penultimate step follows from the Lipschitz continuity of $F(\cdot, \boldsymbol{\mu}, \boldsymbol{\lambda})$ on Y. $\square$

The assumption that the set of optimal dual variables for Problem (P) restricted to any feasible $\mathbf{y} \in Y$ is bounded can be replaced with the less restrictive assumption that there exists an optimal dual variable pair $(\boldsymbol{\mu}^*(\mathbf{y}), \boldsymbol{\lambda}^*(\mathbf{y})) \in \underset{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}}{\arg\max}\ F(\mathbf{y}, \boldsymbol{\mu}, \boldsymbol{\lambda})$ for each $\mathbf{y} \in Y$ such that $\sup_{\mathbf{y}\in Y} \max\{\|\boldsymbol{\mu}^*(\mathbf{y})\|_\infty, \|\boldsymbol{\lambda}^*(\mathbf{y})\|_\infty\} \leq M_\infty$.

A corollary of Theorems 6.5.5 and 6.5.17 is that first-order convergence is guaranteed for the convex relaxation-based reduced-space lower bounding scheme in [76] when first-order pointwise convergent schemes of relaxations on Y are used in its construction. Instead of proving first-order convergence of the lower bounding scheme in [76] at feasible points under the assumption that schemes of bounded optimal dual variables exist, we show that the reduced-space lower bounding scheme in [76] enjoys first-order convergence at any feasible point in the reduced-space under the (less restrictive) assumption that strong duality holds for Problem (P) with $\mathbf{y}$ fixed to the feasible point.

Corollary 6.5.18. Consider Problem (P). Suppose the functions f and g_j, for each $j \in \{1, \cdots, m_I\}$, are Lipschitz continuous on $X \times Y$ with Lipschitz constants M_f, $M_{g,1}, \cdots$, M_{g,m_I}, respectively, and are each of the form (W). Furthermore, suppose functions h_k, $k = 1, \cdots, m_E$, are Lipschitz continuous on $X \times Y$ with Lipschitz constants $M_{h,1}, \cdots, M_{h,m_E}$, respectively, and are each of the form (W_{eq}). Suppose $\mathbf{y}^{\mathrm{f}} \in Y$ is a feasible point in the reduced-space and strong duality holds for Problem (P) when $\mathbf{y}$ is fixed to $\mathbf{y}^{\mathrm{f}}$. Let $(f^{\mathrm{cv}}_{X(Z)\times Z})_{Z\in\mathbb{I}Y}$ and $(g^{\mathrm{cv}}_{j,X(Z)\times Z})_{Z\in\mathbb{I}Y}$, $j = 1, \cdots, m_I$, denote continuous schemes of convex relaxations of f, $g_1, \cdots, g_{m_I}$, respectively, in Y with pointwise convergence orders $\gamma^{\mathrm{cv}}_f \geq 1$, $\gamma^{\mathrm{cv}}_{g,1} \geq 1, \cdots, \gamma^{\mathrm{cv}}_{g,m_I} \geq 1$ and corresponding constants τ^{cv}_f, $\tau^{\mathrm{cv}}_{g,1}, \cdots, \tau^{\mathrm{cv}}_{g,m_I}$. Let $(h^{\mathrm{cv}}_{k,X(Z)\times Z}, h^{\mathrm{cc}}_{k,X(Z)\times Z})_{Z\in\mathbb{I}Y}$, $k = 1, \cdots, m_E$, denote continuous schemes of relaxations of $h_1, \cdots, h_{m_E}$, respectively, in Y with pointwise convergence orders $\gamma_{h,1} \geq 1, \cdots, \gamma_{h,m_E} \geq 1$ and corresponding constants $\tau_{h,1}, \cdots, \tau_{h,m_E}$. Then, the scheme of lower bounding problems $(\mathcal{L}(Z))_{Z\in\mathbb{I}Y}$ proposed in [76] is at least first-order convergent at $\mathbf{y}^{\mathrm{f}}$.

Proof. Let $\left(\boldsymbol{\mu}^{\mathbf{y}^{\mathrm{f}}}, \boldsymbol{\lambda}^{\mathbf{y}^{\mathrm{f}}}\right) \in \underset{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}}{\arg\max}\, F(\mathbf{y}^{\mathrm{f}}, \boldsymbol{\mu}, \boldsymbol{\lambda})$ be an optimal pair of dual variables for $\mathbf{y}$ fixed to $\mathbf{y}^{\mathrm{f}}$ in Problem (P). Suppose $Z \in \mathbb{I}Y$ with $\mathbf{y}^{\mathrm{f}} \in Z$. Similar to the proof of Theorem 6.5.17, we have

$$
\begin{aligned}
&\left| \min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}(Z)} f(\mathbf{x},\mathbf{y}) \;-\; \sup_{(\boldsymbol{\mu},\boldsymbol{\lambda})\in\mathbb{R}_+^{m_I}\times\mathbb{R}^{m_E}} \min_{(\mathbf{x},\mathbf{y})\in X\times Z} \left[f(\mathbf{x},\mathbf{y}) + \boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y}) + \boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y})\right] \right| \\
&\leq \left| F\left(\mathbf{y}^{\mathrm{f}}, \boldsymbol{\mu}^{\mathbf{y}^{\mathrm{f}}}, \boldsymbol{\lambda}^{\mathbf{y}^{\mathrm{f}}}\right) - \sup_{(\boldsymbol{\mu},\boldsymbol{\lambda})\in\mathbb{R}_+^{m_I}\times\mathbb{R}^{m_E}} \min_{\mathbf{y}\in Z} F(\mathbf{y},\boldsymbol{\mu},\boldsymbol{\lambda}) \right| \\
&\leq \left| F\left(\mathbf{y}^{\mathrm{f}}, \boldsymbol{\mu}^{\mathbf{y}^{\mathrm{f}}}, \boldsymbol{\lambda}^{\mathbf{y}^{\mathrm{f}}}\right) - \min_{\mathbf{y}\in Z} F\left(\mathbf{y}, \boldsymbol{\mu}^{\mathbf{y}^{\mathrm{f}}}, \boldsymbol{\lambda}^{\mathbf{y}^{\mathrm{f}}}\right) \right| \\
&\leq \tau^{\mathrm{f}} w(Z),
\end{aligned}
$$

for some constant $\tau^{\mathrm{f}} \geq 0$. The result then holds as a consequence of Theorem 6.5.5 by using $\boldsymbol{\mu}_Z^{\mathbf{y}^{\mathrm{f}}} = \boldsymbol{\mu}^{\mathbf{y}^{\mathrm{f}}}$, $\boldsymbol{\lambda}_Z^{\mathbf{y}^{\mathrm{f}}} = \boldsymbol{\lambda}^{\mathbf{y}^{\mathrm{f}}}$, $\bar{\mu} = \left\|\boldsymbol{\mu}^{\mathbf{y}^{\mathrm{f}}}\right\|_\infty$, and $\bar{\lambda} = \left\|\boldsymbol{\lambda}^{\mathbf{y}^{\mathrm{f}}}\right\|_\infty$ in Theorem 6.5.5. □

The following example shows that the convergence order of the reduced-space dual lower bounding scheme may be as low as one at constrained minima.

Example 6.5.19. Consider the following instance of Problem (P):

$$
\begin{aligned}
\min_{x,y}\ & -xy \\
\text{s.t.}\ & x + y \leq 1, \\
& x \in [-1,1], y \in [0,1].
\end{aligned}
$$

The optimal solution is $(x^*, y^*) = (0.5, 0.5)$ with optimal objective value -0.25. When the inequality constraint is dualized, the following dual function is obtained:

$$
\begin{aligned}
q(\mu) = \min_{x,y}\ & -xy + \mu\,(x + y - 1) \\
\text{s.t.}\ & x \in [-1,1], y \in [0,1].
\end{aligned}
$$

Consider $[y^{\mathrm{L}}, y^{\mathrm{U}}] = [0.5-\varepsilon, 0.5+\varepsilon] =: Z \in \mathbb{I}Y$ with $\varepsilon \in (0, 0.5]$. In order to derive the dual function

$$
q(\mu) = \min_{\substack{x\in[-1,1]\\ y\in[y^{\mathrm{L}},y^{\mathrm{U}}]}} -xy + \mu(x + y - 1)
$$

as an explicit function of μ, we partition the domain of μ to obtain

$$q(\mu) = \begin{cases} (\mu - 1)y^{\mathrm{U}}, & \text{if } \mu \leq y^{\mathrm{L}} \\ \min\{(\mu - 1)y^{\mathrm{U}}, (\mu + 1)y^{\mathrm{L}} - 2\mu\}, & \text{if } y^{\mathrm{L}} \leq \mu \leq y^{\mathrm{U}} \\ (\mu + 1)y^{\mathrm{L}} - 2\mu, & \text{if } \mu \geq y^{\mathrm{U}} \end{cases}$$

when the bounds on x are taken to be $[-1, 1]$ irrespective of the bounds on y. The dual lower bound can therefore be derived as:

$$\sup_{\mu \geq 0} q(\mu) = \frac{(y^{\mathrm{L}} - 1)y^{\mathrm{U}}}{1 + 0.5(y^{\mathrm{U}} - y^{\mathrm{L}})}.$$

Therefore, for $[y^{\mathrm{L}}, y^{\mathrm{U}}] = [0.5 - \varepsilon, 0.5 + \varepsilon]$, the dual lower bound can be derived as

$$\sup_{\mu \geq 0} q(\mu) = \frac{(-0.5 - \varepsilon)(0.5 + \varepsilon)}{1 + \varepsilon} = -\frac{(0.5 + \varepsilon)^2}{1 + \varepsilon}.$$

Consequently,

$$\min_{(x,y) \in \mathcal{F}(Z)} -xy \; - \; \sup_{\mu \geq 0} q(\mu) = -0.25 + \frac{(0.5 + \varepsilon)^2}{1 + \varepsilon} = \frac{0.75\varepsilon + \varepsilon^2}{1 + \varepsilon} \geq 0.75\varepsilon,$$

which implies that the dual lower bounding scheme is at most first-order convergent at y^*.

Remark 6.5.20. Example 6.5.19 provides an instance of Problem (P) for which both the reduced-space dual lower bounding scheme [69] and the reduced-space lower bounding scheme in [76] (this follows from Lemma 6.5.12) are only first-order convergent at the minimizer. Furthermore, for each $y \in [0, 1]$, the reduced-space objective function $v : [0, 1] \to \mathbb{R}$ can be derived as

$$\begin{aligned} v(y) = \min_{x} \;\; & -xy \\ \text{s.t. } \; & x + y \leq 1, \\ & x \in [-1, 1], \end{aligned}$$

which reduces to $v(y) = y^2 - y$. It can be seen that $y^* = 0.5$ is an unconstrained minimum of the reduced-space objective $v(y)$, which implies that at least second-order convergence of the reduced-space lower bounding scheme is typically required at y^* to mitigate clustering [68, 238].

Therefore, neither reduced-space lower bounding scheme can be expected to avoid clustering for this example. Note that this is in contrast to the full-space lower bounding schemes in Section 6.4 which can achieve at least second-order convergence at (x^*, y^*) and thereby mitigate clustering (see Chapter 5).

Note, however, that the use of constraint propagation techniques by reduced-space lower bounding schemes can potentially increase their convergence order as demonstrated by Examples 6.5.21 and 6.5.22. This demonstrates the importance of constraint propagation techniques in reduced-space lower bounding schemes, which has not been emphasized in references [69, 76].

Example 6.5.21. Consider the instance of Problem (P) in Example 6.5.19 with $Z = [y^{\mathrm{L}}, y^{\mathrm{U}}] \subset [0,1]$, $y^{\mathrm{L}} \leq 0.5$, $y^{\mathrm{U}} \geq 0.5$. Suppose we use constraint propagation to derive $X(Z) = [-1, 1-y^{\mathrm{L}}]$. The dual function can be derived as

$$q(\mu) = \begin{cases} \mu(y^{\mathrm{U}} - y^{\mathrm{L}}) + y^{\mathrm{U}}(y^{\mathrm{L}} - 1), & \text{if } \mu \leq y^{\mathrm{L}} \\ \min\{\mu(y^{\mathrm{U}} - y^{\mathrm{L}}) + y^{\mathrm{U}}(y^{\mathrm{L}} - 1), (\mu+1)y^{\mathrm{L}} - 2\mu\}, & \text{if } y^{\mathrm{L}} \leq \mu \leq y^{\mathrm{U}} \\ (\mu+1)y^{\mathrm{L}} - 2\mu, & \text{if } \mu \geq y^{\mathrm{U}} \end{cases},$$

which yields the dual lower bound

$$\sup_{\mu \geq 0} q(\mu) = \frac{(y^{\mathrm{L}} + y^{\mathrm{U}} - y^{\mathrm{L}} y^{\mathrm{U}})(y^{\mathrm{L}} - 2)}{2 + y^{\mathrm{U}} - 2y^{\mathrm{L}}} + y^{\mathrm{L}}.$$

Consider $y^{\mathrm{L}} = 0.5 - \varepsilon$, $y^{\mathrm{U}} = 0.5 + \varepsilon$ for some $\varepsilon \in (0, 0.5)$. The dual lower bound reduces to

$$\sup_{\mu \geq 0} q(\mu) = \frac{-\varepsilon^3 - 4.5\varepsilon^2 - 0.75\varepsilon - 0.375}{1.5 + 3\varepsilon}.$$

Consequently,

$$\begin{aligned} & \min_{(x,y) \in \mathcal{F}(Z)} -xy \ - \ \sup_{\mu \geq 0} q(\mu) \\ = & -0.25 + \frac{\varepsilon^3 + 4.5\varepsilon^2 + 0.75\varepsilon + 0.375}{1.5 + 3\varepsilon} \\ = & -0.25 + \frac{1}{1.5}\left(\varepsilon^3 + 4.5\varepsilon^2 + 0.75\varepsilon + 0.375\right)(1 + 2\varepsilon)^{-1} \end{aligned}$$

$$
\begin{aligned}
&= -0.25 + \frac{1}{1.5}\left(\varepsilon^3 + 4.5\varepsilon^2 + 0.75\varepsilon + 0.375\right)\left(1 - 2\varepsilon + 4\varepsilon^2 + O(\varepsilon^3)\right) \\
&= 3\varepsilon^2 + O(\varepsilon^3) \\
&\leq \tau\varepsilon^2,
\end{aligned}
$$

for some constant $\tau > 0$ (where we may assume that the above inequality holds for $\varepsilon = 0.5$ as well).

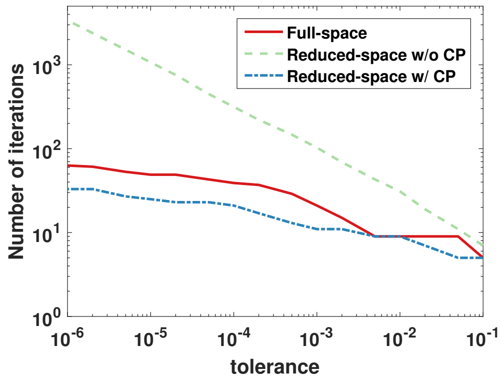
Consider any nondegenerate interval $Z = [y^{\mathrm{L}}, y^{\mathrm{U}}] \subset [0, 1]$ with $0.5 \in Z$ and construct $\bar{Z} \supset Z$ such that $\bar{Z} = [y^* - \varepsilon, y^* + \varepsilon]$ with $\varepsilon = \max\{y^{\mathrm{U}} - y^*, y^* - y^{\mathrm{L}}\}$. We have

$$
\begin{aligned}
&\min_{(x,y)\in\mathcal{F}(Z)} -xy \;-\; \sup_{\mu\geq 0}\min_{(x,y)\in X(Z)\times Z}\left[-xy + \mu g(x,y)\right] \\
\leq &\min_{(x,y)\in\mathcal{F}(\bar{Z})} -xy \;-\; \sup_{\mu\geq 0}\min_{(x,y)\in X(\bar{Z})\times \bar{Z}}\left[-xy + \mu g(x,y)\right] \\
\leq &\,\tau\varepsilon^2 \\
\leq &\,\tau w(Z)^2,
\end{aligned}
$$

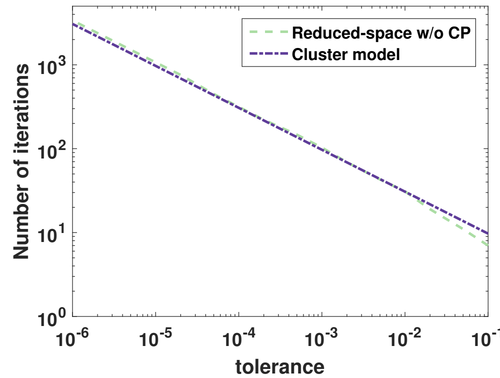
which implies that the reduced-space dual lower bounding scheme with constraint propagation is second-order convergent at y^*.

Figure 6-1 illustrates the performance of the lower bounding schemes considered in this chapter in a bare-bones branch-and-bound framework for Examples 6.5.19 and 6.5.21. The branch-and-bound framework was implemented in MATLAB®, and the (convex) lower bounding problems were solved using the CVX [88] package. The lowest lower bound node selection rule and the interval bisection branching rule (which bisects the domain of the variable whose interval has the largest width) were used by the branch-and-bound algorithm. Since Example 6.5.19 is not particularly challenging, it is assumed that a local solver finds its global solution at the root node of the branch-and-bound tree (i.e., the upper bound is set to the optimal objective value at the root node). In addition, the bounds on x and y were modified to $\left[-1, 1 - \frac{\sqrt{3}}{100}\right]$ and $\left[\frac{\sqrt{2}}{100}, 1\right]$, respectively, to prevent the full-space lower bounding schemes from branching at the optimal solution and (fortuitously) converging early (this modification enables a truer characterization of the convergence rates of the lower bounding schemes).

Figure 6-1a plots the number of iterations of the branch-and-bound algorithm versus the (absolute) termination tolerance for the full-space lower bounding schemes, the reduced-



(a) Comparison of the number of branch-and-bound iterations versus the termination tolerance between the different lower bounding schemes



(b) Comparison of the number of branch-and-bound iterations of the reduced-space lower bounding schemes without constraint propagation with the predictions from the cluster problem model for different termination tolerances

Figure 6-1: (Left Plot) Plots of the number of iterations of the branch-and-bound algorithm versus the absolute termination tolerance for the lower bounding schemes considered in this chapter for Example 6.5.19. The solid line indicates the number of iterations of the full-space lower bounding schemes, the dashed line indicates the number of iterations of the reduced-space lower bounding schemes without constraint propagation, and the dash-dotted line indicates the number of iterations of the reduced-space lower bounding schemes with constraint propagation. (Right Plot) Comparison of the number of iterations of the reduced-space branch-and-bound algorithms without constraint propagation for Example 6.5.19 with the corresponding cluster problem model. The dashed line indicates the number of iterations of the reduced-space lower bounding schemes without constraint propagation, and the dash-dotted line indicates the predicted number of iterations from the cluster problem model.

space lower bounding schemes without constraint propagation (see Example 6.5.19), and the reduced-space lower bounding schemes with constraint propagation (see Example 6.5.21). Note that both full-space (reduced-space) lower bounding schemes considered in this chapter result in the same lower bound for this problem (see Lemma 6.5.12). It can be seen that the full-space lower bounding schemes and the reduced-space lower bounding schemes with constraint propagation perform significantly better than the reduced-space lower bounding schemes without constraint propagation for small tolerances, and that they exhibit a much more favorable scaling with a decrease in the termination tolerance as well. Furthermore, the advantage of using constraint propagation techniques in the reduced-space lower bounding schemes is evident, and its use puts the reduced-space lower bounding schemes at an advantage compared to the full-space lower bounding schemes. Figure 6-1b illustrates that the dependence of the number of iterations on the termination tolerance for the reduced-space lower bounding schemes without constraint propagation is in good agreement with their associated cluster problem models (see [108, Theorem 3] for the details of the cluster problem model). Note that the prediction of the number of iterations from the cluster problem model in Figure 6-1b is obtained by fitting the prefactor in the cluster model (i.e., intercept of the line in the plot; the slope of the line is determined by the cluster model using the estimate of the convergence order of the lower bounding scheme obtained from this chapter) against the number of iterations obtained from the computational experiments. It is worth mentioning at this stage that only basic versions of the lower bounding schemes considered in this chapter have been used to generate Figure 6-1; the performance of the lower bounding schemes may be significantly different if they are implemented within a state-of-the-art branch-and-bound framework that solves additional subproblems to speed up their convergence.

The following example illustrates another instance of Problem (P) for which constraint propagation plays a crucial rule in boosting the convergence order of the convex relaxation-based reduced-space lower bounding scheme in [76].

Example 6.5.22. Consider the following instance of Problem (P):

$$\begin{aligned}
\min_{x,y} \;\; & \exp(x) - 4x + y \\
\text{s.t.} \;\; & x^2 + x\exp(3-y) \leq 10, \\
& x \in [0.5, 2], y \in [-1, 1].
\end{aligned}$$

The optimal solution of the above problem, which is a constrained minimum, is $(x^*, y^*) \approx (1.029, 0.838)$ $\Big($the 'exact' optimal solution can be determined as follows: x^* is the (unique real) root of the function $(4 - \exp(x))(10x - x^3) - x^2 - 10$ in $[0.5, 2]$, and $y^* := 3 - \ln\left(\frac{10-(x^*)^2}{x^*}\right)\Big)$ with optimal objective value approximately equal to -0.480. The reader can verify that (x^*, y^*, μ^*) is a KKT point for Problem (P), where $\mu^* := \frac{1}{x^* \exp(3-y^*)}$. This implies, in particular, that the full-space lower bounding schemes in Section 6.4 can be designed to be at least second-order convergent at (x^*, y^*) (see Theorem 6.4.27 and Corollary 6.4.28). The reader can also verify that second-order convergence of the lower bounding scheme may be sufficient to mitigate the cluster problem around (x^*, y^*) (see Chapter 5).

Since all of the functions in the above instance of Problem (P) are in the form (W), both the reduced-space lower bounding schemes considered in this section can be employed to solve it. The ensuing arguments show that the convex relaxation-based reduced-space lower bounding scheme in [76] is only first-order convergent at y^* when constraint propagation techniques are not used.

Consider $[y^{\mathrm{L}}, y^{\mathrm{U}}] := [y^* - \varepsilon, y^* + \varepsilon] =: Z \in \mathbb{I}Y$ with $0 < \varepsilon \leq 0.1$. The reduced-space lower bounding scheme in [76] yields

$$\begin{aligned}
\mathcal{O}(Z) = \min_{x,y} \ & \exp(x) - 4x + y \\
\text{s.t. } & x^2 + 2\exp(3 - y) + x\exp(3 - y^{\mathrm{L}}) - 2\exp(3 - y^{\mathrm{L}}) \leq 10, \\
& x^2 + 0.5\exp(3 - y) + x\exp(3 - y^{\mathrm{U}}) - 0.5\exp(3 - y^{\mathrm{U}}) \leq 10, \\
& x \in [0.5, 2], y \in [y^{\mathrm{L}}, y^{\mathrm{U}}].
\end{aligned}$$

Note that the point

$$(x^{\mathrm{f}}_Z, y^{\mathrm{f}}_Z) := \left(\frac{\sqrt{(\exp(3 - y^{\mathrm{U}}))^2 + 40 + 2\left(\exp(3 - y^{\mathrm{U}}) - \exp(3 - y^*)\right)} - \exp(3 - y^{\mathrm{U}})}{2}, y^* \right)$$

is feasible for the above lower bounding scheme with objective value $\exp(x^{\mathrm{f}}_Z) - 4x^{\mathrm{f}}_Z + y^{\mathrm{f}}_Z$. Furthermore,

$$\begin{aligned}
& x^{\mathrm{f}}_Z - x^* \\
&= \left(\frac{\sqrt{(\exp(3 - y^{\mathrm{U}}))^2 + 40 + 2\left(\exp(3 - y^{\mathrm{U}}) - \exp(3 - y^*)\right)} - \exp(3 - y^{\mathrm{U}})}{2} - \right. \\
&\qquad \left. \frac{\sqrt{(\exp(3 - y^*))^2 + 40} - \exp(3 - y^*)}{2} \right) \\
&= \frac{\left(\left(\exp(3 - y^{\mathrm{U}})\right)^2 - (\exp(3 - y^*))^2 \right) + 2\left(\exp(3 - y^{\mathrm{U}}) - \exp(3 - y^*)\right)}{2\left(\sqrt{(\exp(3 - y^{\mathrm{U}}))^2 + 40 + 2\left(\exp(3 - y^{\mathrm{U}}) - \exp(3 - y^*)\right)} + \sqrt{(\exp(3 - y^*))^2 + 40} \right)} + \\
&\qquad \frac{\left(\exp(3 - y^*) - \exp(3 - y^{\mathrm{U}})\right)}{2} \\
&= \frac{\left(\exp(3 - y^{\mathrm{U}}) + \exp(3 - y^*) + 2\right)\left(\exp(3 - y^{\mathrm{U}}) - \exp(3 - y^*)\right)}{2\left(\sqrt{(\exp(3 - y^{\mathrm{U}}))^2 + 40 + 2\left(\exp(3 - y^{\mathrm{U}}) - \exp(3 - y^*)\right)} + \sqrt{(\exp(3 - y^*))^2 + 40} \right)} + \\
&\qquad \frac{\left(\exp(3 - y^*) - \exp(3 - y^{\mathrm{U}})\right)}{2}
\end{aligned}$$

$$\geq \frac{\left(\exp(3-y^*)+\exp(3-y^*)+2\right)\left(\exp(3-y^{\mathrm{U}})-\exp(3-y^*)\right)}{2\left(\sqrt{\left(\exp(3-y^*-0.1)\right)^2+40+2\left(\exp(3-y^*-0.1)-\exp(3-y^*)\right)}+\sqrt{\left(\exp(3-y^*)\right)^2+40}\right)}+$$

$$\frac{\left(\exp(3-y^*)-\exp(3-y^{\mathrm{U}})\right)}{2}$$

$$\geq 0.025\left(\exp(3-y^*)-\exp(3-y^{\mathrm{U}})\right)$$

$$=0.025\exp(3-y^*)\varepsilon+o(\varepsilon)$$

$$\geq 0.2\varepsilon+o(\varepsilon).$$

Therefore, we have

$$\begin{aligned}
&\min_{(x,y)\in\mathcal{F}(Z)} f(x,y) \;-\; \min_{(x,y)\in\mathcal{F}^{\mathrm{cv}}(Z)} f^{\mathrm{cv}}_{X(Z)\times Z}(x,y)\\
\geq\;& (\exp(x^*)-4x^*+y^*)-\left(\exp(x^{\mathrm{f}}_Z)-4x^{\mathrm{f}}_Z+y^{\mathrm{f}}_Z\right)\\
=\;& \left(\exp(x^*)-\exp(x^{\mathrm{f}}_Z)\right)+4\left(x^{\mathrm{f}}_Z-x^*\right)\\
=\;& (4-\exp(x^*))\left(x^{\mathrm{f}}_Z-x^*\right)+o\left(\left|x^{\mathrm{f}}_Z-x^*\right|\right)\\
\geq\;& \left(x^{\mathrm{f}}_Z-x^*\right)+o\left(\left|x^{\mathrm{f}}_Z-x^*\right|\right)\\
\geq\;& 0.2\varepsilon+o(\varepsilon)\\
\geq\;& 0.05w(Z)
\end{aligned}$$

for $\varepsilon \ll 1$, which establishes that the reduced-space lower bounding scheme in [76] has at most first-order convergence at y^* (note that first-order convergence of the scheme follows from Corollary 6.5.18). This is rather unfortunate because y^* can be seen to be an unconstrained minimizer of the reduced-space objective function $v : [-1,1] \to \mathbb{R}$, which can be derived (around $y = y^*$) to be

$$v(y)=\exp\left(x^*(y)\right)-4x^*(y)+y, \quad \forall y\in[0.5,1],$$

where $x^* : [0.5,1] \ni y \longmapsto [0.5,2]$ is given by

$$x^*(y) := \frac{\sqrt{\left(\exp(3-y)\right)^2+40}-\exp(3-y)}{2},$$

which implies that at least second-order convergence of the reduced-space lower bounding scheme at y^* is typically required to mitigate clustering [68, 238].

We next show that when constraint propagation is used to infer (exact) bounds for x on Z, second-order convergence of the reduced-space lower bounding scheme in [76] can be achieved. Note that for $[y^{\mathrm{L}}, y^{\mathrm{U}}] := [y^* - \varepsilon, y^* + \varepsilon] =: Z \in \mathbb{I}Y$ with $0 < \varepsilon \leq 0.1$, the best possible (interval) bounds that can be obtained for x are $x \in X(Z) := [x^{\mathrm{L}}_Z, x^{\mathrm{U}}_Z]$ with

$$x^{\mathrm{L}}_Z = 0.5, \quad x^{\mathrm{U}}_Z = \frac{\sqrt{(\exp(3 - y^{\mathrm{U}}))^2 + 40} - \exp(3 - y^{\mathrm{U}})}{2}.$$

The reduced-space lower bounding scheme in [76] with constraint propagation yields

$$\begin{aligned}
\mathcal{O}(Z) = \min_{x,y} \;\; & \exp(x) - 4x + y \\
\text{s.t. } & x^2 + x^{\mathrm{U}}_Z \exp(3 - y) + x \exp(3 - y^{\mathrm{L}}) - x^{\mathrm{U}}_Z \exp(3 - y^{\mathrm{L}}) \leq 10, \\
& x^2 + 0.5 \exp(3 - y) + x \exp(3 - y^{\mathrm{U}}) - 0.5 \exp(3 - y^{\mathrm{U}}) \leq 10, \\
& x \in [0.5, x^{\mathrm{U}}_Z], y \in [y^{\mathrm{L}}, y^{\mathrm{U}}].
\end{aligned}$$

By noticing that the first constraint in the above relaxation of Problem (P) is always active at the solution of the relaxed problem, we can reformulate the reduced-space lower bounding problem as

$$\mathcal{O}(Z) = \min_{y \in [y^{\mathrm{L}}, y^{\mathrm{U}}]} \exp\left(\bar{x}_Z(y)\right) - 4\left(\bar{x}_Z(y)\right) + y,$$

where $\bar{x}_Z : Z \ni y \longmapsto [0.5, x^{\mathrm{U}}_Z]$ is given by

$$\bar{x}_Z(y) := \frac{\sqrt{(\exp(3 - y^{\mathrm{L}}))^2 + 40 + 4x^{\mathrm{U}}_Z\left(\exp(3 - y^{\mathrm{L}}) - \exp(3 - y)\right)} - \exp(3 - y^{\mathrm{L}})}{2}.$$

We have

$$\begin{aligned}
& \bar{x}_Z(y) - x^*(y) \\
= & \left(\frac{\sqrt{(\exp(3 - y^{\mathrm{L}}))^2 + 40 + 4x^{\mathrm{U}}_Z\left(\exp(3 - y^{\mathrm{L}}) - \exp(3 - y)\right)} - \exp(3 - y^{\mathrm{L}})}{2} - \right. \\
& \left. \qquad \frac{\sqrt{(\exp(3 - y))^2 + 40} - \exp(3 - y)}{2} \right)
\end{aligned} \tag{6.11}$$

$$= \frac{\left(\left(\exp(3-y^{\mathrm{L}})\right)^2 - \left(\exp(3-y)\right)^2\right) + 4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}}) - \exp(3-y)\right)}{2\left(\sqrt{\left(\exp(3-y^{\mathrm{L}})\right)^2 + 40 + 4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}}) - \exp(3-y)\right)} + \sqrt{\left(\exp(3-y)\right)^2 + 40}\right)} +$$
$$\frac{\left(\exp(3-y) - \exp(3-y^{\mathrm{L}})\right)}{2}$$
$$= \frac{\left(\exp(3-y^{\mathrm{L}}) + \exp(3-y) + 4x_Z^{\mathrm{U}}\right)\left(\exp(3-y^{\mathrm{L}}) - \exp(3-y)\right)}{2\left(\sqrt{\left(\exp(3-y^{\mathrm{L}})\right)^2 + 40 + 4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}}) - \exp(3-y)\right)} + \sqrt{\left(\exp(3-y)\right)^2 + 40}\right)} -$$
$$\frac{\left(\exp(3-y^{\mathrm{L}}) - \exp(3-y)\right)}{2}. \tag{6.12}$$

We next establish the dependence of the different terms in Equation (6.11) on ε. We first derive an expression for $\exp(3-y^{\mathrm{L}}) + \exp(3-y) + 4x_Z^{\mathrm{U}}$.

$$\begin{aligned}
&\exp(3-y^{\mathrm{L}}) + \exp(3-y) + 4x_Z^{\mathrm{U}} \\
&= \exp(3-y^*+\varepsilon) + \exp(3-y) + 2\sqrt{\left(\exp(3-y^*-\varepsilon)\right)^2 + 40} - 2\exp(3-y^*-\varepsilon) \\
&= \exp(3-y^*) + \exp(3-y) + \varepsilon\exp(3-y^*) + O(\varepsilon^2) + 2\sqrt{\left(\exp(3-y^*)\right)^2\left[1-2\varepsilon+O(\varepsilon^2)\right] + 40} \\
&\qquad - 2\exp(3-y^*)\left[1-\varepsilon+O(\varepsilon^2)\right] \\
&= 2\sqrt{\left(\exp(3-y^*)\right)^2 + 40} + \exp(3-y) - \exp(3-y^*) + 3\exp(3-y^*)\varepsilon - \frac{2\left(\exp(3-y^*)\right)^2\varepsilon}{\sqrt{\left(\exp(3-y^*)\right)^2 + 40}} \\
&\qquad + O(\varepsilon^2).
\end{aligned}$$

Next, we derive an expression for $4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}}) - \exp(3-y)\right)$.

$$\begin{aligned}
&4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}}) - \exp(3-y)\right) \\
&= \left(2\sqrt{\left(\exp(3-y^*-\varepsilon)\right)^2 + 40} - 2\exp(3-y^*-\varepsilon)\right)\left(\exp(3-y^*+\varepsilon) - \exp(3-y)\right) \\
&= \left(2\sqrt{\left(\exp(3-y^*)\right)^2 + 40} - 2\exp(3-y^*)\right)\left(\exp(3-y^*+\varepsilon) - \exp(3-y)\right) + O(\varepsilon^2) \\
&= \left(2\sqrt{\left(\exp(3-y^*)\right)^2 + 40} - 2\exp(3-y^*)\right)\left(\exp(3-y^*) - \exp(3-y) + \exp(3-y^*)\varepsilon\right) + O(\varepsilon^2).
\end{aligned}$$

Finally, we consider the expression $\sqrt{\left(\exp(3-y^{\mathrm{L}})\right)^2 + 40 + 4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}}) - \exp(3-y)\right)} + \sqrt{\left(\exp(3-y)\right)^2 + 40}$.

$$\begin{aligned}
&\sqrt{\left(\exp(3-y^{\mathrm{L}})\right)^2 + 40 + 4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}}) - \exp(3-y)\right)} + \sqrt{\left(\exp(3-y)\right)^2 + 40} \\
&= \sqrt{\left(\exp(3-y^*+\varepsilon)\right)^2 + 40 + 4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}}) - \exp(3-y)\right)} + \sqrt{\left(\exp(3-y)\right)^2 + 40} \\
&= \sqrt{\left(\exp(3-y^*)\right)^2 + 40}\sqrt{1 + \frac{4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}}) - \exp(3-y)\right) + 2\left(\exp(3-y^*)\right)^2\varepsilon + O(\varepsilon^2)}{\left(\exp(3-y^*)\right)^2 + 40}} +
\end{aligned}$$

$$
\begin{aligned}
&\sqrt{(\exp(3-y))^2+40}\\
=&\sqrt{(\exp(3-y^*))^2+40}+\sqrt{(\exp(3-y))^2+40}+\frac{(\exp(3-y^*))^2\,\varepsilon}{\sqrt{(\exp(3-y^*))^2+40}}+\\
&\frac{\left(\sqrt{(\exp(3-y^*))^2+40}-\exp(3-y^*)\right)(\exp(3-y^*)-\exp(3-y)+\exp(3-y^*)\varepsilon)}{\sqrt{(\exp(3-y^*))^2+40}}+O(\varepsilon^2).
\end{aligned}
$$

Substituting the above expressions in Equation (6.11), we get

$$
\begin{aligned}
&\bar{x}_Z(y)-x^*(y)\\
=&\frac{\alpha\left(\exp(3-y^{\mathrm{L}})-\exp(3-y)\right)}{2\left(\sqrt{(\exp(3-y^{\mathrm{L}}))^2+40+4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}})-\exp(3-y)\right)}+\sqrt{(\exp(3-y))^2+40}\right)},
\end{aligned}
$$

with

$$
\begin{aligned}
\alpha :=& \sqrt{(\exp(3-y^*))^2+40}-\sqrt{(\exp(3-y))^2+40}+\exp(3-y)-\exp(3-y^*)-\\
&\frac{\left(\sqrt{(\exp(3-y^*))^2+40}-\exp(3-y^*)\right)(\exp(3-y^*)-\exp(3-y))}{\sqrt{(\exp(3-y^*))^2+40}}+3\exp(3-y^*)\varepsilon-\\
&\frac{3\,(\exp(3-y^*))^2\,\varepsilon}{\sqrt{(\exp(3-y^*))^2+40}}-\frac{\left(\sqrt{(\exp(3-y^*))^2+40}-\exp(3-y^*)\right)\exp(3-y^*)\varepsilon}{\sqrt{(\exp(3-y^*))^2+40}}+O(\varepsilon^2)\\
=&\left(\frac{\exp(3-y^*)+\exp(3-y)}{\sqrt{(\exp(3-y^*))^2+40}+\sqrt{(\exp(3-y))^2+40}}-1\right)(\exp(3-y^*)-\exp(3-y))-\\
&\frac{\left(\sqrt{(\exp(3-y^*))^2+40}-\exp(3-y^*)\right)(\exp(3-y^*)-\exp(3-y))}{\sqrt{(\exp(3-y^*))^2+40}}+3\exp(3-y^*)\varepsilon-\\
&\frac{3\,(\exp(3-y^*))^2\,\varepsilon}{\sqrt{(\exp(3-y^*))^2+40}}-\frac{\left(\sqrt{(\exp(3-y^*))^2+40}-\exp(3-y^*)\right)\exp(3-y^*)\varepsilon}{\sqrt{(\exp(3-y^*))^2+40}}+O(\varepsilon^2)\\
\leq& \hat{\tau}\varepsilon+O(\varepsilon^2)
\end{aligned}
$$

for some $\hat{\tau}\geq 0$ since $y\in Z=[y^{\mathrm{L}},y^{\mathrm{U}}]$ with $w(Z)=O(\varepsilon)$ and each term in the expression for α is $O(\varepsilon)$. Note that $\alpha\geq 0$ (since $\bar{x}_Z(y)\geq x^*(y)$).

Consequently, we have $\forall y \in Z$ that

$$\begin{aligned}
&\bar{x}_Z(y) - x^*(y) \\
\leq\ & \frac{\left(\hat{\tau}\varepsilon + O(\varepsilon^2)\right)\left(\exp(3-y^{\mathrm{L}}) - \exp(3-y)\right)}{2\left(\sqrt{(\exp(3-y^{\mathrm{L}}))^2 + 40 + 4x_Z^{\mathrm{U}}\left(\exp(3-y^{\mathrm{L}}) - \exp(3-y)\right)} + \sqrt{(\exp(3-y))^2 + 40}\right)} \\
\leq\ & \bar{\tau}\varepsilon^2 + O(\varepsilon^3)
\end{aligned}$$

for some $\bar{\tau} \geq 0$, since $\exp(3-y^{\mathrm{L}}) - \exp(3-y)$ is $O(\varepsilon)$. Note that $\bar{x}_Z(y) \geq x^*(y)$, $\forall Z$. Therefore, on intervals $[y^* - \varepsilon, y^* + \varepsilon] =: Z \in \mathbb{I}Y$ with $0 < \varepsilon \leq 0.1$, we have

$$\begin{aligned}
&\min_{(x,y)\in\mathcal{F}(Z)} f(x,y) \ - \min_{(x,y)\in\mathcal{F}^{\mathrm{cv}}(Z)} f^{\mathrm{cv}}_{X(Z)\times Z}(x,y) \\
&= \min_{y\in Z} f(x^*(y),y) - \min_{y\in Z} f(\bar{x}_Z(y),y) \\
&\leq \max_{y\in Z} |f(x^*(y),y) - f(\bar{x}_Z(y),y)| \\
&= \max_{y\in Z} |\exp(x^*(y)) - \exp(\bar{x}_Z(y)) + 4\bar{x}_Z(y) - 4x^*(y)| \\
&= \max_{y\in Z} |(4 - \exp(x^*(y)))\left(\bar{x}_Z(y) - x^*(y)\right) + o\left(\bar{x}_Z(y) - x^*(y)\right)| \\
&\leq \max_{y\in Z} |2\left(\bar{x}_Z(y) - x^*(y)\right) + o\left(\bar{x}_Z(y) - x^*(y)\right)| \\
&\leq 2\bar{\tau}\varepsilon^2 + o\left(\varepsilon^2\right) \\
&\leq \bar{\tau}w(Z)^2
\end{aligned}$$

for $\varepsilon \ll 1$, which establishes second-order convergence of the scheme at y^* when restricted to symmetric intervals around y^*.

Consider any nondegenerate interval $Z = [y^{\mathrm{L}}, y^{\mathrm{U}}] \in \mathbb{I}Y$ with $y^* \in Z$ and $w(Z) \leq 0.1$, and construct $\bar{Z} \supset Z$ such that $\bar{Z} = [y^* - \varepsilon, y^* + \varepsilon]$ with $\varepsilon = \max\{y^{\mathrm{U}} - y^*, y^* - y^{\mathrm{L}}\}$. We have

$$\begin{aligned}
\min_{(x,y)\in\mathcal{F}(Z)} f(x,y) \ - \min_{(x,y)\in\mathcal{F}^{\mathrm{cv}}(Z)} f^{\mathrm{cv}}_{X(Z)\times Z}(x,y) &\leq \min_{(x,y)\in\mathcal{F}(\bar{Z})} f(x,y) \ - \min_{(x,y)\in\mathcal{F}^{\mathrm{cv}}(\bar{Z})} f^{\mathrm{cv}}_{X(Z)\times Z}(x,y) \\
&\leq \bar{\tau}w(\bar{Z})^2 \\
&\leq 4\bar{\tau}w(Z)^2,
\end{aligned}$$

which implies that the convex relaxation-based reduced-space dual lower bounding scheme with constraint propagation is second-order convergent at y^*.

Finally, we show that the reduced-space dual lower bounding scheme in [69] has at least second-order convergence at y^* even when constraint propagation is not used to infer bounds on x. Consider $[y^{\mathrm{L}}, y^{\mathrm{U}}] =: Z \in \mathbb{I}Y$ with $w(Z) \leq 0.1$. The feasible region of the original problem on Z is given by

$$\mathcal{F}(Z) = \left\{(x, y) \in [0.5, 2] \times [y^{\mathrm{L}}, y^{\mathrm{U}}] : x \leq x^*(y)\right\}.$$

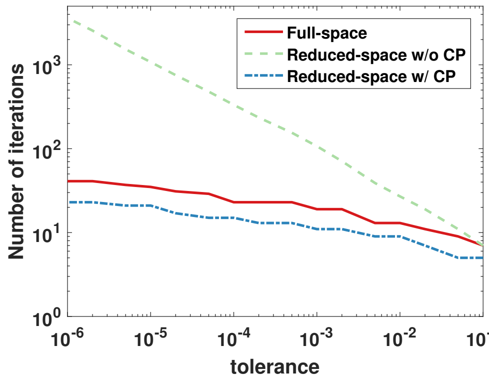
The convex hull of the feasible region on Z is given by

$$\mathrm{conv}(\mathcal{F}(Z)) = \left\{(x, y) \in [0.5, 2] \times [y^{\mathrm{L}}, y^{\mathrm{U}}] : x \leq x_Z^{*,\mathrm{cc}}(y)\right\},$$

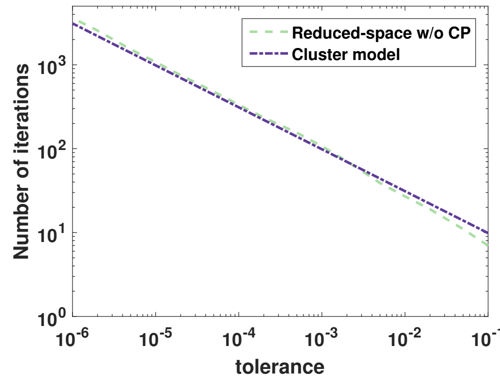
where $x_Z^{*,\mathrm{cc}}$ denotes the concave envelope of x^* on Z. It is not hard to see that we have $d_H\left(\mathcal{F}(Z), \mathrm{conv}(\mathcal{F}(Z))\right) \leq \tilde{\tau} w(Z)^2$ for some $\tilde{\tau} \geq 0$ (this partly follows from the fact that x^* is twice continuously differentiable on $[y^* - 0.1, y^* + 0.1]$ and the fact that $\left(x_Z^{*,\mathrm{cc}}\right)_{Z \in \mathbb{I}Y}$ converges pointwise to x^* with order at least two on $[y^* - 0.1, y^* + 0.1]$). Since the dual lower bounding scheme produces a lower bound that is at least as tight as any convex relaxation-based scheme, we have

$$\begin{aligned}
&\min_{(x,y) \in \mathcal{F}(Z)} f(x, y) \;-\; \sup_{\mu \geq 0} \min_{(x,y) \in X(Z) \times Z} \left[f(x, y) + \mu g(x, y)\right] \\
&\leq \min_{(x,y) \in \mathcal{F}(Z)} f(x, y) - \min_{(x,y) \in \mathrm{conv}(\mathcal{F}(Z))} f(x, y) \\
&= f(x^*, y^*) - f(\tilde{x}_Z, \tilde{y}_Z) \\
&\leq f(\hat{x}_Z, \hat{y}_Z) - f(\tilde{x}_Z, \tilde{y}_Z) \\
&\leq M_f \|(\hat{x}_Z, \hat{y}_Z) - (\tilde{x}_Z, \tilde{y}_Z)\| \\
&\leq M_f \tilde{\tau} w(Z)^2,
\end{aligned}$$

where $(\tilde{x}_Z, \tilde{y}_Z) \in \operatorname*{arg\,min}_{(x,y) \in \mathrm{conv}(\mathcal{F}(Z))} f(x, y)$, the point $(\hat{x}_Z, \hat{y}_Z) \in \mathcal{F}(Z)$ is chosen such that $\|(\hat{x}_Z, \hat{y}_Z) - (\tilde{x}_Z, \tilde{y}_Z)\| \leq \tilde{\tau} w(Z)^2$, and M_f denotes the Lipschitz constant of f on $[0.5, 2] \times [-1, 1]$. Since the Lagrangian dual-based reduced-space lower bounding scheme is at least first-order convergent at y^* from Theorem 6.5.17, it is at least second-order convergent at y^* by analogy to Lemma 6.3.8.



(a) Comparison of the number of branch-and-bound iterations versus the termination tolerance between the convex relaxation-based lower bounding schemes



(b) Comparison of the number of branch-and-bound iterations of the convex relaxation-based reduced-space lower bounding scheme without constraint propagation with the predictions from the cluster problem model for different termination tolerances

Figure 6-2: (Left Plot) Plots of the number of iterations of the branch-and-bound algorithm versus the absolute termination tolerance for the full-space and reduced-space convex relaxation-based lower bounding schemes considered in this chapter for Example 6.5.22. The solid line indicates the number of iterations of the convex relaxation-based full-space lower bounding scheme, the dashed line indicates the number of iterations of the convex relaxation-based reduced-space lower bounding scheme without constraint propagation, and the dash-dotted line indicates the number of iterations of the convex relaxation-based reduced-space lower bounding scheme with constraint propagation. (Right Plot) Comparison of the number of iterations of the convex relaxation-based reduced-space branch-and-bound algorithm without constraint propagation with the corresponding cluster problem model. The dashed line indicates the number of iterations of the convex relaxation-based reduced-space lower bounding scheme without constraint propagation, and the dash-dotted line indicates the predicted number of iterations from the cluster problem model.

Figure 6-2 illustrates the performance of the convex relaxation-based full-space and reduced-space lower bounding schemes in the bare-bones branch-and-bound implementation for Example 6.5.22 (note that we do not consider the Lagrangian dual-based full-space and reduced-space lower bounding schemes for the numerical experiments for this example because we do not have closed-form expressions for the lower bounds obtained using those schemes). Once again, the convex lower bounding problems were solved using the CVX [88] package, and the lowest lower bound node selection rule and the interval bisection branching rule were used by the branch-and-bound algorithm. Since Example 6.5.22 is not particularly challenging, we assume that a local solver finds its global solution at the root node of the branch-and-bound tree (i.e., the upper bound is set to the optimal objective value of the problem at the root node).

Figure 6-2a plots the number of iterations of the branch-and-bound algorithm versus the (absolute) termination tolerance for the full-space convex relaxation-based lower bounding scheme, the reduced-space convex relaxation-based lower bounding scheme without constraint propagation, and the reduced-space convex relaxation-based lower bounding scheme with constraint propagation. It can be seen that the full-space lower bounding scheme and the reduced-space lower bounding scheme with constraint propagation perform significantly better (for small tolerances) and exhibit a much more favorable scaling with a decrease in the termination tolerance compared to the reduced-space lower bounding scheme without constraint propagation. Furthermore, there is a clear advantage in using constraint propagation techniques in the reduced-space lower bounding scheme, and its use makes the performance of the reduced-space lower bounding scheme superior to that of the full-space lower bounding scheme for this example. Figure 6-2b shows that the number of iterations versus the termination tolerance for the reduced-space lower bounding scheme without constraint propagation closely follows the prediction from its associated cluster problem model (see Chapter 5 for the details of the cluster problem model). Note, once again, that the prediction of the number of iterations from the cluster problem model in Figure 6-2b is obtained by fitting the prefactor in the cluster model against the number of iterations obtained from the computational experiments. We wish to reiterate that only basic versions of the convex relaxation-based lower bounding schemes have been used to generate Figure 6-2; the performance of the lower bounding schemes may be significantly different if they are implemented within a state-of-the-art branch-and-bound framework that solves additional subproblems to speed up their convergence.

The following result shows that the reduced-space dual lower bounding scheme is second-order convergent at KKT points even in the absence of constraint propagation when all of the functions in Problem (P) are twice continuously differentiable and separable in $\mathbf{x}$ and $\mathbf{y}$.

Theorem 6.5.23. Consider Problem (P), and suppose f, g_j, $j = 1, \cdots, m_I$, and h_k, $k = 1, \cdots, m_E$, are separable in $\mathbf{x}$ and $\mathbf{y}$. Suppose $\text{int}(X \times Y)$ is nonempty, and f, $\mathbf{g}$, and $\mathbf{h}$ are twice continuously differentiable on $\text{int}(X \times Y)$. Furthermore, suppose there exists $(\mathbf{x}^*, \mathbf{y}^*) \in \text{int}(X \times Y)$, $\boldsymbol{\mu}^* \in \mathbb{R}^{m_I}_+$, $\boldsymbol{\lambda}^* \in \mathbb{R}^{m_E}$ such that $(\mathbf{x}^*, \mathbf{y}^*, \boldsymbol{\mu}^*, \boldsymbol{\lambda}^*)$ is a KKT point for Problem (P). The reduced-space dual lower bounding scheme is at least second-order convergent at $\mathbf{y}^*$.

Proof. Let $L(\mathbf{x},\mathbf{y},\boldsymbol{\mu},\boldsymbol{\lambda}) := f(\mathbf{x},\mathbf{y})+\boldsymbol{\mu}^{\mathrm{T}}\mathbf{g}(\mathbf{x},\mathbf{y})+\boldsymbol{\lambda}^{\mathrm{T}}\mathbf{h}(\mathbf{x},\mathbf{y})$ denote the Lagrangian of Problem (P). Since we are concerned about the convergence order at the reduced-space feasible point $\mathbf{y}^*$, it suffices to show the existence of $\tau \geq 0$ such that for every $Z \in \mathbb{I}Y$ with $\mathbf{y}^* \in Z$,

$$\min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}(Z)} f(\mathbf{x},\mathbf{y}) - \sup_{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}} \min_{(\mathbf{x},\mathbf{y})\in X\times Z} L(\mathbf{x},\mathbf{y},\boldsymbol{\mu},\boldsymbol{\lambda}) \leq \tau w(Z)^2.$$

We have

$$\begin{aligned}
&\sup_{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}} \min_{(\mathbf{x},\mathbf{y})\in X\times Z} L(\mathbf{x},\mathbf{y},\boldsymbol{\mu},\boldsymbol{\lambda}) \\
&\geq \min_{(\mathbf{x},\mathbf{y})\in X\times Z} L(\mathbf{x},\mathbf{y},\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*) \\
&\geq \min_{(\mathbf{x},\mathbf{y})\in X\times Z} \Big[L(\mathbf{x}^*,\mathbf{y},\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*) + \nabla_{\mathbf{x}} L(\mathbf{x}^*,\mathbf{y},\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}^*)\Big] \\
&= \min_{(\mathbf{x},\mathbf{y})\in X\times Z} \Big[L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*) + \nabla_{\mathbf{x}} L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}^*) \\
&\qquad\qquad + \Big(\nabla_{\mathbf{y}}\left(\nabla_{\mathbf{x}} L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}^*)\right)\Big)^{\mathrm{T}}(\mathbf{y}-\mathbf{y}^*) \\
&\qquad\qquad + \nabla_{\mathbf{y}} L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)^{\mathrm{T}}(\mathbf{y}-\mathbf{y}^*) - O(w(Z)^2)\Big] \\
&= \min_{(\mathbf{x},\mathbf{y})\in X\times Z} \left[f(\mathbf{x}^*,\mathbf{y}^*) - O(w(Z)^2)\right] \\
&\geq f(\mathbf{x}^*,\mathbf{y}^*) - O(w(Z)^2).
\end{aligned}$$

Note that we have used the fact that L is partly convex with respect to $\mathbf{x}$ in Step 2, that $L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*) = f(\mathbf{x}^*,\mathbf{y}^*)$, $\nabla_{\mathbf{x}} L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*) = \mathbf{0}$, $\nabla_{\mathbf{y}} L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*) = \mathbf{0}$ in Step 4 since it is assumed that $(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)$ is a KKT point for Problem (P), and that $\nabla_{\mathbf{y}}\Big(\nabla_{\mathbf{x}} L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*)^{\mathrm{T}}(\mathbf{x}-\mathbf{x}^*)\Big) = \mathbf{0}$ in Step 4 by virtue of the assumption that the Lagrangian is separable in $\mathbf{x}$ and $\mathbf{y}$. Therefore,

$$\min_{(\mathbf{x},\mathbf{y})\in\mathcal{F}(Z)} f(\mathbf{x},\mathbf{y}) - \sup_{\boldsymbol{\mu}\geq\mathbf{0},\boldsymbol{\lambda}} \min_{(\mathbf{x},\mathbf{y})\in X\times Z} L(\mathbf{x},\mathbf{y},\boldsymbol{\mu},\boldsymbol{\lambda}) \leq O(w(Z)^2),$$

which establishes the existence of τ for all $Z \in \mathbb{I}Y$ with $\mathbf{y}^* \in Z$ by analogy to Lemma 6.3.8. □

Note that the assumption of separability in Theorem 6.5.23 can be replaced with the weaker assumption that $\nabla^2_{\mathbf{xy}} L(\mathbf{x}^*,\mathbf{y}^*,\boldsymbol{\mu}^*,\boldsymbol{\lambda}^*) = \mathbf{0}$ (equals the zero matrix).

Remark 6.5.24. Similar to Corollary 6.5.18, it can be shown that the reduced-space lower bounding scheme in [76] has second-order convergence at KKT points even in the absence of constraint propagation when all of the functions in Problem (P) are separable in $\mathbf{x}$ and $\mathbf{y}$ and second-order pointwise convergent schemes of relaxations are used. Furthermore, under the above assumption of separability, the reduced-space lower bounding schemes in [76] and [69] can be shown to possess second-order convergence at infeasible points and unconstrained points in the reduced-space under suitable assumptions on the lower bounding schemes (see Remark 6.5.2). Consequently, the convergence properties of the reduced-space lower bounding schemes considered in this section are similar to their counterpart full-space lower bounding schemes in Section 6.4 when all of the functions in Problem (P) are twice continuously differentiable and separable in $\mathbf{x}$ and $\mathbf{y}$. Example 6.4.29 provides an instance wherein the convergence order is exactly two at $\mathbf{y}^*$ under the assumptions of Theorem 6.5.23.

6.6 Conclusion

A definition of convergence order for constrained problems has been introduced. The definition reduces to previously developed notions of convergence order for the case of unconstrained problems. An analysis of the convergence order of some full-space and reduced-space branch-and-bound algorithms has been performed.

It has been shown that convex relaxation-based full-space lower bounding schemes enjoy first-order convergence under mild assumptions and second-order convergence at KKT points when second-order pointwise convergent schemes of relaxations of the objective and the constraints are used. Furthermore, the importance of a sufficiently high convergence order at nearly-feasible points has been demonstrated. Lagrangian dual-based full-space lower bounding schemes have been shown to have at least as large a convergence order as convex relaxation-based lower bounding schemes. In addition, it has been shown that Lagrangian dual-based lower bounding schemes where the dual function is not exactly optimized still enjoy first-order convergence.

The convergence order of the reduced-space convex relaxation-based lower bounding scheme of Epperly and Pistikopoulos has been investigated, and it has been shown that the scheme enjoys first-order convergence under certain assumptions. However, their scheme can have as low as first-order convergence even at unconstrained points which can lead to clustering. It has also been shown that the reduced-space dual lower bounding scheme enjoys first-order convergence and that its convergence order may be as low as one for constrained problems. In that regard, the importance of constraint propagation in boosting the convergence order of reduced-space lower bounding schemes has been demonstrated. Furthermore, it has been shown that when all of the functions in Problem (P) are twice continuously differentiable and separable in $\mathbf{x}$ and $\mathbf{y}$, the above reduced-space lower bounding schemes can achieve second-order convergence at KKT points, at unconstrained points in the reduced-space, and at infeasible points.

Future work involves determining whether full-space lower bounding schemes can achieve second-order convergence on a neighborhood of constrained minima that are KKT points (such a result may be required to mitigate the cluster problem at such constrained minima - see Proposition 5.3.7 in Chapter 5, for instance), analyzing the convergence orders of some other widely-applicable reduced-space lower bounding schemes in the literature (see, for example, [217]), and determining sufficient conditions on the constraint propagation scheme to ensure second-order convergence of reduced-space lower bounding schemes at constrained minima that satisfy certain regularity conditions.

Chapter 7

Conclusion

This thesis has addressed two problems of topical interest in the field of mathematical optimization. The first problem addressed was on the development of algorithms and software for optimization problems under uncertainty that are modeled as two-stage stochastic programs, which has several diverse applications such as the design and operation of chemical process systems, planning and scheduling of energy generation systems, environmental management, supply chain optimization, and portfolio optimization. The prevalence of nonconvexities in chemical process models, in particular, precludes the use of most existing decomposition algorithms and associated software implementations for stochastic programs, thereby acting as a major motivating factor for our work that aimed at developing an efficient and accessible numerical implementation for solving such problems. The second problem tackled was on the development of a theory of convergence order for B&B algorithms for constrained optimization problems in conjunction with an analysis of the cluster problem in constrained global optimization to determine whether candidate B&B algorithms have favorable asymptotic convergence properties. Such an analysis can help explain disparities in the empirical performances of classes of B&B algorithms in the literature, help identify any deficiencies in such algorithms, and possibly provide guidelines for their improvement.

7.1 Summary of contributions

The most tangible results of this thesis are the algorithms and software developed for solving a broad class of two-stage stochastic MINLPs. In Chapter 3, a modified Lagrangian relaxation (MLR) algorithm was developed in which only the nonanticipativity constraints for the continuous complicating variables were dualized, and the resulting dual lower bounding problem was solved in a decomposable manner using NGBD. In addition, tailored, decomposable bounds tightening techniques were developed by building on previous work to accelerate potentially the convergence of the MLR algorithm. From a theoretical standpoint, we proved that the lower bounding problem of the MLR algorithm provides tighter lower bounds than the lower bounding problem of the conventional Lagrangian relaxation (LR) algorithm, and established finite-ε convergence of the MLR algorithm for the case when all of its subproblems are solved in a decomposable manner[1]. Finally, we showed that the MLR algorithm performs favorably compared to the LR algorithm and four commercial general-purpose global optimization software on a tank sizing and scheduling case study from the literature.

In Chapter 4, we detailed our development of the first (known) decomposition software, named `GOSSIP`, for solving two-stage stochastic MINLPs. `GOSSIP` includes implementations of NGBD, LR, and MLR algorithms along with adaptations of advanced techniques from the literature for reformulating and preprocessing user input, constructing relaxations, and domain reduction. `GOSSIP` has been benchmarked on a test library comprising a diverse set of problems that were compiled primarily from the process systems engineering literature. We plan on making `GOSSIP` available, without charge, for academic research soon, and will also make the `GOSSIP` test library accessible to the public.

Chapters 5 and 6 developed a theory of convergence order for B&B algorithms for constrained problems, and analyzed the consequences of a B&B scheme enjoying a particular convergence order on its asymptotic efficiency. In Chapter 6, we developed a notion of convergence order for lower bounding schemes for constrained problems that generalizes the definition for the unconstrained setting. A consequence of a lower bounding scheme possessing a certain convergence order was studied in Chapter 5, where we generalized previous analyses of the cluster problem in unconstrained optimization to the constrained case by conservatively estimating the number of boxes visited by B&B algorithms that use such schemes in vicinities of global minimizers. As a conclusion of this analysis, we showed that first-order convergent lower bounding schemes may be sufficient to mitigate the cluster problem at constrained minima under suitable conditions. This result is in contrast to the

[1] We note that our analysis of the convergence of the MLR algorithm in Chapter 3 can be naturally extended to analyze the convergence of the conventional LR algorithm as well.

unconstrained case where it is known that second-order convergence of the bounding scheme is required to mitigate clustering. For problems with equality constraints, our analysis in Chapter 5 has determined that second-order convergence of the lower bounding scheme is typically required to mitigate clustering at nonisolated minima. We also developed sufficient conditions under which second-order convergence of the lower bounding scheme mitigates the cluster problem.

Our analysis of the convergence orders of lower bounding schemes in Chapter 6 established lower bounds on the convergence orders of convex relaxation-based and Lagrangian duality-based lower bounding schemes in the literature. In particular, we developed sufficient conditions for the convergence order of such schemes at infeasible points to be of a certain order, and established sufficient conditions for such schemes to have first or second-order convergence at feasible points. We determined that the pointwise convergence orders of the schemes of relaxations employed by convex relaxation-based lower bounding schemes plays a crucial role for such lower bounding schemes to achieve a high convergence order. This is a departure from the analysis for the unconstrained case where the Hausdorff convergence order of the schemes of relaxations employed dictates the convergence order of the bounding scheme, and has resulted in some misconceptions in the convergence order literature in the past (see Chapters 5 and 8 of [237], for instance). Chapter 6 also analyzed the convergence orders of some widely applicable reduced-space B&B algorithms in the literature and determined that such methods may suffer from the cluster problem if domain reduction techniques are not employed for the variables whose domains are not partitioned by those algorithms. This analysis emphasizes the importance of domain reduction techniques for reduced-space B&B algorithms to mitigate clustering, whereas such techniques are usually thought to be optional for favorable convergence rates of B&B algorithms in general[2]. We note that our analyses of the cluster problem and convergence order for constrained problems can help explain differences in the performances of interval arithmetic-based full-space B&B algorithms and convex relaxation-based full-space B&B algorithms; this is because interval arithmetic-based full-space B&B algorithms typically possess first-order convergence at constrained minima, whereas convex relaxation-based full-space B&B algorithms employed by state-of-the-art global optimization software usu-

[2]While our analysis indicates that domain reduction techniques are not necessary for full-space B&B algorithms to mitigate clustering, such techniques usually empirically boost the convergence rates of full-space B&B algorithms.

ally enjoy second-order convergence at such points.

7.2 Avenues for future work

There are several questions and topics related to the contributions of this thesis that merit future investigation. On the topic of algorithm development for two-stage stochastic MINLPs, the development of alternative decomposition techniques that do not rely on Lagrangian duality (and, consequently, on nonsmooth optimization codes for solving the dual problem) seems a worthwhile endeavor since Lagrangian duality-based approaches appear to suffer from a few shortcomings, especially when we wish to solve problems to tight termination tolerances. In particular, future work could look to develop global optimization approaches like NGBD for general two-stage stochastic MINLPs that do not involve (explicit) branching on the (complicating) variables to guarantee convergence. It is anticipated, however, that such approaches do not exist for general stochastic programming formulations with continuous complicating variables.

The potential directions along which the development of `GOSSIP` can be continued are virtually endless. Perhaps the most pressing feature that needs to be developed within `GOSSIP` is the option to use a polyhedral relaxation framework when the relaxations constructed by `GOSSIP` are nonlinear. Not only will this feature help improve `GOSSIP`'s robustness while solving practical stochastic programs, but this development will also enable implementations of outer-approximation-based algorithms [72, 79, 118] for solving such problems. A close second is the need to exploit the parallelization capabilities of the algorithms implemented within `GOSSIP` to enable the efficient solution of large-scenario problems. Other features that could greatly enhance `GOSSIP`'s ability to solve stochastic programming applications of interest include: implementation of advanced relaxation strategies (such as the use of reformulation-linearization cuts [209], piecewise linear relaxations [160], and other specialized strategies [11, 120, 162, 164]), incorporating an array of scalable techniques for the generation of feasible points [107, 247], and diversifying `GOSSIP`'s portfolio of software links in a bid to leverage the advantages of multiple state-of-the-art software depending on the application. Of course, the ideal approach to future software development for stochastic programs would involve building on the framework of existing global optimization software so that the advanced strategies that are part of those software can be leveraged to push the envelope on the scale of optimization problems under uncertainty that can be solved to global optimality in practical solution times.

The convergence order analysis in Chapter 6 suggests some immediate avenues for future work. While it was determined therein that the use of domain reduction techniques is crucial towards mitigating the cluster problem faced by reduced-space B&B algorithms, establishing sufficient conditions on the domain reduction techniques so that the resulting reduced-space B&B algorithms can mitigate clustering remains an open problem. On a related note, convergence order analyses of some other widely applicable reduced-space B&B algorithms in the literature (for example, see [217]) is lacking, and remains an important topic for future work, especially considering that such algorithms have great potential in solving challenging process systems engineering applications. Since some of the hypotheses in the statements of the results for the cluster problem and convergence order analyses can be difficult to verify, the development of testable hypotheses for those results merits further work from a practical viewpoint (one approach to resolving this would be to develop particular convergence order results for classes of problems with specific structures). Another open problem arising out of practical considerations (that has also been posed by other researchers; see [38, Section 8], for instance) is the development of a non-asymptotic analysis of the performance of B&B algorithms. As detailed in Chapters 1 and 6, the cluster problem and convergence order analyses in this thesis only reflect on the asymptotic performance of B&B algorithms; therefore, a lower bounding scheme with a high convergence order may not necessarily be efficient/practical. This necessitates the development of theory that can not only capture the performance of lower bounding schemes on small interval subsets of the problem domain, but can also discern the behavior of their bounding error on larger intervals.

Appendix A

A note on the convergence order analysis of a class of reduced-space branch-and-bound algorithms

In this note, we consider the following nonlinear programming formulation (cf. Problem (P) in Chapter 6):

$$\begin{aligned} \min_{\mathbf{x},\mathbf{y}} \ & f(\mathbf{x},\mathbf{y}) \\ \text{s.t.} \ & \mathbf{g}(\mathbf{x},\mathbf{y}) \leq \mathbf{0}, \\ & \mathbf{h}(\mathbf{x},\mathbf{y}) = \mathbf{0}, \\ & \mathbf{x} \in X, \ \ \mathbf{y} \in Y, \end{aligned} \tag{P}$$

where $X \in \mathbb{IR}^{n_x}$ and $Y \in \mathbb{IR}^{n_y}$ are nonempty intervals, functions $f : \bar{X} \times \bar{Y} \to \mathbb{R}$, $\mathbf{g} : \bar{X} \times \bar{Y} \to \mathbb{R}^{m_I}$, and $\mathbf{h} : \bar{X} \times \bar{Y} \to \mathbb{R}^{n_x}$ are continuous, $\bar{X} \subset \mathbb{R}^{n_x}$ and $\bar{Y} \subset \mathbb{R}^{n_y}$ are open sets, and $X \subset \bar{X}$ and $Y \subset \bar{Y}$. Note that the dimension of the codomain of the equality constraint functions $\mathbf{h}$ equals the dimension of X.

When the dimension n_y of the Y-space is significantly less than the dimension n_x of the X-space (as is often the case in chemical process systems-related applications), we may wish to eliminate the equality constraints $\mathbf{h}(\mathbf{x},\mathbf{y}) = \mathbf{0}$ by 'solving for the $\mathbf{x}$ variables in terms of the $\mathbf{y}$ variables, and substitute the resulting solution in Problem (P)' to mitigate the worst-case exponential nature of B&B algorithms for global optimization. However, this is easier said than done for a few reasons: i. there may not exist a unique solution for $\mathbf{x}$ (in X) for each $\mathbf{y} \in Y$, ii. even when a unique solution for $\mathbf{x}$ exists in X for every $\mathbf{y} \in Y$, it may be tedious (or even impossible) to derive a closed-form expression that can be readily substituted in Problem (P), and iii. widely applicable deterministic global optimization frameworks such as branch-and-bound (see Section 2.3.2 of Chapter 2) require (global) relaxation information, and cannot make do with just a numerical solution for $\mathbf{x}$ at each $\mathbf{y} \in Y$.

Assume that for each $\mathbf{y} \in Y$, there exists a unique $\mathbf{x} \in X$ that satisfies the equality constraints $\mathbf{h}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$. By an abuse of notation, we can formally express the aforementioned idea as replacing the variables $\mathbf{x}$ with a closed-form expression of the implicit function $\mathbf{x} : Y \to X$ defined by the equality constraints and solving the resulting reduced-space problem:

$$\begin{aligned} \min_{\mathbf{y} \in Y} \ & f(\mathbf{x}(\mathbf{y}), \mathbf{y}) \\ \text{s.t. } & \mathbf{g}(\mathbf{x}(\mathbf{y}), \mathbf{y}) \leq \mathbf{0}. \end{aligned} \tag{RS}$$

If a scheme of relaxations, $(\mathbf{x}_Z^{\text{cv}}, \mathbf{x}_Z^{\text{cc}})_{Z \in \mathbb{I}Y}$, of the implicit function $\mathbf{x}$ in Y is available, where $\mathbf{x}_Z^{\text{cv}} : Z \to X$ and $\mathbf{x}_Z^{\text{cc}} : Z \to X$ are convex and concave relaxations of $\mathbf{x}$ on Z, respectively, then schemes of relaxations of the 'reduced-space functions' $f(\mathbf{x}(\cdot), \cdot)$ and $\mathbf{g}(\mathbf{x}(\cdot), \cdot)$ can be constructed using schemes of relaxations of the functions f and $\mathbf{g}$ in $X \times Y$ and the generalized McCormick relaxation framework [206, 207]. The above developments can then enable global optimization of the reduced-space Problem (RS) using a reduced-space B&B approach (cf. Section 2.3.2.3 of Chapter 2).

Stuber et al. [217] proposed an approach to construct schemes of relaxations of implicit functions using the framework of generalized McCormick relaxations and the use of a parametric extension of the mean-value theorem in conjunction with parametric interval-Newton methods. Wechsung [237, Chapter 5] proposed a different sensitivities-based interval bounding approach for computing second-order Hausdorff convergent interval bounds on the implicit function. Preliminary computational experiments performed by members of the lab indicate that using the method for constructing schemes of relaxations proposed by Stuber et al. within a reduced-space B&B framework seems to mitigate the cluster problem, whereas employing the interval-based relaxations of implicit functions proposed by Wechsung within such a framework appears to suffer from the cluster problem. Our analysis of the convergence-order of lower bounding schemes for constrained problems in Chapter 6 indicates that the pointwise convergence orders of schemes of relaxations involved in the construction of the lower bounding scheme play a more important role than the corresponding Hausdorff convergence orders in determining the overall convergence order of the lower bounding scheme at constrained minima (which, in turn, determines whether the lower bounding scheme can mitigate the cluster problem). The fact that interval-based schemes of relaxations can have at most first-order pointwise convergence can, along with the analysis by Bompadre and Mitsos [38], therefore potentially help explain the observed unfavorable behavior of the interval-based technique proposed by Wechsung [237, Chapter 5] for relaxing implicit functions within a reduced-space B&B framework. An analysis of the convergence order of the method proposed by Stuber and coworkers [217] is currently lacking.

In the remainder of this note, we consider the case when $\mathbf{h}$ is defined by the parametric linear system of equations $\mathbf{A}(\mathbf{y})\mathbf{x} = \mathbf{b}(\mathbf{y})$, with $\mathbf{A} : Y \to \mathbb{R}^{n_x \times n_x}$ and $\mathbf{b} : Y \to \mathbb{R}^{n_x}$ assumed to be defined by factorable functions. Mitsos et al. [169] proposed an automatable technique to compute schemes of convex and concave relaxations of the implicit function $\mathbf{x}$ for this case that applies McCormick's relaxation technique [154] to a factorable representation of the implicit function obtained via Gaussian elimination (or any other direct solution approach for linear systems of equations). An analysis of the convergence order of their proposed approach follows directly from previous analyses of the convergence order of McCormick relaxations [38, 39] and the analysis in Chapter 6. In particular, it can be shown that the approach of Mitsos and coworkers for this (restrictive) subclass of Problem (P) can mitigate the cluster problem when employed within a reduced-space B&B framework while still enjoying the advantages of dimensionality reduction. Stuber et al. [217, Section 3.3] proposed a Gauss-Seidel-type semi-explicit representation of the implicit function for the parametric linear case, and constructed relaxations of the implicit function using a parametric interval linear solver and the generalized McCormick relaxation technique. While it appears that the theory of convergence order for generalized McCormick relaxations in [199, Chapter 3] might be useful in analyzing the convergence order of the scheme of relaxations of the implicit function that is constructed for this special class of problems, the analysis of the convergence order of the above Gauss-Seidel-type method for constructing schemes of relaxations is nontrivial because the number of iterations of this iterative technique for constructing relaxations is not fixed *a priori*. It is anticipated that the analysis of the convergence order of the above relaxation method for the special case of parametric linear systems will provide insights into the analysis for the general case of parametric nonlinear systems of equations, since (some of) the techniques employed in [217] for this more general case effectively reduce the problem of constructing schemes of relaxations for the implicit function to a corresponding problem for linear parametric systems.

Bibliography

[1] Joaquin Acevedo and Efstratios N Pistikopoulos. Stochastic optimization based algorithms for process synthesis under uncertainty. *Computers & Chemical Engineering*, 22(4):647–671, 1998.

[2] Nilanjan Adhya, Mohit Tawarmalani, and Nikolaos V Sahinidis. A Lagrangian approach to the pooling problem. *Industrial & Engineering Chemistry Research*, 38(5):1956–1972, 1999.

[3] Claire S Adjiman, Stefan Dallwig, Christodoulos A Floudas, and Arnold Neumaier. A global optimization method, αBB, for general twice-differentiable constrained NLPs I. Theoretical advances. *Computers & Chemical Engineering*, 22(9):1137–1158, 1998.

[4] Claire S Adjiman and Christodoulos A Floudas. Rigorous convex underestimators for general twice-differentiable problems. *Journal of Global Optimization*, 9(1):23–40, 1996.

[5] Shabbir Ahmed. Convexity and decomposition of mean-risk stochastic programs. *Mathematical Programming*, 106(3):433–446, 2006.

[6] Ioannis G Akrotirianakis and Christodoulos A Floudas. A new class of improved convex underestimators for twice continuously differentiable constrained NLPs. *Journal of Global Optimization*, 30(4):367–390, 2004.

[7] Faiz A Al-Khayyal and James E Falk. Jointly constrained biconvex programming. *Mathematics of Operations Research*, 8(2):273–286, 1983.

[8] Satyajith Amaran, Tong Zhang, Nikolaos V Sahinidis, Bikram Sharda, and Scott J Bury. Medium-term maintenance turnaround planning under uncertainty for integrated chemical sites. *Computers & Chemical Engineering*, 84:422–433, 2016.

[9] Gustavo Angulo, Shabbir Ahmed, and Santanu S Dey. Improving the integer L-shaped method. *INFORMS Journal on Computing*, 28(3):483–499, 2016.

[10] Egon Balas and Robert Jeroslow. Canonical cuts on the unit hypercube. *SIAM Journal on Applied Mathematics*, 23(1):61–69, 1972.

[11] Xiaowei Bao, Aida Khajavirad, Nikolaos V Sahinidis, and Mohit Tawarmalani. Global optimization of nonconvex problems with multilinear intermediates. *Mathematical Programming Computation*, 7(1):1–37, 2015.

[12] Xiaowei Bao, Nikolaos V Sahinidis, and Mohit Tawarmalani. Multiterm polyhedral relaxations for nonconvex, quadratically constrained quadratic programs. *Optimization Methods & Software*, 24(4-5):485–504, 2009.

[13] Mokhtar S Bazaraa, Hanif D Sherali, and Chitharanjan M Shetty. *Nonlinear programming: theory and algorithms*. John Wiley & Sons, 3rd edition, 2013.

[14] Evelyn ML Beale. On minimizing a convex function subject to linear inequalities. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 173–184, 1955.

[15] Richard Bellman. *Dynamic programming*. Princeton University Press, Princeton, NJ, 1957.

[16] Richard E Bellman and Lotfi A Zadeh. Decision-making in a fuzzy environment. *Management Science*, 17(4):B–141, 1970.

[17] Pietro Belotti, Sonia Cafieri, Jon Lee, and Leo Liberti. Feasibility-based bounds tightening via fixed points. In Weili Wu and Ovidiu Daescu, editors, *Combinatorial Optimization and Applications*, pages 65–76. Springer, 2010.

[18] Pietro Belotti, Christian Kirches, Sven Leyffer, Jeff Linderoth, James Luedtke, and Ashutosh Mahajan. Mixed-integer nonlinear optimization. *Acta Numerica*, 22:1–131, 2013.

[19] Pietro Belotti, Jon Lee, Leo Liberti, François Margot, and Andreas Wächter. Branching and bounds tightening techniques for non-convex MINLP. *Optimization Methods & Software*, 24(4-5):597–634, 2009.

[20] Aharon Ben-Tal, Gideon Eiger, and Vladimir Gershovitz. Global minimization by reducing the duality gap. *Mathematical Programming*, 63(1-3):193–212, 1994.

[21] Aharon Ben-Tal, Laurent El Ghaoui, and Arkadi Nemirovski. *Robust optimization*. Princeton University Press, 2009.

[22] Aharon Ben-Tal, Alexander Goryashko, Elana Guslitzer, and Arkadi Nemirovski. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2):351–376, 2004.

[23] Aharon Ben-Tal and Arkadi Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769–805, 1998.

[24] Aharon Ben-Tal and Arkadi Nemirovski. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. Society for Industrial and Applied Mathematics, 2001.

[25] Jacques F Benders. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, 4(1):238–252, 1962.

[26] Claus Bendtsen and Ole Stauning. FADBAD, a flexible C++ package for automatic differentiation. *Department of Mathematical Modelling, Technical University of Denmark*, 1996.

[27] Maria L Bergamini, Pio Aguirre, and Ignacio Grossmann. Logic-based outer approximation for globally optimal synthesis of process networks. *Computers & Chemical Engineering*, 29(9):1914–1933, 2005.

[28] Timo Berthold. *Heuristic algorithms in global MINLP solvers.* PhD thesis, Technische Universität Berlin, 2014.

[29] Dimitri P Bertsekas. *Dynamic programming and optimal control*, volume 1. Athena Scientific, Belmont, Massachusetts, 4th edition, 1995.

[30] Dimitri P Bertsekas. *Nonlinear programming.* Athena Scientific, Belmont, Massachusetts, 2nd edition, 1999.

[31] Dimitris Bertsimas, David B Brown, and Constantine Caramanis. Theory and applications of robust optimization. *SIAM Review*, 53(3):464–501, 2011.

[32] Dimitris Bertsimas and Constantine Caramanis. Finite adaptability in multistage linear optimization. *IEEE Transactions on Automatic Control*, 55(12):2751–2766, 2010.

[33] Dimitris Bertsimas and Melvyn Sim. The price of robustness. *Operations Research*, 52(1):35–53, 2004.

[34] Binita Bhattacharjee, Panayiotis Lemonidis, William H Green Jr, and Paul I Barton. Global solution of semi-infinite programs. *Mathematical Programming*, 103(2):283–307, 2005.

[35] John R Birge and Francois Louveaux. *Introduction to stochastic programming.* Springer, 2nd edition, 2011.

[36] John R Birge and Francois V Louveaux. A multicut algorithm for two-stage stochastic linear programs. *European Journal of Operational Research*, 34(3):384–392, 1988.

[37] Jerry W Blankenship and James E Falk. Infinitely constrained optimization problems. *Journal of Optimization Theory and Applications*, 19(2):261–281, 1976.

[38] Agustín Bompadre and Alexander Mitsos. Convergence rate of McCormick relaxations. *Journal of Global Optimization*, 52(1):1–28, 2012.

[39] Agustín Bompadre, Alexander Mitsos, and Benoît Chachuat. Convergence analysis of Taylor models and McCormick-Taylor models. *Journal of Global Optimization*, 57(1):75–114, 2013.

[40] Pierre Bonami, Lorenz T Biegler, Andrew R Conn, Gérard Cornuéjols, Ignacio E Grossmann, Carl D Laird, Jon Lee, Andrea Lodi, François Margot, Nicolas Sawaya, and Andreas Wächter. An algorithmic framework for convex mixed integer nonlinear programs. *Discrete Optimization*, 5(2):186–204, 2008.

[41] Pierre Bonami, Gérard Cornuéjols, Andrea Lodi, and François Margot. A feasibility pump for mixed integer nonlinear programs. *Mathematical Programming*, 119(2):331–352, 2009.

[42] Dominik Bongartz and Alexander Mitsos. Deterministic global optimization of process flowsheets in a reduced space using McCormick relaxations. *Journal of Global Optimization*, 69(4):761–796, 2017.

[43] Joseph F Bonnans and Alexander Ioffe. Second-order sufficiency and quadratic growth for nonisolated minima. *Mathematics of Operations Research*, 20(4):801–817, 1995.

[44] Samuel Burer and Adam N Letchford. Non-convex mixed-integer nonlinear programming: A survey. *Surveys in Operations Research and Management Science*, 17(2):97–106, 2012.

[45] Sonia Cafieri, Jon Lee, and Leo Liberti. On convex relaxations of quadrilinear terms. *Journal of Global Optimization*, 47(4):661–685, 2010.

[46] Bruno A Calfa, Ignacio E Grossmann, Anshul Agarwal, Scott J Bury, and John M Wassick. Data-driven individual and joint chance-constrained optimization via kernel smoothing. *Computers & Chemical Engineering*, 78:51–69, 2015.

[47] Yankai Cao and Victor M Zavala. A Scalable Global Optimization Algorithm for Stochastic Nonlinear Programs. *Optimization Online. Available online at http://www.optimization-online.org/DB_HTML/2017/08/6164.html*, pages 1–17, August 2017.

[48] Claus C Carøe and Rüdiger Schultz. Dual decomposition in stochastic integer programming. *Operations Research Letters*, 24(1):37–45, 1999.

[49] B Chachuat. MC++: A versatile library for McCormick relaxations and Taylor models. *http://www3.imperial.ac.uk/people/b.chachuat/research*, 2010.

[50] Abraham Charnes and William W Cooper. Deterministic equivalents for optimizing and satisficing under chance constraints. *Operations Research*, 11(1):18–39, 1963.

[51] Abraham Charnes, William W Cooper, and Gifford H Symonds. Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil. *Management Science*, 4(3):235–263, 1958.

[52] Yang Chen. *Optimal design and operation of energy polygeneration systems*. PhD thesis, Massachusetts Institute of Technology, 2013.

[53] Yang Chen, Thomas A Adams, and Paul I Barton. Optimal design and operation of flexible energy polygeneration systems. *Industrial & Engineering Chemistry Research*, 50(8):4553–4566, 2011.

[54] Yang Chen, Xiang Li, Thomas A Adams, and Paul I Barton. Decomposition strategy for the global optimization of flexible energy polygeneration systems. *AIChE Journal*, 58(10):3080–3095, 2012.

[55] E Ward Cheney and Allen A Goldstein. Newton's method for convex programming and Tchebycheff approximation. *Numerische Mathematik*, 1(1):253–268, 1959.

[56] Naiyuan Chiang, Cosmin G Petra, and Victor M Zavala. Structured nonconvex optimization of large-scale energy systems using PIPS-NLP. In *Power Systems Computation Conference (PSCC), 2014*, pages 1–7. IEEE, 2014.

[57] Frank H Clarke. *Optimization and Nonsmooth Analysis*, volume 5 of *Classics in Applied Mathematics*. SIAM, 1990.

[58] Gianni Codato and Matteo Fischetti. Combinatorial Benders' cuts for mixed-integer linear programming. *Operations Research*, 54(4):756–766, 2006.

[59] Claudia D'Ambrosio, Antonio Frangioni, Leo Liberti, and Andrea Lodi. Experiments with a feasibility pump approach for nonconvex MINLPs. In *SEA*, pages 350–360. Springer, 2010.

[60] Claudia D'Ambrosio, Antonio Frangioni, Leo Liberti, and Andrea Lodi. A storm of feasibility pumps for nonconvex MINLP. *Mathematical Programming*, 136(2):375–402, 2012.

[61] George B Dantzig. Linear programming under uncertainty. In *Stochastic Programming*, pages 1–11. Springer, 2010.

[62] Beman Dawes, David Abrahams, and Rene Rivera. Boost C++ libraries. *http://www.boost.org*, 35:36, 2009.

[63] Erick Delage and Yinyu Ye. Distributionally robust optimization under moment uncertainty with application to data-driven problems. *Operations Research*, 58(3):595–612, 2010.

[64] Moritz Diehl, Boris Houska, Oliver Stein, and Paul Steuermann. A lifting method for generalized semi-infinite programs based on lower level Wolfe duality. *Computational Optimization and Applications*, 54(1):189–210, 2013.

[65] Hatim Djelassi and Alexander Mitsos. A hybrid discretization algorithm with guaranteed feasibility for the global solution of semi-infinite programs. *Journal of Global Optimization*, 68(2):227–253, 2017.

[66] Elizabeth D Dolan and Jorge J Moré. Benchmarking optimization software with performance profiles. *Mathematical Programming*, 91(2):201–213, 2002.

[67] Arne S Drud. CONOPT - a large-scale GRG code. *ORSA Journal on Computing*, 6(2):207–216, 1994.

[68] Kaisheng Du and Ralph B Kearfott. The cluster problem in multivariate global optimization. *Journal of Global Optimization*, 5(3):253–265, 1994.

[69] Miriam Dür and Reiner Horst. Lagrange duality and partitioning techniques in nonconvex global optimization. *Journal of Optimization Theory and Applications*, 95(2):347–369, 1997.

[70] Mirjam Dür. Dual bounding procedures lead to convergent branch–and–bound algorithms. *Mathematical Programming*, 91(1):117–125, 2001.

[71] Mirjam Dür. A class of problems where dual bounds beat underestimation bounds. *Journal of Global Optimization*, 22(1-4):49–57, 2002.

[72] Marco A Duran and Ignacio E Grossmann. An outer-approximation algorithm for a class of mixed-integer nonlinear programs. *Mathematical Programming*, 36(3):307–339, 1986.

[73] Martin Dyer and Leen Stougie. Computational complexity of stochastic programming problems. *Mathematical Programming*, 106(3):423–432, 2006.

[74] Laurent El Ghaoui and Hervé Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications*, 18(4):1035–1064, 1997.

[75] Laurent El Ghaoui, François Oustry, and Hervé Lebret. Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization*, 9(1):33–52, 1998.

[76] Thomas G W Epperly and Efstratios N Pistikopoulos. A reduced space branch and bound algorithm for global optimization. *Journal of Global Optimization*, 11(3):287–311, 1997.

[77] Emre Erdoğan and Garud Iyengar. Ambiguous chance constrained problems and robust optimization. *Mathematical Programming*, 107(1-2):37–61, 2006.

[78] Matteo Fischetti and Michele Monaci. Light robustness. In *Robust and online large-scale optimization*, pages 61–84. Springer, 2009.

[79] Roger Fletcher and Sven Leyffer. Solving mixed integer nonlinear programs by outer approximation. *Mathematical Programming*, 66(1-3):327–349, 1994.

[80] Christodoulos A Floudas and Chrysanthos E Gounaris. A review of recent advances in global optimization. *Journal of Global Optimization*, 45(1):3–38, 2009.

[81] Christodoulos A Floudas, Panos M Pardalos, Claire Adjiman, William R Esposito, Zeynep H Gümüs, Stephen T Harding, John L Klepeis, Clifford A Meyer, and Carl A Schweiger. *Handbook of Test Problems in Local and Global Optimization*, volume 33 of *Nonconvex Optimization and Its Applications*. Springer, 1st edition, 1999.

[82] Christodoulos A Floudas and V Visweswaran. Primal-relaxed dual global optimization approach. *Journal of Optimization Theory and Applications*, 78(2):187–225, 1993.

[83] GAMS Development Corporation. General Algebraic Modeling System (GAMS) Release 24.7.1. Washington, DC, USA, 2016.

[84] Edward P Gatzke, John E Tolsma, and Paul I Barton. Construction of convex relaxations using automated code generation techniques. *Optimization and Engineering*, 3(3):305–326, 2002.

[85] Arthur M Geoffrion. Generalized Benders decomposition. *Journal of Optimization Theory and Applications*, 10(4):237–260, 1972.

[86] Philip E Gill, Walter Murray, and Michael A Saunders. SNOPT: An SQP algorithm for large-scale constrained optimization. *SIAM review*, 47(1):99–131, 2005.

[87] Alexandre Goldsztejn, Ferenc Domes, and Brice Chevalier. First order rejection tests for multiple-objective optimization. *Journal of Global Optimization*, 58(4):653–672, 2014.

[88] Michael Grant and Stephen Boyd. CVX: Matlab software for disciplined convex programming, version 2.1. http://cvxr.com/cvx, March 2014.

[89] Ignacio E Grossmann, Keshava P Halemane, and Ross E Swaney. Optimization strategies for flexible chemical processes. *Computers & Chemical Engineering*, 7(4):439–462, 1983.

[90] Ignacio E Grossmann, Jagadisan Viswanathan, Aldo Vecchietti, Ramesh Raman, and Erwin Kalvelagen. GAMS/DICOPT: A discrete continuous optimization package. *GAMS Development Corporation, Washington, DC*, 2002.

[91] Monique Guignard and Siwhan Kim. Lagrangean decomposition: a model yielding stronger Lagrangean bounds. *Mathematical Programming*, 39(2):215–228, 1987.

[92] Gurobi Optimization Inc. Gurobi Optimizer Reference Manual, 2017.

[93] Eldon Hansen. *Topics in interval analysis*. Clarendon Press Oxford, 1969.

[94] Pierre Hansen, Brigitte Jaumard, and Shi-Hui Lu. An analytical approach to global optimization. *Mathematical Programming*, 52(1-3):227–254, 1991.

[95] Pierre Hansen and Nenad Mladenović. Variable neighborhood search. In *Search methodologies*, pages 313–337. Springer, 2014.

[96] Iiro Harjunkoski, Tapio Westerlund, Ray Pörn, and Hans Skrifvars. Different transformations for solving non-convex trim-loss problems by MINLP. *European Journal of Operational Research*, 105(3):594–603, 1998.

[97] Stuart M Harwood and Paul I Barton. Lower level duality and the global solution of generalized semi-infinite programs. *Optimization*, 65(6):1129–1149, 2016.

[98] Rainer Hettich and Kenneth O Kortanek. Semi-infinite programming: theory, methods, and applications. *SIAM Review*, 35(3):380–429, 1993.

[99] Hassan Hijazi and Leo Liberti. Constraint qualification failure in action. *Operations Research Letters*, 44(4):503–506, 2016.

[100] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Convex Analysis and Minimization Algorithms II: Advanced theory and bundle methods*, volume 306 of *Grundlehren der mathematischen Wissenschaften*. Springer-Verlag, 1st edition, 1993.

[101] Reiner Horst and Hoang Tuy. *Global optimization: Deterministic approaches*. Springer-Verlag Berlin Heidelberg, 3rd edition, 1996.

[102] John K. Hunter. *An introduction to real analysis*. Department of Mathematics, University of California at Davis, 2014.

[103] IBM. IBM ILOG CPLEX. High-performance mathematical programming engine. http://www-01.ibm.com/software/integration/optimization/cplex/, 2017.

[104] Alexander Ioffe. On sensitivity analysis of nonlinear programs in Banach spaces: the approach via composite unconstrained optimization. *SIAM Journal on Optimization*, 4(1):1–43, 1994.

[105] Ruiwei Jiang and Yongpei Guan. Data-driven chance constrained stochastic program. *Mathematical Programming*, 158(1-2):291–327, 2016.

[106] Fritz John. Extremum problems with inequalities as subsidiary conditions. In *Traces and emergence of nonlinear programming*, pages 197–215. Springer, 2014.

[107] Jia Kang, Yankai Cao, Daniel P Word, and Carl D Laird. An interior-point method for efficient solution of block-structured NLP problems using an implicit Schur-complement decomposition. *Computers & Chemical Engineering*, 71:563–573, 2014.

[108] Rohit Kannan and Paul I Barton. The cluster problem in constrained global optimization. *Journal of Global Optimization*, 69(3):629–676, 2017.

[109] Rohit Kannan and Paul I Barton. Convergence-order analysis of branch-and-bound algorithms for constrained problems. *Journal of Global Optimization*, in press.

[110] Napsu Karmitsa, Adil Bagirov, and Marko M Mäkelä. Empirical and theoretical comparisons of several nonsmooth minimization methods and software. Technical report, TUCS Technical Report 959, Turku Centre for Computer Science, October, 2009.

[111] Ramkumar Karuppiah and Ignacio E Grossmann. Global optimization of multiscenario mixed integer nonlinear programming models arising in the synthesis of integrated water networks under uncertainty. *Computers & Chemical Engineering*, 32(1):145–160, 2008.

[112] Ramkumar Karuppiah and Ignacio E Grossmann. A Lagrangean based branch-and-cut algorithm for global optimization of nonconvex mixed-integer nonlinear programs with decomposable structures. *Journal of Global Optimization*, 41(2):163–186, 2008.

[113] William Karush. Minima of functions of several variables with inequalities as side conditions. *Master thesis, University of Chicago*, 1939.

[114] Ralph B Kearfott. Empirical evaluation of innovations in interval branch and bound algorithms for nonlinear systems. *SIAM Journal on Scientific Computing*, 18(2):574–594, 1997.

[115] Ralph B Kearfott. Some observations on exclusion regions in branch and bound algorithms. *Journal of Global Optimization*, 62(2):229–241, 2015.

[116] Ralph B Kearfott and Kaisheng Du. The cluster problem in global optimization: the univariate case. In *Validation Numerics*, volume 9 of *Computing Supplementum*, pages 117–127. Springer, 1993.

[117] James E Kelley, Jr. The cutting-plane method for solving convex programs. *Journal of the Society for Industrial and Applied Mathematics*, 8(4):703–712, 1960.

[118] Padmanaban Kesavan, Russell J Allgor, Edward P Gatzke, and Paul I Barton. Outer approximation algorithms for separable nonconvex mixed-integer nonlinear programs. *Mathematical Programming*, 100(3):517–535, 2004.

[119] Aida Khajavirad and Jeremy J Michalek. A deterministic Lagrangian-based global optimization approach for quasiseparable nonconvex mixed-integer nonlinear programs. *Journal of Mechanical Design*, 131(5):051009, 2009.

[120] Aida Khajavirad, Jeremy J Michalek, and Nikolaos V Sahinidis. Relaxations of factorable functions with convex-transformable intermediates. *Mathematical Programming*, 144(1-2):107–140, 2014.

[121] Aida Khajavirad and Nikolaos V Sahinidis. Convex envelopes of products of convex and component-wise concave functions. *Journal of Global Optimization*, 52(3):391–409, 2012.

[122] Aida Khajavirad and Nikolaos V Sahinidis. Convex envelopes generated from finitely many compact convex sets. *Mathematical Programming*, pages 1–38, 2013.

[123] Kamil A Khan. *Sensitivity analysis for nonsmooth dynamic systems*. PhD thesis, Massachusetts Institute of Technology, 2015.

[124] Kamil A Khan, Harry A J Watson, and Paul I Barton. Differentiable McCormick relaxations. *Journal of Global Optimization*, 67(4):687–729, 2017.

[125] Kibaek Kim and Victor M Zavala. Algorithmic innovations and software for the dual decomposition method applied to stochastic mixed-integer programs. *Mathematical Programming Computation*, 2017.

[126] Krzysztof C Kiwiel. An aggregate subgradient method for nonsmooth convex minimization. *Mathematical Programming*, 27(3):320–341, 1983.

[127] Rudolf Krawczyk and Karl Nickel. Die zentrische form in der Intervallarithmetik, ihre quadratische Konvergenz und ihre Inklusionsisotonie. *Computing*, 28(2):117–137, 1982.

[128] Daniel Kuhn, Wolfram Wiesemann, and Angelos Georghiou. Primal and dual linear decision rules in stochastic and robust optimization. *Mathematical Programming*, 130(1):177–209, 2011.

[129] Harold W Kuhn and Albert W Tucker. Nonlinear programming. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 481–492. University of California Press, 1951.

[130] Alexei Kuntsevich and Franz Kappel. SolvOpt: The solver for local nonlinear optimization problems. *Institute for Mathematics, Karl-Franzens University of Graz*, 1997.

[131] Carl D Laird and Lorenz T Biegler. Large-scale nonlinear programming for multi-scenario optimization. In *Modeling, simulation and optimization of complex processes*, pages 323–336. Springer, 2008.

[132] Gilbert Laporte and Francois V Louveaux. The integer L-shaped method for stochastic integer programs with complete recourse. *Operations Research Letters*, 13(3):133–142, 1993.

[133] Nikolaos H Lappas and Chrysanthos E Gounaris. Multi-stage adjustable robust optimization for process scheduling under uncertainty. *AIChE Journal*, 62(5):1646–1667, 2016.

[134] Dan Li and Xiang Li. Domain reduction for Benders decomposition based global optimization. *Computers & Chemical Engineering*, 93:248–265, 2016.

[135] Dan Li and Xiang Li. A new optimization model and a customized solution method for natural gas production network design and operation. *AIChE Journal*, 63(3):933–948, 2017.

[136] Xiang Li, Emre Armagan, Asgeir Tomasgard, and Paul I Barton. Stochastic pooling problem for natural gas production network design and operation under uncertainty. *AIChE Journal*, 57(8):2120–2135, 2011.

[137] Xiang Li, Yang Chen, and Paul I Barton. Nonconvex generalized Benders decomposition with piecewise convex relaxations for global optimization of integrated process design and operation problems. *Industrial & Engineering Chemistry Research*, 51(21):7287–7299, 2012.

[138] Xiang Li, Arul Sundaramoorthy, and Paul I Barton. Nonconvex generalized Benders decomposition. In Themistocles M Rassias, Christodoulos A Floudas, and Sergiy Butenko, editors, *Optimization in Science and Engineering*, pages 307–331. Springer, 2014.

[139] Xiang Li, Asgeir Tomasgard, and Paul I Barton. Nonconvex generalized Benders decomposition for stochastic separable mixed-integer nonlinear programs. *Journal of Optimization Theory and Applications*, 151(3):425–454, 2011.

[140] Xiang Li, Asgeir Tomasgard, and Paul I Barton. Decomposition strategy for the stochastic pooling problem. *Journal of Global Optimization*, 54(4):765–790, 2012.

[141] Leo Liberti, Giacomo Nannicini, and Nenad Mladenović. A good recipe for solving MINLPs. In *Matheuristics*, pages 231–244. Springer, 2009.

[142] Leo Liberti and Constantinos C Pantelides. Convex envelopes of monomials of odd degree. *Journal of Global Optimization*, 25(2):157–168, 2003.

[143] Leo Liberti and Constantinos C Pantelides. An exact reformulation algorithm for large nonconvex NLPs involving bilinear terms. *Journal of Global Optimization*, 36(2):161–189, 2006.

[144] Xiaoxia Lin, Stacy L Janak, and Christodoulos A Floudas. A new robust optimization approach for scheduling under uncertainty: I. Bounded uncertainty. *Computers & Chemical Engineering*, 28(6):1069–1085, 2004.

[145] Youdong Lin and Linus Schrage. The global solver in the LINDO API. *Optimization Methods & Software*, 24(4-5):657–668, 2009.

[146] Jeff Linderoth, Alexander Shapiro, and Stephen Wright. The empirical behavior of sampling methods for stochastic programming. *Annals of Operations Research*, 142(1):215–241, 2006.

[147] Marco Locatelli and Fabio Schoen. On convex envelopes for bivariate functions over polytopes. *Mathematical Programming*, 144(1-2):65–91, 2014.

[148] James Luedtke and Shabbir Ahmed. A sample approximation approach for optimization with probabilistic constraints. *SIAM Journal on Optimization*, 19(2):674–699, 2008.

[149] Andreas Lundell, Joakim Westerlund, and Tapio Westerlund. Some transformation techniques with applications in global optimization. *Journal of Global Optimization*, 43(2):391–405, 2009.

[150] Thomas L Magnanti and Richard T Wong. Accelerating Benders decomposition: Algorithmic enhancement and model selection criteria. *Operations Research*, 29(3):464–484, 1981.

[151] Marko M Mäkelä. Multiobjective proximal bundle method for nonconvex nonsmooth optimization: Fortran subroutine MPBNGC 2.0. *Reports of the Department of Mathematical Information Technology, Series B. Scientific Computing, B*, 13, 2003.

[152] Costas D Maranas and Christodoulos A Floudas. Global optimization in generalized geometric programming. *Computers & Chemical Engineering*, 21(4):351–369, 1997.

[153] Günter Mayer. Epsilon-inflation in verification algorithms. *Journal of Computational and Applied Mathematics*, 60(1):147–169, 1995.

[154] Garth P McCormick. Computability of global solutions to factorable nonconvex programs: Part I - Convex underestimating problems. *Mathematical Programming*, 10(1):147–175, 1976.

[155] Clifford A Meyer and Christodoulos A Floudas. Trilinear monomials with mixed sign domains: Facets of the convex and concave envelopes. *Journal of Global Optimization*, 29(2):125–155, 2004.

[156] Clifford A Meyer and Christodoulos A Floudas. Trilinear monomials with positive or negative domains: Facets of the convex and concave envelopes. In *Frontiers in Global Optimization*, pages 327–352. Springer, 2004.

[157] Clifford A Meyer and Christodoulos A Floudas. Convex envelopes for edge-concave functions. *Mathematical Programming*, 103(2):207–224, 2005.

[158] Clifford A Meyer and Christodoulos A Floudas. Global optimization of a combinatorially complex generalized pooling problem. *AIChE Journal*, 52(3):1027–1037, 2006.

[159] Ruth Misener and Christodoulos A Floudas. Global optimization of large-scale generalized pooling problems: quadratically constrained MINLP models. *Industrial & Engineering Chemistry Research*, 49(11):5424–5438, 2010.

[160] Ruth Misener and Christodoulos A Floudas. Global optimization of mixed-integer quadratically-constrained quadratic programs (MIQCQP) through piecewise-linear and edge-concave relaxations. *Mathematical Programming*, pages 1–28, 2012.

[161] Ruth Misener and Christodoulos A Floudas. GloMIQO: Global mixed-integer quadratic optimizer. *Journal of Global Optimization*, 57(1):3–50, 2013.

[162] Ruth Misener and Christodoulos A Floudas. ANTIGONE: Algorithms for coNTinuous/Integer Global Optimization of Nonlinear Equations. *Journal of Global Optimization*, 59(2-3):503–526, 2014.

[163] Ruth Misener and Christodoulos A Floudas. A framework for globally optimizing mixed-integer signomial programs. *Journal of Optimization Theory and Applications*, 161(3):905–932, 2014.

[164] Ruth Misener, James B Smadbeck, and Christodoulos A Floudas. Dynamically generated cutting planes for mixed-integer quadratically constrained quadratic programs and their incorporation into GloMIQO 2. *Optimization Methods and Software*, 30(1):215–249, 2015.

[165] Ruth Misener, Jeffrey P Thompson, and Christodoulos A Floudas. APOGEE: Global optimization of standard, generalized, and extended pooling problems via linear and logarithmic partitioning schemes. *Computers & Chemical Engineering*, 35(5):876–892, 2011.

[166] Sumit Mitra, Pablo Garcia-Herreros, and Ignacio E Grossmann. A cross-decomposition scheme with integrated primal–dual multi-cuts for two-stage stochastic programming investment planning problems. *Mathematical Programming*, 157(1):95–119, 2016.

[167] Alexander Mitsos. Global solution of nonlinear mixed-integer bilevel programs. *Journal of Global Optimization*, 47(4):557–582, 2010.

[168] Alexander Mitsos. Global optimization of semi-infinite programs via restriction of the right-hand side. *Optimization*, 60(10-11):1291–1308, 2011.

[169] Alexander Mitsos, Benoît Chachuat, and Paul I Barton. McCormick-based relaxations of algorithms. *SIAM Journal on Optimization*, 20(2):573–601, 2009.

[170] Alexander Mitsos, Panayiotis Lemonidis, and Paul I Barton. Global solution of bilevel programs with a nonconvex inner program. *Journal of Global Optimization*, 42(4):475–513, 2008.

[171] Ramon E Moore. Interval arithmetic and automatic error analysis in digital computing. Technical report, Department of Mathematics, Stanford University, 1962.

[172] Ramon E Moore, Ralph B Kearfott, and Michael J Cloud. *Introduction to Interval Analysis*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2009.

[173] Jaromił Najman, Dominik Bongartz, Angelos Tsoukalas, and Alexander Mitsos. Erratum to: Multivariate McCormick relaxations. *Journal of Global Optimization*, 68(1):219–225, 2017.

[174] Jaromił Najman and Alexander Mitsos. Convergence analysis of multivariate McCormick relaxations. *Journal of Global Optimization*, 66(4):597–628, 2016.

[175] Angelia Nedic and Dimitri P Bertsekas. Incremental subgradient methods for nondifferentiable optimization. *SIAM Journal on Optimization*, 12(1):109–138, 2001.

[176] Arkadi Nemirovski and Alexander Shapiro. Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17(4):969–996, 2006.

[177] Arnold Neumaier. *Interval methods for systems of equations*, volume 37 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 1990.

[178] Arnold Neumaier. Complete search in continuous global optimization and constraint satisfaction. *Acta Numerica*, 13:271–369, 2004.

[179] Mario VF Pereira and Leontina MVG Pinto. Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming*, 52(1-3):359–375, 1991.

[180] Warren B Powell. A unified framework for optimization under uncertainty. In *Optimization Challenges in Complex, Networked and Risky Systems*, pages 45–83. INFORMS, 2016.

[181] András Prékopa. *Stochastic Programming*, volume 324 of *Mathematics and Its Applications*. Springer Netherlands, 1995.

[182] Yash Puranik and Nikolaos V Sahinidis. Domain reduction techniques for global NLP and MINLP optimization. *Constraints*, 22(3):338–376, 2017.

[183] Liqun Qi and Jie Sun. A nonsmooth version of Newton's method. *Mathematical Programming*, 58(1-3):353–367, 1993.

[184] Helmut Ratschek and Jon Rokne. *Computer methods for the range of functions*. Mathematics and its Applications. Ellis Horwood Ltd., 1984.

[185] Steffen Rebennack. Combining sampling-based and scenario-based nested benders decomposition methods: application to stochastic dual dynamic programming. *Mathematical Programming*, 156(1-2):343–389, 2016.

[186] Steffen Rebennack, Josef Kallrath, and Panos M Pardalos. Optimal storage design for a multi-product plant: A non-convex MINLP formulation. *Computers & Chemical Engineering*, 35(2):255–271, 2011.

[187] Anatoliy D Rikun. A convex envelope formula for multilinear functions. *Journal of Global Optimization*, 10(4):425–437, 1997.

[188] Ralph T Rockafellar. *Convex Analysis*. Princeton University Press, 1970.

[189] Ralph T Rockafellar. Optimization under uncertainty. Technical report, Lecture Notes, University of Washington, 2001.

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

[191] Günter Rote. The convergence rate of the sandwich algorithm for approximating convex functions. *Computing*, 48(3-4):337–361, 1992.

[192] Walter Rudin. *Principles of mathematical analysis*. McGraw-Hill New York, 3rd edition, 1976.

[193] Siegfried M Rump. INTLAB - interval laboratory. In *Developments in reliable computing*, pages 77–104. Springer, 1999.

[194] Hong S Ryoo and Nikolaos V Sahinidis. Global optimization of nonconvex NLPs and MINLPs with applications in process design. *Computers & Chemical Engineering*, 19(5):551–566, 1995.

[195] Hong S Ryoo and Nikolaos V Sahinidis. A branch-and-reduce approach to global optimization. *Journal of Global Optimization*, 8(2):107–138, 1996.

[196] Nikolaos V Sahinidis. Optimization under uncertainty: state-of-the-art and opportunities. *Computers & Chemical Engineering*, 28(6):971–983, 2004.

[197] Ali M Sahlodin and Benoît Chachuat. Convex/concave relaxations of parametric ODEs using Taylor models. *Computers & Chemical Engineering*, 35(5):844–857, 2011.

[198] Martin WP Savelsbergh. Preprocessing and probing techniques for mixed integer programming problems. *ORSA Journal on Computing*, 6(4):445–454, 1994.

[199] Spencer D Schaber. *Tools for dynamic model development*. PhD thesis, Massachusetts Institute of Technology, 2014.

[200] Martin Schechter. *Principles of functional analysis*, volume 36. American Mathematical Society, 2nd edition, 2001.

[201] Hermann Schichl, Mihály C Markót, and Arnold Neumaier. Exclusion regions for optimization problems. *Journal of Global Optimization*, 59(2-3):569–595, 2014.

[202] Hermann Schichl and Arnold Neumaier. Exclusion regions for systems of equations. *SIAM Journal on Numerical Analysis*, 42(1):383–408, 2004.

[203] Anita Schöbel and Daniel Scholz. The theoretical and empirical rate of convergence for geometric branch-and-bound methods. *Journal of Global Optimization*, 48(3):473–495, 2010.

[204] Stefan Scholtes. *Introduction to piecewise differentiable equations*. SpringerBriefs in Optimization. Springer-Verlag New York, 1st edition, 2012.

[205] Daniel Scholz. Theoretical rate of convergence for interval inclusion functions. *Journal of Global Optimization*, 53(4):749–767, 2012.

[206] Joseph K Scott. *Reachability analysis and deterministic global optimization of differential-algebraic systems*. PhD thesis, Massachusetts Institute of Technology, 2012.

[207] Joseph K Scott, Matthew D Stuber, and Paul I Barton. Generalized McCormick relaxations. *Journal of Global Optimization*, 51(4):569–606, 2011.

[208] Hanif D Sherali and Warren P Adams. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM Journal on Discrete Mathematics*, 3(3):411–430, 1990.

[209] Hanif D Sherali and Warren P Adams. *A reformulation-linearization technique for solving discrete and continuous nonconvex problems*, volume 31. Springer Science & Business Media, 2013.

[210] Hanif D Sherali, Evrim Dalkiran, and Leo Liberti. Reduced RLT representations for nonconvex polynomial programming problems. *Journal of Global Optimization*, 52(3):447–469, 2012.

[211] Hanif D Sherali and Barbara MP Fraticelli. A modification of Benders' decomposition algorithm for discrete subproblems: An approach for stochastic programs with integer recourse. *Journal of Global Optimization*, 22(1-4):319–342, 2002.

[212] Naum Z Shor. *Minimization methods for non-differentiable functions*, volume 3 of *Springer Series in Computational Mathematics*. Springer-Verlag, 1st edition, 1985.

[213] Edward M B Smith and Constantinos C Pantelides. A symbolic reformulation/spatial branch-and-bound algorithm for the global optimisation of nonconvex MINLPs. *Computers & Chemical Engineering*, 23(4):457–478, 1999.

[214] Pantelis Sopasakis and Davide Giraudo. Basic properties of the point-to-set distance function. Mathematics Stack Exchange. http://math.stackexchange.com/questions/107478/basic-properties-of-the-point-to-set-distance-function (Version: 2012-02-10. Accessed: 2017-04-11).

[215] Allen L Soyster. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations Research*, 21(5):1154–1157, 1973.

[216] Oliver Stein. How to solve a semi-infinite optimization problem. *European Journal of Operational Research*, 223(2):312–320, 2012.

[217] Matthew D Stuber, Joseph K Scott, and Paul I Barton. Convex and concave relaxations of implicit functions. *Optimization Methods and Software*, 30(3):1–37, 2015.

[218] Anirudh Subramanyam, Chrysanthos E Gounaris, and Wolfram Wiesemann. K-adaptability in two-stage mixed-integer robust optimization. *arXiv preprint arXiv:1706.07097*, 2017.

[219] Arul Sundaramoorthy, Xiang Li, James MB Evans, and Paul I Barton. Capacity planning under clinical trials uncertainty in continuous pharmaceutical manufacturing, 2: Solution method. *Industrial & Engineering Chemistry Research*, 51(42):13703–13711, 2012.

[220] Bora Tarhan, Ignacio E Grossmann, and Vikas Goel. Stochastic programming approach for the planning of offshore oil or gas field infrastructure under decision-dependent uncertainty. *Industrial & Engineering Chemistry Research*, 48(6):3078–3097, 2009.

[221] Mohit Tawarmalani, Jean-Philippe P Richard, and Chuanhui Xiong. Explicit convex and concave envelopes through polyhedral subdivisions. *Mathematical Programming*, 138(1–2):531–577, 2013.

[222] Mohit Tawarmalani and Nikolaos V Sahinidis. Convex extensions and envelopes of lower semi-continuous functions. *Mathematical Programming*, 93(2):247–263, 2002.

[223] Mohit Tawarmalani and Nikolaos V Sahinidis. *Convexification and global optimization in continuous and mixed-integer nonlinear programming: theory, algorithms, software, and applications*, volume 65. Springer, 2002.

[224] Mohit Tawarmalani and Nikolaos V Sahinidis. Global optimization of mixed-integer nonlinear programs: A theoretical and computational study. *Mathematical Programming*, 99(3):563–591, 2004.

[225] Mohit Tawarmalani and Nikolaos V Sahinidis. A polyhedral branch-and-cut approach to global optimization. *Mathematical Programming*, 103(2):225–249, 2005.

[226] Francisco Trespalacios and Ignacio E Grossmann. Review of Mixed-Integer Nonlinear and Generalized Disjunctive Programming Methods. *Chemie Ingenieur Technik*, 86(7):991–1012, 2014.

[227] Angelos Tsoukalas and Alexander Mitsos. Multivariate McCormick relaxations. *Journal of Global Optimization*, 59(2-3):633–662, 2014.

[228] Zsolt Ugray, Leon Lasdon, John Plummer, Fred Glover, James Kelly, and Rafael Martí. Scatter search and local NLP solvers: A multistart framework for global optimization. *INFORMS Journal on Computing*, 19(3):328–340, 2007.

[229] Ronald J Van Iwaarden. *An improved unconstrained global optimization algorithm.* PhD thesis, University of Colorado at Denver, 1996.

[230] Tony J Van Roy. A cross decomposition algorithm for capacitated facility location. *Operations Research*, 34(1):145–163, 1986.

[231] Richard M Van Slyke and Roger Wets. L-shaped linear programs with applications to optimal control and stochastic programming. *SIAM Journal on Applied Mathematics*, 17(4):638–663, 1969.

[232] Phebe Vayanos, Daniel Kuhn, and Berç Rustem. A constraint sampling approach for multi-stage robust optimization. *Automatica*, 48(3):459–471, 2012.

[233] Stefan Vigerske and Ambros Gleixner. SCIP: global optimization of mixed-integer nonlinear programs in a branch-and-cut framework. *Optimization Methods and Software*, pages 1–31, 2017 (in press).

[234] Xuan-Ha Vu, Hermann Schichl, and Djamila Sam-Haroud. Interval propagation and search on directed acyclic graphs for numerical constraint solving. *Journal of Global Optimization*, 45(4):499–531, 2009.

[235] Andreas Wächter and Lorenz T Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Mathematical Programming*, 106(1):25–57, 2006.

[236] Jean-Paul Watson, David L Woodruff, and William E Hart. PySP: modeling and solving stochastic programs in Python. *Mathematical Programming Computation*, 4(2):109–149, 2012.

[237] Achim Wechsung. *Global optimization in reduced space.* PhD thesis, Massachusetts Institute of Technology, 2014.

[238] Achim Wechsung, Spencer D Schaber, and Paul I Barton. The cluster problem revisited. *Journal of Global Optimization*, 58(3):429–438, 2014.

[239] Achim Wechsung, Joseph K Scott, Harry AJ Watson, and Paul I Barton. Reverse propagation of McCormick relaxations. *Journal of Global Optimization*, 63(1):1–36, 2015.

[240] Wolfram Wiesemann, Daniel Kuhn, and Melvyn Sim. Distributionally robust convex optimization. *Operations Research*, 62(6):1358–1376, 2014.

[241] Yu Yang and Paul I Barton. Integrated crude selection and refinery optimization under uncertainty. *AIChE Journal*, 62(4):1038–1053, 2016.

[242] Yu Yang and Paul I Barton. A stochastic programming approach for refinery optimization integrated with a nonlinear crude distillation unit model. To be submitted.

[243] Yu Yang, Phebe Vayanos, and Paul I Barton. Chance-constrained optimization for refinery blend planning under uncertainty. *Industrial & Engineering Chemistry Research*, 56(42):12139–12150, 2017.

[244] Fengqi You, John M Wassick, and Ignacio E Grossmann. Risk management for a global supply chain planning under uncertainty: models and algorithms. *AIChE Journal*, 55(4):931–946, 2009.

[245] Dajun Yue and Fengqi You. Planning and scheduling of flexible process networks under uncertainty with stochastic inventory: MINLP models and algorithm. *AIChE Journal*, 59(5):1511–1532, 2013.

[246] Juan M Zamora and Ignacio E Grossmann. A branch and contract algorithm for problems with concave univariate, bilinear and linear fractional terms. *Journal of Global Optimization*, 14(3):217–249, 1999.

[247] Victor M Zavala, Carl D Laird, and Lorenz T Biegler. Interior-point decomposition approaches for parallel solution of large-scale nonlinear parameter estimation problems. *Chemical Engineering Science*, 63(19):4834–4845, 2008.

[248] Bo Zeng and Long Zhao. Solving two-stage robust optimization problems using a column-and-constraint generation method. *Operations Research Letters*, 41(5):457–461, 2013.

[249] Qi Zhang, Ignacio E Grossmann, Clara F Heuberger, Arul Sundaramoorthy, and Jose M Pinto. Air separation with cryogenic energy storage: optimal scheduling considering electric energy and reserve markets. *AIChE Journal*, 61(5):1547–1558, 2015.

[250] Keith Zorn and Nikolaos V Sahinidis. Global optimization of general nonconvex problems with intermediate polynomial substructures. *Journal of Global Optimization*, 59(2-3):673–693, 2014.

[251] Keith Zorn and Nikolaos V Sahinidis. Global optimization of general nonconvex problems with intermediate polynomial substructures. *Journal of Global Optimization*, 59(2-3):673–693, 2014.

[252] Jikai Zou, Shabbir Ahmed, and Xu Andy Sun. Stochastic dual dynamic integer programming. *Optimization Online. Available online at http://www.optimization-online.org/DB_HTML/2016/05/5436.html*, pages 1–30, 2016.