Solving Constrained Piecewise Linear Optimization Problems by Exploiting the Abs-Linear Approach

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat.)

im Fach Mathematik

eingereicht an der

Mathematisch-Naturwissenschaftlichen Fakultät

der Humboldt-Universität zu Berlin

von

M. Sc. Timo Kreimeier

Präsidentin der Humboldt-Universität zu Berlin:
Prof. Dr. Julia von Blumenthal

Dekanin der Mathematisch-Naturwissenschaftlichen Fakultät:
Prof. Dr. Caren Tischendorf

Tag der Einreichung: 24. Februar 2023

Gutachter/innen:	1. Prof. Dr. Andrea Walther
	2. Prof. Dr. Frauke Liers
	3. Prof. Dr. Marc Steinbach
Tag der mündlichen Prüfung:	16. Juni 2023

Acknowledgments

Foremost, I want to thank my supervisor, Prof. Dr. Andrea Walther. Firstly, for offering me a PhD position and thus giving me the opportunity to research in the field of nonsmooth optimization, which turned out to be very exciting for me. On the other hand, I would like to thank her for always giving me the best support during the whole time of my PhD and for always being open to a helpful exchange.

Furthermore, I would like to thank Prof. Dr. Frauke Liers and Prof. Dr. Marc Steinbach for their willingness to review this thesis and to become part of my doctoral committee. I would also like to thank the entire committee for their interest in my work and for spending their time.

Many thanks also go to Prof. Dr. Andreas Griewank, who was significantly involved in the initial ideas of this work. It was a special pleasure to discuss with him the world of abs-smooth functions.

In addition, I would like to thank my working group and the former members, as well as the members of the befriended working groups, both in Berlin and in the short time I spent in Paderborn, for the always friendly and nice working atmosphere. It was always great to exchange ideas about the most different things. Special thanks to Franz Bethke, Veronika Schulze and Olga Weiß.

Thanks also to all the proofreaders of this thesis for their helpful feedback.

Last but not least, I would like to thank my family and friends. The biggest thanks go to my parents, who always supported me on my way, even if they didn't understand what I am doing most of the time. Nevertheless, they are and always have been there for me.

Thank you all very much!

Abstract

This thesis presents an algorithm for solving finite-dimensional optimization problems with a piecewise linear objective function and piecewise linear constraints. For this purpose, it is assumed that the functions are in the so-called Abs-Linear Form, a matrix-vector representation. Using this form, the domain space can be decomposed into polyhedra, so that the nonsmoothness of the piecewise linear functions can coincide with the edges of the polyhedra. For the class of abs-linear functions, necessary and sufficient optimality conditions that can be verified in polynomial time are given for both the unconstrained and the constrained case.

For unconstrained piecewise linear optimization problems, Andrea Walther and Andreas Griewank already presented a solution algorithm called the Active Signature Method (ASM) in 2019. Building on this method and combining it with the idea of the Active Set Method to handle inequality constraints, a new algorithm called the Constrained Active Signature Method (CASM) for constrained problems emerges. Both algorithms explicitly exploit the piecewise linear structure of the functions by using the Abs-Linear Form. Part of the analysis of the algorithms is to show finite convergence to local minima of the respective problems as well as an efficient solution of the saddle point systems occurring in each iteration of the algorithms.

The numerical performance of CASM is illustrated by several examples. The test problems cover academic problems, including bi-level and linear complementarity problems, as well as application problems from gas network optimization and inventory problems.

Keywords: nonsmooth optimization, optimality conditions, Abs-Linearization, Abs-Linear Form, constrained optimization, quadratic overestimation method, CASM, gas network optimization

Zusammenfassung

In dieser Arbeit wird ein Algorithmus zur Lösung von endlichdimensionalen Optimierungsproblemen mit stückweise linearer Zielfunktion und stückweise linearen Nebenbedingungen vorgestellt. Dabei wird angenommen, dass die Funktionen in der sogenannten Abs-Linear Form, einer Matrix-Vektor-Darstellung, vorliegen. Mit Hilfe dieser Form lässt sich der Urbildraum in Polyeder zerlegen, so dass die Nichtglattheiten der stückweise linearen Funktionen mit den Kanten der Polyeder zusammenfallen können. Für die Klasse der abs-linearen Funktionen werden sowohl für den unbeschränkten als auch für den beschränkten Fall notwendige und hinreichende Optimalitätsbedingungen bewiesen, die in polynomialer Zeit verifiziert werden können.

Für unbeschränkte stückweise lineare Optimierungsprobleme haben Andrea Walther und Andreas Griewank bereits 2019 mit der Active Signature Method (ASM) einen Lösungsalgorithmus vorgestellt. Aufbauend auf dieser Methode und in Kombination mit der Idee der aktiven Mengen Strategie zur Behandlung von Ungleichungsnebenbedingungen entsteht ein neuer Algorithmus mit dem Namen Constrained Active Signature Method (CASM) für beschränkte Probleme. Beide Algorithmen nutzen die stückweise lineare Struktur der Funktionen explizit aus, indem sie die Abs-Linear Form verwenden. Teil der Analyse der Algorithmen ist der Nachweis der endlichen Konvergenz zu lokalen Minima der jeweiligen Probleme sowie die Betrachtung effizienter Berechnung von Lösungen der in jeder Iteration der Algorithmen auftretenden Sattelpunktsysteme.

Die numerische Performanz von CASM wird anhand verschiedener Beispiele demonstriert. Dazu gehören akademische Probleme, einschließlich bi-level und lineare Komplementaritätsprobleme, sowie Anwendungsprobleme aus der Gasnetzwerkoptimierung und dem Einzelhandel.

Stichworte: Nichtglatte Optimierung, Optimalitätsbedingungen, Abs-Linearisierung, Abs-Linear Form, beschränkte Optimierung, quadratische Überschätzungsmethode, CASM, Gasnetzwerkoptimierung

Contents

1 Introduction **1**

1.1 Motivation . . . 1

1.2 Contributions . . . 3

1.3 Structure of this Dissertation . . . 4

2 Aspects of Smooth Optimization **7**

2.1 Optimality Conditions for Smooth Constrained Optimization . . . 7

2.2 Active Set Method . . . 10

2.2.1 Equality Constrained Problem . . . 11

2.2.2 Strategy of Active Sets for Inequality Constraints . . . 13

2.2.3 Finite Convergence . . . 16

2.2.4 Degeneracy, Cycling, and Further Remarks . . . 17

2.2.5 Finding a Feasible Starting Point: Phase I Methods . . . 18

2.3 Penalty-Approaches . . . 20

2.3.1 General Penalty Method . . . 20

2.3.2 l_1-Penalty Method . . . 22

2.3.3 Quadratic Penalty Approach . . . 23

3 The Abs-Linear Optimization Problem: Theory and Solution Method **25**

3.1 Abs-Linear Form and Abs-Linearization . . . 26

3.2 Optimality Conditions . . . 31

3.3 Active Signature Method . . . 39

3.3.1 Computing a Descent Direction for Given σ . . . 40

3.3.2 Computing a Step Size β . . . 42

3.3.3 Checking the Optimality . . . 44

3.3.4 The Overall Algorithm . . . 44

3.3.5 Convergence Analysis of the Active Signature Method . . . 45

3.4 Optimization Methods Based on Abs-Linearization . . . 51

4 The Constrained Abs-Linear Optimization Problem: Theory and Solution Method **53**

4.1 Optimality Conditions . . . 56

4.2 Constrained Active Signature Method . . . 63

4.2.1 Computing a Descent Direction for Given σ and ω . . . 64

4.2.2 Computing a Step Size β . . . 68
4.2.3 Checking the Optimality . . . 70
4.2.4 The Overall Algorithm . . . 71
4.2.5 Convergence Analysis of the Constrained Active Signature Method . 72
4.3 Penalty Approaches for Piecewise Linear Optimization Problems . . . 83

5 Numerical Results **87**
5.1 Aspects of Implementation . . . 87
5.2 Academical Examples . . . 89
5.3 The Gas Transport Problem . . . 97
5.3.1 GasLib-Instances . . . 100
5.4 Piecewise Linear Regression in Retail . . . 110

6 Conclusions and Outlook **113**
6.1 Summary . . . 113
6.2 Future Research Directions . . . 114

Bibliography **119**

List of Figures

3.1 Plot of the Hill-function 27
3.2 Signature domains of the Hill-function 30
3.3 (Non) signature optimal points for Hill-function 33
3.4 Illustration of the step sizes β^z 42
3.5 Scheme of Active Signature Method (ASM) 46

4.1 Feasible signature domains for constrained Hill-function 55
4.2 Illustration of the two different step sizes β^z and β^H 69
4.3 Scheme of Constrained Active Signature Method (CASM) 73

5.1 Illustration of the iteration sequences generated by Constrained Active Signature Method (CASM) for the Hill-function 89
5.2 Illustration of the iteration sequences generated by CASM for the constrained HUL-function 92
5.4 Optimization history of CASM for the GasLib-11 cascade 101
5.3 Topology of GasLib-11 101
5.5 Topology of GasLib-40 102
5.6 Optimization history of CASM for the GasLib-40 cascade 103
5.7 Comparison of the third optimization for nonrobust feasible GasLib-40 with and without warm start 104
5.8 Topology of GasLib-134 105
5.9 Optimization history of CASM for the GasLib-134 cascade 106
5.10 Topology of GasLib-582 107
5.11 Optimization history of CASM for GasLib-582 with feasible starting point . 109
5.12 Optimization history of Active Signature Method (ASM) for GasLib-582 with nonfeasible starting point 109
5.13 Optimization history for the Phase I and CASM of the GasLib-582 using the zero vector as nonfeasible starting point 110
5.14 Optimization history of CASM for the optimization problem (5.8) 112

6.1 Illustration of the iteration sequences generated by Successive Abs-Linear MINimization (SALMIN) for Example 6.1 117

List of Tables

5.1 Optimization history of CASM for the Hill-function 90

5.2 Optimization history of CASM for the constrained HUL-function 91

5.3 Number of iterations generated by CASM for the constrained Rosenbrock-Nesterov II example with different values of dimension n 94

5.4 Optimization history of CASM for the linear bi-level problem 96

5.5 Complexity of GasLib-11 instances for their respective optimizations and iterations needed by CASM 102

5.6 Complexity of GasLib-40 instances for their respective optimizations and iterations needed by CASM 103

5.7 Complexity of GasLib-134 instances for their respective optimizations and iterations needed by CASM 105

5.8 Comparison of the complexity of GasLib-582 instances for their respective optimizations and iterations needed by CASM and ASM with a l_1-penalty approach 108

5.9 Comparison of the complexity of the optimization problems (5.8) and (5.9) as well as iterations needed by CASM 112

Glossary, Acronyms and Symbols

Glossary

s.t. subject to.

e.g., exempli gratia (Latin) *or* for example (English).

i.e., id est (Latin) *or* that is to say (English).

w.l.o.g. without loss of generality.

Acronyms

AD algorithmic differentiation.

ASM Active Signature Method.

CASM Constrained Active Signature Method.

GPM General Penalty Method.

KKT Karush-Kuhn-Tucker.

LICQ linear independence constraint qualification.

LIKQ linear independence kink qualification.

MIP mixed-integer programming.

SALMIN Successive Abs-Linear MINimization.

SQP sequential quadratic programming.

Symbols

Δx	search direction.
β	step length.
$\mathcal{C}^d_{\text{abs}}$	class of abs-smooth functions.
$\mathcal{F}$	feasible Set.
$\mathcal{F}_\sigma$	feasible signature domain.
$\mathcal{I}(x)$	index set of active inequality constraints in point x.
$\mathcal{P}_\sigma$	signature domain.
λ, δ, ν	Lagrange-multipliers (except in section 5.3).
$\mathbb{N}$	set of natural numbers.
$\mathbb{R}$	set of real numbers.
∇	gradient operator.
σ, Σ	signature vector and matrix, respectively.
φ	nonsmooth (target) function.
f	smooth or abs-normal respectively abs-linear (target) function.
g, A, B, C	vector and matrices of equality constrains, respectively.
h, D, E, F	vector and matrices of inequality constrains, respectively.
m, p	number of equality and inequality constraints, respectively.
n	dimension of the variable space.
s	number of switching variables.
x_i	ith component of the vector x.

1 Introduction

1.1 Motivation

Solution methods for piecewise linear optimization problems, i.e., mathematical optimization problems where both the objective function and the equality and inequality constraints are piecewise linear, are receiving increasing attention and can be motivated by several applications. According to [96], a mathematical function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be piecewise linear if there exists a $k \in \mathbb{N}$ and a finite set of affine linear functions $f_i(x)$, $i = 1, \ldots, k$ such that the inclusion $f(x) \in \{f_1(x), \ldots, f_k(x)\}$ holds for all $x \in \mathbb{R}^n$. This piecewise composition leads to points in the domain of the function at which the function itself is no longer differentiable and therefore difficult for derivative based optimization algorithms. For illustration, a comparison between smooth and nonsmooth optimization is helpful.

In smooth optimization, linearization techniques are often used to generate local linear or quadratic models of the original problem. The classical approaches, such as trust-region methods, sequential quadratic programming (SQP) or Newton's methods, require gradients of the original problem [32, 87]. Such methods usually rely on the Taylor expansion to produce an approximation of the functions. Thus, depending on the distance to the evolutionary point, the models have local errors of an appropriate order [35]. For smooth functions, gradients exist from a theoretical point of view and can be computed numerically using algorithmic differentiation (AD). Examples of tools for algorithmic differentiation, also called automatic differentiation, are the packages ADOL-C [106], TAPENADE [53] or CoDiPack [94] available for C/C++, and ADiMAT [14] for Matlab, just to name a few. For a more detailed list and more information about AD, see e.g., [16, 43].

However, unlike the smooth functions, these considerations do not transfer so easily to nonsmooth functions. Here it is not a reasonable assumption that derivative information is available at every point and thus one cannot directly generate a local model with the Taylor expansion. In his article [41], published in 2013, Andreas Griewank addressed precisely this problem. There he presented an idea of a piecewise linearization for Lipschitz continuous and thus not necessarily smooth functions. According to Rademacher's theorem Lipschitz continuous functions are differentiable almost everywhere, i.e., the points in which f is not differentiable form a set of Lebesgue measure zero [27]. In addition, a second-order approximation property can be shown for this piecewise linearization approach. The focus is on the nonsmooth absolute value function abs as well as the min and max functions, which can also be expressed by the absolute value. Furthermore, Andreas Griewank describes the possibility to generalize algorithmic differentiation for this problem class.

Therefore, the focus of this dissertation is the handling of piecewise linear functions in the optimization context, i.e., the class of functions arising from linearization. Because of the second-order approximation property due to linearization, on the one hand, the natural motivation arises to consider such an optimization algorithm as an inner solver for general algorithms for nonsmooth problems, e.g., SALMIN [28]. Additionally, it can be used to solve directly local models for the original problem [50, 77]. On the other hand, there are concrete application areas for piecewise linear problems like, e.g., train time tabling [31], shallow training of neural networks with the Rectified Linear Unit (ReLU) as activation function [40, 101] or inventory problems [4].

Another large class of applications are the mixed-integer (non)linear optimization problems [11]. In [38] it is described how piecewise linear functions can be used to solve such problems. However, it is known that it is NP-hard to solve mixed-integer or piecewise linear optimization problems [65]. In the latter this is also the case when exact penalty methods are used to handle the piecewise linear constraints [60].

Mixed-integer optimization problems may arise, e.g., in gas network optimization. A recent approach is based on the idea of solving series of mixed-integer linear relaxations [6]. As described in more detail later in this thesis, this results in piecewise linear optimization problems. This, as well as the various application examples given before, motivate the development of algorithms for the special class of piecewise linear optimization problems.

For unconstrained piecewise linear optimization problems, Andrea Walther and Andreas Griewank introduced the ASM in 2019 [45]. The idea of this method is to decompose the domain into polyhedra such that the nonsmoothness of the function lies on the edges of the polyhedra. Therefore, the objective function is converted into the so-called Abs-Linear Form [41], which is a matrix-verctor-based representation of the function. Subsequently, an appropriately linearized optimization problem with a quadratic regularization term arises on each polyhedron and, beginning from a starting point and its associated polyhedron, a sequence of these subproblems is solved. Unsatisfied optimality conditions are used to decide on which neighboring polyhedra to optimize further, until finally a local minimum is found. To verify this, an optimality condition can be derived, which can be checked as a matrix-vector product with polynomial effort.

The goal of this work is to develop an optimization algorithm based on the ASM, which, in addition to the piecewise linear objective function, also allows for piecewise linear equality and inequality constraints. For this purpose, the idea of the polyhedral decomposition of ASM is extended by handling the constraints in an active set sense. Since the constraints are also piecewise linear, they in addition to the objective function influence the polyhedral decomposition into potentially more polyhedra than in the unconstrained case of the same objective function. This new algorithm is called Constrained Active Signature Method (CASM) and optimality conditions are also proven for it, which can be verified with polynomial effort. Afterwards, its numerical performance is illustrated using various examples of applications, both academic and real world. In the latter, the main focus is on the previously described subproblems from gas network optimization but also a piecewise linear regression problem motivated by inventory problems.

1.2 Contributions

A substantial part of this thesis is devoted to the development of an optimization algorithm for solving constrained piecewise linear optimization problems and testing it on various application examples. The main contributions of this thesis can be summarized as follows:

- State the already known Active Signature Method (ASM) from [45], collecting various results from different publications on the class of abs-linear functions and supplementing them by details missing so far.
- Show necessary and sufficient optimality conditions for (constrained) piecewise linear optimization problems, which can be tested with polynomial effort.
- Development of the Constrained Active Signature Method (CASM) for constrained piecewise linear optimization problems combining ASM and Active Set Method.
- Proof of finite convergence to local optima for ASM and CASM.
- Numerical performance tests for CASM using acdemic and real application problems with a focus on subproblems arising from gas network optimization.
- For relaxed subproblems from the gas network optimization, as alternative to MIP solver, using CASM that allows a warm start strategy such that it can profit from the results obtained for coarser relaxations.

The theoretical results, the solution method CASM as well as the numerical results in this dissertation are based essentially on the following articles and proceedings contributions, to which the author has given a substantial contribution:

- Timo Kreimeier, Andrea Walther, and Andreas Griewank. An active signature method for constrained abs-linear minimization. Submitted to Computational Optimization and Applications. Available at https://opus4.kobv.de/opus4-trr154/frontdoor/index/index/docId/474, 2022.
- Timo Kreimeier, Martina Kuchlbauer, Frauke Liers, Michael Stingl, and Andrea Walther. Towards the solution of robust gas network optimization problems using the constrained active signature method. In Christina Büsing and Arie M. C. A. Koster, editors, *Network Optimization INOC 2022*, 2022.

In addition, the following two reports were drafted during the authors PhD phase, in which the author also participated:

- Timo Kreimeier, Henning Sauter, Tom Streubel, Caren Tischendorf, and Andrea Walther. Solving least-squares collocated differential algebraic equations by successive abs-linear minimization - a case study on gas network simulation. Available at https://opus4.kobv.de/opus4-trr154/frontdoor/index/index/docId/473, 2021.
- Timo Kreimeier, Sebastian Pokutta, Andrea Walther, and Zev Woodstock. On a Frank-Wolfe approach for abs-smooth functions. Available at https://opus4.kobv.de/opus4-trr154/frontdoor/index/index/docId/499, 2022.

1.3 Structure of this Dissertation

The main focus of this work is to present an optimization algorithm, called CASM, which can be used to solve piecewise linear optimization problems with piecewise linear constraints. For this purpose, various application fields were already shown in Chapter 1 as motivation, where piecewise linear optimization problems can occur (see Section 1.1). For the upcoming derivation of CASM, some needed basics such as the excerpts from KKT theory (Section 2.1) and the Acitve Set Method, (Section 2.2) including Phase I Methods for determining feasible starting points, are given in Chapter 2. These introductory aspects conclude with Section 2.3 on penalty approaches, which provide an alternative to the explicit treatment of constraints.

In Chapter 3, the Active Signature Method (ASM) as already published in [45] is presented and supplemented by further details. For this purpose, the concepts of Abs-Linear Forms is introduced in Section 3.1 and it is briefly described how an Abs-Linear Form can be generated for arbitrary abs-smooth functions using abs-linearization. After presenting optimality conditions for abs-linear problems in Section 3.2 and adding a more detailed proof, the ASM is then the subject of Section 3.3. As a special highlight of the optimality condition it should be pointed out that it can be verified in polynomial time. The ASM is an algorithm for solving unconstrained piecewise linear optimization problems. Essentially, it consists of three components namely computing a descent direction, determining a step size along that direction, and checking optimality. After the algorithm is then given as pseudocode, its finite convergence will be shown. At the end of the chapter, optimization methods based on abs-linearization are briefly discussed (see Section 3.4).

As an extension of ASM, Chapter 4 then focuses on the Constrained Active Signature Method (CASM), which considers additional piecewise linear constraints by using an active set strategy. The structure of this chapter is congenial to that of Chapter 3. First, Section 4.1 establishes the counterpart pendants of the necessary and sufficient optimality conditions for the constrained case, which can again be verified in polynomial time. The CASM itself is then the subject of Section 4.2, which, analogous to ASM, consists of the three parts of determining the direction of descent, computing the step size, and checking optimality. Subsection 4.2.4 is used to state the overall algorithm, and afterwards its finite convergence will be shown as well. This chapter concludes then with a brief discussion of penalty approaches as an alternative to explicit handling of constraints in Section 4.3.

The numerical results for the performance of CASM on different example problems, also in comparison to other solvers like Gurobi [52], are presented in Chapter 5. At first some aspects concerning the implementation of CASM, e.g., exploiting the sparsity of matrices and choosing a suitable solver for a occurring linear system of equations, are discussed. Thereafter, with respect to dimension, rather smaller and academic examples are computed (see Section 5.2). These also include in dimension scalable or bi-level problems as well as those involving linear complementarity constraints. A major focus is then in Section 5.3 on application examples from gas network optimization. Here, a subproblem from an adaptive bundle method is computed for different GasLib instances and scenarios. The possibility of a warm start strategy for mixed integer problems is also demonstrated. As a last application, a piecewise linear balancing problem is solved, which is known from optimization problems in the retail industry (see Section 5.4).

Finally, this thesis concludes with Chapter 6, in which the main results are summarized (see Section 6.1) and an outlook on possible future research directions is given (see Section 6.2).

2

Aspects of Smooth Optimization

This chapter should introduce the necessary mathematical basics from smooth constrained optimization. For this purpose, the chapter is structured as follows: Section 2.1 introduces a class of optimality conditions, the so-called Karush-Kuhn-Tucker (KKT)-conditions. Subsequently, in Section 2.2 the Active Set Method is explained, which can be used to solve quadratic optimization problems with linear equality and inequality constraints. Therefore, the theory motivating the algorithm is described, the algorithm itself is stated as a pseudocode, and statements about its convergence are given. As an alternative to the explicit handling of the constraints, as it is done by the Active Set Method, in Section 2.3 a passage in which penalty approaches are presented is following. These are methods which combine the constraints as penalty terms to the objective function and thus create an unconstrained optimization problem from the constrained one.

2.1 Optimality Conditions for Smooth Constrained Optimization

One of the most widely used and well-known necessary optimality conditions are considered to be the KKT-conditions, named after the US mathematicians William Karush, Harold William Kuhn and Albert William Tucker. Historically, these were first shown in his master's thesis by William Karush in 1939 [64], but did not gain more popularity until a conference paper was published by Harold W. Kuhn and Albert William Tucker in 1950 [72]. However, the formulations and results given here are from [37, Chapter 2].

Definition 2.1 (KKT-conditions)**.** *Consider the optimization problem*

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & f(x) \\ s.t. \quad & 0 = G(x)\ , \\ & 0 \geq H(x)\ , \end{aligned} \tag{2.1}$$

with $f : \mathbb{R}^n \to \mathbb{R}$*,* $G : \mathbb{R}^n \to \mathbb{R}^m$ *and* $H : \mathbb{R}^n \to \mathbb{R}^p$ *all continuously differentiable. The conditions*

$$\begin{aligned} \nabla_x L(x, \delta, \nu) &= 0 \ , \\ G(x) &= 0 \ , \\ H(x) &\le 0 \ , \\ \nu^\top H(x) &= 0 \ , \\ \nu &\ge 0 \ . \end{aligned}$$

are called KKT-conditions *of* (2.1)*, where* $\nabla_x L(x, \delta, \nu)$ *denotes the derivative of the* Lagrange function

$$L(x, \delta, \nu) = f(x) + \sum_{i=1}^{m} \delta_i G_i(x) + \sum_{i=1}^{p} \nu_i H_i(x)$$

with respect to the argument x *and multipliers* $\delta \in \mathbb{R}^m$ *and* $\nu \in \mathbb{R}^p$*. Every point* $(x^*, \delta^*, \nu^*) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$ *that fulfills the KKT-conditions is called* KKT-point *of* (2.1) *and the componenents of* δ^* *and* ν^* *or often the vectors themselves are called* Lagrange-multipliers*.*

In the later parts of this thesis special attention will be paid to piecewise linear functions and especially to those as constraints. Therefore, first a result based on the KKT-conditions for the linear case is given here, which will also be useful for the piecewise linear consideration.

Theorem 2.2 (KKT-conditions for linear constrained problems)**.** *Let* x^* *be a local minimizer of the linear constrained problem*

$$\begin{aligned} \min_{x \in \mathbb{R}^n} \quad & f(x) \\ s.t. \quad & 0 = g + Ax \ , \\ & 0 \ge h + Dx \ , \end{aligned} \tag{2.3}$$

with $f : \mathbb{R}^n \to \mathbb{R}$ *continuously differentiable,* $g \in \mathbb{R}^m$*,* $h \in \mathbb{R}^p$*,* $A \in \mathbb{R}^{m \times n}$ *and* $D \in \mathbb{R}^{p \times n}$*. Then, there exist Lagrange-multipliers* $\delta^* \in \mathbb{R}^m$ *and* $\nu^* \in \mathbb{R}^p$ *such that* (x^*, δ^*, ν^*) *fulfill the*

KKT-conditions

$$\nabla f(x^*) + (\delta^*)^\top A + (\nu^*)^\top D = 0\ , \tag{2.4a}$$
$$g + Ax^* = 0\ , \tag{2.4b}$$
$$h + Dx^* \leq 0\ , \tag{2.4c}$$
$$(\nu^*)^\top (h + Dx^*) = 0\ , \tag{2.4d}$$
$$\nu^* \geq 0\ . \tag{2.4e}$$

Proof. See [37, Satz 2.42]. □

Moreover, if the objective function is assumed to be convex, the KKT-conditions are even sufficient optimality conditions in the case of linear constraints.

Theorem 2.3 (KKT-conditions as necessary and sufficient optimality conditions)**.** *Consider the optimization problem* (2.3) *and additionally assume that f is convex. Then $x^* \in \mathbb{R}^n$ is a (local = global) minimizer of* (2.3)*, if and only if there exist Lagrange-multiplier $\delta^* \in \mathbb{R}^m$ and $\nu^* \in \mathbb{R}^p$ such that the tuple (x^*, δ^*, ν^*) is a KKT-point of* (2.3)*.*

Proof. See [37, Korollar 2.47]. □

For both linear and nonlinear constraints, if one adds further regularity assumptions, the so-called LICQ, it can be shown that the Lagrange multipliers must even be unique. The definition of the LICQ will given next, followed by corresponding result.

Definition 2.4 (LICQ)**.** *Let $x \in \mathbb{R}^n$ be a feasible point of the optimization problem* (2.1) *and $\mathcal{I}(x) = \{i \mid H_i(x) = 0\}$ the corresponding set of active inequality constraints. Then x fulfills the* linear independence constraint qualification (LICQ) *if the gradients*

$$\nabla G_i(x) \text{ for } i = 1, \ldots, m \quad \text{and} \quad \nabla H_i(x) \text{ for } i \in \mathcal{I}(x)$$

are linear independent.

Theorem 2.5 (KKT-conditions for constrained problems with C^1-functions)**.** *Let x^* be a local minimizer of the optimization problem* (2.1) *which fulfills the LICQ. Then there exist unique Lagrange multipliers $\delta^* \in \mathbb{R}^m$ and $\nu^* \in \mathbb{R}^p$ such that the tuple (x^*, δ^*, ν^*) is a KKT-point of* (2.1)*.*

Proof. See [37, Satz 2.41]. □

In addition to LICQ, there are nowadays countless further regularity assumptions as well as optimality conditions based on them, see, e.g., [1, 37, 46, 82, 99] to only name a few. In the context of this thesis, however, the one mentioned earlier suffice for the upcoming discussions.

For the later convergence analysis, the concept of the descent direction will be important. Therefore, the following two definitions as well as the Lemma 2.8 are taken from [2, Chapter 1 and 3]. From the naming it is already self-explaining what descent direction means: It concerns a direction, in which the function value decreases.

Definition 2.6 (Descent direction). *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function and $x \in \mathbb{R}^n$. A vector $d \in \mathbb{R}^n$ is called* descent direction *of f in direction d, if there exists a $\bar{t} > 0$ such that $f(x + td) < f(x)$ holds for all $t \in\ ]0, \bar{t}[$.*

To characterize descent directions one can use directional derivatives. These are defined as follows:

Definition 2.7 (Directional derivative). *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function and $x, d \in \mathbb{R}^n$. If the limit*

$$f'(x; d) := \lim_{t \searrow 0} \frac{1}{t} \left(f(x + td) - f(x) \right)$$

exists then $f'(x; d)$ is called directional derivative *of f in the point x in the direction d.*

In case of convex functions the relation of descent direction and directional derivative is now described by the following lemma:

Lemma 2.8. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function and $x, d \in \mathbb{R}^n$. Then the following statements are equivalent*

i) d is a descent direction of f in x,

ii) $f'(x; d) < 0$.

Proof. See [2, Satz 3.2.2]. □

2.2 Active Set Method

The Active Set Method has been developed to solve quadratic optimization problems with linear equality and inequality constraints. It belongs to the standard methods of quadratic optimization and is accordingly also discussed in many lectures on nonlinear optimization. It can also be found in many standard textbooks on this topic. This chapter is also essentially based on the two books [37, 87].

To avoid confusion with the Active Signature Method (ASM), no acronym is used for the Active Set Method in this thesis. However, to make the conceptual similarities to the ASM more recognizable, the notation from the corresponding books is adapted to the setting considered here.

The basic idea of the Active Set Method is to solve a sequence of equality-constrained optimization problems where the active inequality constraints at the current iterate are added to the equality constraints. Theoretical considerations then provide information about which of the inequality constraints are added as active constraints for the next iteration or which become inactive again. The handling of the active or inactive inequality constraints is then described by a set. Therefore, this is also an explanation for the naming of the Active Set Method.

2.2.1 Equality Constrained Problem

Consider in a first step only the following equality constrained optimization problem

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & a^\top x + \frac{1}{2} x^\top Q x \\ \text{s.t.} \quad & 0 = g + Ax \ , \end{aligned} \tag{2.5}$$

with $Q \in \mathbb{R}^{n\times n}$ symmetric, $a \in \mathbb{R}^n, g \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m\times n}$. Hence, in this case there are n independent optimization variables and m equality constraints. Note, that w.l.o.g. a constant shift in the objective function has been omitted here. Assume $x^* \in \mathbb{R}^n$ to be a local minimizer, then by Theorem 2.2 there exists at least one Lagrange-multiplier $\delta^* \in \mathbb{R}^m$ such that the pair (x^*, δ^*) fulfills the KKT-conditions, i.e.,

$$\begin{aligned} a^\top + x^\top Q + \delta^\top A &= 0 \ , \\ g + Ax &= 0 \ . \end{aligned}$$

This is a linear system of equations to determine a KKT-point and directly yields the following result:

Theorem 2.9. *A pair $(x^*, \delta^*) \in \mathbb{R}^n \times \mathbb{R}^m$ is a KKT-point of Eq. (2.5), if and only if*

(x^*, δ^*) *is a solution of the system of linear equations*

$$\begin{pmatrix} Q & A^\top \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ \delta \end{pmatrix} = - \begin{pmatrix} a \\ g \end{pmatrix} . \tag{2.6}$$

Proof. See [37, Satz 5.1]. □

In the case of a positive semi-definite matrix Q the target function is convex and thus, by Theorem 2.3, the KKT-conditions of (2.5) are fully equivalent to the optimization problem itself. Thus, a solution of an equality constrained optimization problem can be reduced to the solution of corresponding systems of linear equations (2.6).

By elementary rearrangement the following statement follows directly from Theorem 2.9, which is already a preparation for the Active Set Method. Therefore, substituting x by $x + \Delta x$ in Eq. (2.6), where x now denotes a feasible and fixed point for the quadratic optimization problem (2.5) and $\Delta x \in \mathbb{R}^n$ a correction term, which plays the role of a search direction. Hence, it follows

$$\begin{aligned} & \begin{pmatrix} Q & A^\top \\ A & 0 \end{pmatrix} \begin{pmatrix} x + \Delta x \\ \delta \end{pmatrix} = - \begin{pmatrix} a \\ g \end{pmatrix} \\ \Leftrightarrow & \begin{pmatrix} Q & A^\top \\ A & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \delta \end{pmatrix} = - \begin{pmatrix} a \\ g \end{pmatrix} - \begin{pmatrix} Q & A^\top \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} -a - Qx \\ -g - Ax \end{pmatrix} \\ \Leftrightarrow & \begin{pmatrix} Q & A^\top \\ A & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \delta \end{pmatrix} = \begin{pmatrix} -\nabla f(x) \\ 0 \end{pmatrix} , \end{aligned}$$

where in the last rearrangement $\nabla f(x) = a + Qx$ and the feasibility of x for (2.5) is used. This immediately provides the following theorem:

Theorem 2.10. *Let $\bar{x} \in \mathbb{R}^n$ be a feasible point of the optimization problem (2.5). Then, $(x^*, \delta^*) \in \mathbb{R}^n \times \mathbb{R}^m$ is a KKT-point of (2.5), if and only if, $x^* = \bar{x} + \Delta x^*$ and $(\Delta x^*, \delta^*)$ is a solution of the system of linear equations*

$$\begin{pmatrix} Q & A^\top \\ A & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \delta \end{pmatrix} = \begin{pmatrix} -\nabla f(\bar{x}) \\ 0 \end{pmatrix} ,$$

with $f(x) = a^\top x + \frac{1}{2} x^\top Q x$.

Proof. See [37, Satz 5.2]. □

2.2.2 Strategy of Active Sets for Inequality Constraints

Having initially only the equality constraints taken into account, the inequality constraints are now added to the optimization problem (2.5). Thus, the optimization problem considered in the Active Set Method, is described by

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & a^\top x + \frac{1}{2} x^\top Q x \\ \text{s.t.} \quad & 0 = g + Ax \;, \\ & 0 \geq h + Dx \;, \end{aligned} \tag{2.7}$$

again with $Q \in \mathbb{R}^{n\times n}$ symmetric and positive semi-definite, $a \in \mathbb{R}^n, g \in \mathbb{R}^m, A \in \mathbb{R}^{m\times n}$ and furthermore $h \in \mathbb{R}^p$ and $D \in \mathbb{R}^{p\times n}$. As mentioned at the beginning of the chapter, to solve the actual problem (2.7) a sequence of equality-constrained problems is solved using Theorem 2.9. This is followed by only adding the inequality constraints which are active at the current iterate to the equality constraints. The so-called working set $\mathcal{W}$ is introduced to approximate these index set of active inequality constraints in the point x given by

$$\mathcal{I}(x) = \left\{ i \in \{1, \dots, p\} \mid e_i^\top (h + Dx) = 0 \right\} \;,$$

where $e_i \in \mathbb{R}^p$ denotes the ith unit vector. Then, the projection onto the approximation $\mathcal{W}$ of the active inequality constraints is defined as $P_{\mathcal{W}} \equiv (e_i^\top)_{i\in\mathcal{W}} \in \mathbb{R}^{|\mathcal{W}|\times p}$.

Using these notations, in every iteration of the Active Set Method the optimization problem

$$\begin{aligned} \min_{x\in\mathbb{R}^n} \quad & a^\top x + \frac{1}{2} x^\top Q x \\ \text{s.t.} \quad & 0 = g + Ax \;, \\ & 0 = P_{\mathcal{W}}(h + Dx) \end{aligned} \tag{2.8}$$

has to be solved, which is by Theorem 2.9 equivalent to setting $x = \bar{x} + \Delta x$ for the solution Δx to

$$\begin{pmatrix} Q & A^\top & (P_{\mathcal{W}}D)^\top \\ A & 0 & 0 \\ P_{\mathcal{W}}D & 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \delta \\ \nu_{\mathcal{W}} \end{pmatrix} = \begin{pmatrix} -\nabla f(\bar{x}) \\ 0 \\ 0 \end{pmatrix} \;, \tag{2.9}$$

with $\nu_{\mathcal{W}} \equiv P_{\mathcal{W}}\nu \in \mathbb{R}^{|\mathcal{W}|}$ and $\bar{x}$ denoting the current point. Starting from a solution $(\Delta x, \delta, \nu_{\mathcal{W}})$ of (2.9), two main questions arise. First, how to determine the new iterate including the calculation of the step length, i.e., how far to go in the direction of Δx and second, how does the working set change.

To begin with, because of the complementarity condition Eq. (2.4d) of the KKT-conditions, ν_i is set to zero for all index $i \notin \mathcal{W}$. Now distinguish the two cases $\Delta x = 0$ and $\Delta x \neq 0$. If the first one holds true, the current iterate is a candidate for a local minimzer of the whole problem (2.7), at least it is a local minimizer of (2.8). To be a local minimizer of (2.7) by Theorem 2.3 it is necessary and sufficient that the KKT-conditions (2.4) are fulfilled. Therefore, the four conditions, namely the stationarity condition (2.4a) the two primal feasibility (2.4b) and (2.4c) as well as the complementary condition (2.4d), are fulfilled via the construction above. Thus, only the dual feasibility condition (2.4e) has to be checked. If this is also fulfilled the algorithm stops. If not, any index j for which $\nu_j < 0$ dropping the jth inequality constraint usually will yield a direction Δx along which the algorithm can make progress as can be seen by the following theorem.

Theorem 2.11. *Suppose that the point $\bar{x}$ satisfies the first-order conditions for the equality-constrained problem with the working set $\mathcal{W}$, i.e.,*

$$P_{\mathcal{W}} D \nu_{\mathcal{W}} = -Q\bar{x} - a$$

is satisfied along with $P_{\mathcal{W}}(h + D\bar{x}) = 0$. Further, assume that the matrix

$$\begin{pmatrix} A \\ P_{\mathcal{W}} D \end{pmatrix}$$

has full row rank, i.e., the active constraints are linearly independent, and finally suppose that there exists an index $j \in \mathcal{W}$ such that $\nu_j < 0$. Let Δx be the solution obtained by dropping the constraint j and solving the system

$$\begin{pmatrix} Q & A^\top & \left(P_{\mathcal{W}\setminus\{j\}} D\right)^\top \\ A & 0 & 0 \\ P_{\mathcal{W}\setminus\{j\}} D & 0 & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \delta \\ \nu_{\mathcal{W}\setminus\{j\}} \end{pmatrix} = \begin{pmatrix} -\nabla f(\bar{x}) \\ 0 \\ 0 \end{pmatrix}.$$

Then Δx is a feasible direction for jth inequality constraint, i.e., $e_j^\top D \Delta x \leq 0$

Proof. See [87, Theorem 16.5]. □

Here, the common way is to chose the most negative multiplier, which is motivated by the sensitivity analysis given in [87, Chapter 12].

Algorithm 1 Active Set Method

Require: Feasible start point $x \in \mathbb{R}^n, n \in \mathbb{N}, m, p \in \mathbb{N} \cup \{0\}, a \in \mathbb{R}^n, Q = Q^\top \in \mathbb{R}^{n \times n}$ symmetric and positive semi-definite, $g \in \mathbb{R}^m$, $h \in \mathbb{R}^p$, $A \in \mathbb{R}^{m \times n}$, $D \in \mathbb{R}^{p \times n}$, $\mathcal{W} := \{i \mid e_i^\top (h + Dx) = 0\}$

loop
  Set $\nu_i = 0$ for $i \notin \mathcal{W}$ und compute $(\Delta x, \delta, \nu_{\mathcal{W}})$ by solving (2.9)
  if $\Delta x = 0$ **then**
    if $\nu \geq 0$ **then** ▷ Termination
      return x
    else ▷ Drop Constraint
      Determine index j such that $\nu_j = \min\{\nu_i \mid i \in \mathcal{W}\}$
      Set $x^+ = x$ and $\mathcal{W}^+ = \mathcal{W} \setminus \{j\}$
  else
    if $x + \Delta x$ is feasible for (2.3) **then** ▷ Full step
      Set $x^+ = x + \Delta x$ and $\mathcal{W}^+ = \mathcal{W}$
    else ▷ Add Constraint
      Determine index j and step length β via Eq. (2.10)
      Set $x^+ = x + \beta \Delta x$ and $\mathcal{W} = \mathcal{W} \cup \{j\}$
  Set $x = x^+$ and $\mathcal{W} = \mathcal{W}^+$

Now consider the case where $\Delta x \neq 0$. The simple case here is when updating the current iterate x by adding Δx still yields a feasible point for the original problem (2.7). Then exactly this step is applied and the working set remains unchanged. The more complicated case is when a complete step in the direction of Δx no longer yields a feasible point. In this case a step length up to the first blocking inequality constraint must be determined. For the current iterate x the step length β is defined as

$$\beta \equiv \min_{i \notin \mathcal{W}} \left(\frac{-e_i^\top (h + Dx)}{e_i^\top (D\Delta x)} \,\middle|\, e_i^\top (D\Delta x) < 0 \right) \tag{2.10}$$

and the first index for which the minimum is attained is denoted by j. Using the step length the update of the current iterate and the working set is given by

$$x^+ := x + \beta \Delta x \,,$$
$$\mathcal{W}^+ := \mathcal{W} \cup \{j\} \,.$$

When all these considerations are combined, they lead to the Active Set Method as given in Algorithm 1.

2.2.3 Finite Convergence

For the Active Set Method, it is possible to show finite convergence under relatively mild assumptions. First, the following result is needed for this.

Theorem 2.12. *Let a solution Δx of the linear equation system (2.9), which is obtained form the current iterate $\bar{x}$, be given. Suppose, that Δx is nonzero and satisfies the second-order sufficient condition for optimality for that problem, i.e., $Z^\top QZ$ is positive definite for the matrix Z whose columns are a basis of the null space of the constraints of (2.8). Then, the function f, defined by*

$$f(x) := a^\top x + \frac{1}{2}x^\top Qx$$

is strictly decreasing along the direction Δx.

Proof. See [87, Theorem 16.6]. □

Now, according to [87, page 477 f.] the argumentation to show the finite convergence of Algorithm 1 is described. Assume that for strictly convex quadratic programs the Active Set Method determines a positive step size β in every iteration and furthermore that the direction Δx as the solution of the system (2.9) is nonzero. Then, the argumentation is as follows:

If $\Delta x = 0$ by Theorem 2.12 the current iterate $\bar{x}$ is the unique global minimizer of (2.8), i.e., for the current working set $\mathcal{W}$. If at least one of the Lagrange-multiplier of ν is negative, $\bar{x}$ is not the solution of the original problem (2.7), but then combining Theorems 2.11 and 2.12 leads to the conclusion that the direction Δx computed after a constraint is dropped will be a strict descent direction for f. Taking the assumption $\beta > 0$ into account, it follows that the value of f is lower than $f(\bar{x})$ at all subsequent iterations. Since x is actually the unique global minimizer for the current working set, the algorithm can never return to it, after the value of f has decrease once. In summary, these considerations provide two findings: First, after a working set has been left once, it cannot be reached again, and, second, after a constraint is dropped, there is a strict descent in the function value.

Furthermore, the algorithm encounters an iterate for which $\Delta x = 0$ at most on every $(p + 1)$th iterations, because either there is a constraint to add, which can occur at most p times and is discussed next or there is a full step, in which case the minimizer for the current working set is reached as seen by the equivalence of (2.8) and (2.9) and the next iterate leads to $\Delta x = 0$. Note, in case of linearly independent equality and active inequality constraints, there could be at most $p \leq n$ considered constraints.

It is easy to see that the cases of adding constraints or dropping constraints can only occur finitely often in sequence. Since there are only finitely many constraints, namely p of them, the case that an element is added to the working set can occur at most p times in a row. The same applies to the dropping out of constraints. Here it is also known that in the case of a real step, i.e., $\beta \Delta x \neq 0$, this working set is no longer reached. Note that the assumption of the existence of a positive step size is important here, since otherwise there may be a cycling between adding and dropping the same constraint. More about this is given in the following section.

In summary, there can only be a positive step size or a zero step size Δx each finitely often in series. And since there can only be a finite number of different working sets that cannot be reached again after leaving them, the algorithm can only run through the phase for $\Delta x = 0$ finitely often. This provides the finite convergence of the Active Set Method.

2.2.4 Degeneracy, Cycling, and Further Remarks

In this short subsection some further remarks about the Active Set Method are made, which should be mentioned in this context, but are not considered in detail here.

First of all it should be noted that scaling the constraints can influence the sequence of iterations. If one considers several constraints and scales some but not all of them with an arbitrary factor, these constraints remain mathematically equivalent. However, this may can have an influence on the value of the Lagrange multipliers when solving Eq. (2.9). I.e. possibly other inequality constraints are deactivated as it would be the case otherwise.

Further it should be mentioned that, for example, if LICQ is not satisfied in the optimal point or if the optimal point also coincides with the solution of the unconstrained problem, degeneracies may occur. It can happen that the Active Set Method cycles [80, 87, 113].

Finally, it should be mentioned that there are now a large number of Active Set Methods or those that build on a similar idea of handling constraints. Likewise, there are dual or primal-dual Active Set Methods in addition to the primal one. For this purpose, reference is made to [5, 23, 34, 57].

2.2.5 Finding a Feasible Starting Point: Phase I Methods

One assumption that is made for the Active Set Method, which has not been discussed yet, is the assumption of an existing feasible starting point. Up to now it was always assumed that a feasible starting point exists, but of course the question arises how to determine such a point if it is not yet known for the given problem for other reasons. One approach to this is called the *Phase I Method* and is motivated by the Simplex Algorithm, which was developed to solve linear constrained optimization problems [85, 86]. Thus, it is also immediately clear that for both the Simplex Algorithm and the Active Set Method, the feasible sets are described by linear functions and thus the same or similar approaches of a Phase I are motivated. Some approaches to this will now be briefly described and like the previous chapters, the current one is also essentially based on [87]. Alternatively, there are other sources, some of which are also closely related or also developed specifically for the field of operation research, e.g., [10, 17, 98] to name only a few.

A first approach for a Phase I Method can be described by the following linear optimization problem, where $\tilde{x}$ is a random starting point:

$$\begin{aligned} \min_{x\in\mathbb{R}^n,\pi\in\mathbb{R}^{m+p}} \quad & e^\top \pi \\ \text{s.t.} \quad & 0 = g_i + a_i^\top x + \theta_i \pi_i \;, \qquad i = 1,\dots,m \;, \\ & 0 \geq h_i + d_i^\top x + \theta_i \pi_i \;, \qquad i = m+1,\dots,m+p \;, \\ & 0 \leq \pi \;, \end{aligned}$$

where $a_i^\top$ and $d_i^\top$ denotes the ith row of the matrices A and D, respectively, and $e = (1,\dots,1)^\top$, $\theta = -\operatorname{sgn}(g_i + a_i^\top \tilde{x})$ for $i = 1,\dots,m$ and $\theta_i = -1$ for $i = m+1,\dots,m+p$. A feasible initial point for this problem is then

$$\begin{aligned} & x = \tilde{x} \;, \quad \pi_i = |g_i + a_i^\top \tilde{x}| \;, \text{ for } i = 1,\dots,m \;, \\ & \pi_i = \max(h_i + d_i^\top \tilde{x}, 0) \;, \text{ for } i = m+1,\dots,m+p \;. \end{aligned}$$

If the objective value of this problem is zero then the corresponding solution point $(x, 0)$ is a feasible starting point for the original problem (2.7) and the other way around if the original problem has a feasible point, then the optimal value in this Phase I-formulation is zero.

A second approach for a Phase I Method is the penalty-based, so called *Big M*-Method. In fact, this is no longer a strict Phase I Method in the sense that only a feasible starting point is searched for, but the objective function itself is also included in the optimization problem. By adding a scalar-valued variable ρ as a measure for the violation of the constraints to the target function, the following optimization problem is to be solved:

$$\begin{aligned}
\min_{x\in\mathbb{R}^n,\rho\in\mathbb{R}} \quad & a^\top x + \frac{1}{2}x^\top Qx + M\rho \\
\text{s.t.} \quad & \rho \geq g_i + a_i^\top x \;, && i = 1,\dots,m \;, \\
& \rho \geq -\left(g_i + a_i^\top x\right) \;, && i = 1,\dots,m \;, \\
& \rho \geq h_i + d_i^\top x \;, && i = m+1,\dots,m+p \;, \\
& \rho \geq 0 \;,
\end{aligned} \tag{2.11}$$

where M is a large positive value. For this problem, any random $\tilde{x}$ can be part of a feasible starting point, which is completed by an appropriate choice of ρ. This parameter must then be chosen large enough depending on $\tilde{x}$ so that all constraints are satisfied. A critical aspect of this approach is the appropriate choice of M. Here, a heuristic is to increase M and re-solve the problem until ρ becomes zero. If $\rho = 0$, then the problem formulation (2.11) is equivalent to the original (2.7). As will be seen in the next section, this approach is an l_∞-penalty approach.

By introducing slack variables $u, v \in \mathbb{R}^m, w \in \mathbb{R}^p$ and an l_1-penalization of the violation of the constraint, the problem (2.11) can also be formulated as

$$\begin{aligned}
\min_{x\in\mathbb{R}^n,u,v\in\mathbb{R}^m,w\in\mathbb{R}^p} \quad & a^\top x + \frac{1}{2}x^\top Qx + Me_m^\top(u+v) + Me_p^\top w \\
\text{s.t.} \quad & 0 = g + Ax + u - v \;, \\
& 0 \geq h + Dx - w \;, \\
& 0 \leq u, v, w \;,
\end{aligned}$$

where e_m is the vector $(1,\dots,1)^\top$ of length m and analog for e_p. Again, the problem for $u, v, w = 0$ is consistent with the original problem (2.7) and, as in the previous approach, the choice of M is nontrivial.

All these approaches have in common that they are formulated as optimization problems, which in turn fall into the class of problems that can be solved by the solver itself. However, in contrast to the original problem, a feasible starting point can always be easily specified.

2.3 Penalty-Approaches

In the previous section, the Active Set Method was described, a method which explicitly handles constraints, but was considered here only for quadratic problems with linear constraints. In contrast to this, there are the so-called penalty methods for more general optimization problems. Characteristic of these approaches is that the constraints are linked to the objective function and a violation of feasibility is penalized. This has the consequence that a constrained optimization problem becomes an unconstrained one and thus solvers can be used which are especially designed for unconstrained problems. Therefore, they provide an alternative way to solve constrained optimization problems without explicitly dealing with constraints.

Since the basic idea of penalty methods is a standard approach to constrained optimization, standard textbooks are again given as references. Thus, the following section is essentially based on the books [37, 61, 87].

2.3.1 General Penalty Method

For this section, consider the general optimization problem

$$\min_{x \in \mathbb{R}^n} \quad f(x) \quad \text{s.t.} \quad x \in \mathcal{F} \;, \tag{2.12}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is the continuous target function and, at first, the feasible set $\mathcal{F} \subseteq \mathbb{R}^n$ is closed. In this setting, an indication function can then be defined as

$$r : \mathbb{R}^n \to \mathbb{R}_+, \text{ where } \begin{array}{ll} r(x) > 0 & \text{for } x \notin \mathcal{F} \;, \\ r(x) = 0 & \text{for } x \in \mathcal{F} \;. \end{array}$$

Here, the function r indicates the violation of feasibility by a positive function value. If the feasible set is given in the form

$$\mathcal{F} := \{x \in \mathbb{R}^n \mid G(x) = 0, H(x) \leq 0\} \;, \tag{2.13}$$

where $G : \mathbb{R}^n \to \mathbb{R}^m$ and $H : \mathbb{R}^n \to \mathbb{R}^p$ are continuous functions, an indication function can be defined as

$$r(x) := \sum_{i=1}^{m} |G_i(x)|^\kappa + \sum_{i=1}^{p} \left(H_i^+(x)\right)^\kappa \;,$$

Algorithm 2 General Penalty Method

Require: Random $\rho > 0$

loop
  Determine (or approximate) a local minimum x^+ of $\Phi(x, \rho)$
  if $x^+ \in \mathcal{F}$ **then**
    return x^+
  else
    Choose $\rho^+ \geq 2\rho$
  Set $x = x^+$ and $\rho = \rho^+$

with $\kappa > 0$ and $H_i^+(x) := \max(0, H_i(x))$. If now a *penalty parameter* $\rho > 0$ is added, the resulting and so called *penalty function* is composed as

$$\Phi(x, \rho) := f(x) + \rho \cdot r(x) \tag{2.14}$$

and is a weighted sum of the objective function and the indication function. There is now the hope that solving the unconstrained problem

$$\min_{x \in \mathbb{R}^n} \Phi(x, \rho) \tag{2.15}$$

for a ρ large enough will give a good approximation of the solution for the constrained problem (2.12). Therefore, in a *General Penalty Method (GPM)*, a sequence of problems (2.15) is solved, where the penalty parameter is successively increased until the respective iterate represents a feasible point of (2.12). But not only this implicit assumption motivates this approach. In general it can be shown that a minimun of the original problem can be found. For this purpose, first the GPM is presented in Algorithm 2 and therefore there is the following result:

Theorem 2.13. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous function, x^* a strictly local minimum of* (2.12) *with $\mathcal{F}$ nonempty, convex and closed and $r : \mathbb{R}^n \to \mathbb{R}$ an indication function. Then, there exists a $\rho_0 > 0$ such that for all $\rho > \rho_0$ the function $\Phi(x, \rho) := f(x) + \rho r(x)$ assume a local minimum $x(\rho)$ for which*

$$\lim_{\rho \to \infty} x(\rho) = x^*$$

holds. Consequently, the GPM converges to a local minimum of f provided ρ goes to infinity.

Proof. See [61, Satz 11.1.5]. □

Implicitly, it is assumed that solving the problem (2.15) is easier for small ρ. In fact, there exist simple examples showing that penalty methods with a too large choice of the penalty parameter ρ do not perform well. In this context, using the eigenvalues of the Hessian of the penalty function, one can show that these matrices are ill-conditioned, making the subproblems occurring in line 2 of Algorithm 2 difficult to solve numerically, e.g., [37, Chapter 5.2] or [87, Chapter 17.1].

To conclude the general case, the following remarks are relevant: First, there are counterexamples that even if (2.15) for $\rho = \rho_k \to \infty$ has a sequence of local minima $x(\rho_k)$ converging to a point $\tilde{x} \in \mathcal{F}$, then $\tilde{x}$ is not necessarily a local minimizer of (2.12) with $\mathcal{F}$ nonempty, convex and closed, because in Theorem 2.13 a strict local minimizer is assumed. Second, differentiability is a critical aspect. Thus, a smooth optimization problem can become a nonsmooth problem depending on the choice of κ and therefore, it would be useful to keep the differentiability. Third, the GPM theoretically generates an infinite sequence. Hence, it would be ideal if there is a value $\rho^* > 0$ such that a local minimizer x^* of (2.12) with $\mathcal{F}$ nonempty, convex and closed is also local minimal for any unbounded formulation (2.15) with $\rho \geq \rho^*$.These are called exact penalty functions, for which the reader is referred to the literature at this point [37, 61, 87].

In the following, a concrete penalty method will be discussed, i.e., for a concrete choice of κ. Since in the context of this work it is in particular about piecewise linear problems, here the focus is on the l_1-penalty method.

2.3.2 l_1-Penalty Method

To begin with, the l_1*-penalty function* for the optimization problem (2.12) with the feasible set $\mathcal{F}$ given by (2.13) is defined as

$$\Phi(x, \rho) := f(x) + \rho \sum_{i=1}^{m} |G_i(x)| + \rho \sum_{i=1}^{p} H_i^+(x) \ ,$$

again with $H_i^+(x) := \max(0, H_i(x))$. Thus, for the choice of $\kappa = 1$, this is consistent with the formulation from (2.14). The name l_1-penalty function is also naturally motivated by the l_1-norm as the sum of all absolute values of the vector components.

Therefore, some directly apparent advantages and disadvantages are given: A critical aspect is certainly, as also noted before, that by this formulation the objective function becomes nonsmooth. So for an original smooth problem, smooth solvers are generally no longer usable. If the function was already nonsmooth before, this consideration is of course less critical. However, a clear advantage is that the l_1-penalty function is an exact penalty function. This means that for a suitable choice of the penalty parameter only one optimization, i.e., only one loop pass in the Algorithm 2 is necessary. Other penalty functions, such as the quadratic penalty function, which will be briefly described in the next subsection, are not exact. This has the consequence that the penalty parameter must be increased in the worst case even infinitely often, until a solution for the original problem is found. This extremely useful property of the l_1-penalty function is stated in the following theorem.

Theorem 2.14. *Suppose that x^* is a strict local solution of the nonlinear problem (2.1) at which the conditions of Theorem 2.5 are satisfied with Lagrange multipliers $\delta^* \in \mathbb{R}^m$ and $\nu^* \in \mathbb{R}^p$. Then x^* is a local minimizer of the l_1-penalty function for the optimization problem (2.1) for all $\rho > \rho^*$, where*

$$\rho^* = ||(\delta^*, \nu^*)||_\infty = \max\{|\delta_i^*|, \nu_j \mid i = 1, \dots, m\ , j = 1, \dots p\}.$$

Proof. See [87, Theorem 17.3]. □

Note that in contrast to the optimization problem (2.12) with the feasibly set given by (2.13), the Theorem 2.14 assumes differentiability of the functions in addition to continuity.

In [3] the exact l_1-penalty method for constrained nonsmooth optimization problems was also considered, but with the additional assumptions that the functions are invex. Simply speaking, invex functions are understood to be a generalization of convex functions, but not so general as to include the class of piecewise linear functions.

2.3.3 Quadratic Penalty Approach

For the optimization problem (2.12) with the feasible set $\mathcal{F}$ defined as in (2.13) the *quadratic penalty function* is defined as

$$\Phi(x, \rho) := f(x) + \frac{\rho}{2} \sum_{i=1}^{m} G_i^2(x) + \frac{\rho}{2} \sum_{i=1}^{p} \left(H_i^+(x)\right)^2 \ ,$$

again with $H_i^+(x) := \max(0, H_i(x))$. As usual, the penalty parameter is scaled with $\frac{1}{2}$ for differentiation reasons.

However, since the focus is to be on piecewise linear and not piecewise smooth functions, these approaches will not be considered further in the remainder of this thesis. However, they are meant to clarify that there are further penalty functions beyond the l_1-penalty function. For more details on these and other approaches, the reader is referred to [37, 61, 87].

3

The Abs-Linear Optimization Problem: Theory and Solution Method

The focus of this chapter is the derivation of an optimization algorithm for unconstrained piecewise linear optimization problems, which has already been published in [45]. Therefore, Section 3.1 makes the transition from the contents of the previous chapter to nonsmooth optimization. For this purpose, abs-smooth functions are first introduced, and optimality conditions for abs-linear functions are shown in Section 3.2. Abs-linear functions, as a subclass of abs-smooth functions, are compositions of linear functions and absolute value evaluations, in particular continuous piecewise linear functions. Therefore, the nondifferentiability is generated only by the absolute value. The main aspect of this chapter, the Active Signature Method (ASM) as an optimization algorithm for an abs-linear function, is part of Section 3.3. In this context, the ASM represents the essential basis for the algorithm for solving constrained piecewise optimization problems (see Chapter 4) developed in the context of this dissertation. Optimality conditions, the derivation of the algorithm itself and convergence statements are given. Afterwards, the chapter concludes with a short overview of other optimization methods that have been and are still being developed for abs-smooth functions in Section 3.4.

It should be noted that the notation has been adjusted in comparison to [45] because the Abs-Linear Form is used in a slightly different representation. The reason for this is that in Chapter 4, in addition to the objective function, piecewise linear constraints are also considered, which should have a representation as general as possible. All required definitions in this thesis are adapted accordingly, hence the notation is self-consistent in this work.

3.1 Abs-Linear Form and Abs-Linearization

After introducing selected aspects of smooth optimization in the previous chapter and using the l_1-penalty approach as a first nonsmooth method, the focus will now be on relevant aspects of nonsmooth optimization.

As mentioned in the introduction (see Chapter 1), there are classes of nonsmooth functions for which derivative information can be generated using AD. To describe one of these classes, suppose that the functions are given by an evaluation program. Here, the nonsmoothness can be represented by the absolute value function, as well as the min- and max-functions, which themselves can be represented by the abs-function via

$$\min(u,v) = \frac{1}{2}\left(u + v - |u-v|\right) \tag{3.1}$$

$$\text{and} \quad \max(u,v) = \frac{1}{2}\left(u + v + |u-v|\right) \; . \tag{3.2}$$

For details on this evaluation process, there are a number of sources, see, e.g., [29, 30, 41, 43] and it leads to the following definition of the class of these functions:

Definition 3.1 ($\mathcal{C}^d_{\text{abs}}(\mathbb{R}^n)$-functions)**.** *For any $d \in \mathbb{N}$, the set of functions $\varphi : \mathbb{R}^n \to \mathbb{R},\ y = \varphi(x)$, defined by an* Abs-Smooth Form

$$\begin{aligned} y &= f(x,z) \; , \\ z &= F(x,z,|z|) \; , \end{aligned} \tag{3.3}$$

with $f \in \mathcal{C}^d(\mathbb{R}^{n+s},\mathbb{R})$ and $F \in \mathcal{C}^d(\mathbb{R}^{n+s+s},\mathbb{R}^s)$, such that z_i is determined only by the values of z_j, for $1 \le j < i$, is denoted by $\mathcal{C}^d_{\text{abs}}(\mathbb{R}^n)$. The components z_i, $1 \le i \le s$, of z are called switching variables*, the whole vector z* switching vector*, and Eq. (3.3)* switching equation*.*

Here and throughout, $|z|$ denotes the component-wise modulus of a vector z. As in [48], assume that the switching variable which does not impose nonsmoothness is located in the last component of z and that only the first $\tilde{s} \in \{s-1, s\}$ components $z_1, \dots, z_{\tilde{s}}$ are arguments of the absolute value.

For piecewise linear functions as a subclass of $\mathcal{C}^d_{\text{abs}}(\mathbb{R}^n)$-functions the Abs-Smooth Form can be given explicitly in the so-called Abs-Linear Form. All further contributions in this thesis rely on this form.

Definition 3.2 (Abs-Linear Form, switching vector)**.** *A continuous piecewise linear func-*

tion $\varphi : \mathbb{R}^n \to \mathbb{R}$ is in Abs-Linear Form *if $y \equiv \varphi(x)$ is given by*

$$y = d + a^\top x + b^\top z \, , \tag{3.4}$$
$$z = c + Zx + Mz + L|z| \, , \tag{3.5}$$

with $x \in \mathbb{R}^n$ the argument vector, $z \in \mathbb{R}^s$ the vector of switching variables*, called* switching vector*, and constants $d \in \mathbb{R}$, $a \in \mathbb{R}^n$, $b, c \in \mathbb{R}^s$, $Z \in \mathbb{R}^{s\times n}$, $L, M \in \mathbb{R}^{s\times s}$, where the last two matrices are strictly lower triangular. Eq. (3.5) is called* switching system*.*

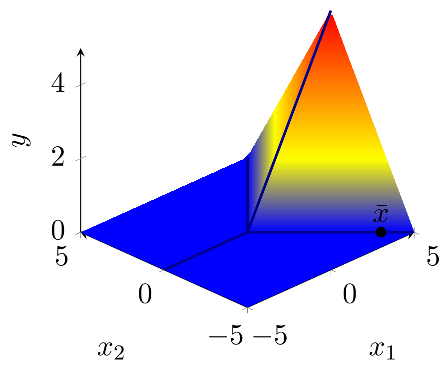

Figure 3.1: Plot of the Hill-function

In context of optimization w.l.o.g. it is assumed that $d = 0$. Using the representations of the min and max as given in Eqs. (3.1) and (3.2) via the abs-function and [96, Proposition 2.2.2] an important fact is that every piecewise linear function can be represented in an Abs-Linear Form, see also [48, Lemma 2].

To illustrate the different concepts introduced in this thesis, starting with the Abs-Linear Form, the following simple example will be used throughout.

Example 3.3 (Abs-Linear Form of Hill-function)**.** *Consider the function*

$$\varphi(x)\colon \mathbb{R}^2 \to \mathbb{R}, x \mapsto \max\{0, x_1 - |x_2|\} \, . \tag{3.6}$$

A plot of this function is given in Figure 3.1, and because of its shape will be referred to as Hill-function *in the following. Note that this is a nonsmooth and nonconvex function, which in particular is already piecewise linear. The points where the function is not differentiable in the classical sense are marked by dark blue lines. However, the function is differentiable for $x_1 < 0$ with $x_2 = 0$, but this line is relevant for the polyhedral decomposition that will be discussed later. This graphical illustration also motivates the term* kink *for the nondifferentiable segments in the following. According to Rademacher's theorem [27, Theorem 3.1.6], these are always sets of measure zero.*

Using Eq. (3.2) the Hill-function can be rewritten as

$$\varphi(x) = \frac{1}{2}\left(x_1 - |x_2| + \big|x_1 - |x_2|\big|\right) \, .$$

From this, it can also be seen that nestings of absolute values are also included in the *function class considered here. In [41, 44], this nesting property and the number of nestings is called* switching depth*. The Abs-Smooth Form for the Hill-function is given by*

$$y = f(x,z) := \frac{1}{2}x_1 + z_3 \ , \qquad z = F(x,z,|z|) := \begin{pmatrix} x_2 \\ x_1 - |z_1| \\ -\frac{1}{2}|z_1| + \frac{1}{2}|z_2| \end{pmatrix} \tag{3.7}$$

and therefore, in particular the full Abs-Linear Form by

$$y = 0 + \begin{bmatrix} \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \ ,$$

$$z = \begin{bmatrix} x_2 \\ x_1 - |z_1| \\ -\frac{1}{2}|z_1| + \frac{1}{2}|z_2| \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \begin{bmatrix} |z_1| \\ |z_2| \\ |z_3| \end{bmatrix} \ .$$

In his outstanding article [41], Andreas Griewank has studied piecewise linearizations of abs-smooth functions. The basic idea, well known from smooth analysis, is that one linearizes functions or applies a Taylor expansion to get an approximation. For the purpose of abs-smooth functions as in Definition 3.1 with $d \geq 1$ consider the following vectors and matrices:

$$\begin{aligned} a &= \frac{\partial}{\partial x} f(x,z) \in \mathbb{R}^n \ , & b &= \frac{\partial}{\partial z} f(x,z) \in \mathbb{R}^s \ , \\ Z &= \frac{\partial}{\partial x} F(x,z,w) \in \mathbb{R}^{s \times n} \ , & & \\ M &= \frac{\partial}{\partial z} F(x,z,w) \in \mathbb{R}^{s \times s} & & \text{strictly lower triangular} \ , \\ L &= \frac{\partial}{\partial w} F(x,z,w) \in \mathbb{R}^{s \times s} & & \text{strictly lower triangular} \ . \end{aligned}$$

Using these vectors and matrices it is possible to define an Abs-Linear Form of abs-smooth functions, which represents a piecewise linear approximation of the abs-smooth function. Given an abs-smooth function $\varphi : \mathbb{R}^n \to \mathbb{R}$ and a base point $\mathring{x} \in \mathbb{R}^n$ from which the approximation is constructed, the resulting and so-called *Abs-Linearization* is denoted by $\Delta\varphi(\mathring{x};\cdot) : \mathbb{R}^n \to \mathbb{R}$. For this purpose, since the argument of the function $\Delta\varphi(\mathring{x};\cdot)$ is rather to be understood as a direction, the argument vector x in the representation given by Eqs. (3.4) and (3.5) is replaced by a direction vector $\Delta x \in \mathbb{R}^n$.

Numerically, an Abs-Linear Form of an abs-smooth function can be generated using appropriate variants of AD (see Section 1.1). A special property of the approximation of abs-smooth functions by their abs-linearization is that it is a second-order approximation.

Theorem 3.4. *Suppose, f is abs-smooth on $D \subset K \subset \mathbb{R}^n$, D open, K closed and convex. Then there exists some $\gamma > 0$ such that for all $x, \mathring{x} \in D$*

$$||\varphi(x) - \varphi(\mathring{x}) - \Delta\varphi(\mathring{x}, x - \mathring{x})|| \leq \gamma ||x - \mathring{x}||^2 .$$

Proof. See [41, Proposition 1]. □

An optimization algorithm that uses exactly this abs-linearization is called Successive Abs-Linear MINimization (SALMIN), [28, 29]. It solves unconstrained optimization problems with abs-smooth objective functions using the abs-linearization as a local model over which the optimization is performed. A solver for the resulting inner abs-linear subproblem is the Active Signature Method (ASM) by Andreas Griewank and Andrea Walther [45]. It is presented in the following section. The algorithm was developed to solve unconstrained piecewise linear optimization problems. Here, the nonsmoothness of the objective function is explicitly exploited by the polyhedral structure that arises as described next. Therefore, the remaining chapter is essentially based on [45], but the notation is adapted to fit this thesis, as it is already stated in [48] or [70]. Furthermore, significant details are added in the derivation and in the proofs.

To begin with, in the further course it will be useful to distinguish the switching variables into active and inactive ones.

Definition 3.5 (Active switching variables)**.** *A switching variable z_i is called active at x if $z_i(x) = 0$. The active switching set $\alpha(x)$ collects all indices of* active switching variables *that explicitly depend on x, i.e.,*

$$\alpha(x) \equiv \{i \in \{1, \ldots, \tilde{s}\} \mid z_i(x) = 0\} .$$

The projection onto the active components of *$z(x)$ is defined as $P_\alpha \equiv (e_i^\top)_{i \in \alpha} \in \mathbb{R}^{|\alpha| \times s}$ with e_i denoting the ith unit vector of appropriate size.*

Example 3.6 (Active switching variables for Hill-function)**.** *Consider the Hill-function with its Abs-Smooth Form (3.7) at the point $\bar{x} = (4; -4)^\top$. Then the set of active switching variables in $\bar{x}$ is a singleton set $\alpha(\bar{x}) = \{2\}$. This can be seen in Figure 3.1, because the point $\bar{x}$ lies exactly on one of the kinks (in the right corner of the plot). Thus, this one is active for $\bar{x}$.*

In [44] it is shown, that it is possible to decompose $\mathbb{R}^n$ into polyhedra using the signatures of the switching vector.

Definition 3.7 (Signature vector and signature matrix)**.** *Let a $\mathcal{C}^d_{\text{abs}}(\mathbb{R}^n)$-function be given in Abs-Smooth Form. For each $x \in \mathbb{R}^n$ the* signature vector *is defined by*

$$\sigma(x) \equiv (\text{sgn}(z_1(x)), \ldots, \text{sgn}(z_s(x))) \in \{-1, 0, 1\}^s .$$

The corresponding signature matrix *is given by $\Sigma(x) = \text{diag}(\sigma(x))$. A signature vector $\sigma(x)$ is called* definite, *if no component vanishes, i.e., $\sigma(x) \in \{-1, 1\}^s$. This situation is denoted by $0 \notin \sigma(x)$. Otherwise it is called* indefinite.

Since frequently fixed signature vectors are considered in this thesis, the dependence on x is explicitly stated if there is one. Based on these signature vectors, it is possible to decompose $\mathbb{R}^n$ into polyhedra as follows.

Definition 3.8 ((Extended) Signature domain)**.** *For a fixed $\sigma \in \{-1, 0, 1\}^s$, define*

$$\mathcal{P}_\sigma \equiv \{x \in \mathbb{R}^n \mid \text{sgn}(z(x)) = \sigma\} \subset \overline{\mathcal{P}}_\sigma \equiv \{x \in \mathbb{R}^n \mid \Sigma z(x) = |z(x)|\} .$$

The set $\mathcal{P}_\sigma$ is called signature domain *and the set $\overline{\mathcal{P}}_\sigma$* extended signature domain.

Note that here and throughout the symbol $\subset$ denotes a subset relation that also allows equality of sets.

For the Hill-function the signature domains are illustrated in the following example.

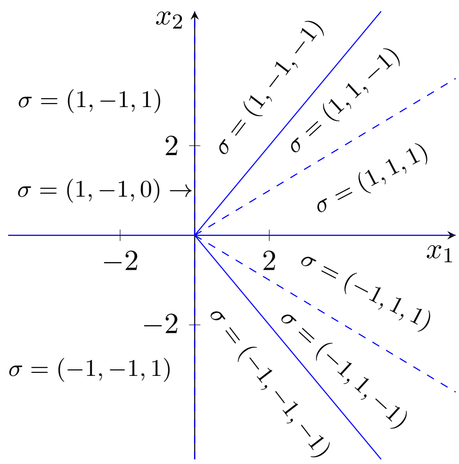

Figure 3.2: Signature domains of the Hill-function

Example 3.9 (Signature vectors and domains for Hill-function)**.** *Again, consider the Hill-function with its Abs-Smooth Form (3.7). Figure 3.2 shows the decomposition of the $\mathbb{R}^2$ in the polyhedra, i.e., the signature domains $\mathcal{P}_\sigma$, labeled by their corresponding definite signature vectors σ. The blue lines correspond to polyhedra with indefinite signature vectors, were one of them is marked by $\sigma = (1, -1, 0)$ on the left half of Figure 3.2 to give also an example for the indefinite ones. Here, those kinks that do not cause nonsmoothness in the sense that they are arguments of the absolute value are represented by the dashed lines.*

The properties of these domains have been discussed in detail in papers such as [44, 45, 49]. One of these characteristics is that the domains $\mathcal{P}_\sigma$ are given as inverse images of the corresponding σ and represent a disjoint decomposition of $\mathbb{R}^n$ into relatively open polyhedra. This polyhedral structure also motivates the identifier $\mathcal{P}$ for these sets. A second one is, that the boundaries of the polyhedra $\mathcal{P}_\sigma$ are exactly the set where the corresponding abs-smooth function potentially has its nonsmooth points. In the remainder, this characteristic will be further clarified by graphical illustrations of examples. Last but not least, it is also possible to define a partial ordering as follows

$$\sigma \prec \tilde{\sigma} \quad \Longleftrightarrow \quad \sigma_l^2 \leq \tilde{\sigma}_l \sigma_l \quad \text{for } 1 \leq l \leq s \quad \Longleftrightarrow \quad \overline{\mathcal{P}}_\sigma \subseteq \overline{\mathcal{P}}_{\tilde{\sigma}} \ .$$

As stated in [45, 70], the aim of the ASM is to determine a local minimizer of a piecewise linear objective function $\varphi : \mathbb{R}^n \to \mathbb{R}$ given in Abs-Linear Form (cf. Definition 3.2). Using this form of the objective function φ, one obtains the following equivalent *abs-linear optimization problem*

$$\begin{aligned} \min_{x \in \mathbb{R}^n, z \in \mathbb{R}^s} \quad & a^\top x + b^\top z \\ \text{s.t.} \quad & z = c + Zx + Mz + L|z| \ . \end{aligned} \tag{ALOP}$$

At this point, it should be stated that whenever a minimum is mentioned in the remainder of this work, a local minimum is meant, unless it is explicitly specified otherwise.

3.2 Optimality Conditions

First, necessary and sufficient optimality conditions for (ALOP) are to be derived. These are not given in the original source [45] in detail, however, were already further developed in the appendix of [48] as well as in [70], with additional constraints. In addition, in the context of her PhD thesis Lisa Hegerhorst-Schultchen considers also optimality conditions for abs-smooth optimization problems under the supervision of Marc Steinbach [54, 55, 56].

There, necessary and sufficient first and second order optimality conditions were established for unconstrained and constrained abs-smooth nonlinear optimization problems. For this purpose, optimality conditions were first considered for individual branch problems, and these results were then combined to show the optimality conditions for the general problem. Furthermore, it has been shown that the class of abs-smooth nonlinear optimization problems is equivalent to the class of mathematical programs with equilibrium constraints

(MPEC), and additionally, an equivalence of the corresponding regularity conditions can be shown. These are, on the one hand, the MPEC-LICQ and, on the other hand, the regularity conditions for abs-smooth problems, the so-called LIKQ, which will also be introduced in this thesis in Definition 3.12 for the unconstrained case and in Definition 4.6 for the constrained case.

In contrast to the results in the thesis of Lisa Hegerhorst-Schultchen, this thesis will show necessary and sufficient optimality conditions for the optimization problem (ALOP) explicitly. On the one hand, their representation also directly implies that they can be verified in polynomial time. On the other hand, the connection to the optimality criteria and thus to the termination criteria of the algorithms given later in this thesis can easily be seen.

Therefore, consider for each fixed $\sigma \in \{-1, 0, 1\}^s$ the optimization problem

$$\min_{x\in\mathbb{R}^n, z\in\mathbb{R}^s} \quad a^\top x + b^\top z \tag{3.8a}$$

$$\text{s.t.} \quad z = c + Zx + Mz + L\Sigma z \ , \tag{3.8b}$$

$$0 = (I_s - |\Sigma|)z \ , \tag{3.8c}$$

$$0 \leq \Sigma z \ ,$$

where I_s denotes the identity matrix of dimension s. As described in [45], the last two constraints restrict the feasible set to corresponding polyhedra $\mathcal{P}_\sigma$. Optimal points of these are defined as follows:

Definition 3.10 (Signature optimal point)**.** *Let an optimization problem of the form* (ALOP) *be given. Consider a fixed signature vector $\sigma \in \{-1, 0, 1\}^s$. A minimizer $x_\sigma \in \mathcal{P}_\sigma$ of the optimization problem* (3.8) *is called a* signature optimal point *of the original, unconstrained optimization problem* (ALOP)*.*

Note, that for many or even most σ the polyhedra $\mathcal{P}_\sigma$ do not contain minimizers, in which case the solutions of (3.8) lie on their relative boundary.

Example 3.11 ((Non) signature optimal points for Hill-function)**.** *For the Hill-function given by Eq.* (3.6) *consider the three points $\bar{x} = (4, -4)^\top$, $\tilde{x} = (4, 2.5)^\top$ and $\hat{x} = (0, 0)^\top$ as marked in Figure 3.3. The point $\bar{x}$ has the signature vector $\sigma(\bar{x}) = (-1, 0, -1)^\top \equiv \bar{\sigma}$. When optimizing over $\overline{\mathcal{P}}_{\bar{\sigma}}$, $\bar{x}$ itself is a minimizer and therefore a signature optimal point for the Hill-function. Note, due to the fact, that the function is piecewise linear, in this case the signature optimal points are not unique. The point $\tilde{x}$ is not a signature optimal point and* *has the signature vector* $\sigma(\tilde{x}) = (1, -1, -1)^\top \equiv \tilde{\sigma}$. *When optimizing over* $\overline{\mathcal{P}}_{\tilde{\sigma}}$ *one could obtain the minimizer* $\hat{x}$*, which is still not a unique minimizer. This point has the signature vector* $\sigma(\hat{x}) = (0, 0, 0) = \hat{\sigma} \neq \tilde{\sigma}$. *Hence, one has* $\hat{x} \notin \mathcal{P}_{\tilde{\sigma}}$*, but* $\hat{x}$ *is signature optimal on the polyhedron* $\mathcal{P}_{\hat{\sigma}}$ *that contains only* $\hat{x}$. *The polyhedron* $\mathcal{P}_{\tilde{\sigma}}$ *contains no signature optimal points.*

To simplify the notation, define

$$\tilde{Z} := (I_s - M - L\Sigma)^{-1} Z \qquad \text{and} \qquad \tilde{c} := (I_s - M - L\Sigma)^{-1} c \, . \tag{3.9}$$

Combining the two Eqs. (3.8b) and (3.8c) and substituting the linear constraint into the target function (3.8a) yields the optimization problem

$$\begin{aligned} \min_{x \in \mathbb{R}^n, z \in \mathbb{R}^s} \quad & a^\top x + b^\top |\Sigma| z \\ \text{s.t.} \quad & 0 = |\Sigma| z - \tilde{c} - \tilde{Z} x \, , \\ & 0 \leq \Sigma z \, . \end{aligned} \tag{3.10a}$$

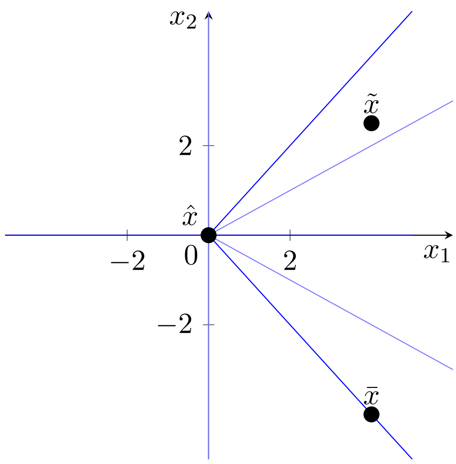

Figure 3.3: (Non) signature optimal points for Hill-function

Since this is a linear constrained optimization problem, by [87, Lemma 12.7] the set of feasible directions at x coincides with the tangent cone at x. This ensures the existence of not necessarily unique Lagrange multipliers without further regularity conditions. As written in [70], the goal is to derive optimality conditions that can be verified in polynomial time. If this can be shown, it is a significant advantage over typical theoretical optimality conditions of nonsmooth optimization. Many concepts of these optimality conditions are based on some kind of subgradients. At nondifferentiable points these subgradients are often given by multi-element sets, the subdifferentials. In the worst case, they can even contain infinitely many elements. A simple example is given by the absolute value function. One of the best known subgradient concepts is given by the Clarke derivative, [18]. An optimality condition here requires 0 to be an element of the subdifferential. In the case of the absolute value function, however, the subdifferential at the minimum $x = 0$ is given by the interval $[-1, 1]$. However, such quantities are difficult to determine numerically, and thus the optimality condition is difficult to test. Moreover, this condition does not allow to distinguish between minima and maxima, since the subdifferential of the negative absolute value function at $x = 0$ is also given by $[-1, 1]$. However, this thesis will completely lack such generalized derivation concepts, so at this point it is referred to the literature for further details, e.g., [18, 58, 92, 93].

Returning to the existence of Lagrange multipliers, a regularity assumption ensuring their uniqueness was introduced in [44] and is called linear independence kink qualification (LIKQ). It is a generalization of LICQ (cf. Definition 2.4) from smooth to nonsmooth optimization. For the piecewise linear case it is defined as:

Definition 3.12 (LIKQ (unconstrained case))**.** *Let an optimization problem of the form* (ALOP) *and a signature vector $\sigma \in \{-1, 0, 1\}^s$ be given. Then, the* linear independence kink qualification (LIKQ) *holds at a point $x \in \mathcal{P}_\sigma$ if the active Jacobian*

$$\mathcal{J}_\sigma = P_\alpha \tilde{Z}$$

with $\alpha \equiv \alpha(x)$ has full row rank $|\alpha|$. This requires in particular that $||\sigma||_1 \geq s - n$ so that $\mathcal{J}_\sigma$ is a short, thick matrix.

For a small example of this definition, consider once more the Hill-function.

Example 3.13 (LIKQ for Hill-function)**.** *Using Eq.* (3.9) *one obtains for the Hill-function given in its Abs-Linear Form as stated in Example 3.3 at the point $\bar{x} = (4, -4)^\top$ the active Jacobian*

$$\mathcal{J}_{(-1,0,-1)^\top} = P_{\alpha(\bar{x})}\tilde{Z} = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 0 & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \end{bmatrix} ,$$

which has obviously full row rank $|\alpha(\bar{x})| = 1$. Thus, LIKQ is fulfilled in $\bar{x}$. In particular, since the first $\tilde{\alpha} = 2$ rows of $\tilde{Z}$ are linearly independent, it even holds that $P_{\alpha(x)}\tilde{Z}$ has full row rank for all $x \in \mathbb{R}^n$. Thus, for this example, LIKQ is satisfied at every point.

To increase the possibility for linear independence kink qualification (LIKQ) to ensure uniqueness of the Lagrange multipliers (see Theorem 3.14), one can assume that $z_s(x^*) \neq 0$ holds for x^* as a solution of (ALOP). This can be achieved by a constant shift, since the Abs-Linear Form is not unique.

Note that due to the slightly different definition of the Active Switching Set $\alpha(x)$ in

Definition 3.5 compared to [44], the question naturally arises as whether the definition of LIKQ in this work is the same as the original definition. By the fact that according to the Definition 3.1 the arguments of the absolute value are still represented by the first $\tilde{s}$ switching variables, and thus also the arising nonsmoothness, exactly these components are also considered in the Definition 3.12.

Now consider a signature optimal point x_σ. In the optimization problem (3.10) all functions are continuous. Due to this fact and the relation $\Sigma z = |z|$, the components z_i, with $i \notin \alpha$, of the switching vector z determined by Eq. (3.10a) will not drop to zero in an open neighborhood $U(x_\sigma)$ of x_σ. Using the identity $\Sigma z = \Sigma|\Sigma|z$, in the neighborhood $U(x_\sigma)$ the optimization problem (3.10) is equivalent to

$$\begin{aligned} \min_{x \in U(x_\sigma)} \quad & a^\top x + b^\top |\Sigma|(\tilde{c} + \tilde{Z}x) \\ \text{s.t.} \quad & 0 = P_\alpha(\tilde{c} + \tilde{Z}x) \ . \end{aligned} \tag{3.11a}$$

This concludes the preliminary considerations to show the necessary and sufficient optimality conditions similar to [48, 70], which verifies the optimality of a signature optimal point in polynomial time.

Theorem 3.14 (Necessary and sufficient optimality conditions for (ALOP))**.** *Let an optimization problem of the form* (ALOP) *and a signature vector $\sigma \in \{-1, 0, 1\}^s$ be given. Assume that x_σ is signature optimal for* (ALOP) *and that LIKQ holds at x_σ. Then x_σ is a local minimizer of* (ALOP) *if and only if there exists a unique Lagrange multiplier $\lambda \in \mathbb{R}^s$, such that*

$$0 = a^\top + b^\top |\Sigma| \tilde{Z} - \lambda^\top P_\alpha^\top P_\alpha \tilde{Z} \ , \tag{3.12}$$

$$0 = b^\top |\Sigma| + \lambda^\top |\Sigma| \tag{3.13}$$

$$\text{and} \qquad |P_\alpha(b + \lambda)| \leq -P_\alpha \tilde{L}^\top \lambda \tag{3.14}$$

with $\tilde{L}$ given by $\tilde{L} = (I_s - M - L\Sigma)^{-1} L$.

Proof. Since x_σ is signature optimal for (ALOP) per definition it is a minimizer of (3.8) for the given signature vector σ. As seen by the conversions above, x_σ is thus also a minimizer of (3.11). Therefore, due to the assumtion of LIKQ once obtains from the KKT-theory (cf. Theorem 2.5) that there exist a unique Lagrange multiplier $\check{\lambda} \in \mathbb{R}^{|\alpha|}$ associated with the equality constraint (3.11a) such that

$$0 = a^\top + b^\top |\Sigma| \tilde{Z} + \check{\lambda}^\top P_\alpha \tilde{Z} \; .$$

Hence, each vector $\lambda \in \mathbb{R}^s$ such that $\check{\lambda} = -P_\alpha \lambda$ fulfills Eq. (3.12).

For local minimality of (ALOP) it is necessary and sufficient that $(x_\sigma, z(x_\sigma))$ is a minimizer of (3.8) on all extended signature domains $\mathcal{P}_{\tilde{\sigma}}$ with definite $\tilde{\sigma} \succ \sigma$. Any such $\tilde{\sigma}$ can be written as $\tilde{\sigma} = \sigma + \gamma$ with $\gamma \in \{-1, 0, 1\}^s$ structurally orthogonal to σ such that for $\Gamma \equiv \mathrm{diag}(\gamma)$ the matrix equations

$$\tilde{\Sigma} = \Sigma + \Gamma \quad \text{and} \quad \Sigma\Gamma = 0 = |\Sigma|\Gamma \tag{3.15}$$

hold true. Using this decomposition in Eq. (3.8b) for the expression $z(x) = z_{\tilde{\sigma}}(x)$ for $x \in \overline{\mathcal{P}}_{\tilde{\sigma}}$ yields

$$z_{\tilde{\sigma}}(x) = z_{\sigma+\gamma}(x) = (I_s - M - L\Sigma - L\Gamma)^{-1} (c + Zx) = \left(I_s - \tilde{L}\Gamma\right)^{-1} \left(\tilde{c} + \tilde{Z}x\right) \; . \tag{3.16}$$

Now x_σ must be the minimizer on $\overline{\mathcal{P}}_{\tilde{\sigma}}$, i.e., solves the smooth optimization problem

$$\min_{x \in \mathbb{R}^n, z \in \mathbb{R}^s} \quad a^\top x + b^\top (|\Sigma| + |\Gamma|)\, z$$

$$\text{s.t.} \quad 0 = \left(I_s - \tilde{L}\Gamma\right) z - \tilde{c} - \tilde{Z}x \; , \tag{3.17a}$$

$$0 \le P_\alpha \Gamma z \; . \tag{3.17b}$$

Notice that the inequalities are imposed only on the sign constraints that are active at x_σ since the strict inequalities are maintained in a neighborhood of x_σ due to the continuity of $z(x)$. Applying once more the KKT-theory to (3.17), there must exist Lagrange multipliers $\lambda \in \mathbb{R}^s$ associated with reformulated switching system (3.17a) and $0 \le \mu \in \mathbb{R}^{|\alpha|}$ associated with the sign condition (3.17b) such that

$$0 = a^\top - \lambda^\top \tilde{Z} \quad \text{and} \tag{3.18}$$

$$0 = b^\top (|\Sigma| + |\Gamma|) + \lambda^\top \left(I_s - \tilde{L}\Gamma\right) - \mu^\top P_\alpha \Gamma \; . \tag{3.19}$$

Multiplying the last equation from the right by $|\Sigma|\tilde{Z}$ and using the identity $\Sigma = \Sigma|\Sigma|$ as well as the matrix equation (3.15) yields

$$0 = b^\top |\Sigma| \tilde{Z} + \lambda^\top |\Sigma| \tilde{Z} \; . \tag{3.20}$$

By exploiting the decomposition of the identity given by

$$I_s = |\Sigma| + P_\alpha^\top P_\alpha \ , \tag{3.21}$$

due to the definition of Σ and P_α and adding Eq. (3.20) to Eq. (3.18) one obtains

$$0 = a^\top + b^\top |\Sigma| \tilde{Z} - \lambda^\top P_\alpha^\top P_\alpha \tilde{Z} \ .$$

Hence, it follows that the Lagrange multiplier $\lambda \in \mathbb{R}^s$ fulfills Eq. (3.12) with $\check{\lambda} = -P_\alpha \lambda$. Due to the kink qualification LIKQ, one also has that the components $P_\alpha \lambda \in \mathbb{R}^{|\alpha|}$ are determined uniquely. The remaining components of $\lambda \in \mathbb{R}^s$ can be obtained by multiplying Eq. (3.19) this time only with $|\Sigma|$ from the right and using Eq. (3.15) yielding

$$0 = b^\top |\Sigma| + \lambda^\top |\Sigma|$$

and thus also Eq. (3.13).

To obtain the third condition (3.14) the following identities regarding the projection matrix P_α are needed:

$$P_\alpha^\top P_\alpha = \Gamma\Gamma = |\Gamma| \ , \quad P_\alpha P_\alpha^\top = I_{|\alpha|} \quad \text{and thus} \quad \Gamma\Gamma P_\alpha^\top = P_\alpha^\top \ . \tag{3.22}$$

Using this identities and again multiplying Eq. (3.19) from the right this time with $\Gamma P_\alpha^\top$ yields

$$0 = b^\top \Gamma P_\alpha^\top + \lambda^\top \left(\Gamma P_\alpha^\top - \tilde{L} P_\alpha^\top \right) - \mu^\top \ . \tag{3.23}$$

By reformulating this equation and using $\mu \geq 0$, it follows that

$$-\left(b^\top + \lambda^\top \right) \Gamma P_\alpha^\top = -\lambda^\top \tilde{L} P_\alpha^\top - \mu^\top \leq -\lambda^\top \tilde{L} P_\alpha^\top \ .$$

Now the key observation is that this condition is linear in Γ and is strongest for the choice $\gamma_i = -\operatorname{sgn}\left(b^\top + \lambda^\top \right)_i$ for $i \in \alpha$ yielding the inequality

$$|P_\alpha \left(b + \lambda \right)| \leq -P_\alpha \tilde{L}^\top \lambda$$

and thus Eq. (3.14) which completed the necessary optimality condtions.

Next, consider the sufficient optimality conditions. For this, again consider Eq. (3.19) and multiply it from the right-hand side by $\Gamma P_\alpha^\top$. By using Eq. (3.22), Eq. (3.15) and Eq. (3.14) one obtains

$$\mu^\top = b^\top \Gamma P_\alpha^\top + \lambda^\top \left(I_s - \tilde{L}\Gamma\right) \Gamma P_\alpha^\top = \left(b^\top + \lambda^\top\right) \Gamma P_\alpha^\top - \lambda^\top \tilde{L} P_\alpha^\top \geq 0 \tag{3.24}$$

and therefore the feasibility. To show Eq. (3.18) one can use Eq. (3.21), Eq. (3.13) multiplied from the right-hand side by $\tilde{Z}$ and Eq. (3.12) to obtain

$$\lambda^\top \tilde{Z} = \lambda^\top \left(|\Sigma| + P_\alpha^\top P_\alpha\right) \tilde{Z} = \lambda^\top |\Sigma| \tilde{Z} + \lambda^\top P_\alpha^\top P_\alpha \tilde{Z} = -b^\top |\Sigma| \tilde{Z} + a^\top + b^\top |\Sigma| \tilde{Z} = a^\top \ .$$

Now only Eq. (3.19) remains to be shown. Therefore, using Eq. (3.21), Eq. (3.19) holds true if and only if

$$0 = b^\top \left(|\Sigma| + |\Gamma|\right) + \lambda^\top \left(|\Sigma| + P_\alpha^\top P_\alpha - \tilde{L}\Gamma\right) - \mu^\top P_\alpha \Gamma$$

holds true. Using Eq. (3.13) the last equation is equivalent to

$$0 = b^\top |\Gamma| + \lambda^\top \left(P_\alpha^\top P_\alpha - \tilde{L}\Gamma\right) - \mu^\top P_\alpha \Gamma \ .$$

Again, multiplying the last equation from the right-hand side by $\Gamma P_\alpha^\top$ and using Eq. (3.22) yields

$$\mu^\top = \left(b^\top + \lambda^\top\right) \Gamma P_\alpha^\top - \lambda^\top \tilde{L} P_\alpha^\top \ .$$

Now, defining the Lagrange multiplier μ as above, the multiplier satisfies Eq. (3.19). To see this, insert μ into Eq. (3.19) and use Eq. (3.22), (3.13) and (3.21) to obtain

$$\begin{aligned} 0 &= b^\top \left(|\Sigma| + |\Gamma|\right) + \lambda^\top \left(I_s - \tilde{L}\Gamma\right) - \left(b^\top + \lambda^\top\right) |\Gamma| + \lambda^\top \tilde{L}\Gamma \\ &= -\lambda^\top |\Sigma| + \lambda^\top I_s - \lambda^\top |\Gamma| \\ &= -\lambda^\top |\Sigma| + \lambda^\top \left(|\Sigma| + P_\alpha^\top P_\alpha\right) - \lambda^\top |\Gamma| \ . \end{aligned}$$

Since $P_\alpha^\top P_\alpha = |\Gamma|$ holds according to Eq. (3.22), this is a true statement.

The uniqueness of the Lagrange multiplier follows directly from Eqs. (3.12) and (3.13). Due to the full rank by the assumptions of LIKQ, Eq. (3.12) can be used to uniquely determine the components λ_i for all $i \in \alpha(x_\sigma)$. By Eq. (3.13), for the remaining $i \in \alpha^C$, the complement of α, with $\Sigma_{ii} = \pm 1$ immediately follows $\lambda_i = -b_i$. □

As mentioned above, a particular achievement of this theorem is that it allows the optimality of a given Lagrange multiplier to be tested by simple matrix-vector multiplication in polynomial time. For this, the two so-called *tangential stationarity* conditions Eqs. (3.12), (3.13) and the *normal growth* condition Eq. (3.14) must hold. These terms have been used in [44, 45, 47, 48], among others.

3.3 Active Signature Method

After the theoretical preliminary considerations, the algorithm for solving problems of the form (ALOP), the Active Signature Method (ASM), shall now be derived. This section is therefore essentially based on [45], continued using the adapted notation of this thesis, and extended by recent considerations from [70].

The basic idea to solve the problem (ALOP) is to decompose the $\mathbb{R}^n$ into polyhedra as sketched in Subsection 3.1 using the signature vectors $\sigma(x)$ and to optimize a penalized version of the objective function on these domains, switching from one polyhedron to the next one in an appropriate way. To achieve this behavior, information about the structure of nonsmoothness is exploited for each $x \in \mathbb{R}^n$ and the corresponding $z = z(x)$. This strategy can be seen as a generalization of the usual active set quadratic optimization method to find a stationary point on each polyhedron.

Consider the optimization problem (3.10) where a quadratic penalty term with a positive definite matrix $Q = Q^\top \in \mathbb{R}^{n\times n}$ is added. This penalty term ensures that the objective function is bounded from below, and thus the existence of a minimum is guaranteed, even if some or all polyhedra from the polyhedral decomposition of $\mathbb{R}^n$ are unbounded. This leads to the following optimization problem

$$\begin{aligned} \min_{x\in\mathbb{R}^n, z\in\mathbb{R}^s} \quad & a^\top x + b^\top |\Sigma| z + \frac{1}{2} x^\top Q x \\ \text{s.t.} \quad & 0 = |\Sigma| z - \tilde{c} - \tilde{Z} x \;, \\ & 0 \le \Sigma z \;, \end{aligned} \tag{3.25}$$

with $\tilde{Z}$ and $\tilde{c}$ defined as in Eq. (3.9). Due to the fixed signature vector, this is a quadratic optimization problem with linear constraints. Of course, such problems can be solved with standard QP solvers, such as KNITRO [15], SCIP [13], Gurobi [52], or IPOPT [104], just to name a few. However, the developers of ASM wanted to explicitly exploit the structure of the signature vectors to achieve the best possible performance for the overall algorithm.

Based on the optimization problem (3.25), the ASM has three main components: First, the computation of a search direction, second, the step size in this direction, and finally, the optimality check and, in case of nonoptimality, the use of these conditions to switch from one polyhedron to the next one and thus to update σ.

3.3.1 Computing a Descent Direction for Given σ

To determine a descent direction, KKT-theory as stated in Theorem 2.2 is applied to the optimization problem (3.25). With Lagrange multipliers $\lambda \in \mathbb{R}^n$ and $\mu \in \mathbb{R}^s$, this provides the following necessary optimality conditions:

$$0 = a^\top + x^\top Q - \lambda^\top \tilde{Z} \ , \tag{3.26}$$
$$0 = b^\top |\Sigma| + \lambda^\top |\Sigma| - \mu^\top \Sigma \ , \tag{3.27}$$
$$0 = |\Sigma| z - \tilde{c} - \tilde{Z} x \ ,$$
$$0 \le \Sigma z \ , \quad 0 \le \mu \ , \quad 0 = \mu^\top \Sigma z \ . \tag{3.28}$$

As was done in [45], multiplying Eq. (3.27) from the right by Σ and using Eq. (3.28) yields

$$0 \le \mu^\top |\Sigma| = b^\top \Sigma + \lambda^\top \Sigma \ . \tag{3.29}$$

If $\sigma = \operatorname{sgn}(z)$ and hence the underlying x is stationary, Eq. (3.29) reduces because of the required complementarity $\mu^\top |\Sigma| z = 0$ to

$$0 = b^\top \Sigma + \lambda^\top \Sigma \ .$$

Hence, if the signature optimal point x_σ exists, it must satisfy the system of linear equation

$$\begin{bmatrix} Q & 0 & -\tilde{Z}^\top \\ 0 & 0 & |\Sigma| \\ -\tilde{Z} & |\Sigma| & 0 \end{bmatrix} \begin{bmatrix} x \\ z \\ \lambda \end{bmatrix} = \begin{bmatrix} -a \\ -|\Sigma| b \\ \tilde{c} \end{bmatrix} \ , \tag{3.30}$$

where the middle row is once more scaled by Σ to make the resulting Jacobian fully symmetric.

There is a lot of literature on solving saddle point systems, e.g., [113]. Besides, in [45, Lemma 3.1] a simplified way to solve such a saddle point system for the original problem representation has already been derived. Therefore, the lemma and the proof will be adapted in the following.

Lemma 3.15 (Partitioned solution (unconstrained case)). *The solution of the equations system* (3.30) *can be reduced to solving the symmetric semidefinite linear system*

$$\left[|\overline{\Sigma}|\tilde{Z}Q^{-1}\tilde{Z}^\top|\overline{\Sigma}|\right]\lambda = |\overline{\Sigma}|\tilde{Z}Q^{-1}\left[a+\tilde{Z}^\top|\Sigma|b\right] - |\overline{\Sigma}|\tilde{c} \tag{3.31}$$

for the nontrivial entries of $|\overline{\Sigma}|\lambda$ *where* $|\overline{\Sigma}| \equiv I_s - |\Sigma|$ *denotes the complementary orthogonal projection to* $|\Sigma|$. *The system is uniquely solvable exactly when the rows of the* $s \times n$ *matrix* $\tilde{Z}$ *that correspond to active kinks are linearly independent, i.e., LIKQ holds. For* $\tilde{\lambda}$ *as a solution of the reduced system* (3.31) *the dual and primal variables are then obtained as*

$$\begin{aligned}
\lambda &= |\overline{\Sigma}|\tilde{\lambda} - \Sigma b\,, \\
x &= Q^{-1}\left(\tilde{Z}^\top\lambda - a\right)\,, \\
z &= \tilde{c} + \tilde{Z}x\,.
\end{aligned}$$

Proof. If λ is given, the expressions for x can be read off directly from the original system (3.30) as well as the components of z which belongs to inactive kinks. The remaining components can simply be set to zero.

So the main task consists in calculating λ. Therefore, using Q to eliminate the block $\tilde{Z}$ on the bottom left of (3.30) yields

$$\begin{bmatrix} Q & 0 & -\tilde{Z}^\top \\ 0 & 0 & |\Sigma| \\ 0 & |\Sigma| & -\tilde{Z}Q^{-1}\tilde{Z}^\top \end{bmatrix}\begin{bmatrix} x \\ z \\ \lambda \end{bmatrix} = \begin{bmatrix} -a \\ -|\Sigma|b \\ \tilde{\tilde{c}} \end{bmatrix}\,,$$

with $\tilde{\tilde{c}} = \tilde{c} - \tilde{Z}Q^{-1}a$. Thus, the second row as a projection on $|\Sigma|$ gives $|\Sigma|\lambda = -|\Sigma|b$, which makes it easy to calculate the components belonging to the inactive kinks. To find the remaining components, consider the projection of the third line through $|\overline{\Sigma}|$

$$|\overline{\Sigma}|\left[|\Sigma|z - \tilde{Z}Q^{-1}\tilde{Z}^\top\left(|\Sigma| + |\overline{\Sigma}|\right)\lambda\right] = |\overline{\Sigma}|\tilde{\tilde{c}}\,.$$

Skipping the sign to the right-hand side and using $|\overline{\Sigma}|\cdot|\Sigma| = (I_s - |\Sigma|)|\Sigma| = 0$ as well as $|\Sigma|\lambda = -|\Sigma|b$ yields

$$\left[|\overline{\Sigma}|\tilde{Z}Q^{-1}\tilde{Z}^\top\right]\left(-|\Sigma|b + |\overline{\Sigma}|\lambda\right) = -|\overline{\Sigma}|\tilde{\tilde{c}}\,.$$

Finally, after moving the expression in $|\Sigma|b$ to the right-hand side, resubstituting $\tilde{\tilde{c}}$ from above yields

$$\left[|\overline{\Sigma}|\tilde{Z}Q^{-1}\tilde{Z}^\top|\overline{\Sigma}|\right]\lambda = -|\overline{\Sigma}|\left[\tilde{c} - \tilde{Z}Q^{-1}a\right] + |\overline{\Sigma}|\tilde{Z}Q^{-1}\tilde{Z}^\top|\Sigma|b$$
$$= |\overline{\Sigma}|\tilde{Z}Q^{-1}\left[a + \tilde{Z}^\top|\Sigma|b\right] - |\overline{\Sigma}|\tilde{c}\,,$$

which completes the proof. □

Numerically, the solutions obtained by Lemma 3.15 must yield $|\overline{\Sigma}|z = 0$ up to rounding. After checking this, the nonbasic components of z can be set exactly to zero, while λ is generally dense without any sign constraints.

Now in the following $(\hat{x}, \hat{z}, \lambda)$ denotes the solution of the saddle point system (3.30). This system is in principle similar to a classical Newton step. Thus, the search direction from a current point (x, z) in direction $(\hat{x}, \hat{z})$ results naturally in

$$\Delta x = \hat{x} - x \quad \text{and} \quad \Delta z = \hat{z} - z\,.$$

3.3.2 Computing a Step Size β

After calculating a search direction, the next question is how far to go in that direction. On the one hand, there are several optimization algorithms that explicitly compute a step size, but on the other hand, there are also methods that use a step size heuristic, since it is not always easy to compute. For this, reference is made to standard literature such as [8, 37, 61].

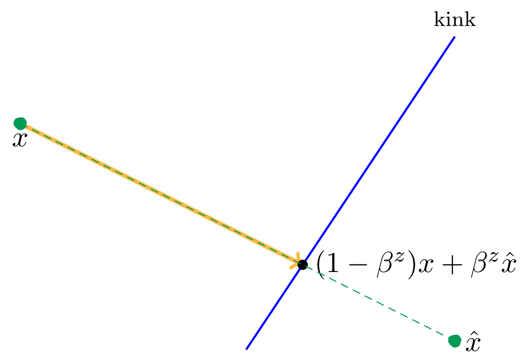

Figure 3.4: Illustration of the step sizes β^z

Since the functions in the context of this section are (piecewise) linear, it is quite easy to determine a step size for the ASM as described in [45]. If $(\hat{x}, \hat{z})$ are given as parts of the solution of Eq. (3.30), then it is necessary to check whether $\sigma(\hat{x}) = \sigma(x)$ still holds and thus $\hat{x}$ still lies in the same polyhedron as x or whether it is left. In the first case, a full step can be made along the direction Δx. In the latter case, however, there is a blocking constraint generated by the polyhedron, i.e., a blocking kink. A graphical illustration of this is shown in Figure 3.4, where the blue line represents a kink and the yellow arrow indicates the step size. Leaving a polyhedron results in a sign change in at least one of the components of z. The step size, denoted by β^z, can then be calculated by

$$\beta^z = \inf_{1 \le l \le s} \left\{ \beta_l^z \equiv \frac{-z_l}{\hat{z}_l - z_l} \middle| (\hat{z}_l - z_l)\sigma_l < 0 \right\} \in]0, \infty] \; . \tag{3.32}$$

Compared to [45], the factor z_l is replaced by σ_l in the sign test. Since only the sign is important, the formulation in Eq. (3.32) is equivalent to the original formulation, but numerically more stable. Note that only the first $\tilde{s}$ components represent the nonsmoothness and thus the kinks that impose nonsmoothness. Therefore, it is possible to stop at points where the function is smooth, but these points lie on a boundary of a polyhedron. If $\beta^z < \infty$ the first index for which the minimum is attained is denoted by j^z. If $\beta^z \le 1$, the step size is calculated to reach exactly a kink. In this case the corresponding index j^z is used to update the vector σ. The new sigma is set as $\sigma^+ = \sigma - \sigma_{j^z} e_{j^z}$ with the j^zth unit vector e_{j^z} which means setting $\sigma_{j^z} = 0$. Thus, the vector σ is restricted.

Note that due to the fact that for a given signature vector σ and a point $x \in \mathcal{P}_\sigma$ one has that for $\sigma_i \neq 0$ also $z_i(x) \neq 0$ holds. Thus, for σ with $\sigma_i \neq 0$ one has $z_i(x) \neq 0$ and $\beta^z > 0$ must hold.

In order to bound the step size, the current step size β is set to

$$\beta = \min\{\beta^z, 1\} \; , \tag{3.33}$$

where the upper bound 1 on β ensures with the update of the current iterate

$$x^+ = (1 - \beta)x + \beta\hat{x} = x + \beta\Delta x$$

that the next iterate is still contained in $\overline{\mathcal{P}}_\sigma$. In the case of a full step, i.e. $\beta = 1$, obviously $x^+ = \hat{x}$ and $\sigma(x^+) = \sigma$ holds. Since in this case the vector remains unchanged, the point x^+ is then also called *signature stationary*.

One could avoid stopping at points that lie on kinks that do not represent nonsmoothness by simply looking at the first $\tilde{s}$ indices in the infimum (3.32) and also checking whether such a kink is crossed. If this is the case, one must additionally change the corresponding entries of σ.

3.3.3 Checking the Optimality

If a signature optimal point is found on the current polyhedron $\mathcal{P}_\sigma$, it is necessary to check if it is already a minimizer of the original problem (ALOP) with the added regularization term. If such a minimizer is detected, the algorithm can terminate. Otherwise, the optimization must be continued on a neighboring polyhedron $\mathcal{P}_{\tilde{\sigma}}$ with definite $\tilde{\sigma} \succ \sigma$. As in the proof of the Theorem 3.14, such a $\tilde{\sigma}$ can be decomposed into $\sigma + \gamma$ with $|\sigma|^\top |\gamma| = 0$. Then replacing Σ in the optimality conditions (3.26)-(3.28) by the corresponding $\Sigma + \Gamma$ and using Eq. (3.16) yields that the primal feasibility condition and the dual equality constraint are still satisfied by the solution $(\hat{x}, \hat{z}, \lambda)$. The only difference is that Eq. (3.29) has as many new nontrivial components as γ, which can be written as

$$0 \leq \mu^\top |\Gamma| = b^\top \Gamma + \lambda^\top \Gamma - \lambda^\top \tilde{L} |\Gamma|$$

similar to Eq. (3.24). This optimality condition is violated if and only if there is at least one index k such that $\gamma = -\operatorname{sgn}(b_k + \lambda_k)\, e_k$ satisfies

$$0 > -\left| b^\top + \lambda^\top \right| e_k - \lambda^\top \tilde{L} e_k \quad \text{and} \quad \sigma_k = 0\,. \tag{3.34}$$

Eq. (3.34) represents a violation of the normal growth condition, as can be seen from Theorem 3.14. Thus, if the optimality condition is violated, i.e., Eq. (3.34) holds true for at least one k, one possible strategy is to choose the index k for which the right-hand side of Eq. (3.34) is minimal. This is a heuristic known, e.g., from Active Set Methods [87]. Then, by updating $\sigma_k = -\operatorname{sgn}(b_k + \lambda_k)$, the next signature vector, denoted by σ^+, has one component less that equals zeros. This can be interpreted as releasing a kink in that one does not insist anymore that the corresponding absolute value is evaluated at zero. Thus, the vector σ is relaxed.

3.3.4 The Overall Algorithm

A pseudocode for the ASM is given in Algorithm 3 and a schematic representation in Figure 3.5.

The Active Signature Method (ASM) given in Algorithm 3 combines all the previous steps in a suitable way. Therefore, given an optimization problem of the form (ALOP), a positive definite matrix $Q = Q^\top \in \mathbb{R}^{n \times n}$ and a starting point $x \in \mathbb{R}^n$, the associated switching and signature vectors are determined and then the saddle point system (3.30) is solved (cf. line 2 of Algorithm 3). Afterwards, in line 3 the step size β is computed

Algorithm 3 Active Signature Method (ASM)

Given: Optimization problem of the form (ALOP), $Q = Q^\top \in \mathbb{R}^{n\times n}$ positive definite and starting point $x \in \mathbb{R}^n$
Compute: $z := z(x)$ via Eq. (3.5), $\sigma := \sigma(z(x))$

1: **loop**
2: Compute $(\hat{x}, \hat{z}, \lambda)$ by solving Eq. (3.30) ▷ Solve saddle point system
3: Compute β^z via Eq. (3.32) and β via Eq. (3.33) ▷ Compute step size
4: Set $(x^+, z^+) = (1-\beta)(x, z) + \beta(\hat{x}, \hat{z})$ ▷ Update iterate
5: **if** $\beta^z = \beta$ **then** Restrict σ ▷ Add kink
6: **if** $\beta = 1$ **then** ▷ x^+ is signature stationary
7: **if** Eq. (3.34) holds true **then** ▷ x^+ is signature optimal
8: Relax σ, set $\beta = 0$ ▷ Drop kink
9: **else** ▷ x^+ is local optimal
10: **return** (x^+, z^+) ▷ Problem solved
11: Set $(x, z) = (x^+, z^+)$

as in Eq. (3.33). If now $\beta^z = \beta$ holds, i.e., $\beta^z \leq 1$ and thus a kink is reached, this kink is added by restricting σ (cf. line 5). If $\beta = 1$, then optimality is checked. Note that by construction, at least one of the two if conditions must be true in each iteration. If the current point is not optimal, Eq. (3.34) is used to modify, i.e., relax, a component of σ. Otherwise the algorithm terminates and has found a local minimum (cf. line 7–10). In case of nonoptimality the algorithm continues with the next iterate (cf. line 11).

For numerical results the reader should refer to the corresponding paper [45], where also other aspects, such as more efficient implementation of the update of the saddle point system are discussed.

3.3.5 Convergence Analysis of the Active Signature Method

The goal now is to show the convergence of ASM. The paper [45], in which this method was proposed, only briefly mentions that the algorithm terminates after a finite number of steps. The formal proof of this will now be given. First, an auxiliary result is needed which shows a descent of the objective function in certain cases. This lemma has already been formulated and proven in [48, Corollary A1]. However, the proof in this thesis is used to fill in missing details and to extend it to the regularized problem. Therefore, consider

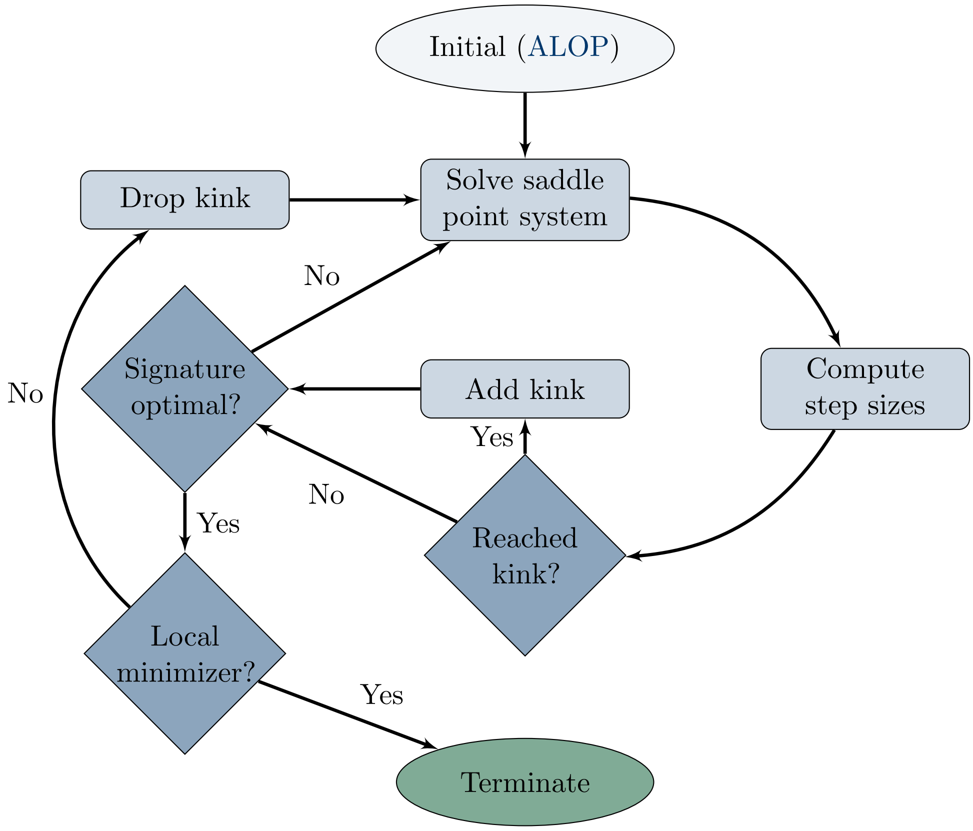

Figure 3.5: Scheme of Active Signature Method (ASM)

the following problem

$$\begin{aligned} \min_{x\in\mathbb{R}^n, z\in\mathbb{R}^s} \quad & a^\top x + b^\top |\Sigma| z + \frac{1}{2} x^\top Q x \\ \text{s.t.} \quad & 0 = |\Sigma| z - \tilde{c} - \tilde{Z} x \ , \end{aligned}$$

which is exactly the one related to the saddle point system (3.30), and is identical to (3.25) if one omits the sign condition on the switching equation described by the last inequality condition of (3.25). Substituting the equality constraint into the target function and the identity $|\Sigma| z = z$ on the corresponding polyhedron $\mathcal{P}_\sigma$ (cf. Eq. (3.8c)), the following function is obtained

$$f(x) := a^\top x + b^\top |\Sigma| \left(\tilde{c} + \tilde{Z} x \right) + \frac{1}{2} x^\top Q x \ . \tag{3.35}$$

To show the descent of this objective function, the next auxiliary result will be required.

Lemma 3.16. *Assume that $x \in \mathcal{P}_\sigma$ and σ is the signature vector for a given switching equation*

$$z_\sigma(x) = \tilde{c} + \tilde{Z}x \ ,$$

with $\tilde{c}$ and $\tilde{Z}$ defined as in Eq. (3.9). Further assume that $\gamma \in \{-1, 0, 1\}^s$ is structurally orthogonal to σ, i.e., $\sigma^\top \gamma = 0$ and $||\gamma||_1 = 1$. Then for $\tilde{\sigma} := \sigma + \gamma$ it holds

$$z_{\tilde{\sigma}}(x) = z_{\sigma+\gamma}(x) = \left(I_s - \tilde{L}\Gamma\right)^{-1} \left(\tilde{c} + \tilde{Z}x\right) = \tilde{c} + \tilde{Z}x \ . \tag{3.36}$$

Proof. The first two equalities directly followed from Eq. (3.16). Now let i denote the component where $\gamma_i \neq 0$. Since γ is orthogonal to σ it must hold that $\sigma_i = 0$. In the first brackets in Eq. (3.36) the matrix $\Gamma = \text{diag}(\gamma)$ acts as a projection onto the ith column of $\tilde{L}$. Since $\tilde{L}$ is strictly lower triangular, $\left(I_s - \tilde{L}\Gamma\right)$ is also lower triangular with ones on the diagonal and other entries only in the ith column below of the diagonal element. Therefore, it holds that $\left(I_s - \tilde{L}\Gamma\right)^{-1} = \left(I_s + \tilde{L}\Gamma\right)$. Thus, using $z = |\Sigma| z = \tilde{c} + \tilde{Z}x$ and $\Gamma\Sigma = 0$ yields

$$\left(I_s - \tilde{L}\Gamma\right)^{-1} \left(\tilde{c} + \tilde{Z}\right) = \left(I_s + \tilde{L}\Gamma\right) |\Sigma| z = \tilde{c} + \tilde{Z}x \ .$$

□

Now it is possible to show results regarding the descent direction in the nonoptimal case, as given in [48, Corollary A1]

Lemma 3.17 (Descent direction in the nonoptimal case for ASM)**.** *Let $x \in \mathbb{R}^n$ and the optimization problem (3.25) be given and suppose that LIKQ holds in x. If tangential stationarity is violated, i.e.,*

$$0 \neq a^\top + b^\top |\Sigma| \tilde{Z} + x^\top Q - \lambda^\top P_\alpha^\top P_\alpha \tilde{Z} \ ,$$

there exists some direction $d \in \mathbb{R}^n$ such that $P_\alpha \tilde{Z} d = 0$ but

$$\left(a^\top + b^\top |\Sigma| \tilde{Z} + x^\top Q\right) d < 0 \ , \tag{3.37}$$

and the target function defined in (3.35) is decreasing in direction τd, i.e., $f(x + \tau d) < f(x)$ for $\tau \gtrsim 0$. If tangential stationarity holds but normal growth fails there exists at least one

$i \in \alpha$ with $|b_i + \lambda_i| > -\lambda^\top \tilde{L} e_i$. Defining $\gamma = -\operatorname{sgn}(b_i + \lambda_i) e_i \in \mathbb{R}^s$, any d satisfying

$$\left(I_s - \tilde{L}\Gamma\right)^{-1} \tilde{Z} d = \gamma \tag{3.38}$$

is a descent direction.

Proof. In the first case, since tangential stationarity is violated, the point x is not a minimizer on the current polyhedron $\mathcal{P}_{\sigma(x)}$. Therefore, let x_σ be a minimizer on $\overline{\mathcal{P}}_{\sigma(x)}$ and denote by $d \coloneqq x_\sigma - x$ a corresponding direction. Since both points, x and x_σ lie on the same extended signature domain, they have the same index set of active switching variables α and thus $P_\alpha \tilde{Z} d = 0$ holds. Furthermore, since f is convex due to the positive definite matrix Q, the vector d is a descent direction in that f is decreasing in direction τd for $\tau \in]0, 1]$.

Moreover, $x + \tau d \in \mathcal{P}_\sigma$ holds true for $\tau \gtrsim 0$, since by $P_\alpha \tilde{Z} d = 0$ the components of $z(x + \tau d)$ with indices in α stay zero and the others vary only slightly. Then, computing the directional derivative of f at x in direction d and using the symmetry of Q yields

$$\begin{aligned} f'(x; d) &= \lim_{t \searrow 0} \frac{1}{t} \left(a^\top (x + td) + b^\top |\Sigma| \left(\tilde{c} + \tilde{Z}(x + td) \right) + \frac{1}{2} (x + td)^\top Q (x + td) \right. \\ &\qquad \left. - \left(a^\top x + b^\top |\Sigma| \left(\tilde{c} + \tilde{Z} x \right) + \frac{1}{2} x^\top Q x \right) \right) \\ &= \lim_{t \searrow 0} \frac{1}{t} \left(t a^\top d + t b^\top |\Sigma| \tilde{Z} d + t x^\top Q d + \frac{1}{2} t^2 d^\top Q d \right) \\ &= \left(a^\top + b^\top |\Sigma| \tilde{Z} + x^\top Q \right) d \,. \end{aligned} \tag{3.39}$$

Using Lemma 2.8, d being a descent direction of f in x is equivalent to $f'(x; d) < 0$. This together with Eq. (3.39) proves Eq. (3.37).

Otherwise, λ is well defined and one can choose $i \in \alpha$ with $|b_i + \lambda_i| > -\lambda^\top \tilde{L} e_i$. Setting $\gamma = \gamma_i e_i$ with $\gamma_i = -\operatorname{sgn}(b_i + \lambda_i)$, one obtains for d with $\left(I_s - \tilde{L}\Gamma\right)^{-1} \tilde{Z} d = \gamma$ that $x + \tau d \in \mathcal{P}_{\sigma + \gamma}$. To see this, consider Eq. (3.16) and Lemma 3.16

$$\begin{aligned} z_{\sigma+\gamma}(x + \tau d) &= \left(I_s - \tilde{L}\Gamma\right)^{-1} \left(\tilde{c} + \tilde{Z} \left(x + \tau d \right) \right) \\ &= \left(I_s - \tilde{L}\Gamma\right)^{-1} \left(\tilde{c} + \tilde{Z} x \right) + \tau \left(I_s - \tilde{L}\Gamma\right)^{-1} \tilde{Z} d \\ &= z_\sigma(x) + \tau \gamma \,. \end{aligned}$$

Thus, since $\sigma_i = 0$ it holds that $\operatorname{sgn}(z_{\sigma+\gamma}(x + \tau d)) = \operatorname{sgn}(z_\sigma(x)) + \operatorname{sgn}(\tau\gamma) = \sigma + \gamma$ yielding

$x + \tau d \in \mathcal{P}_{\sigma+\gamma}$. On that polyhedron the Lagrange multiplier vector μ is also well defined by Eq. (3.23). It can be re-sorted to

$$\mu^\top = \left(b^\top + \lambda^\top\right) \Gamma P_\alpha^\top - \lambda^\top \tilde{L} P_\alpha^\top \ ,$$

and in this case by assumption for the ith component it follows

$$\mu_i = (b_i + \lambda_i)\gamma_i - \lambda^\top \tilde{L} e_i = -|b_i + \lambda_i| - \lambda^\top \tilde{L} e_i < 0 \ .$$

Now consider the directional derivative of the objective on $\mathcal{P}_{\sigma+\gamma}$ at x in direction τd:

$$\begin{aligned}
&f'(x;\tau d)\\
&= \lim_{t\searrow 0} \frac{1}{t} \Bigg(a^\top(x{+}t\tau d) + b^\top(|\Sigma|{+}|\Gamma|) \left(I_s{-}\tilde{L}\Gamma\right)^{-1} \left(\tilde{c}{+}\tilde{Z}\,(x{+}t\tau d)\right)\\
&\qquad + \frac{1}{2}(x{+}t\tau d)^\top Q(x{+}t\tau d) - a^\top x - b^\top\,(|\Sigma|{+}|\Gamma|) \left(I_s{-}\tilde{L}\Gamma\right)^{-1} \left(\tilde{c}{+}\tilde{Z}x\right) - \frac{1}{2}x^\top Q x \Bigg)\\
&= \lim_{t\searrow 0} \frac{1}{t} \left(t\tau a^\top d + t\tau b^\top\,(|\Sigma| + |\Gamma|) \left(I_s - \tilde{L}\Gamma\right)^{-1} \tilde{Z} d + t\tau x^\top Q d + \frac{1}{2} t^2\tau^2 d^\top Q d \right)\\
&= \tau a^\top d + \tau x^\top Q d + \tau b^\top\,(|\Sigma| + |\Gamma|) \left(I_s - \tilde{L}\Gamma\right)^{-1} \tilde{Z} d \ .
\end{aligned}$$

Finally, substituting Eqs. (3.26) (which does not depend on the $\sigma + \gamma$ decomposition but includes the regularization term), (3.19) and (3.38) as well as using $\mu_i < 0$ the last expression then yields

$$f'(x;\tau d) = \tau \left(\lambda^\top \tilde{Z} d - \lambda^\top \tilde{Z} d + \mu^\top P_\alpha \Gamma \left(I_s - \tilde{L}\Gamma\right)^{-1} \tilde{Z} d \right) = \tau \mu_i \gamma_i^2 < 0 \ .$$

Therefore, there is again a descent, which completes the proof. □

Note that in the case of $Q = 0$ Lemma 3.17 simplifies exactly to the purely (piecewise) linear case considered in the appendix of [48]. Using this lemma, it is now possible to show the finite convergence of Algorithm 3, which reduces the result of [70, Theorem 4.2] to the unconstrained case.

Theorem 3.18. *Suppose that an optimization problem of the form* (ALOP) *is given, LIKQ holds at every point $x \in \mathbb{R}^n$ and let $Q = Q^\top \in \mathbb{R}^{n\times n}$ be a positive definite matrix. Then, Algorithm 3 terminates for any starting point $x \in \mathbb{R}^n$ after finitely many iterations at a* *minimizer of the quadratically penalized optimization problem*

$$\begin{aligned} \min_{x\in\mathbb{R}^n, z\in\mathbb{R}^s} \quad & a^\top x + b^\top z + \frac{1}{2} x^\top Q x \\ s.t. \quad & z = c + Zx + Mz + L|z| \, . \end{aligned} \tag{3.40}$$

Proof. Algorithm 3 starts by solving the system of equations (3.30). If the solution returns the same point $\hat{x}$ as the current point x, then x is a signature optimal point, the step size is $\beta = 1$, and the algorithm tests for optimality. If this test is successful, the algorithm terminates. If it is not successful, then, as seen in the proof of Lemma 3.17, there exists at least one index $i \in \alpha$ such that $\mu_i < 0$. Furthermore, it follows from Lemma 3.17 and the fact that $\beta > 0$, that after adjusting the signature vector, there exists a new descent direction. Lemma 4.11 shows that solving the system of equations (3.30) also yields a descent direction. Since the algorithm always finds a signature optimal point before changing the polyhedron and thus obtains the best possible function value on this polyhedron, the algorithm cannot reach the same polyhedron again because of the descent after leaving a kink. In addition, there are only finitely many closed polyhedra, at most 2^s.

Now consider the situation where there is no signature optimum point yet. Then the solutions of the saddle point system (3.30) provide a search direction. The calculated step size will then either return $\beta = 1$, where the optimum is reached on the current polyhedron and the next iteration returns a zero step, or (maybe also and) a kink is added. The latter can be repeated at most s times - due to LIKQ - until all linearly independent kinks have been added. Again, the next iteration will return a zero step.

So the key arguments are that nonzero steps always yield a descent, an optimum is found on a polyhedron after finitely many iterations, and there are only finitely many polyhedra. Hence, this leads to convergence after finitely many iterations. □

To see that the ASM also provides a solution to the original unregularized problem (ALOP), it is referred to Theorem 4.13 in the following chapter.

Furthermore, it is important to mention again that the optimization problem (3.40) always has a global minimum due to the quadratic regularization, but the algorithm only guarantees to find a local minimum.

3.4 Optimization Methods Based on Abs-Linearization

Since the main focus in the previous parts of this thesis was on the ASM, which is a substantial foundation for the further results in this thesis, this section shall be used to give a short overview of other methods based on the Abs-Linear Form respectively the abs-linearization.

One approach that has been developed in recent years, is the so-called *Successive Abs-Linear MINimization (SALMIN)*. It was mainly developed by Sabrina Fiege during her PhD thesis [28] together with her supervisors Andrea Walther and Andreas Griewank [29, 30, 49]. SALMIN is designed to optimize respectively minimize continuous nonsmooth objective functions without constraints. For this purpose it is assumed that the nonsmoothness is only caused by the evaluation of the absolute value. In simple terms, the method consists of generating an abs-linear model for each current iterate and then optimizing this abs-linearization. This approach is motivated by smooth optimization, where quadratic models of the objective function that provide a good approximation are generated and then successively minimized on them. To generate the abs-linear model, as already mentioned, several AD tools can be used. As a solver for the abs-linear model, another method was originally developed, see [28], which was then replaced by ASM due to its higher efficiency.

Therefore, to keep it brief and denoting the vector of optimization variables by x the optimization approach of SALMIN can be described by the updating step

$$x^{+} = x + \arg\min_{\Delta x} \left\{ \Delta F(x; \Delta x) + \frac{\rho}{2} ||\Delta x||^2 \right\} , \tag{3.41}$$

in each iteration. Here, ρ denotes a penalty parameter, for which also adaptation strategies exists, and Δx the search direction as described in the path before Theorem 3.4.

In the article [69], to which the author of this work also contributed, the SALMIN approach was used to solve differential algebraic equations (DAEs) in combination with a least-squares collocation. This problem was motivated by gas networks with regulating elements, whose mathematical modeling results in DAEs. In contrast to classical approaches, more collocation points were chosen than there are degrees of freedom for the collocation polynomials. The resulting system of equations is nonsmooth due to the regulating elements in the gas networks. Therefore, using a least squares formulation, the SALMIN approach can be used to solve this system.

The idea of transferring SALMIN from finite-dimensional to infinite-dimensional optimization was also developed by Olga Weiß [109] under the supervision of Andrea Walther.

Since this would lead too far away from the subject and the necessary notation without having a direct relation to the content of this thesis, the interested reader is referred to the corresponding work [107, 109, 110, 111, 112].

Another algorithm based on the ASM combined with the so-called *Frank-Wolfe* (or *conditional gradient*) algorithm [36, 75] is currently under development. Its goal is to optimize an abs-smooth function over a compact and convex set [68]. Since the author of this thesis is also involved in the development and the approach can be seen as an extension of the results of this thesis, the reader is referred to the outlook, Section 6.2, and, of course, to [68] for more details. Furthermore, there are also works on the Frank-Wolfe algorithms in the infinite-dimensional setting with nonsmoothness, e.g., [73, 89].

4

The Constrained Abs-Linear Optimization Problem: Theory and Solution Method

In the previous chapter, especially in Section 3.3, the focus was mainly on unconstrained optimization problems. In this chapter, however, the results and algorithms of Chapters 2 and 3 will be combined and extended to develop an algorithm which can solve constrained piecewise linear optimization problems, i.e., with piecewise linear constraints as well. In addition to academic test problems, there are of course real-world application problems where such optimization problems arise. As a motivation, gas networks are mentioned and for further details as well as numerical results, reference is made to the next Chapter 5.

Several aspects presented in the following, notably the solution approach, the Constrained Active Signature Method (CASM), have already been published in [70] and are further extended here.

As in the unconstrained case, assume that the objective function, but now also the constraints are given in Abs-Linear Form (see Definition 3.2). Since it can happen that some components of the switching vectors belong to more than one of the occurring functions, assume that the switching vectors of all functions are combined into one switching system

$$z = c + Zx + Mz + L|z| \ ,$$

as introduced already in Eq. (3.3). Thus, the following *constrained abs-linear optimization problem* (CALOP) is the starting point for further investigations:

$$\begin{aligned} \min_{x \in \mathbb{R}^n, z \in \mathbb{R}^s} \quad & a^\top x + b^\top z \\ \text{s.t.} \quad & 0 = g + Ax + Bz + C|z| \ , \\ & 0 \geq h + Dx + Ez + F|z| \ , \\ & z = c + Zx + Mz + L|z| \ , \end{aligned} \tag{CALOP}$$

where $g \in \mathbb{R}^m$, $h \in \mathbb{R}^p$, $A \in \mathbb{R}^{m\times n}$, $B, C \in \mathbb{R}^{m\times s}$, $D \in \mathbb{R}^{p\times n}$ and $E, F \in \mathbb{R}^{p\times s}$. Therefore, (CALOP) is a single objective optimization problem with m equality, p inequality constraints and s switching variables. Since, the constraints should be allowed to be as general as possible, there are no further restrictions made on the matrices C and F. For later use, define

$$\begin{aligned}
&f : \mathbb{R}^n \times \mathbb{R}^s \to \mathbb{R}, && (x, z) \mapsto a^\top x + b^\top z \;, \\
&G : \mathbb{R}^n \times \mathbb{R}^s \times \mathbb{R}^s \to \mathbb{R}^m, && (x, z, |z|) \mapsto g + Ax + Bz + C|z| \;, \qquad (4.1)\\
\text{and}\quad &H : \mathbb{R}^n \times \mathbb{R}^s \times \mathbb{R}^s \to \mathbb{R}^p, && (x, z, |z|) \mapsto h + Dx + Ez + F|z| \;. \qquad (4.2)
\end{aligned}$$

By adding the constraints, the concept of signature domains (cf. Definition 3.8) can be extended, or more precisely, the respective sets can be constrained.

Definition 4.1 (Feasible (extended) signature domain)**.** *For a fixed signature vector* $\sigma \in \{-1, 0, 1\}^s$*, define*

$$\mathcal{F}_\sigma \equiv \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{l} G(x, z(x), \Sigma z(x)) = 0, \\ H(x, z(x), \Sigma z(x)) \le 0, \\ \operatorname{sgn}(z(x)) = \sigma, \end{array} \right\} \subseteq \overline{\mathcal{F}}_\sigma \equiv \left\{ x \in \mathbb{R}^n \;\middle|\; \begin{array}{l} G(x, z(x), |z(x)|) = 0, \\ H(x, z(x), |z(x)|) \le 0, \\ \Sigma z(x) = |z(x)| \end{array} \right\} .$$

The set $\mathcal{F}_\sigma$ *is called* feasible signature domain *and* $\overline{\mathcal{F}}_\sigma$ *the* feasible extended signature domain*.*

Note that Definition 4.1 does not depend on the explicit representation of the constraints as in Eq. (4.1) and Eq. (4.2), but can applied in general to arbitrary equality and inequality constraint functions G and H. Obviously, for the feasible (extended) signature domains, the subset relations $\mathcal{F}_\sigma \subseteq \mathcal{P}_\sigma$ and $\overline{\mathcal{F}}_\sigma \subseteq \overline{\mathcal{P}}_\sigma$ hold, and with appropriate constraints, $\mathcal{F}_\sigma$ can also be empty. To illustrate this extended definition, consider again the Hill-function from the previous chapter.

Example 4.2 (Feasible signature domains for Hill-function)**.** *In the previous examples, the Hill-function was always considered as an unconstrained function, see Example 3.3. Now add the following constraint*

$$|z_3| = \left| -\frac{1}{2}|z_1| + \frac{1}{2}\big|x_1 - |x_2|\big| \right| \le 2 \;, \tag{4.3}$$

which in the sense of Eq. (4.2) *can be formulated as*

$$H(x,z,|z|) = -2 + \begin{bmatrix} 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} |z_1| \\ |z_2| \\ |z_3| \end{bmatrix} \leq 0 \, .$$

Now, $|z_3|$ *contributes explicitly to the evaluation of the abs-linear constraint. Thus, in comparison to the unconstrained case, further kinks that introduce nonsmoothness via the constraints are added. Figure 4.1 shows the polyhedra* $\mathcal{P}$ *labeled by the corresponding definite signature vectors* σ*. All points that lie inside or on the edges of the red area are feasible.*

Similar to the switching vectors, signature vectors can be defined for the inequality constraints. From the point of view of the Active Set Method (see Section 2.2), this is a way of representing the active set $\mathcal{W}$.

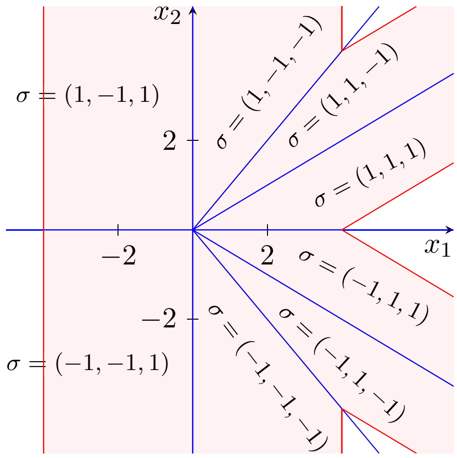

Figure 4.1: Feasible signature domains for constrained Hill-function

Definition 4.3 (Signature vector and signature matrix of inequality constraints)**.** *Let* $x \in \mathbb{R}^n$ *be given such that it fulfills the equality and inequality constraints of* (CALOP)*. Then the* signature vector of the inequality constraints *is defined as*

$$\omega(x) \equiv \operatorname{sgn}(H(x,z,|z|)) \in \{-1, 0\}^p \, .$$

The ith inequality constraint is called active *if* $\omega_i(x) = 0$ *and* inactive *otherwise. The* signature matrix of the inequality constraints *is denoted by* $\Omega(x) = \operatorname{diag}(\omega(x))$*. Furthermore,* $\mathcal{I} \equiv \mathcal{I}(x)$ *collects the indices of the active inequality constraints at* x*. The* projection onto the active components of $H(x)$ *is defined as* $P_{\mathcal{I}} \equiv (e_i^\top)_{i \in \mathcal{I}} \in \mathbb{R}^{|\mathcal{I}| \times p}$ *with* e_i *denoting the ith unit vector of appropriate size.*

After this short preparation, the rest of this chapter is structured very similar to the previous one, except that the constraints are now considered everywhere. First, in Section 4.1 the optimality conditions are shown, which can still be verified in polynomial time as already in the unconstrained case. Then, in Section 4.2, CASM is presented. For this purpose, the section is divided into the components of computing a descent direction (see Subsection 4.2.1) and a step size (see Subsection 4.2.2), testing the optimality (see Subsection 4.2.3), the algorithm itself as pseudocode (see Subsection 4.2.4) and finally the convergence analysis in Subsection 4.2.5. As already mentioned, many of the following sections are taken from [70], the paper related to CASM. This chapter then concludes in Section 4.3 with a brief discussion of the possible use of penalty approaches and Phase I Methods in the context of piecewise linear optimization problems.

4.1 Optimality Conditions

As in the unconstrained case, necessary and sufficient optimality conditions for (CALOP) are shown first. Therefore, consider for each fixed $\sigma \in \{-1, 0, 1\}$ the optimization problem

$$\begin{aligned} \min_{x\in\mathbb{R}^n, z\in\mathbb{R}^s} \quad & a^\top x + b^\top z \\ \text{s.t.} \quad & 0 = g + Ax + Bz + C\Sigma z \;, \\ & 0 \geq h + Dx + Fz + G\Sigma z \;, \\ & z = c + Zx + Mz + L\Sigma z \;, && \text{(4.4a)} \\ & 0 = (I_s - |\Sigma|)z \;, && \text{(4.4b)} \\ & 0 \leq \Sigma z \;, \end{aligned}$$

where again I_s denotes the identity matrix of dimension s. Therefore, as an extension of Definition 3.10, feasible optimal points are defined as follows:

Definition 4.4 (Feasible signature optimal point)**.** *Let an optimization problem of the form* (CALOP) *be given. Consider a fixed signature vector* $\sigma \in \{-1, 0, 1\}^s$*. A minimizer* $x_\sigma \in \mathcal{F}_\sigma$ *of the optimization problem* (4.4) *is called* feasible signature optimal point *of the original, constrained optimization problem* (CALOP)*.*

Note that compared to the signature optimal points for (ALOP), for the feasible signature optimal points only the feasibility is added as a new property.

In the same way as in the unconstrained case, it is possible to combine the two constraints Eqs. (4.4a) and (4.4b) and using the notation from Eq. (3.9) to reformulate problem (4.4) into the equivalent problem

$$
\begin{aligned}
\min_{x\in\mathbb{R}^n, z\in\mathbb{R}^s} \quad & a^\top x + b^\top |\Sigma| z \\
\text{s.t.} \quad & 0 = g + Ax + Bz + C\Sigma z \ , \\
& 0 \geq h + Dx + Fz + G\Sigma z \ , \\
& 0 = |\Sigma| z - \tilde{c} - \tilde{Z} x \ , \\
& 0 \leq \Sigma z \ .
\end{aligned} \tag{4.5}
$$

Arguing as in Section 3.2, since this optimization problem has only linear constraints by fixing the signature Σ, the set of feasible directions at x coincides with the tangent cone at x, see [87, Lemma 12.7] and therefore no further constraint qualifications are needed to ensure the existence of Lagrange multipliers. However, this does not guarantee their uniqueness. But since the goal for the constrained case is also to derive optimality conditions that can be verified in polynomial time, it is necessary to ensure their uniqueness. Otherwise, any dependence on the signature vectors would lead to combinatorial complexity in 2^s. As regularity conditions, the LIKQ for unconstrained problems was given in Definition 3.12. In the context of her PhD thesis, Lisa Hegerhorst-Schultchen extended this concept for constrained optimization problems under the supervision of Marc Steinbach [54, 55, 56]. In the present work, a version is defined that is specially adapted to the notation used in the thesis and that exploits the piecewise linearity.

To derive these optimality conditions, the optimization problem (4.5) is analyzed for a feasible signature optimal point x_σ in more detail. Due to the continuity of all involved functions and the relation $\Sigma z = |z|$, the components z_i, with $i \notin \alpha$, of the vector z determined by the penultimate constraint of (4.5) will not drop to zero in an open neighborhood $U(x_\sigma)$ of x_σ. As a consequence, in combination with the identity $\Sigma z = \Sigma|\Sigma| z$, in this neighborhood $U(x_\sigma)$ the optimization problem (4.5) is equivalent to

$$
\begin{aligned}
\min_{x\in U(x_\sigma)} \quad & a^\top x + b^\top |\Sigma| (\tilde{c} + \tilde{Z} x) \\
\text{s.t.} \quad & 0 = g + Ax + B|\Sigma|(\tilde{c} + \tilde{Z} x) + C\Sigma(\tilde{c} + \tilde{Z} x) \ , & \text{(4.6a)} \\
& 0 \geq h + Dx + E|\Sigma|(\tilde{c} + \tilde{Z} x) + F\Sigma(\tilde{c} + \tilde{Z} x) \ , & \text{(4.6b)} \\
& 0 = P_\alpha(\tilde{c} + \tilde{Z} x) \ . & \text{(4.6c)}
\end{aligned}
$$

Definition 4.5 (Active Jacobian)**.** *Consider for the constrained optimization problem* (CALOP) *and a given signature vector $\sigma \in \{-1, 0, 1\}^s$ a feasible point x_σ. The* active

Jacobian *of the equivalent problem* (4.6) *is given by*

$$\mathcal{J}_\sigma \equiv \begin{bmatrix} A + B|\Sigma|\tilde{Z} + C\Sigma\tilde{Z} \\ P_{\mathcal{I}}(D + E|\Sigma|\tilde{Z} + F\Sigma\tilde{Z}) \\ P_\alpha \tilde{Z} \end{bmatrix} \in \mathbb{R}^{(m+|\mathcal{I}|+|\alpha|)\times n} .$$

Now, the required kink qualification can be stated for the setting considered in this thesis.

Definition 4.6 (LIKQ (constrained case))**.** *Let a constrained optimization problem of the form* (CALOP) *and a signature vector* $\sigma \in \{-1, 0, 1\}^s$ *be given. Then the* linear independence kink qualification (LIKQ) *holds at a feasible point* $x \in \mathcal{F}_\sigma$ *if the active Jacobian* $\mathcal{J}_\sigma$ *has full row rank* $m + |\mathcal{I}| + |\alpha|$.

Example 4.7 (LIKQ for constrained Hill-function)**.** *Using Eq.* (3.9) *one obtains for the constrained Hill-function given in its Abs-Linear Form as stated in Example 3.3 with the constraint* (4.3) *at the point* $\bar{x} = (4, -4)^\top$ *the active Jacobian*

$$\mathcal{J}_{(-1,0,1)^\top} = \begin{bmatrix} P_{\mathcal{I}(\bar{x})}\left(D + E|\Sigma|\tilde{Z} + F\Sigma\tilde{Z}\right) \\ P_{\alpha(\bar{x})}\tilde{Z} \end{bmatrix} = \begin{bmatrix} 0 & -\frac{1}{2} \\ 1 & 1 \end{bmatrix} ,$$

which has obviously full row rank $|\mathcal{I}(\bar{x})| + |\alpha(\bar{x})| = 2$. *Thus, LIKQ is fulfilled in* $\bar{x}$. *Contrary to Example 3.13, LIKQ is no longer satisfied at every point, because, e.g., for the origin all three switching variables are active, so the three lines of* $P_{\alpha(0)}\tilde{Z}$ *are no longer linearly independent.*

To shorten the notation for the further course, define for the equality and inequality parts of the active Jacobian the following designators

$$\mathcal{J}_{\sigma,G} := A + B|\Sigma|\tilde{Z} + C\Sigma\tilde{Z} \quad \text{and} \quad \mathcal{J}_{\sigma,H} := D + E|\Sigma|\tilde{Z} + F\Sigma\tilde{Z} . \tag{4.7}$$

In order to consider only the terms which belong to the switching variable z, another z is added in the subscript. This results in the following short notations

$$\mathcal{J}_{\sigma,G,z} := B|\Sigma| + C\Sigma \quad \text{and} \quad \mathcal{J}_{\sigma,H,z} := E|\Sigma| + F\Sigma \tag{4.8}$$

and obviously the identities

$$\mathcal{J}_{\sigma,G} = A + \mathcal{J}_{\sigma,G,z}\tilde{Z} \quad \text{and} \quad \mathcal{J}_{\sigma,H} = D + \mathcal{J}_{\sigma,H,z}\tilde{Z} . \tag{4.9}$$

Thus, all preliminary considerations from the unconstrained case are extended and therefore it is also possible to show the necessary and sufficient optimality conditions from Theorem 3.14 for the constrained optimization problem (CALOP). These were also already published in [70] and are included here accordingly.

Theorem 4.8 (Necessary and sufficient optimality condition for (CALOP))**.** *Let a constrained optimization problem of the form* (CALOP) *and a signature vector* $\sigma \in \{-1, 0, 1\}^s$ *be given. Assume that* x_σ *is feasible signature optimal for* (CALOP) *and that the LIKQ holds at* x_σ*. Then* x_σ *is a local minimizer of* (CALOP) *if and only if there exist unique Lagrange multipliers* $\delta \in \mathbb{R}^m$*,* $0 \leq \nu \in \mathbb{R}^p$ *and* $\lambda \in \mathbb{R}^s$*, such that*

$$0 = a^\top + b^\top|\Sigma|\tilde{Z} + \delta^\top \mathcal{J}_{\sigma,G} + \nu^\top \mathcal{J}_{\sigma,H} - \lambda^\top P_\alpha^\top P_\alpha \tilde{Z} \;, \tag{4.10}$$

$$0 = b^\top|\Sigma| + \delta^\top \mathcal{J}_{\sigma,G,z} + \nu^\top \mathcal{J}_{\sigma,H,z} + \lambda^\top|\Sigma| \tag{4.11}$$

and

$$|P_\alpha(b + B^\top\delta + E^\top\nu + \lambda)| \leq P_\alpha \left(C^\top\delta + F^\top\nu - \tilde{L}^\top\lambda\right) \tag{4.12}$$

with $\tilde{L}$ *given by* $\tilde{L} = (I_s - M - L\Sigma)^{-1}L$ *.*

Proof. Since x_σ is feasible signature optimal for (CALOP) per definition it is also a minimizer of (4.4) for the given signature vector σ. As seen by the conversions above, x_σ is thus also a minimizer of (4.6). Therefore, due to the assumption LIKQ one obtains from the KKT-theory (cf. Theorem 2.5) that there exist unique Lagrange multiplier $\delta \in \mathbb{R}^m$, $0 \leq \nu \in \mathbb{R}^p$ and $\check{\lambda} \in \mathbb{R}^{|\alpha|}$ associated with the equality constraint (4.6a), the inequality constraint (4.6b) and the equality constraint given by the reformulated switching system (4.6c) such that

$$0 = a^\top + b^\top|\Sigma|\tilde{Z} + \delta^\top \mathcal{J}_{\sigma,G} + \nu^\top \mathcal{J}_{\sigma,H} + \check{\lambda}^\top P_\alpha \tilde{Z} \;.$$

Hence together with $\delta \in \mathbb{R}^m$ and $0 \leq \nu \in \mathbb{R}^p$, each vector $\lambda \in \mathbb{R}^s$ such that $\check{\lambda} = -P_\alpha\lambda$ fulfills Eq. (4.10).

For local minimality of (CALOP) it is necessary and sufficient that $(x_\sigma, z(x_\sigma))$ is a minimizer of (4.4) on all feasible extended signature domains $\overline{\mathcal{F}}_{\tilde{\sigma}}$ with definite $\tilde{\sigma} \succ \sigma$. Any such $\tilde{\sigma} \succ \sigma$ can be written as $\tilde{\sigma} = \sigma + \gamma$ with $\gamma \in \{-1, 0, 1\}^s$ structurally orthogonal to σ such that for $\Gamma = \text{diag}(\gamma)$ the matrix equations Eq. (3.15) still holds. Therefore, it is still possible to express $z(x) = z_{\tilde{\sigma}}(x)$ for $x \in \overline{\mathcal{P}}_\sigma$ as stated in Eq. (3.16).

Now x_σ must be the minimizer on $\overline{\mathcal{P}}_{\tilde{\sigma}}$, i.e., solves the smooth constrained optimization problem

$$\begin{aligned}
\min_{x\in\mathbb{R}^n, z\in\mathbb{R}^s} \quad & a^\top x + b^\top \left(|\Sigma| + |\Gamma|\right) z \\
\text{s.t.} \quad & 0 = g + Ax + B\left(|\Sigma| + |\Gamma|\right) z + C\left(\Sigma + \Gamma\right) z \;, && \text{(4.13a)}\\
& 0 \geq h + Dx + E\left(|\Sigma| + |\Gamma|\right) z + C\left(\Sigma + \Gamma\right) z \;, && \text{(4.13b)}\\
& 0 = \left(I_s - \tilde{L}\Gamma\right) z - \tilde{c} - \tilde{Z}x \;, && \text{(4.13c)}\\
& 0 \leq P_\alpha \Gamma z \;. && \text{(4.13d)}
\end{aligned}$$

Again, notice that the inequalities are only imposed on the sign constraints that are active at x_σ since the strict inequalities are maintained in a neighborhood of x_σ due to the continuity of $z(x)$. Applying once more the KKT-theory to problem (4.13), there must exist Lagrange multipliers $\delta \in \mathbb{R}^m$, $0 \leq \nu \in \mathbb{R}^p$, $\lambda \in \mathbb{R}^s$ and $0 \leq \mu \in \mathbb{R}^{|\alpha|}$ associated with the equality constraint (4.13a), the inequality constraint (4.13b), the reformulated switching system (4.13c) and the sign condition (4.13d) such that

$$0 = a^\top + \delta^\top A + \nu^\top D - \lambda^\top \tilde{Z} \quad \text{and} \tag{4.14}$$

$$\begin{aligned}
0 = {} & b^\top \left(|\Sigma| + |\Gamma|\right) + \delta^\top \left(B\left(|\Sigma| + |\Gamma|\right) + C\left(\Sigma + \Gamma\right)\right) \\
& + \nu^\top \left(E\left(|\Sigma| + |\Gamma|\right) + F\left(\Sigma + \Gamma\right)\right) + \lambda^\top \left(I_s - \tilde{L}\Gamma\right) - \mu^\top P_\alpha \Gamma \;.
\end{aligned} \tag{4.15}$$

Multiplying the last equation from the right by $|\Sigma|\tilde{Z}$ and using the identity $\Sigma = \Sigma|\Sigma|$ as well as the matrix equations (3.15) and the notation introduced in Eq. (4.8), yields

$$\begin{aligned}
0 &= b^\top |\Sigma| \tilde{Z} + \delta^\top \left(B|\Sigma| + C\Sigma\right) \tilde{Z} + \nu^\top \left(E|\Sigma| + F\Sigma\right) \tilde{Z} + \lambda^\top |\Sigma| \tilde{Z} \\
&= b^\top |\Sigma| \tilde{Z} + \delta^\top \mathcal{J}_{\sigma,G,z} \tilde{Z} + \nu^\top \mathcal{J}_{\sigma,H,z} \tilde{Z} + \lambda^\top |\Sigma| \tilde{Z} \;.
\end{aligned} \tag{4.16}$$

Now, by adding Eq. (4.16) to Eq. (4.14) and using the decomposition of the identity given by Eq. (3.21) and the notation for the Jacobian given by Eq. (4.9) one obtains

$$0 = a^\top + b^\top |\Sigma| \tilde{Z} + \delta^\top \mathcal{J}_{\sigma,G} + \nu^\top \mathcal{J}_{\sigma,H} - \lambda^\top P_\alpha^\top P_\alpha \tilde{Z} \;.$$

Hence, it follows that the Lagrange multipliers $\delta \in \mathbb{R}^m$, $\nu \in \mathbb{R}^p$ and $\lambda \in \mathbb{R}^s$ fulfill Eq. (4.10) with $\check{\lambda} = -P_\alpha \lambda$. Due to the kink qualification LIKQ, one also has that the vectors $\delta \in \mathbb{R}^m$ as well as the components $P_{\mathcal{I}}\nu \in \mathbb{R}^{|\mathcal{I}|}$ and $P_\alpha \lambda \in \mathbb{R}^{|\alpha|}$ are determined uniquely. The remaining components of $\nu \in \mathbb{R}^p$ can be set to zero and those of $\lambda \in \mathbb{R}^s$ can be obtained by multiplying Eq. (4.15), this time only with $|\Sigma|$ from the right and using Eq. (3.15), yielding

$$0 = b^\top |\Sigma| + \delta^\top \mathcal{J}_{\sigma,G,z} + \nu^\top \mathcal{J}_{\sigma,H,z} + \lambda^\top |\Sigma|$$

and thus Eq. (4.11).

To derive the third condition (4.12) again multiply Eq. (4.15) from the right, this time with $\Gamma P_\alpha^\top$ and use the identities regarding to the projection matrix P_α given by Eq. (3.22) to obtain

$$\begin{aligned} 0 = b^\top \Gamma P_\alpha^\top + \delta^\top \left(B\Gamma P_\alpha^\top + C P_\alpha^\top \right) + \nu^\top \left(E\Gamma P_\alpha^\top + F P_\alpha^\top \right) \\ + \lambda^\top \left(\Gamma P_\alpha^\top - \tilde{L} P_\alpha^\top \right) - \mu^\top \ . \end{aligned} \tag{4.17}$$

Reformulating this equation and using $\mu \geq 0$ yields

$$\begin{aligned} -\left(b^\top + \delta^\top B + \nu^\top E + \lambda \right) \Gamma P_\alpha^\top &= \left(\delta^\top C + \nu^\top F - \lambda^\top \tilde{L} \right) P_\alpha^\top - \nu^\top \\ &\leq \left(\delta^\top C + \nu^\top F - \lambda^\top \tilde{L} \right) P_\alpha^\top \ . \end{aligned}$$

As in the unconstrained case, the key observation is that this condition is linear in Γ and is strongest for the choice $\gamma_i = -\operatorname{sgn}\left(b^\top + \delta^\top B + \nu^\top E + \lambda^\top \right)_i$ for $i \in \alpha$ yielding the inequality

$$\left| P_\alpha \left(b + B^\top \delta + E^\top \nu + \lambda \right) \right| \leq P_\alpha \left(C^\top \delta + F^\top \nu - \tilde{L}^\top \lambda \right)$$

and thus Eq. (4.12) which completed the necessary optimality conditions.

Now consider the sufficient optimality condition. For this, consider again Eq. (4.15) and multiply this from the right by $\Gamma P_\alpha^\top$. Using Eq. (3.22), Eq. (3.15) and Eq. (4.12) one obtains

$$\begin{aligned} \mu^\top &= \left(b^\top + \delta^\top \left(B|\Gamma| + C\Gamma \right) + \nu^\top \left(E|\Gamma| + F\Gamma \right) \right) \Gamma P_\alpha^\top + \lambda^\top \left(I_s - \tilde{L}\Gamma \right) \Gamma P_\alpha^\top \\ &= \left(b^\top + \delta^\top B + \nu^\top E + \lambda^\top \right) \Gamma P_\alpha^\top + \left(\delta^\top C + \nu^\top F - \lambda^\top \tilde{L} \right) P_\alpha^\top \geq 0 \end{aligned} \tag{4.18}$$

and therefore feasibility. To show Eq. (4.14) use Eq. (3.21), Eq. (4.11) multiplied from the right by $\tilde{Z}$ and Eq. (4.10) together with Eq. (4.9) to obtain

$$\begin{aligned}\lambda^\top \tilde{Z} &= \lambda^\top \left(|\Sigma| + P_\alpha^\top P_\alpha\right) \tilde{Z} \\ &= \lambda^\top |\Sigma| \tilde{Z} + \lambda^\top P_\alpha^\top P_\alpha \tilde{Z} \\ &= -b^\top |\Sigma| \tilde{Z} - \delta^\top \mathcal{J}_{\sigma,G,z} \tilde{Z} - \nu^\top \mathcal{J}_{\sigma,H,z} \tilde{Z} + a^\top + b^\top |\Sigma| \tilde{Z} + \delta^\top \mathcal{J}_{\sigma,G} + \nu^\top \mathcal{J}_{\sigma,H} \\ &= a^\top + \delta^\top A + \nu^\top D \,.\end{aligned}$$

Now only Eq. (4.15) remains to be shown. Therefore, using Eq. (3.21), Eq. (4.15) holds true if and only if

$$\begin{aligned}0 = {} & b^\top \left(|\Sigma| + |\Gamma|\right) + \delta^\top \left(B\left(|\Sigma| + |\Gamma|\right) + C\left(\Sigma + \Gamma\right)\right) \\ & + \nu^\top \left(E\left(|\Sigma| + |\Gamma|\right) + F\left(\Sigma + \Gamma\right)\right) + \lambda^\top \left(|\Sigma| + P_\alpha^\top P_\alpha - \tilde{L}\Gamma\right) - \mu^\top P_\alpha \Gamma\end{aligned}$$

holds true. Using Eq. (4.11) and Eq. (4.8) the last equation is equivalent to

$$0 = b^\top |\Gamma| + \delta^\top \left(B|\Gamma| + C\Gamma\right) + \nu^\top \left(E|\Gamma| + F\Gamma\right) + \lambda^\top \left(P_\alpha^\top P_\alpha - \tilde{L}\Gamma\right) - \mu^\top P_\alpha \Gamma \,.$$

Again, multiplying the last equation from the right by $\Gamma P_\alpha^\top$ and using Eq. (3.22) yields

$$\mu^\top = -\lambda^\top \tilde{L} P_\alpha^\top + \left(b^\top + \delta^\top B + \nu^\top E + \lambda^\top\right) \Gamma P_\alpha^\top + \left(\delta^\top C + \nu^\top F\right) P_\alpha^\top \,.$$

Now, by defining the Lagrange multiplier μ as above, it satisfies Eq. (4.15). To see this, insert μ into Eq. (4.15) and use Eq. (3.22) to obtain

$$\begin{aligned}0 = {} & b^\top \left(|\Sigma| + |\Gamma|\right) + \delta^\top \left(B\left(|\Sigma| + |\Gamma|\right) + C\left(\Sigma + \Gamma\right)\right) \\ & + \nu^\top \left(E\left(|\Sigma| + |\Gamma|\right) + F\left(\Sigma + \Gamma\right)\right) + \lambda^\top \left(I_s - \tilde{L}\Gamma\right) \\ & - \left(-\lambda^\top \tilde{L} P_\alpha^\top + \left(b^\top + \delta^\top B + \nu^\top E + \lambda^\top\right) \Gamma P_\alpha^\top + \left(\delta^\top C + \nu^\top F\right) P_\alpha^\top\right) P_\alpha \Gamma \\ = {} & b^\top \left(|\Sigma| + |\Gamma|\right) + \delta^\top \left(B\left(|\Sigma| + |\Gamma|\right) + C\left(\Sigma + \Gamma\right)\right) \\ & + \nu^\top \left(E\left(|\Sigma| + |\Gamma|\right) + F\left(\Sigma + \Gamma\right)\right) + \lambda^\top \left(I_s - \tilde{L}\Gamma\right) \\ & + \lambda^\top \tilde{L}\Gamma - \left(b^\top |\Gamma| + \delta^\top B|\Gamma| + \nu^\top E|\Gamma| + \lambda^\top |\Gamma| + \delta^\top C\Gamma + \nu^\top F\Gamma\right) \\ = {} & b^\top |\Sigma| + \delta^\top \left(B|\Sigma| + C\Sigma\right) + \nu^\top \left(E|\Sigma| + F\Sigma\right) + \lambda^\top I_s - \lambda^\top |\Gamma|\end{aligned}$$

Using Eq. (4.11) and Eq. (3.21) it follows that

$$0 = -\lambda^\top |\Sigma| + \lambda^\top |\Sigma| + \lambda^\top P_\alpha^\top P_\alpha - \lambda^\top |\Gamma|$$

must be true. Using Eq. (3.22), one sees that this is a true statement and thus completes the proof.

For the uniqueness of the Lagrange multipliers, see again the first paragraph of this proof. There it was stated that the Lagrange multipliers δ and ν as well as the components λ_i belonging to the index set $\alpha(x_\sigma)$ are unique. Finally, for the remaining $i \in \alpha^C$, the complement of α, the components λ_i can be uniquely determined by Eq. (4.11). □

This theorem shows that the optimality conditions for the unconstrained case (cf. Theorem 3.14) can be extended to the constrained case. It is still possible to test optimality by simple matrix-vector multiplications in polynomial time. Moreover, as expected, one sees that Theorem 4.8 coincides with Theorem 3.14 when the constraints are removed. Again, the conditions given in Eqs. (4.10) and (4.11) can be called *tangential stationarity* condition and Eq. (4.12) can be called *normal growth* condition.

4.2 Constrained Active Signature Method

After showing that the optimality of a point x for an optimization problem of the form (CALOP) can be tested in polynomial time, the goal now is to present a corresponding algorithm that can solve such optimization problems. As mentioned above, this algorithm has already been published in [70], and essential parts are adopted, as well as some additional details are added here.

In very simplified terms, the algorithm, called Constrained Active Signature Method (CASM), is a suitable combination of ASM and the Active Set Method. Analogous to ASM, the $\mathbb{R}^n$ is decomposed into different polyhedra, defined by the signature vector, and a penalized version of the objective function is optimized over them. The inequality constraints are handled as in the Active Set Method and are therefore considered as active or inactive. Thus, as before, the information about the nonsmoothness, now also that of the constraints, is explicitly exploited by the switching vector $z = z(x)$. Note, that on the one hand, on a polyhedron all constraints are linear, because if they would cause a nonsmoothness, a new polyhedron would be located there. On the other hand, since no further assumptions, such as compactness or convexity, are made on the feasible set, it is still possible that the polyhedra themselves can be unbounded even when adding constraints. Thus, it is still necessary to add the regularization term in order to guarantee the existence of a minimizer. Therefore, consider the optimization problem (4.5) with the addition of this regularization term, again with a positive definite matrix $Q = Q^\top \in \mathbb{R}^{n\times n}$, to obtain the following opitmization problem

$$
\begin{aligned}
\min_{x\in\mathbb{R}^n, z\in\mathbb{R}^s} \quad & a^\top x + b^\top |\Sigma| z + \frac{1}{2} x^\top Q x && \text{(4.19a)}\\
\text{s.t.} \quad & 0 = g + Ax + B|\Sigma|z + C\Sigma z \;, && \text{(4.19b)}\\
& 0 \geq h + Dx + E|\Sigma|z + F\Sigma z \;, && \text{(4.19c)}\\
& 0 = |\Sigma|z - \tilde{c} - \tilde{Z}x \;, && \text{(4.19d)}\\
& 0 \leq \Sigma z \;, && \text{(4.19e)}
\end{aligned}
$$

with $\tilde{Z}$ and $\tilde{c}$ defined as in Eq. (3.9). Due to the fixed signature vector, this is still a quadratic optimization problem with linear constraints.

Based on the optimization problem (4.19), CASM has three main components: First, the computation of a search direction, second, the step size in this direction, and finally, the checking of optimality and, in case of nonoptimality, the use of these conditions to drop active inequality constraints by updating ω or switching from one polyhedron to the next and thus updating σ.

4.2.1 Computing a Descent Direction for Given σ and ω

To determine a descent direction, the KKT-theory is applied to the optimization problem (4.19). With Lagrange multipliers $\delta \in \mathbb{R}^m$, $\nu \in \mathbb{R}^p$, $\lambda \in \mathbb{R}^n$ and $\mu \in \mathbb{R}^s$ associated with the equality constraint (4.19b), inequality constraint (4.19c), the switching equation (4.19d) and the sign condition (4.19e), this provides the following necessary optimality conditions:

$$
\begin{aligned}
& 0 = a^\top + x^\top Q + \delta^\top A + \nu^\top D - \lambda^\top \tilde{Z} \;, && \text{(4.20)}\\
& 0 = b^\top |\Sigma| + \delta^\top \mathcal{J}_{\sigma,G,z} + \nu^\top \mathcal{J}_{\sigma,H,z} + \lambda^\top |\Sigma| - \mu^\top \Sigma \;, && \text{(4.21)}\\
& 0 = g + Ax + B|\Sigma|z + C\Sigma z \;, \\
& 0 \geq h + Dx + E|\Sigma|z + F\Sigma z \;, \\
& 0 = |\Sigma|z - \tilde{c} - \tilde{Z}x \;, \\
& 0 \leq \Sigma z \;, \quad 0 \leq \mu \;, \quad 0 = \mu^\top \Sigma z \;, && \text{(4.22)}\\
& 0 \leq \nu \;, \qquad 0 = \nu^\top \left(h + Dx + E|\Sigma|z + F\Sigma z\right) \;. && \text{(4.23)}
\end{aligned}
$$

Now, the proceeding is quite similar to the unconstrained case in Section 3.3.1: Multiplying Eq. (4.21) from the right by Σ and using Eq. (4.22) yields

$$0 \le \mu^\top |\Sigma| = b^\top \Sigma + \delta^\top \mathcal{J}_{\sigma,G,z}\Sigma + \nu^\top \mathcal{J}_{\sigma,H,z}\Sigma + \lambda^\top \Sigma \,. \tag{4.24}$$

If $\sigma = \operatorname{sgn}(z)$ so that the underlying x is stationary, Eq. (4.24) reduces, because of the required complementarity $\mu^\top |\Sigma| z = 0$, to

$$0 = b^\top \Sigma + \delta^\top \mathcal{J}_{\sigma,G,z}\Sigma + \nu^\top \mathcal{J}_{\sigma,H,z}\Sigma + \lambda^\top \Sigma \,.$$

Hence, with $\omega = \operatorname{sgn}(H(x,z,|z|))$ and $\Omega = \operatorname{diag}(\omega)$ denoting the projection onto the inactive inequality constraints, if the feasible signature optimal point x_σ exists, it must satisfy the system of linear equation

$$\begin{bmatrix} Q & 0 & -\tilde{Z}^\top & A^\top & D^\top \\ 0 & 0 & \Sigma & \Sigma\mathcal{J}_{\sigma,G,z}^\top & \Sigma\mathcal{J}_{\sigma,H,z}^\top \\ -\tilde{Z} & |\Sigma| & 0 & 0 & 0 \\ A & \mathcal{J}_{\sigma,G,z} & 0 & 0 & 0 \\ \overline{\Omega} D & \overline{\Omega}\mathcal{J}_{\sigma,H,z} & 0 & 0 & \Omega \end{bmatrix} \begin{bmatrix} x \\ z \\ \lambda \\ \delta \\ \nu \end{bmatrix} = - \begin{bmatrix} a \\ \Sigma b \\ -\tilde{c} \\ g \\ \overline{\Omega} h \end{bmatrix} ,$$

where $\overline{\Omega} = I_p - |\Omega| = I_p + \Omega$ forces the inactive inequalities to vanish. The matrix Ω in the lower right corner ensures that ν is zero for the inactive inequality constraints. For this reason, one can also multiply the rightmost blocks in the first and second rows from the right with $\overline{\Omega}$ and the second row from the left with Σ to make the resulting Jacobian fully symmetric and obtain

$$\begin{bmatrix} Q & 0 & -\tilde{Z}^\top & A^\top & D^\top\overline{\Omega} \\ 0 & 0 & |\Sigma| & \mathcal{J}_{\sigma,G,z}^\top & \mathcal{J}_{\sigma,H,z}^\top\overline{\Omega} \\ -\tilde{Z} & |\Sigma| & 0 & 0 & 0 \\ A & \mathcal{J}_{\sigma,G,z} & 0 & 0 & 0 \\ \overline{\Omega} D & \overline{\Omega}\mathcal{J}_{\sigma,H,z} & 0 & 0 & \Omega \end{bmatrix} \begin{bmatrix} x \\ z \\ \lambda \\ \delta \\ \nu \end{bmatrix} = - \begin{bmatrix} a \\ |\Sigma| b \\ -\tilde{c} \\ g \\ \overline{\Omega} h \end{bmatrix} . \tag{4.25}$$

In comparison to Lemma 3.15, a similar result can be shown for the constrained case, i.e., it is possible to reduce this saddle point system to determine a solution with less effort.

Lemma 4.9 (Partitioned solution (constrained case))**.** *The solution of the equation sys-*

tem (4.25) can be reduced to solving a symmetric semidefinite linear system

$$\begin{bmatrix} -|\overline{\Sigma}|\tilde{Z}Q^{-1}\tilde{Z}^\top|\overline{\Sigma}| & |\overline{\Sigma}|\tilde{Z}Q^{-1}\mathcal{J}_{\sigma,G}^\top & |\overline{\Sigma}|\tilde{Z}Q^{-1}\mathcal{J}_{\sigma,H}^\top\overline{\Omega} \\ \mathcal{J}_{\sigma,G}Q^{-1}\tilde{Z}^\top|\overline{\Sigma}| & -\mathcal{J}_{\sigma,G}Q^{-1}\mathcal{J}_{\sigma,G}^\top & -\mathcal{J}_{\sigma,G}Q^{-1}\mathcal{J}_{\sigma,H}^\top\overline{\Omega} \\ \overline{\Omega}\mathcal{J}_{\sigma,H}Q^{-1}\tilde{Z}^\top|\overline{\Sigma}| & -\overline{\Omega}\mathcal{J}_{\sigma,H}Q^{-1}\mathcal{J}_{\sigma,G}^\top & -\overline{\Omega}\mathcal{J}_{\sigma,H}Q^{-1}\mathcal{J}_{\sigma,H}^\top\overline{\Omega} \end{bmatrix} \begin{bmatrix} \lambda \\ \delta \\ \nu \end{bmatrix} = \begin{bmatrix} |\overline{\Sigma}|\hat{c} \\ \hat{g} \\ \overline{\Omega}\hat{h} \end{bmatrix} , \quad (4.26)$$

with

$$\begin{aligned} \hat{c} &= \tilde{c} - \tilde{Z}Q^{-1}\left[a + \tilde{Z}^\top|\Sigma|b\right] , \\ \hat{g} &= \mathcal{J}_{\sigma,G}Q^{-1}\left[a + \tilde{Z}^\top|\Sigma|b\right] - \mathcal{J}_{\sigma,G,z}\tilde{c} - g , \\ \hat{h} &= \mathcal{J}_{\sigma,H}Q^{-1}\left[a + \tilde{Z}^\top|\Sigma|b\right] - \mathcal{J}_{\sigma,H,z}\tilde{c} - h , \end{aligned}$$

for the nontrivial entries of $|\overline{\Sigma}|\lambda$ and $\overline{\Omega}\nu$, where $|\overline{\Sigma}| \equiv I_s - |\Sigma|$ denotes the complementary orthogonal projection to $|\Sigma|$. The system is uniquely solvable exactly when the rows of the $(s+m+p) \times n$ matrix

$$\begin{bmatrix} |\overline{\Sigma}|\tilde{Z} \\ A + B|\overline{\Sigma}|\tilde{Z} + C\Sigma\tilde{Z} \\ \overline{\Omega}\left(D + E|\overline{\Sigma}|\tilde{Z} + F\Sigma\tilde{Z}\right) \end{bmatrix} Q^{-1}\tilde{Z}^\top$$

that correspond to active kinks and active constraints are linearly independent, i.e., LIKQ holds. For $(\tilde{\lambda}, \delta, \tilde{\nu})$ as a solution of the reduced system (4.26) the dual and primal variables are then easily obtained as

$$\begin{aligned} \lambda &= |\overline{\Sigma}|\tilde{\lambda} - \left[|\Sigma|b + \mathcal{J}_{\sigma,G,z}^\top\delta + \mathcal{J}_{\sigma,H,z}^\top\overline{\Omega}\nu\right] , \\ \nu &= \overline{\Omega}\tilde{\nu} , \\ x &= Q^{-1}\left(\tilde{Z}^\top\lambda - a - A^\top\delta - D^\top\overline{\Omega}\nu\right) , \\ z &= \tilde{Z}x + \tilde{c} . \end{aligned}$$

Proof. If λ, δ and ν are given, again the expressions for x can be read off directly from the original system (4.25) as well as the component of z which belongs to the inactive kinks.

Similar to the unconstrained setting, the main task consists in calculating the Lagrange multipliers λ, δ and ν. Therefore, the second row as a projection on $|\Sigma|$ gives

$$|\Sigma|\lambda = -\left[|\Sigma|b + \mathcal{J}_{\sigma,G,z}^\top\delta + \mathcal{J}_{\sigma,H,z}^\top\overline{\Omega}\nu\right] , \quad (4.27)$$

which makes it again easy to calculate the components belonging to the inactive kinks when δ and ν are given. In the same manner the projection on Ω gives $\Omega\nu = 0$, which means setting the components belonging to the inactive inequality constraints to zero. Using Q to eliminate the three last blocks in the first column and since $\Omega\nu = 0$ ignoring the matrix in the right lower corner, yields

$$\begin{bmatrix} Q & 0 & -\tilde{Z}^\top & A^\top & D^\top\overline{\Omega} \\ 0 & 0 & |\Sigma| & \mathcal{J}_{\sigma,G,z}^\top & \mathcal{J}_{\sigma,H,z}^\top\overline{\Omega} \\ 0 & |\Sigma| & -\tilde{Z}Q^{-1}\tilde{Z}^\top & \tilde{Z}Q^{-1}A^\top & \tilde{Z}Q^{-1}D^\top\overline{\Omega} \\ 0 & \mathcal{J}_{\sigma,G,z} & AQ^{-1}\tilde{Z}^\top & -AQ^{-1}A^\top & -AQ^{-1}D^\top\overline{\Omega} \\ 0 & \overline{\Omega}\mathcal{J}_{\sigma,H,z} & \overline{\Omega}DQ^{-1}\tilde{Z}^\top & -\overline{\Omega}DQ^{-1}A^\top & -\overline{\Omega}DQ^{-1}D^\top\overline{\Omega} \end{bmatrix} \begin{bmatrix} x \\ z \\ \lambda \\ \delta \\ \nu \end{bmatrix} = - \begin{bmatrix} a \\ |\Sigma|b \\ \tilde{\tilde{c}} \\ \tilde{g} \\ \overline{\Omega}\tilde{h} \end{bmatrix},$$

with $\tilde{\tilde{c}} = \tilde{Z}Q^{-1}a - \tilde{c}$, $\tilde{g} = g - AQ^{-1}a$ and $\tilde{h} = h - DQ^{-1}a$.

Next, considering the projection of the third row through $-|\overline{\Sigma}|$ by using $|\overline{\Sigma}||\Sigma| = 0$. Moreover, adding the third row multiplied by $-\mathcal{J}_{\sigma,G,z}$ from the left-hand side to the fourth row and the third row multiplied by $-\overline{\Omega}\mathcal{J}_{\sigma,H,z}$ from the left-hand side to the fifth row yields the three equations

$$\begin{aligned} |\overline{\Sigma}|\tilde{Z}Q^{-1}\tilde{Z}^\top\lambda - |\overline{\Sigma}|\tilde{Z}Q^{-1}A^\top\delta - |\overline{\Sigma}|\tilde{Z}Q^{-1}D^\top\overline{\Omega}\nu = |\overline{\Sigma}|\tilde{\tilde{c}}\,, \\ \Big(\mathcal{J}_{\sigma,G,z}\tilde{Z}Q^{-1}\tilde{Z}^\top + AQ^{-1}\tilde{Z}^\top\Big)\lambda - \Big(\mathcal{J}_{\sigma,G,z}\tilde{Z}Q^{-1}A^\top + AQ^{-1}A^\top\Big)\delta \\ - \Big(\mathcal{J}_{\sigma,G,z}\tilde{Z}Q^{-1}D^\top\overline{\Omega} + AQ^{-1}D^\top\overline{\Omega}\Big)\nu = \tilde{\tilde{g}}\,, \\ \overline{\Omega}\Big(\mathcal{J}_{\sigma,H,z}\tilde{Z}Q^{-1}\tilde{Z}^\top + DQ^{-1}\tilde{Z}^\top\Big)\lambda - \overline{\Omega}\Big(\mathcal{J}_{\sigma,H,z}\tilde{Z}Q^{-1}A^\top + DQ^{-1}A^\top\Big)\delta \\ -\overline{\Omega}\Big(\mathcal{J}_{\sigma,H,z}\tilde{Z}Q^{-1}D^\top\overline{\Omega} + DQ^{-1}D^\top\overline{\Omega}\Big)\nu = \overline{\Omega}\tilde{\tilde{h}}\,, \end{aligned}$$

with $\tilde{\tilde{g}} = \mathcal{J}_{\sigma,G,z}\tilde{\tilde{c}} - \tilde{g}$ and $\tilde{\tilde{h}} = \mathcal{J}_{\sigma,H,z}\tilde{\tilde{c}} - \tilde{h}$. To simplify the expressions, use Eq. (4.9) to obtain

$$|\overline{\Sigma}|\tilde{Z}Q^{-1}\tilde{Z}^\top\lambda - |\overline{\Sigma}|\tilde{Z}Q^{-1}A^\top\delta - |\overline{\Sigma}|\tilde{Z}Q^{-1}D^\top\overline{\Omega}\nu = |\overline{\Sigma}|\tilde{\tilde{c}}\,, \tag{4.28}$$

$$\mathcal{J}_{\sigma,G}Q^{-1}\tilde{Z}^\top\lambda - \mathcal{J}_{\sigma,G}Q^{-1}A^\top\delta - \mathcal{J}_{\sigma,G}Q^{-1}D^\top\overline{\Omega}\nu = \tilde{\tilde{g}}\,, \tag{4.29}$$

$$\overline{\Omega}\mathcal{J}_{\sigma,H}Q^{-1}\tilde{Z}^\top\lambda - \overline{\Omega}\mathcal{J}_{\sigma,H}Q^{-1}A^\top\delta - \overline{\Omega}\mathcal{J}_{\sigma,H}Q^{-1}D^\top\overline{\Omega}\nu = \overline{\Omega}\tilde{\tilde{h}}\,, \tag{4.30}$$

Now, using $\lambda = \big(|\Sigma| + |\overline{\Sigma}|\big)\,\lambda = |\Sigma|\lambda + |\overline{\Sigma}|\lambda$ and substituting Eq. (4.27) in Eqs. (4.28), (4.29) and (4.30) as well as some reformulations yields the reduced symmetric semidefinite linear system

$$\begin{bmatrix} -|\overline{\Sigma}|\tilde{Z}Q^{-1}\tilde{Z}^\top|\overline{\Sigma}| & |\overline{\Sigma}|\tilde{Z}Q^{-1}\mathcal{J}_{\sigma,G}^\top & |\overline{\Sigma}|\tilde{Z}Q^{-1}\mathcal{J}_{\sigma,H}^\top\overline{\Omega} \\ \mathcal{J}_{\sigma,G}Q^{-1}\tilde{Z}^\top|\overline{\Sigma}| & -\mathcal{J}_{\sigma,G}Q^{-1}\mathcal{J}_{\sigma,G}^\top & -\mathcal{J}_{\sigma,G}Q^{-1}\mathcal{J}_{\sigma,H}^\top\overline{\Omega} \\ \overline{\Omega}\mathcal{J}_{\sigma,H}Q^{-1}\tilde{Z}^\top|\overline{\Sigma}| & -\overline{\Omega}\mathcal{J}_{\sigma,H}Q^{-1}\mathcal{J}_{\sigma,G}^\top & -\overline{\Omega}\mathcal{J}_{\sigma,H}Q^{-1}\mathcal{J}_{\sigma,H}^\top\overline{\Omega} \end{bmatrix} \begin{bmatrix} \lambda \\ \delta \\ \nu \end{bmatrix} = \begin{bmatrix} |\overline{\Sigma}|\hat{c} \\ \hat{g} \\ \overline{\Omega}\hat{h} \end{bmatrix},$$

with

$$\hat{c} = -\tilde{\tilde{c}} - \tilde{Z}Q^{-1}\tilde{Z}^\top|\Sigma|b\,, \quad \hat{g} = \tilde{\tilde{g}} + \mathcal{J}_{\sigma,G}Q^{-1}\tilde{Z}^\top|\Sigma|b \quad \text{and} \quad \hat{h} = \tilde{\tilde{h}} + \mathcal{J}_{\sigma,H}Q^{-1}\tilde{Z}^\top|\Sigma|b\,.$$

□

Numerically and similar to Lemma 3.15, the solutions obtained by Lemma 4.9 must yield $|\overline{\Sigma}|z = 0$ and $\overline{\Omega}\nu = 0$ up to rounding. After that has been checked the nonbasic components of z can be set exactly to zero, while λ is generally dense without any sign constraints.

Comparing the effort to solve the two systems of equations, it can be seen that the system matrices in both Eq. (4.25) and Eq. (4.26) are symmetric and quadratic. The matrix from Eq. (4.25) has a total of $n + 2s + m + p$ rows and columns, but it also contains some zero blocks, which makes the matrix potentially quite sparse, depending on the respective dimensions n, s, m and p. In comparison, the matrix from Eq. (4.26) has a total of only $s + m + p$ rows and columns, which makes the size completely independent of the dimension n. However, this matrix no longer contains any zero blocks, which makes it potentially more densely filled, depending on the nature of the individual matrices given by the optimization problem (CALOP) itself. In summary, it is not possible to determine which of the two systems of equations is more efficient to solve. This may depend on the specific application, or may be numerically influenced by a well-chosen solver.

In the following $(\hat{x}, \hat{z}, \lambda, \delta, \nu)$ denotes the solution of the saddle point system (4.25). This system is still in principle similar to a classical Newton step. Thus, the search direction from a current point (x, z) in direction $(\hat{x}, \hat{z})$ will be again denoted with

$$\Delta x = \hat{x} - x \quad \text{and} \quad \Delta z = \hat{z} - z\,. \tag{4.31}$$

4.2.2 Computing a Step Size β

As before, when a search direction has been calculated, the question is how far to go in that direction. The first option is still the length until a kink is reached. With $(\hat{x}, \hat{z})$ as part of the solution of the system of equations (4.25), the calculation for this is unchanged

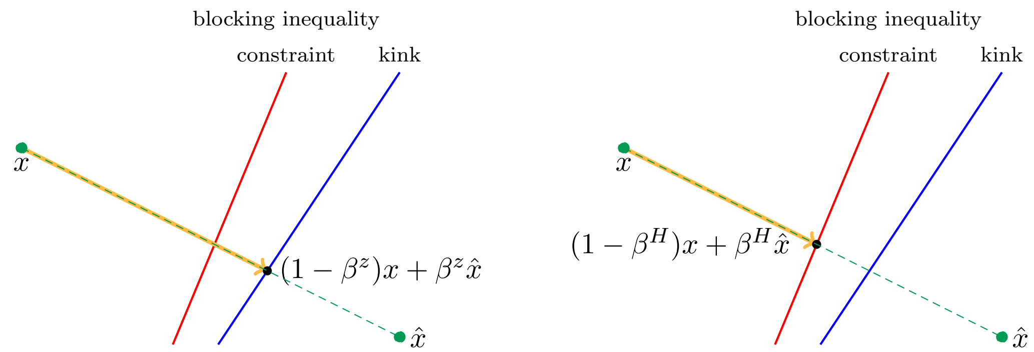

Figure 4.2: Illustration of the two different step sizes β^z and β^H

compared to the unconstrained case and with β^z as given in Eq. (3.32). Furthermore, let j^z denote the index for which the minimum was first assumed in Eq. (3.32).

Now, by adding inequality constraints, it can also happen that they block the search direction. In this case, the idea is to calculate the step size in such a way that one ends up exactly on the first blocking constraint, as it is also done in the Active Set Method, see [87] or Subsection 2.2.2. For this, in a similar way as β^z, the step size β^H related to the constraints can be calculated as

$$\beta^H = \inf_{1 \leq l \leq p} \left\{ \beta_l^H \equiv \frac{H_l}{H_l - \hat{H}_l} \middle| (\hat{H}_l - H_l)\omega_l < 0 \right\} , \tag{4.32}$$

where $H \equiv H(x, z, \Sigma z)$, $\hat{H} \equiv H(\hat{x}, \hat{z}, \Sigma \hat{z})$ and l denote the lth component of H and $\hat{H}$, respectively. Similar to the first step size, denote by j^H the smallest index for which the minimum is attained. A graphical illustration of both step sizes is shown in Figure 4.2, where on the left-hand side β^z is sketched and on the right-hand side β^H. The blue lines represent a kink, the red line an inequality constraint that becomes active and the yellow arrow indicates the corresponding step size.

In order to bound the step, the current step size β is set to

$$\beta = \min\{\beta^z, \beta^H, 1\} . \tag{4.33}$$

Here, the upper bound 1 on β ensures with the current iterate

$$x^+ = (1 - \beta)x + \beta\hat{x}$$

that the next iterate is still in $\mathcal{F}_\sigma$. In the case of a full step, i.e., $\beta = 1$, obviously $x^+ = \hat{x}$

and $\sigma(x^+) = \hat{x}$ holds. Since the vector is therefore kept, the point x^+ is also called feasible signature stationary.

Furthermore, by determining β, it can be seen that the case is chosen that occurs first. I.e., if an inequality constraint is reached first, the step is performed only up to the constraint, and similarly, if a kink is reached first, the step is performed only up to the kink. If the step sizes β^z and β^H are equal, then both, a kink and an inequality constraint can also be added.

Given the step size β, the next question remains how to update σ, but now also how to update ω, i.e., on which feasible polyhedron to optimize next and which constraints to be considered. In case of adding a constraint, it follows directly from the considerations above. For this, the corresponding index j^H is used to update ω. The new one is set as $\omega^+ = \omega + e_{j^H}$ with the j^Hth unit vector, which means to set $\omega_{j^H} = 0$. Thus, the vector ω is restricted. Since all active constraints are encoded in ω, one must have $\beta^H > 0$. Activating kinks is unchanged as for ASM in Subsection 3.3.2. The index j^z is used to set $\sigma_{j^z} = 0$ and therefore restrict the signature vector σ.

4.2.3 Checking the Optimality

The case of leaving a constraint is relatively simple. If no feasible optimal point is found due to the fact that a constraint is still active but should be inactive, at least one component of ν must be negative. Therefore, $\nu \geq 0$ does not hold, and it is one possible heuristic to choose the component for which $\nu \geq 0$ is most violated and drop the corresponding constraint. Hence, the associated entry of ω is set to -1 to relax ω. This is the same procedure as in the Active Set Method described in Section 2.2.

Finally, there is the case of releasing an active kink, i.e., relaxing σ. This is the most complex case, but the basic idea has already been described in Subsection 3.3.3 and now needs to be extended under consideration of the constraints. In this case, a feasible signature optimal point is found on the current feasible polyhedron $\mathcal{F}_\sigma$, and it is necessary to check if it is already a minimizer of the original problem (CALOP) with the added regularization term. If such a minimizer is detected, the algorithm can terminate. Otherwise, the optimization must be continued on a neighboring polyhedron $\mathcal{P}_{\tilde{\sigma}}$ with definite $\tilde{\sigma} \prec \sigma$. As in the proofs of the Theorems 3.14 and 4.8, such a $\tilde{\sigma}$ can be decomposed into $\sigma + \gamma$ with $|\sigma|^\top |\gamma| = 0$. Then replacing Σ in the optimality conditions (4.20)-(4.23) by the corresponding $\Sigma + \Gamma$ and using Eq. (3.16) yields that the primal feasibility condition and the dual equality constraint are still satisfied by the solution $(\hat{x}, \hat{z}, \lambda, \delta, \nu)$. The only change is that Eq. (4.24) contains as many new nontrivial components as γ, which can be written as

$$\begin{aligned} 0 \le \mu^\top |\Gamma| &= b^\top \Gamma + \delta^\top (B\Gamma + C|\Gamma|) + \nu^\top (E\Gamma + F|\Gamma|) + \lambda^\top \left(I_s - \tilde{L}\Gamma\right)\Gamma \\ &= \left(b^\top + \delta^\top B + \nu^\top E + \lambda^\top\right)\Gamma + \left(\delta^\top C + \nu^\top F - \lambda^\top \tilde{L}\right)|\Gamma| \,, \end{aligned}$$

similar to Eq. (4.18). This optimality condition is violated if and only if there is at least one index such that $\gamma = -\operatorname{sgn}\left(b_k + \delta^\top B e_k + \nu^\top E e_k + \lambda_k\right) e_k$ satisfies

$$0 > \left(\delta^\top C + \nu^\top F - \lambda^\top \tilde{L}\right) e_k - \left|b^\top + \delta^\top B + \nu^\top E + \lambda^\top\right| e_k \quad \text{and} \quad \sigma_k = 0 \,, \tag{4.34}$$

which represents a violation of the normal growth condition. Thus, if the optimality condition is violated, i.e., Eq. (4.34) holds true for at least one k, one possible strategy is to choose the index k for which the right-hand side of Eq. (4.34) is minimal. This is a well-known heuristic, e.g., from Active Set Methods [87]. Then, by updating $\sigma_k = -\operatorname{sgn}\left(b_k + \delta^\top B e_k + \nu^\top E e_k \lambda_k\right)$, the next signature vector, denoted by σ^+, has one component less that equals zeros. This can be interpreted as releasing a kink, since one no longer insist on evaluating the corresponding absolute value at zero.

4.2.4 The Overall Algorithm

A pseudocode for the Constrained Active Signature Method (CASM) is given in Algorithm 4 and a schematic representation is given in Figure 4.3.

CASM combines now all these previous steps in an appropriate way. This means, given an optimization problem of the form (CALOP), a positive definite matrix $Q = Q^\top \in \mathbb{R}^{n\times n}$ and a feasible starting point $x \in \mathcal{F}$, the associated switching and signature vector as well as the signature vector for the active inequality constraints are determined and then the saddle point system (4.25) is solved (cf. line 2). Afterwards, in line 3 the step size β is computed as in Eq. (4.33). If now $\beta^H = \beta$ holds, i.e., $\beta^H \le 1$ and thus a blocking inequality constraint is reached, this one is added by restricting ω (cf. line 5). The same applies to β^z and the corresponding σ (cf. line 6), as already seen for ASM.

If $\beta = 1$, then optimality is checked. Note that by construction, at least one of the three *if*-conditions must be true in every iteration. The optimality is verified in line 7–14. If $\beta < 1$ the current iterate can not be feasible signature stationary. Hence, for x being a minimizer, it is necessary that $\beta = 1$ holds (cf. line 7). Then the optimality check is starting by checking the sign condition on the Lagrange multiplier ν. If there is at least one negative component of ν, it is required to drop a corresponding constraint (cf. line 9).

Algorithm 4 Constrained Active Signature Method (CASM)

Given: Optimization problem of the form (CALOP), $Q = Q^\top \in \mathbb{R}^{n\times n}$ positive definite and feasible start point $x \in \mathcal{F}$
Compute: $z := z(x)$ via Eq. (3.5), $\sigma := \sigma(x)$, $\omega := \omega(x)$

1: **loop**
2: Compute $(\hat{x}, \hat{z}, \lambda, \delta, \nu)$ by solving Eq. (4.25) ▷ Solve saddle point system
3: Compute β^z via Eq. (3.32), β^H via Eq. (4.32) and β via Eq. (4.33)
4: Set $(x^+, z^+) = (1-\beta)(x, z) + \beta(\hat{x}, \hat{z})$ ▷ Update iterate
5: **if** $\beta^H = \beta$ **then** Restrict ω ▷ Add constraint
6: **if** $\beta^z = \beta$ **then** Restrict σ ▷ Add kink
7: **if** $\beta = 1$ **then** ▷ x^+ is feasible signature stationary
8: **if** $\nu \not\geq 0$ **then**
9: Relax ω ▷ Drop constraint
10: **else** ▷ x^+ is feasible signature optimal
11: **if** Eq. (4.34) holds true **then**
12: Relax σ ▷ Drop kink
13: **else** ▷ x^+ is local optimal
14: **return** (x^+, z^+) ▷ Problem solved
15: Set $(x, z) = (x^+, z^+)$

Otherwise, when line 10 is reached, a feasible signature optimal point has been found and the last optimality condition must hold for local optimality. If this is not yet the case, a kink is dropped (cf. line 12), otherwise the algorithm terminates in line 14 with a local minimum. In case of nontermination the algorithm continues with the next iterate (cf. line 15).

4.2.5 Convergence Analysis of the Constrained Active Signature Method

Similar to ASM, now the goal is to show the convergence of CASM. In the paper [70], in which this method was proposed, the proof is already given and is taken over here. As for ASM, it is also possible to show for CASM that the algorithm terminates after a finite number of steps. Therefore, first the results of Lemma 3.17 have to be adapted to the constraint setting. Thus, consider the following problem

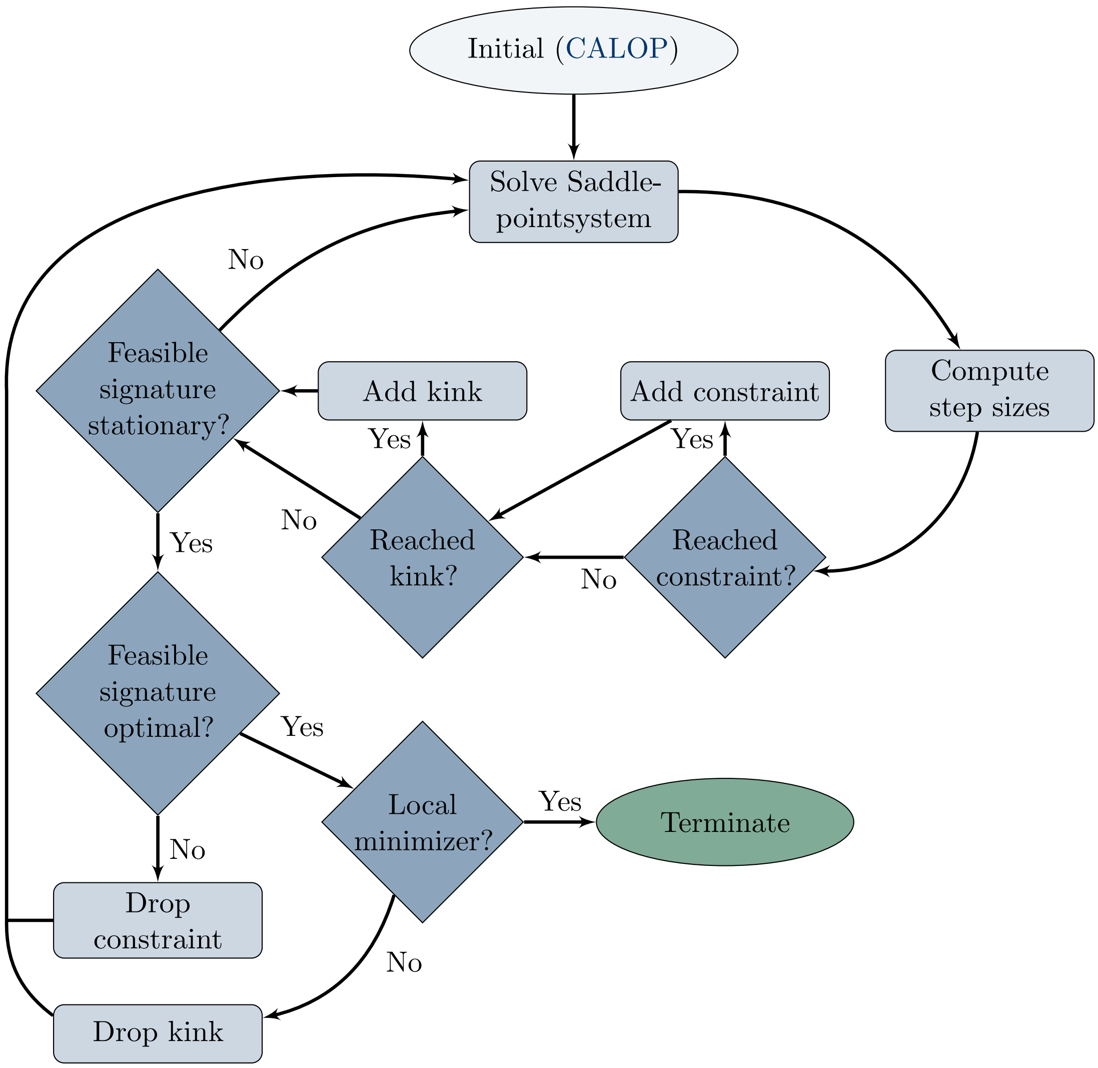

Figure 4.3: Scheme of Constrained Active Signature Method (CASM)

$$
\begin{aligned}
\min_{x\in\mathbb{R}^n, z\in\mathbb{R}^s} \quad & a^\top x + b^\top |\Sigma| z + \frac{1}{2} x^\top Q x \\
\text{s.t.} \quad & 0 = g + Ax + B|\Sigma|z + C\Sigma z \,, \\
& 0 = \overline{\Omega}\left(h + Dx + E|\Sigma|z + F\Sigma z\right) \,, \\
& 0 = |\Sigma|z - \tilde{c} - \tilde{Z}x \,,
\end{aligned}
\tag{4.35}
$$

which is exactly the one related to the saddle point system (4.25) ignoring the Ω in the right lower corner, and is identical to (4.19) if one omits the sign condition on the switching equation described by the last inequality condition of (4.19). In the same way as in Section 3.3.5 it is possible to derive the same objective function as in Eq. (3.35).

Now it is possible to show results regarding the descent direction in the nonoptimal case.

Lemma 4.10 (Descent direction in the nonoptimal case for CASM)**.** *Let $x \in \mathcal{F}_\sigma$ and the optimization problem (4.19) be given and suppose that LIKQ holds in x.*

1. *If tangential stationarity is violated, i.e.,*

$$0 \neq a^\top + b^\top |\Sigma| \tilde{Z} + x^\top Q + \delta^\top \mathcal{J}_{\sigma,G} + \nu^\top \mathcal{J}_{\sigma,H} - \lambda^\top P_\alpha^\top P_\alpha \tilde{Z} \,,$$

   there exists some direction $d \in \mathbb{R}^n$ such that $P_\alpha \tilde{Z} d = 0$, $\mathcal{J}_{\sigma,G} d = 0$ and $\overline{\Omega}\mathcal{J}_{\sigma,H} d = 0$ but

$$\left(a^\top + b^\top |\Sigma| \tilde{Z} + x^\top Q\right) d < 0 \,, \tag{4.36}$$

   and the target function defined in (3.35) is decreasing in direction τd i.e., $f(x+\tau d) < f(x)$ for $\tau \gtrsim 0$.

2. *If tangential stationarity holds but $\nu \geq 0$ fails, there exists at least one $i \in \mathcal{I}$ with $\nu_i < 0$. Defining $\upsilon = -e_i \in \mathbb{R}^p$ and $\Upsilon = \operatorname{diag}(\upsilon)$, any d satisfying*

$$P_\alpha \tilde{Z} d = 0 \,, \quad \mathcal{J}_{\sigma,G} d = 0 \quad \textit{and} \quad \left(\overline{\Omega} + \Upsilon\right) \mathcal{J}_{\sigma,H} d = \upsilon \tag{4.37}$$

   is a descent direction.

3. *If tangential stationarity and $\nu \geq 0$ holds but normal growth fails, there exists at least one $i \in \alpha$ with $|b_i + \delta^\top B e_i + \nu^\top E e_i + \lambda_i| > \left(\delta^\top C + \nu^\top F - \lambda^\top \tilde{L}\right) e_i$. Defining*

$\gamma = -\operatorname{sgn}(b_i + \delta^\top B e_i + \nu^\top E e_i + \lambda_i) e_i \in \mathbb{R}^s$, *any* d *satisfying*

$$Ad + \mathcal{J}_{\sigma+\gamma,G,z}\left(I_s - \tilde{L}\Gamma\right)^{-1}\tilde{Z}d = 0\ , \tag{4.38}$$

$$\overline{\Omega}\left(Dd + \mathcal{J}_{\sigma+\gamma,H,z}\left(I_s - \tilde{L}\Gamma\right)^{-1}\tilde{Z}d\right) = 0 \tag{4.39}$$

$$\text{and} \quad \left(I_s - \tilde{L}\Gamma\right)^{-1}\tilde{Z}d = \gamma \tag{4.40}$$

is a descent direction.

Proof. 1. Since tangential stationarity is violated, the point x is not a minimizer on the current feasible polyhedron $\mathcal{F}_{\sigma(x)}$ for the given index set of active inequality constraints $\mathcal{I}(x)$. Therefore, let x_σ be a minimizer on $\overline{\mathcal{F}}_{\sigma(x)}$ with the same index set, i.e., $\mathcal{I}(x) = \mathcal{I}(x_\sigma)$, and denote by $d \coloneqq x_\sigma - x$ a corresponding direction. Both points, x and x_σ lie on the same feasible extended signature domain, thus $\mathcal{J}_{\sigma,G}d = 0$ holds. Moreover, they have the same index sets of active switching variables α and thus $P_\alpha \tilde{Z}d = 0$ holds. In addition, they have the same index set of active inequality constraints $\mathcal{I}$ and thus, also $\overline{\Omega}\mathcal{J}_{\sigma,H}d = 0$ holds.

Next, it is to show, that $x + \tau d$ is still feasible. Therefore, it is to prove that $x + \tau d \in \mathcal{F}_\sigma$ for $\tau \gtrsim 0$. Using $P_\alpha \tilde{Z}d = 0$ the components of $z_\sigma(x + \tau d)$ are equal to zero and the others vary only slightly. Hence, the signature of $x + \tau d$ coincides with the one of x. For the feasibility consider the notation given in Eqs. (4.7) and (4.8) as well as $z_\sigma(x) = \tilde{c} + \tilde{Z}x$ to obtain for the equality constraint

$$g + Ax + B|\Sigma|z + C\Sigma z = g + Ax + \mathcal{J}_{\sigma,G,z}\left(\tilde{c} + \tilde{Z}x\right) = g + \mathcal{J}_{\sigma,G,z}\tilde{c} + \mathcal{J}_{\sigma,G}x$$

and analogously for the inequality constraint

$$\overline{\Omega}\left(h + Dx + E|\Sigma|z + F\Sigma z\right) = \overline{\Omega}\left(h + \mathcal{J}_{\sigma,H,z}\tilde{c} + \mathcal{J}_{\sigma,H}x\right)\ .$$

Thus, for obtaining the feasibility it is to show that the equality and the projection of the active inequality constraints are still zero. Therefore, using the assumptions regarding the direction d, one obtains

$$g + \mathcal{J}_{\sigma,G,z}\tilde{c} + \mathcal{J}_{\sigma,G}(x + \tau d) = g + \mathcal{J}_{\sigma,G,z}\tilde{c} + \mathcal{J}_{\sigma,G}x + \tau\mathcal{J}_{\sigma,G}d = 0\ ,$$

due to the feasibility of x and analogously

$$\overline{\Omega}\left(h + \mathcal{J}_{\sigma,H,z}\tilde{c} + \mathcal{J}_{\sigma,H}(x + \tau d)\right) = \overline{\Omega}\left(h + \mathcal{J}_{\sigma,H,z}\tilde{c} + \mathcal{J}_{\sigma,G}x + \tau\mathcal{J}_{\sigma,H}d\right) = 0\ .$$

Thus, the components with indices in $\mathcal{I}$ stay zero. The remaining ones only vary slightly for $\tau \gtrsim 0$.

The same computation and arguments as in the proof of Lemma 3.17 show descent in the function value and Eq. (4.36).

2. There are two statements to show, first the feasibility and second the descent in the function value. For the feasibility of $x + \tau d \in \mathcal{F}_\sigma$ for $\tau \gtrsim 0$ in Lemma 3.17 and the first statement of this lemma it was already shown that $x+\tau d \in \mathcal{P}_\sigma$ and the equality constraints are still fulfilled. Hence, it remains only to prove that the inequality constraints hold. Using Eq. (4.8) to consider the relevant lines of the inequalities and $z_\sigma(x) = \tilde{c} + \tilde{Z}x$ yields

$$\left(\overline{\Omega} + \Upsilon\right)\left(h + Dx + Ez + F\Sigma Z\right) = \left(\overline{\Omega} + \Upsilon\right)\left(h + Dx + \mathcal{J}_{\sigma,H,z}\left(\tilde{c} + \tilde{Z}x\right)\right) .$$

Therefore, to obtain the signs at the point $x + \tau d$ the last expression provides

$$\begin{aligned}
&\left(\overline{\Omega} + \Upsilon\right)\left(h + D\left(x + \tau d\right) + \mathcal{J}_{\sigma,H,z}\left(\tilde{c} + \tilde{Z}\left(x + \tau d\right)\right)\right)\\
&= \left(\overline{\Omega} + \Upsilon\right)\left(h + Dx + \mathcal{J}_{\sigma,H,z}\left(\tilde{c} + \tilde{Z}x\right)\right) + \tau\left(\overline{\Omega} + \Upsilon\right)\left(Dd + \mathcal{J}_{\sigma,H,z}\tilde{Z}d\right)\\
&= \tau\left(\overline{\Omega} + \Upsilon\right)\mathcal{J}_{\sigma,H}d\\
&= \tau\upsilon ,
\end{aligned}$$

where for the second equality one uses that the active constraints at x are zero and for the third equality the assumption (4.37). Examining the sign, all active inequality constraints stay active with the exception of the ith one, which turns negative. Taking this into account and the fact that the inactive constraints vary only slightly for $\tau \gtrsim 0$ and stay negative, the feasibility of $x + \tau d$ follows.

To show that τd is a descent direction consider the optimization problem that results after dropping the ith constraint

$$\begin{aligned}
\min_{x\in\mathbb{R}^n, z\in\mathbb{R}^s} \quad & a^\top x + b^\top|\Sigma|z + \frac{1}{2}x^\top Qx\\
\text{s.t.} \quad & 0 = g + Ax + B|\Sigma|z + C\Sigma z ,\\
& 0 = \left(\overline{\Omega} + \Upsilon\right)\left(h + Dx + E|\Sigma|z + F\Sigma z\right) ,\\
& 0 = P_\alpha\left(z - \tilde{c} - \tilde{Z}x\right) .
\end{aligned}$$

Note that the switching equation is multiplied by P_α, since these are only the relevant components because the remaining components of the Lagrange multiplier λ are set to zero. The stationarity condition of the KKT-theory yields for the Lagrange multipliers $\delta \in \mathbb{R}^m$, $\nu \in \mathbb{R}^p$ and $\lambda \in \mathbb{R}^{|\alpha|}$ and the notation introduced in Eq. (4.8) the two equations

$$
\begin{aligned}
0 &= a^\top + x^\top Q + \delta^\top A \nu^\top \left(\overline{\Omega} + \Upsilon\right) D + \lambda^\top P_\alpha \tilde{Z} \ , \\
0 &= b^\top |\Sigma| + \delta^\top \mathcal{J}_{\sigma,G,z} + \nu^\top \left(\overline{\Omega} + \Upsilon\right) \mathcal{J}_{\sigma,H,z} + \lambda^\top P_\alpha \ .
\end{aligned}
$$

Substituting these two equations in the expression of the directional derivative given in Eq. (3.39) yields

$$
\begin{aligned}
f'(x; \tau d) &= \tau \left(a^\top + x^\top Q + b^\top |\Sigma| \tilde{Z}\right) d \\
&= \tau \Big(-\delta^\top A - \nu^\top \left(\overline{\Omega} + \Upsilon\right) D + \lambda^\top P_\alpha \tilde{Z} \\
&\qquad -\delta^\top \mathcal{J}_{\sigma,G,z} \tilde{Z} - \nu^\top \left(\overline{\Omega} + \Upsilon\right) \mathcal{J}_{\sigma,H,z} \tilde{Z} - \lambda^\top P_\alpha \tilde{Z}\Big) d \\
&= \tau \left(-\delta^\top \mathcal{J}_{\sigma,G} d - \nu^\top \left(\overline{\Omega} + \Upsilon\right) \mathcal{J}_{\sigma,H} d\right) \\
&= -\tau \nu_i \upsilon < 0 \ ,
\end{aligned}
$$

where for the last equality the assumptions (4.37). are used.

3. Once more, there are two statements to show, first the feasibility and second descent in the function value. For the first one it is to show that $x + \tau d \in \mathcal{F}_{\sigma+\gamma}$. The fact that $x + \tau d \in \mathcal{P}_{\sigma+\gamma}$ was already proven in Lemma 3.17. For the equality constraint consider Eq. (4.8) as well as Eq. (3.16) to obtain

$$
\begin{aligned}
&g + Ax + B\left(|\Sigma| + |\Gamma|\right) z_{\sigma+\gamma}(x) + C\left(\Sigma + \Gamma\right) z_{\sigma+\gamma}(x) \\
&= g + Ax + \mathcal{J}_{\sigma+\gamma,G,z} \left(I_s - \tilde{L}\Gamma\right)^{-1} \left(\tilde{c} + \tilde{Z}x\right)
\end{aligned} \tag{4.41}
$$

and analogously for the inequality constraint

$$
\begin{aligned}
&\overline{\Omega}\left(h + Dx + E\left(|\Sigma| + |\Gamma|\right) z_{\sigma+\gamma}(x) + F\left(\Sigma + \Gamma\right)\right) z_{\sigma+\gamma} \\
&= \overline{\Omega}\left(h + Dx + \mathcal{J}_{\sigma+\gamma,H,z} \left(I_s - \tilde{L}\Gamma\right)^{-1} \left(\tilde{c} + \tilde{Z}x\right)\right) \ .
\end{aligned} \tag{4.42}
$$

Thus, showing that the equality and the projection of the active inequality constraints are still zero yields the feasibility. Therefore, consider the representation for the equality constraint in Eq. (4.41) at the point $x + \tau d$, yielding

$$\begin{aligned} & g + A(x+\tau d) + \mathcal{J}_{\sigma+\gamma,G,z}\left(I_s - \tilde{L}\Gamma\right)^{-1}\left(\tilde{c} + \tilde{Z}(x+\tau d)\right) \\ = & \, g + Ax + \mathcal{J}_{\sigma+\gamma,G,z}\left(I_s - \tilde{L}\Gamma\right)^{-1}\left(\tilde{c} + \tilde{Z}x\right) \qquad (4.43) \\ & + \tau\left(Ad + \mathcal{J}_{\sigma+\gamma,G,z}\left(I_s - \tilde{L}\Gamma\right)^{-1}\tilde{Z}d\right) \\ = & \, \tau\left(Ad + \mathcal{J}_{\sigma+\gamma,G,z}\left(I_s - \tilde{L}\Gamma\right)^{-1}\tilde{Z}d\right) = 0\,, \end{aligned}$$

where the contribution in line (4.43) is zero because of the feasibility of x. The last equation follows from the assumption (4.38). Analogously, by using assumption (4.39) it follows for the projection onto the active inequality constraints that Eq. (4.42) holds at $x + \tau d$ and thus the components with indices in $\mathcal{I}$ stay zero and the remaining ones vary only slightly for $\tau \gtrsim 0$.

Finally, the descent in the function value has to be shown. On the polyhedron $\mathcal{P}_{\sigma+\gamma}$ the Lagrange multipliers are well defined and μ is given by Eq. (4.17) and can re-sorted be to

$$\mu^\top = \left(b^\top + \delta^\top B + \nu^\top E + \lambda^\top\right)\Gamma P_\alpha^\top + \left(\delta^\top C + \nu^\top F - \lambda^\top \tilde{L}\right)P_\alpha^\top\,,$$

but in this case by assumption for the ith component it follows

$$\begin{aligned} \mu_i &= \gamma_i\left(b + \delta^\top B + \nu^\top E + \lambda\right)e_i + \left(\delta^\top C + \nu^\top F - \lambda^\top \tilde{L}\right)e_i \\ &= -\left|b + \delta^\top B + \nu^\top E + \lambda\right|e_i + \left(\delta^\top C + \nu^\top F - \lambda^\top \tilde{L}\right)e_i < 0\,. \end{aligned}$$

Next, consider the directional derivative of the objective on $\mathcal{P}_{\sigma+\gamma}$ at x in direction τd. The first steps are the same as in Lemma 3.17 yielding immediately

$$f'(x;\tau d) = \tau\left(a^\top d + x^\top Q d + b^\top\left(|\Sigma| + |\Gamma|\right)\left(I_s - \tilde{L}\Gamma\right)^{-1}\tilde{Z}d\right)\,.$$

Then, substituting Eqs. (4.20) (which does not depend on the $\sigma + \gamma$ decomposition but includes the regularization term), (4.15) and using the assumption (4.40) on d as well as

$\mu_i < 0$ and the notation introduced in Eq. (4.8) the last expression yields

$$\begin{aligned} f'(x;\tau d) = \tau \Bigg(& \left(\lambda^\top \tilde{Z} - \delta^\top A - \nu^\top D\right) d - \delta^\top \mathcal{J}_{\sigma+\gamma,G,z} \left(I_s - \tilde{L}\Gamma\right)^{-1} \tilde{Z} d \\ & -\nu^\top \mathcal{J}_{\sigma+\gamma,H,z} \left(I_s - \tilde{L}\Gamma\right)^{-1} \tilde{Z} d - \lambda^\top \tilde{Z} d + \mu^\top P_\alpha \Gamma \left(I_s - \tilde{L}\Gamma\right)^{-1} \tilde{Z} d \Bigg) \\ = & -\tau\delta^\top \left(A + \mathcal{J}_{\sigma+\gamma,G,z} \left(I_s - \tilde{L}\Gamma\right)^{-1} \tilde{Z}\right) d \\ & - \tau\nu^\top \left(D + \mathcal{J}_{\sigma+\gamma,H,z} \left(I_s - \tilde{L}\Gamma\right)^{-1} \tilde{Z}\right) d \\ & + \tau\mu^\top P_\alpha \Gamma \left(I_s - \tilde{L}\Gamma\right)^{-1} \tilde{Z} d \\ = & \ \tau \mu_i \gamma_i^2 < 0 \ . \end{aligned}$$

Therefore, there is again a descent, which completes the proof. □

Consequently, the lemma has shown that there exists a descent direction in all cases where optimality does not hold. The following lemma demonstrates further that solving the saddle point system (4.25) yields such a descent direction under mild assumption. As a preparation, the optimization problem (4.35) is rewritten as an optimization problem in the search direction. This means, with the help of the search directions defined in Eq. (4.31), it is possible to reformulate problem (4.35). Therefore, denote by $f_Q(x,z) = a^\top x + b^\top |\Sigma| z + \frac{1}{2} x^\top Q x$ the target function and $\bar{x}$ and $\bar{z}$ the current point to obtain

$$\begin{aligned} & f_Q(\bar{x} + \Delta x, \bar{z} + \Delta z) = a^\top (\bar{x} + \Delta x) + b^\top |\Sigma| (\bar{z} + \Delta z) + \frac{1}{2} (\bar{x} + \Delta x)^\top Q (\bar{x} + \Delta x) \\ = & \underbrace{a^\top \bar{x} + b^\top |\Sigma| \bar{z} + \frac{1}{2} \bar{x}^\top Q \bar{x}}_{= f_Q(\bar{x},\bar{z}) = \text{const.}} + a^\top \Delta x + b^\top |\Sigma| \Delta z + \frac{1}{2} \Delta x^\top Q \Delta x + \underbrace{\frac{1}{2} \bar{x}^\top Q \Delta x + \frac{1}{2} \Delta x^\top Q \bar{x}}_{= \bar{x}^\top Q \Delta z} \\ = & (a^\top + \bar{x}^\top Q) \Delta x + b^\top |\Sigma| \Delta z + \frac{1}{2} \Delta x^\top Q \Delta x + f_Q(\bar{x}, \bar{z}). \end{aligned} \tag{4.44}$$

Defining $\varphi(\Delta x, \Delta z) := (a + Q\bar{x})^\top \Delta x + b^\top |\Sigma| \Delta z + \frac{1}{2} \Delta x^\top Q \Delta x$ then yields the problem

$$\begin{aligned} \min_{(\Delta x, \Delta z) \in \mathbb{R}^{n+s}} \quad & \varphi(\Delta x, \Delta z) \\ \text{s.t.} \quad & 0 = A \Delta x + B |\Sigma| \Delta z + C \Sigma \Delta z \ , \\ & 0 = \overline{\Omega} (D \Delta x + E |\Sigma| \Delta z + F \Sigma \Delta z) \ , \\ & 0 = |\Sigma| \Delta z - \tilde{Z} \Delta x \ . \end{aligned} \tag{4.45}$$

The only difference of this problem compared to (4.35) is that here it is minimized along the search direction for fixed $\bar{x}$ and $\bar{z}$, whereas in the original problem (4.35) it is searched for the point where the minimum is attained. Thus, as in [70], it can be shown that solving the saddle point system (4.25) yields a descent direction.

Lemma 4.11. *Suppose that* $(\Delta x^*, \Delta z^*)$ *is a solution of* (4.45) *with* $\Delta x^* \neq 0$ *and let the zero vector be no solution of* (4.45)*. Then the objective function* $f_Q(\cdot,\cdot)$ *is strictly decreasing along the direction* $(\Delta x^*, \Delta z^*)$*. If LIKQ also holds, then this direction is unique.*

Proof. Since the zero vector is a feasible point but no solution of (4.45) and $(\Delta x^*, \Delta z^*)$ is a solution, one has that

$$\varphi(\Delta x^*, \Delta z^*) < \varphi(0) \quad \Rightarrow \quad (a + Q\bar{x})^\top \Delta x^* + b^\top |\Sigma| \Delta z^* + \frac{1}{2}\Delta x^* Q \Delta x^* < 0 \; . \quad (4.46)$$

Since Q is positive definite, one has $\frac{1}{2}(\Delta x^*)^\top Q \Delta x^* > 0$ and it follows with Eq. (4.46):

$$(a + Q\bar{x})^\top \Delta x^* + b^\top |\Sigma| \Delta z^* < 0.$$

Therefore, using Eq. (4.44) one obtains

$$\begin{aligned} f_Q(\bar{x} + \alpha\Delta x^*, \bar{z} + \alpha\Delta z^*) &= f_Q(\bar{x}, \bar{z}) + \underbrace{\alpha(a + Q\bar{x})^\top \Delta x^* + \alpha b^\top |\Sigma| \Delta z^* + \frac{1}{2}\alpha^2 \Delta x^* Q \Delta x^*}_{<0} \\ &< f_Q(\bar{x}, \bar{z}) \; , \end{aligned}$$

for all $\alpha > 0$ sufficiently small. The uniqueness follows from the assumption that LIKQ holds true (see Section 4.1). □

Using the last two lemmata, it is now possible to show the convergence of Algorithm 4. The result is similar to the one for the unconstrained case in Theorem 3.18 and has already been formulated and shown in [70].

Theorem 4.12. *Suppose that an optimization problem of the form* (CALOP) *is given, LIKQ holds at every feasible point and let* $Q = Q^\top \in \mathbb{R}^{n \times n}$ *be a positive definite matrix. Then, Algorithm 4 terminates for any feasible starting point* $x \in \mathbb{R}^n$ *after finitely many* *iterations at a minimizer of the quadratically penalized optimization problem*

$$\begin{aligned}
\min_{x\in\mathbb{R}^n, z\in\mathbb{R}^s} \quad & a^\top x + b^\top z + \frac{1}{2}x^\top Q x \\
s.t. \quad & 0 = g + Ax + Bz + C|z| \ , \\
& 0 \leq h + Dx + Fz + E|z| \ , \\
& z = c + Zx + Mz + L|z| \ .
\end{aligned}$$

Proof. The essential argumentation is very similar to the proof of Theorem 3.18: It has to be shown that on every polyhedron an optimum is found after finitely many steps, then finite convergence follows again from the finite number of polyhedra.

Algorithm 4 starts by solving the system of equations (4.25). If the solution returns the same point $\hat{x}$ as the current point x, then x is a feasible signature stationary point, the step size $\beta = 1$ and the algorithm tests for optimality. If this test is successful, the algorithm terminates. If it is not successful, then, as seen in the proof of Lemma 4.10, there exists at least one index $i \in \alpha$ such that $\mu_i < 0$, or $j \in \mathcal{I}$ such that $\nu_j < 0$. Further, it follows by Lemma 4.10 and the fact that $\beta > 0$, after adjusting the signature vector or the signature vector of inequality constraints, that there exists a new descent direction afterwards.

Since the algorithm always finds a feasible signature optimal point before changing the polyhedron and thus assumes the best possible function value on this feasible polyhedron, the algorithm cannot reach the same polyhedron again because of the descent after leaving a kink. Now there are only finitely many closed polyhedra, namely at most 2^s. In a similar manner the algorithm always finds a feasible signature stationary point before dropping an active constraint. Therefore, the best possible function value for the currently active inequality constraints on the current polyhedron is found. As seen in Lemma 4.10 and by the fact that $\beta > 0$ there is a descent after dropping an active inequality constraint and thus the algorithm cannot reach the set of active inequality constraints on the current polyhedron again. There are only finitely many combinations for the active sets, namely at most 2^p.

Now consider the situation where there is no feasible signature stationary point yet. Then the solutions of the saddle point system (4.25) provides a search direction. The computed step size will then either return $\beta = 1$ *or*, and maybe *and*, an inactive inequality constraint or a kink will be added. The first case means that the optimum for the current active inequality constraints on the current polyhedron has been reached and the next iteration will return a zero step. If this is not the case, adding constraints and kinks can only occur repeatedly at most $p + s$ times, until all - because of LIKQ - linearly independent inequality constraints and kinks have been added. Also in this case the next iteration will yield a zero step.

Thus, as in the unconstrained case an optimum is found on a polyhedron after finitely many iterations, and there are only finitely many polyhedra. This leads to convergence after finitely many iterations. □

Because of the regularizer, Algorithm 4 (as well as Algorithm 3) naturally does not necessarily find a solution of the original problem (CALOP). In practice, however, the matrix Q is to be chosen with very small entries in absolute value, e.g., a multiple of the unit matrix qI, for a $q \gtrsim 0$ very small. Theorem 4.8 can then be used to check the optimality of the found point with respect to the problem (CALOP). If it is not yet optimal but the optimization problem is bounded on the feasible set, one can use the found point as a new starting point and set Q to zero or tend it to zero. Otherwise, the solution cannot be given anyway. However, the numerical results in Chapter 5 show that such cases do not occur in practice so far. Nevertheless, this is also formally proven in the following theorem, which of course can also be applied to ASM, if one assumes artificial box constraints that have no influence on the minima but bound the feasible set.

Theorem 4.13. *Let f be bounded from below on the feasible set given by* (CALOP)*, denoted by $\mathcal{F}$. Further assume that $\mathcal{F}$ is bounded and LIKQ holds at every feasible point. Then, for $Q \to 0$ the solutions generated by Algorithm 4 converge to a solution of the optimization problem* (CALOP)*.*

Proof. Since $\mathcal{F}$ is bounded and f is bounded from below it attains a minimum on $\mathcal{F}$. All the optimality conditions (4.21) to (4.23) are independent from Q and therefore coincide with the corresponding optimality conditions of (CALOP) (cf. Theorem 4.8). Thus, the only optimality condition that depends on Q is stated in Eq. (4.20). For reasons of continuity, if Q tends to zero, Eq. (4.20) converges to the same optimality condition as given in Eq. (4.14). Thus, the solution generated by Algorithm 4 coincides with the solution of (CALOP). □

To conclude this section, the regularization term should be discussed once more. With the choice of $\frac{1}{2}x^\top Qx$ used here, the regularization is done around zero. However, it can also be in the interest of the user to do this not around zero but around another point. This can be the case, for example, if the user can expect beforehand that the solution is to be expected "very far away" from zero. It is also possible to move the regularizer around the current iteration, so that it adapts in each iteration. If $\bar{x}$ is the point around which the regularizer is to be shifted, the regularization term $\frac{1}{2}(x - \bar{x})^\top Q(x - \bar{x})$ results. The theory is not significantly influenced by this shift, only the first block of the right-hand side in the saddle point system (4.25) changes to the term $a - Q\bar{x}$ instead of just a.

4.3 Penalty Approaches for Piecewise Linear Optimization Problems

As described in the previous sections (see Subsection 4.2.4), Algorithm 4 is a feasible point method, i.e., a feasible starting point is required and any iteration point generated by the algorithm is feasible. If no feasible starting point is available, it is necessary to determine one. Phase I Methods, very similar to those described in Section 2.2.5, can be used for this purpose. However, for piecewise linear functions in combination with l_1-penalty approaches (see Section 2.3), problems may arise. Examining this in more detail is the purpose of this section. Other publications that have focused on (exact) Penalty Methods for nonsmooth functions are, e.g., [60, 90, 91, 108].

In a wider sense, all Phase I Methods from Section 2.2.5 can be interpreted as penalty-like methods in the sense of Section 2.3, since in all cases the violation of constraints is penalized by a penalty term. Thus, discussions of Phase I Methods and Penalty Methods are closely related in this context.

By combining algorithms for unconstrained problems with Penalty Methods, Theorem 2.13 shows under which conditions a solution to the corresponding constrained optimization problem can be found. This is also true for Algorithm 3 given in this paper. Here, however, two essential aspects play a role. First, for Theorem 2.13 to be valid, the feasible set of the constrained optimization problem must be convex and closed. Another aspect is that the theorem "only" gives results concerning local minima. Now, for example, compare Algorithm 3 combined with a l_1-penalty approach with Algorithm 4 and assume that the feasible set described by the constraints satisfies the conditions from Theorem 2.13. By this way, the theorem ensures that the solutions coincide and a local minimum is reached.

However, local minimality can lead to problems with the Phase I Methods of Subsection 2.2.5, at least when these methods are applied to piecewise linear problems. For convex functions, such as those considered in the Active Set Method in Algorithm 1, these methods are also applicable, since convexity ensures that any local minimum is already a global minimum. If the feasible set then satisfies the conditions in Theorem 2.13, this is equivalent to saying that the solution of the Phase I Methods also provides a feasible starting point for the actual problem.

Of course, in the piecewise linear case, the feasible set can be convex and closed. However, the danger is much greater here, because of the kinks, that this is no longer true for arbitrary problems. Therefore, in the general case, there is a very high risk that a Phase I Method or even a l_1-penalty method will terminate in a local minimum without finding a feasibility point.

For example, consider the equality constraint $\max(x-1, -|x|+2) = 0$, which has the only feasible point given by $\check{x} = -2$. Using the l_1-penalty approach to determine this feasible point, one obtains the minimization problem

$$\min_{x \in \mathbb{R}} \quad \rho \left|\max(x-1, -|x|+2)\right|,$$

for $\rho > 0$. Regardless of the choice of penalty parameter ρ, this optimization problem has a global minimizer at $\check{x} = -2$ and a local, but no global, minimizer at $\hat{x} = 1.5$. Therefore, using any starting point $x \geq 1.5$ for Algorithm 3 or 4 yields the solution $\hat{x}$ and thus a nonfeasible point for the constraint itself.

Nevertheless, there are cases where are Penalty Methods or Phase I Methods are successful. The first one has already been discussed indirectly. If one additionally assumes that for the corresponding problems every local minimum is also a global minimum, then Phase I Methods also provide feasible starting points. Of course, this is especially true for convex functions, but not only. Still, this is a relatively strong assumption. An alternative is to run a Phase I Method with a quadratic objective function on each polyhedron. This way one has a quadratic optimization problem with only linear constraints on each polyhedron. If at the end the penalty term is zero, then one has a feasible starting point. If not, the entire polyhedron must not contain any feasible point. However, this is very close to a globalization strategy, which becomes very expensive depending on the number of polyhedra.

To conclude the theoretical discussion in this thesis on the topic of Penalty Methods, the choice of the penalty parameter is now briefly discussed. A possible choice of this parameter was already given in Theorem 2.14. For this, however, the Lagrange multipliers have to be determined. Illustratively, any of the violation of the constraints must penalize more than the function value decreases. The gradient of the objective function of the optimization problem (4.19) as a function of σ can be determined by substituting the switching equation (4.19d) into the objective function (4.19a), thus obtaining Eq. (3.35).

The largest slope over all polyhedra is then obtained by

$$\mathcal{J}_f^{\max} := \max_{\sigma \in \{-1,0,1\}^s} ||f'(x)||_\infty := \max_{\sigma \in \{-1,0,1\}^s} ||a + \tilde{Z}^\top |\Sigma| b + Qx||_\infty \ .$$

Using the active Jacobian (cf. Definition 4.5) the gradients of the constraints can be determined. Therefore, the smallest value is of interest. To ensure that it does not vanish, assume that the active Jacobian $\mathcal{J}_\sigma$ for all $\sigma \in \{-1, 0, 1\}$ contains no zero lines. Then define the smallest slope of the constraints as

$$\mathcal{J}^{\min} := \min_{\sigma \in \{-1,0,1\}^s} ||\mathcal{J}_\sigma||_\infty \ .$$

To compensate the slope of the objective function by the penalty parameter ρ^* and the constraints, the following must be fulfilled

$$\mathcal{J}_f^{\max} \leq \rho^* \mathcal{J}^{\min} \ .$$

Thus, in case that every local minimizer is a global one, any penalty parameter $\rho > \rho^*$ with

$$\rho^* \geq \frac{\mathcal{J}_f^{\max}}{\mathcal{J}^{\min}}$$

yields the exactness of the l_1-penalty function in that Algorithm 3 with the l_1-penalty function for (CALOP) yields a local minimum of (CALOP).

5

Numerical Results

In order to conclude the main part of this thesis, the numerical performance of Algorithm 4, hereafter always referred to as CASM, will be demonstrated in the following chapter for different types of examples. For this purpose, this chapter is divided into four sections. Section 5.1 briefly discusses aspects of the implementation of CASM. First simple and academic examples are then considered in Section 5.2. These include problems that are scalable in dimension as well as linear complementary and bi-level problems. This is followed by Section 5.3 where optimization problems with application background are solved. These are subproblems that arise in the optimization of gas networks and can be approximated by constrained piecewise linear problems. Different network sizes are considered here as example instances and different aspects are considered in each case. The final Section 5.4 considers a piecewise linear regression problem, where the parameters of the regression problem are determined for data from the retail trade.

5.1 Aspects of Implementation

For the numerical results in the following sections, CASM is implemented in Matlab version R2021a [84]. The examples are performed on a Lenovo ThinkPad T490 with Ubuntu 20.04.5, 16 GB RAM and an Intel(R) Core(TM) i7-8565U processor.

Since quantitative statements about runtimes are naturally associated with uncertainties, for example due to background processes, the used programming language, or different implementations, some more qualitative anomalies are listed here. During the implementation a number of options occurred, which have influence on the performance of CASM regarding the stability or run time.

Since CASM does not ensure LIKQ in every iteration, the saddle point systems (4.25) or (4.26) may not be uniquely solvable. To deal with this initially, e.g., the solver `mldivide` (respectively \) was used. However, it happened that individual entries of the solution vector were no longer presentable numbers. From the mathematical point of view this is also reasonable. For example, if LIKQ is not fulfilled, there can be an infinite number of solutions, where entries are numerically also outside of the representable numbers. The same behavior was observed if the system matrix was inverted with `inv` to solve the system of equations. In the end, the solver `lsqminnorm` turned out to be the most stable one. According to the Matlab documentation [84], `lsqminnorm` returns an array X that solves the linear equation $AX = B$ and minimizes the value of $\|AX - B\|$. If several solutions exist to this problem, then `lsqminnorm` returns the solution that minimizes $\|X\|$. As a result, it no longer occurs that solutions are computed where entries cannot be represented numerically.

Two other factors that had a significant impact on the computation time were the precise updating of the saddle point matrix (4.25), respectively (4.26), and the exploitation of sparsity. In all examples given here, most of the matrices and also in some cases the vectors are very sparse, i.e., many entries are zero. For such a situation Matlab has the function `sparse`, which according to the documentation converts a full matrix into sparse form by squeezing out any zero elements. If a matrix contains many zeros, converting the matrix to sparse storage saves memory. Of course, the use of this function depends largely on the problem itself, but as already mentioned, in practice most matrices that occur are sparse.

Now consider the update of the saddle point matrix: In the first implementations, the entire system matrix was completely rebuilt in each iteration. But since the matrix in Eq. (4.25) changes only slightly in each iteration, because at most one component of σ and ω is changed, it is sufficient to adjust only the corresponding rows and columns and not to rebuild the whole matrix. Unlike the system matrix from Eq. (4.25), this is no longer so easy for the one from Eq. (4.26): By the transformation resulting from reduction of the system matrix (cf. Lemma 4.9) the blocks are no longer simply multiplied from one side with Σ or Ω but these are now in the middle of the blockterms. The consequence is that a change in Σ respectively Ω has influence on a larger number of entries, which can also no longer be localized so easily. For this reason, the first variant was used for the numerical examples in this thesis, i.e., the sparse saddle point system (4.25) was solved using the function `lsqminnorm`. Then, in each iteration, only the corresponding rows and columns of the matrix blocks were updated after a change in the signature vectors σ and/or ω.

These two points, the exploitation of sparsity and the precise updating of the system matrix have reduced the runtime to about less than 5% of the original one, where this was not taken into account. This made it possible to solve significantly more complex problems, which otherwise would have taken several days.

5.2 Academical Examples

In this section, some rather simpler, academic examples are now considered and the behavior of CASM is illustrated. The examples in this section have already been considered in the article in which CASM was presented and are reproduced here, see [70, Section 5]. The first example is the Hill-function, which has already been considered several times in the course of this thesis.

Example 5.1 (Hill-function)**.** *Consider again the constrained optimization problem stated in Example 4.2 which is given by*

$$\begin{aligned} \min_{x\in\mathbb{R}^2} \quad & \max\{0, x_1 - |x_2|\} \\ s.t. \quad & \left| -\frac{1}{2}|z_1| + \frac{1}{2}\left|x_1 - |x_2|\right| \right| \leq 2\,. \end{aligned}$$

Figure 5.1 shows two different sequences of iterates generated by CASM, for two different starting values $x^0 = (8,3)$ and $\tilde{x}^0 = (8,-5)$. For the regularization term $Q = qI_2$ is chosen with $q = 10^{-10}$. Once more, the resulting kinks are given by the blue lines and the feasible set is marked by the red area. In Table 5.1, the individual iterates, the signature vectors and signature vectors of the inequality constraints are given.

In the sequence labeled by x^i in Figure 5.1, it can be seen that the constraints are not touched at all. This is different for the sequence marked with $\tilde{x}^i$. There, the constraint becomes active in the first iteration. Then the function is minimized along this constraint and yields the locally minimal point $\tilde{x}^3 = (4,-4)$. Numerically, in this example the choice of the parameter q plays a role. If one chooses q one order of magnitude higher, then

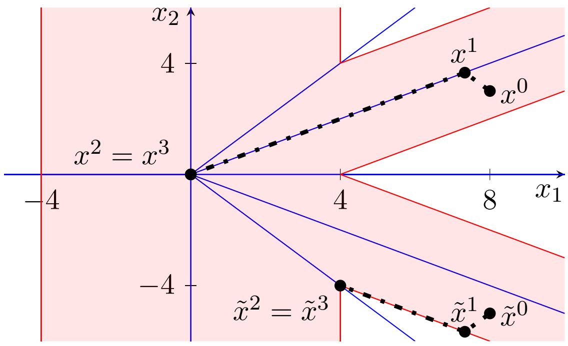

Figure 5.1: Illustration of the iteration sequences generated by CASM for the Hill-function

Iteration	x^i	σ^i	ω^i	$\tilde{x}^i$	$\tilde{\sigma}^i$	$\tilde{\omega}^i$
0	(8.00, 3.00)	(1, 1, 1)	-1	(8.00, -5.00)	(-1, 1, -1)	-1
1	(7.33, 3.66)	(1, 1, 0)	-1	(7.33, -5.66)	(-1, 1, -1)	0
2	(0.00, 0.00)	(1, 0, 0)	-1	(4.00, -4.00)	(-1, 0, -1)	0
3	(0.00, 0.00)	(1, 0, 0)	-1	(4.00, -4.00)	(-1, 0, -1)	0

Table 5.1: Optimization history of CASM for the Hill-function

another step is also made to the point $(0,0)$*. This is due to the choice of* ε*-parameters in the implementation and the* `lsqminnorm` *solver of Matlab. Of course, both* $(0,0)$ *and* $(4,-4)$ *are solutions. The example thus shows that different starting points do not necessarily yield the same minimum point.*

The next example is a convex piecewise linear function for which it has already been shown in [59] that a steepest descent approach with exact line search for the starting point $x^0 = (9, -2.5)$ yields zigzag behavior and convergence to a nonstationary point. Numerical results using ASM for the unconstrained problem have already been presented in [45, Chapter 4].

Example 5.2 (Constrained HUL)**.** *In [59], Hiriart-Urruty and Lemaréchal considered the piecewise linear and convex function* $\varphi : \mathbb{R}^2 \to \mathbb{R}$*,*

$$\varphi(x) = \max\{\max\{-100, 2x_1 + 5|x_2|\}, 3x_1 + 2|x_2|\} \; . \tag{5.1}$$

To test CASM, the feasible starting point $x^0 = (9, -2.5)$ *is chosen and the two constraints*

$$\begin{aligned} H_1(x) &= -0.25x_1 - x_2 - 10 \le 0 \; , \\ H_2(x) &= 2 - 0.2|x_1 + 9| - |x_2 + 1| \le 0 \; , \end{aligned} \tag{5.2}$$

are added. After reformulation of the max *functions in Eq. (5.1) by means of the absolute value, this optimization problem requires the six switching variables and the target function*

$$\begin{aligned} z_1 &= x_2 \; , & z_2 &= -100 - 2x_1 - 5|z_1| \; , \\ z_3 &= -50 - 2x_1 + 0.5|z_1| + 0.5|z_2| \; , & z_4 &= x_1 + 9 \; , \\ z_5 &= x_2 + 1 \; , & z_6 &= 2.25|z_1| + 0.25|z_2| + 0.5|z_3| \; , \\ y &= -25 + 2x_1 + z_6 \; , \end{aligned}$$

Iteration	x^i	σ^i	ω^i
0	(9.00, -2.50)	(-1, 1, 1, 1, -1, 1)	(-1, -1)
1	(6.75, -1.00)	(-1, 1, 1, 1, 0, 1)	(-1, -1)
2	(3.00, -1.00)	(-1, 1, 0, 1, 0, 1)	(-1, -1)
3	(3.00, -1.00)	(-1, 1, 0, 1, 1, 1)	(-1, -1)
4	(0.00, 0.00)	(0, 1, 0, 1, 1, 1)	(-1, -1)
5	(0.00, 0.00)	(0, 1, -1, 1, 1, 1)	(-1, -1)
6	(- 4.00, 0.00)	(0, 1, -1, 1, 1, 1)	(-1, 0)
7	(- 4.00, 0.00)	(1, 1, -1, 1, 1, 1)	(-1, 0)
8	(- 9.00, 1.00)	(1, 1, -1, 0, 1, 1)	(-1, 0)
9	(- 9.00, 1.00)	(1, 1, -1, -1, 1, 1)	(-1, 0)
10	(-14.00, 0.00)	(0, 1, -1, -1, 1, 1)	(-1, 0)
11	(-14.00, 0.00)	(0, 1, -1, -1, 1, 1)	(-1, 0)
12	(-40.00, 0.00)	(0, 1, -1, -1, 1, 1)	(0, -1)
13	(-40.00, 0.00)	(1, 1, -1, -1, 1, 1)	(0, -1)
14	(-66.67, 6.67)	(1, 0, -1, -1, 1, 1)	(0, -1)
15	(-66.67, 6.67)	(1, 0, -1, -1, 1, 1)	(0, -1)

Table 5.2: Optimization history of CASM for the constrained HUL-function

e.g., one has $n = 2$, $s = 6$, $m = 0$ and $p = 2$. Using CASM, 15 iterations are needed as stated in Table 5.2, which also illustrates the deactivation and the activation of the switches and the inequality constraint during the optimization.

Figure 5.2 shows a plot of the resulting kinks originating from the objective function and from the constraints (blue lines). The inequality constraints are marked by the red lines and the red area represents the feasible set. Finally, the iterates generated by CASM are denoted by the black dots.

This example illustrates one of the advantages of CASM. The feasible set is a nonconvex set because the constraint (5.2) cuts out a rhombus. Since the remaining set is still path-connected, CASM can find a trajectory that avoids the hole created by the rhombus, provided that no local minimum is reached.

The next example is particularly interesting because it is scalable in dimension. Furthermore, it is known for having only one unique global minimum but no further local minima, instead it has many stationary points where many optimization algorithms get stuck [45]. This example has been also already studied in [70] and is also reported here.

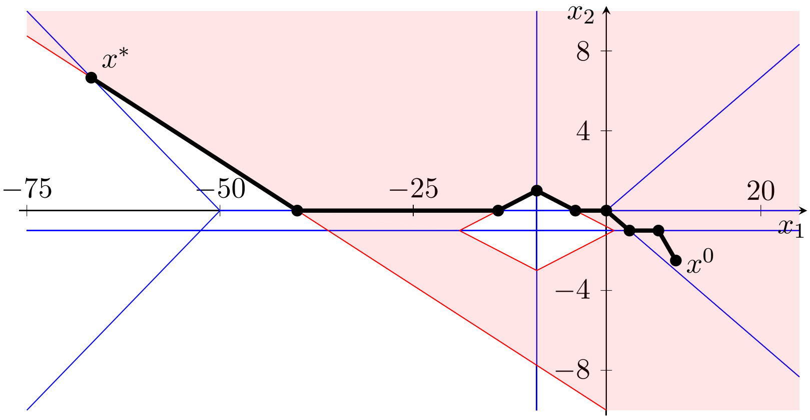

Figure 5.2: Illustration of the iteration sequences generated by CASM for the constrained HUL-function

Example 5.3 (Constrained Rosenbrock-Nesterov II)**.** *According to [51], Nesterov suggested the Rosenbrock-like test function*

$$\varphi : \mathbb{R}^n \to \mathbb{R}, \quad \varphi(x) = \frac{1}{4}|x_1 - 1| + \sum_{i=1}^{n-1} |x_{i+1} - 2|x_i| + 1|$$

that is piecewise linear and nonconvex. It has the unique global minimizer $x^* = (1, 1, \ldots, 1) \in \mathbb{R}^n$ *and* $2^{n-1} - 1$ *other Clarke stationary points, where none of them is a local minimizer. For the starting point*

$$x_1^0 = -1 \quad \textit{and} \quad x_i^0 = 1 \ , \ \textit{for } 2 \le i \le n \ ,$$

the paper [45] contains numerical results and comparisons to other solvers showing that nonsmooth optimization algorithms may get stuck at one of these stationary points that are no minimizers. From the literature [45, 51] it is known that the selected starting point is particularly well chosen, since most algorithms run through all stationary points first. A different starting point would then yield the same iteration sequence in the further course after reaching a stationary point. Since constrained problems are considered in this thesis, the piecewise linear constraint

$$\sum_{i=1}^{n} |x_i - 1| \ge \frac{1}{2n}$$

is added. Hence, there is an n*-dimensional rhombus around the global optimum which is* *cut out of the* $\mathbb{R}^n$*. The remaining* $2^{n-1}-1$ *stationary points are still feasible. To derive an abs-linear representation of this constrained optimization problem,* $s = 3n-1$ *switching variables are defined, namely*

$$z_i = x_i \quad \textit{for} \quad 1 \le i < n\,, \qquad z_{n+i} = x_{i+1} - 1 \quad \textit{for} \quad 0 \le i < n\,,$$
$$z_{2n+i} = x_{i+2} - 2\,|z_{i+1}| + 1 \quad \textit{for} \quad 0 \le i < n-1\,, \qquad z_{3n-1} = \frac{1}{4}|z_n| + \sum_{i=0}^{n-2} |z_{2n+i}|\,.$$

Hence, they yield the matrices and vectors

$$Z = \begin{bmatrix} I_{n-1} & 0 \\ \multicolumn{2}{c}{I_n} \\ 0 & I_{n-1} \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{s\times n}, \quad M = 0\,, \quad L = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ -2\,I_{n-1} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{4} & 0 & \mathbf{1}^\top & 0 \end{bmatrix} \in \mathbb{R}^{s\times s},$$
$$h = \tfrac{1}{2n}\,, \quad D = 0\,, \quad E = 0\,, \quad F = (\underbrace{0,\ldots,0}_{n-1}, \underbrace{-1,\ldots,-1}_{n}, \underbrace{0,\ldots,0}_{n})\,,$$
$$a = 0 \in \mathbb{R}^n, \quad b = e_{3n-1} \in \mathbb{R}^s, \quad c = (\underbrace{0,\ldots,0}_{n-1}, \underbrace{-1,\ldots,-1}_{n}, \underbrace{1,\ldots,1}_{n-1}, 0)\,,$$

with $\mathbf{1} \in \mathbb{R}^{n-1}$ *as the vector with* 1 *in every component. Consider the point*

$$x_i^* = 1 - \frac{2^{i-1}}{2^n - 1} \cdot \frac{1}{2n} \in (0,1) \quad \textit{for } 1 \le i \le n\,.$$

Then one has

$$\sigma(x^*) = (\underbrace{1,\ldots,1}_{n-1}, \underbrace{-1,\ldots,-1}_{n}, \underbrace{0,\ldots,0}_{n})^\top \Rightarrow \alpha = \{2n,\ldots,3n-2\} \quad \textit{and} \quad \omega(x^*) = 0\,.$$

Note that the index $3n-1$ *is not contained in* α *according to Definition 3.5, since* z_{3n-1} *does not occur as an argument of an absolute value. Then, one obtains*

$$\begin{bmatrix} P_{\mathcal{I}}(D + E|\Sigma|\tilde{Z} + F\Sigma\tilde{Z}) \\ P_\alpha \tilde{Z} \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & \cdots & \cdots & 1 \\ -2 & 1 & 0 & & & 0 \\ 0 & -2 & 1 & 0 & & 0 \\ & \ddots & \ddots & \ddots & \ddots & 0 \\ & & \ddots & \ddots & \ddots & 0 \\ & & & 0 & -2 & 1 \end{bmatrix} \in \mathbb{R}^{(n-1)\times n},$$

n	1	2	3	4	5	6	7	8	9	10	11	12
#	2	5	14	27	64	117	238	439	856	1685	3382	6807

n	13	14	15	16	17	18	19	20
#	13592	26285	42994	82995	131096	262173	605342	1119907

Table 5.3: Number of iterations generated by CASM for the constrained Rosenbrock-Nesterov II example with different values of dimension n

such that LIKQ holds. The optimality conditions stated in Theorem 4.8 require

$$\nu^\top = \lambda^\top P_\alpha^\top P_\alpha \tilde{Z} \, , \qquad \nu^\top = -\lambda^\top |\Sigma| \qquad \textit{and} \qquad |P_\alpha \lambda| \leq -P_\alpha \tilde{L}^\top \lambda \, .$$

These conditions hold for $\nu = 0 \in \mathbb{R}$ and $\lambda = 0 \in \mathbb{R}^s$. Hence, x_* is a minimizer.

For varying values of n, the number of iterations required by CASM is shown in Table 5.3. As can be seen from the iteration counts, the number of visited polyedra is much less than the total number of polyhedra with definite signatures given by 2^s. For a comparison, the MPBNGC solver, [81], was also applied to solve this problem. MPBNGC is a multiobjective proximal bundle method for nonconvex, nonsmooth and generally constrained minimization. For $n = 1$, seven iterations are needed. Already for $n = 2$, the solver gets stuck at a non optimal stationary point. The same can be observed for larger values of n.

Another class of problems that can be put into the form of (CALOP) are linear complementarity problems. An example of such a problem is the following one.

Example 5.4. *Consider the linear complementarity problem (LCP) given by*

$$Mx + q \geq 0 \quad \textit{and} \quad x^\top (Mx + q) = 0 \tag{5.3}$$

for $0 \leq x \in \mathbb{R}^n$, $M \in \mathbb{R}^{n \times n}$ and $q \in \mathbb{R}^n$. In [12], the LCP is formulated as a system of piecewise linear equations

$$\min(x, Mx + q) = 0 \, , \tag{5.4}$$

where the minimum operator acts componentwise. In the same paper, the authors present an algorithm that can be viewed as a semismooth Newton method and show nonconvergence for a special choice of the matrix M. They pointed out that the problem has a unique *solution for any* $q \in \mathbb{R}^n$ *if and only if* M *is a* $\boldsymbol{P}$*-matrix, i.e.,* M *has positive principal minors* $\det(M_{II}) > 0$ *for all nonempty* $I \subset \{1, \ldots, n\}$.

To solve Eq. (5.4) *with CASM the problem* (5.3) *is reformulated as*

$$\min_{x \in \mathbb{R}^n} \sum_{i=1}^{n} |\min(x_i, (Mx + q)_i)| \,.$$

For this problem, consider the two instances with $n = 3$ *respectively* $n = 4$ *and the corresponding matrices for* M*:*

$$M_3 \equiv \begin{bmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} \quad \text{and} \quad M_4 \equiv \begin{bmatrix} 1 & 0 & \frac{1}{2} & \frac{4}{3} \\ \frac{4}{3} & 1 & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{4}{3} & 1 & 0 \\ 0 & \frac{1}{2} & \frac{4}{3} & 1 \end{bmatrix} .$$

In both cases, the right-hand side is given by $q = \mathbf{1}$ *as the vector with* 1 *in every component of appropriate dimension as considered also in [12]. As starting point the first unit vector in* $\mathbb{R}^n$ *is used as proposed in [12]. Then, CASM needs five iterations in both cases, i.e., for* M_3 *and* M_4*, respectively, to reach the solution* $\mathbf{0}$ *as zero vector of the appropriate dimension. In [12, Proposition 3.7] it is shown that the algorithm proposed in that paper does not converge but generates a circle of three respectively four reoccurring iterates.*

As a final academic example in this section, a bi-level problem is considered. Bi-level problems have the structure that a lower-level optimization problem must be solved and its solution has an impact on an upper-level optimization problem [24]. They play an important role in many real-world applications, see, e.g., [9, 24, 26, 78], and are closely related to linear complementarity problems as in Example 5.4. In the bi-level problem considered here all functions appearing as objective functions of the upper and lower level as well as all constraints are linear. To convert such problems to the setting of this thesis the KKT-theory (cf. Theorem 2.3) is applied to the lower level to obtain a set of equalities and inequalities representing the necessary and sufficient optimality conditions for the lower level. Therefore, the lower level problem is replaced by these conditions in form of additional constraints. Hence, the Lagrange multipliers from the lower problem also become optimization variables. However, the resulting complementarity condition is no longer a linear function. Thus, for the application of CASM, they can be reformulated analogous to Eq. (5.4) as a piecewise linear constraints.

Example 5.5. *Consider the following linear bi-level problem taken from [102, Chapter 7]:*

$$\begin{aligned}
\min_{x,y\in\mathbb{R}^2} \quad & 3x_1 + 2x_2 + y_1 + y_2 \\
s.t. \quad & x_1 + x_2 + y_1 + y_2 \leq 4 \ , \\
& y \in \underset{\tilde{y}\in\mathbb{R}^2}{\arg\min} \quad 4\tilde{y}_1 + \tilde{y}_2 \\
& \qquad\quad s.t. \qquad 3x_1 + 5x_2 + 6\tilde{y}_1 + 2\tilde{y}_2 \geq 15 \ , \\
& x \geq 0, \ y \geq 0 \ .
\end{aligned}$$

Using the starting point

$$x = (2.5, 1.5) \ , \quad y = (0, 0) \ , \quad \textit{and} \quad \mu = (0, 4, 1) \ ,$$

where μ represents the Lagrange multiplier resulting from the lower level problem as described above, Table 5.4 shows information about the iterates when solving this problem with CASM. In [102], a structurally quite different algorithm is used to solve the problem, making it difficult to compare the effort. Both algorithms perform some preparatory work in that a pre-solve is performed before applying the algorithm proposed in [102] and a feasible starting point has to be determined for CASM. Subsequently, the algorithm presented in [102] requires three iterations, each of which requires the solution of two linear programs. CASM needed six iterations, where a system of equations with a 27×27 system matrix must be solved in each iteration. Both algorithms attain the same solution.

i	x^i	y^i	μ^i	σ^i	ω^i
0	(2.5, 1.5)	(0, 0)	(0.0, 4.0, 1.0)	(0, 1, 1)	(0, -1, -1, 0, 0, 0, 0, -1, -1)
1	(2.5, 1.5)	(0, 0)	(0.0, 4.0, 1.0)	(0, 1, 1)	(-1, -1, -1, 0, 0, 0, 0, -1, -1)
2	(0.0, 3.0)	(0, 0)	(0.0, 4.0, 1.0)	(0, 1, 1)	(-1, 0, -1, 0, 0, 0, 0, -1, -1)
3	(0.0, 3.0)	(0, 0)	(0.0, 4.0, 1.0)	(0, 1, 1)	(-1, 0, -1, 0, -1, 0, 0, -1, -1)
4	(0.0, 3.0)	(0, 0)	(0.0, 4.0, 1.0)	(0, 1, 1)	(-1, 0, -1, 0, -1, 0, -1, -1, -1)
5	(0.0, 3.0)	(0, 0)	(0.0, 4.0, 1.0)	(1, 1, 1)	(-1, 0, -1, 0, -1, 0, -1, -1, -1)
6	(0.0, 3.0)	(0, 0)	(0.5, 4.0, 0.0)	(1, 1, 0)	(-1, 0, -1, 0, -1, 0, -1, -1, 0)

Table 5.4: Optimization history of CASM for the linear bi-level problem

5.3 The Gas Transport Problem

In contrast to the previous section, the following one will now deal with a real-world application example, namely gas transport problems. The results given here are mainly derived from a cooperation with Martina Kuchlbauer, during her PhD phase under the supervision of Frauke Liers and Michael Stingl, and have already been published jointly in the conference paper [67]. The derivation of the optimization problem and the corresponding numerical results will be restated and extended by results on larger dimensioned problems.

The problem considered here is called the *stationary robust gas transport problem*. It is an optimization problem under uncertainties in demand and physical parameters, which is nonconvex in node pressure and flow along the pipes. This section is neither about modeling nor about other solution methods. For details about the mathematical modeling and other solution methods the reader is referred to [66]. The focus here is the problem formulation and afterwards an overview of numerical results, which are generated with the help of CASM. In the following, the arising robust gas transport problem from the point of view of the network operator is modeled [6].

The gas network is described by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{A})$, where the set $\mathcal{V}$ denotes the nodes. In addition, the arcs model pipes and compressors ($\mathcal{A} = \mathcal{A}_{pi} \cup \mathcal{A}_c$), resulting in a corresponding incidence matrix $M \in \{-1, 0, 1\}^{|\mathcal{V}| \times |\mathcal{A}|}$. The gas flow is denoted by $q \in \mathbb{R}^{|\mathcal{A}|}$, where its sign indicates the flow's direction. Furthermore, squared pressure values are denoted by $\pi \in \mathbb{R}^{|\mathcal{V}|}$. To ensure uniqueness of the physical states, the pressure value is fixed at one so-called root node.

The aim of the stationary robust gas transport problem is to find an optimal control that is robustly protected against the perturbation of physical parameters and a minimum-cost control of compressors. Therefore, by $w(\Delta)$ the costs of a control Δ of compressors are denoted. As a first set of constraints all demands should be satisfied and as a second set for all arcs none of the physical conditions should be violated. The control of active elements can be modeled as here-and-now variables at the first stage and the realization of physical states as wait-and-see variables at the second stage. The realization of physical states takes place after uncertain parameters realize themselves. As already mentioned, the uncertain parameters are the demand and the physical parameters, for the latter more precisely the pressure loss coefficients, which are given by the uncertainty in the frictions of the pipes. For every possible realization of the pressure loss coefficients, physical feasibility of the gas transport has to be maintained by the network operator. In order to let a value of Δ_a cause a pressure increase of Δ_a at the compressor $a \in \mathcal{A}_c$, a linear compressor model is used. This yields the robust optimization problem:

$$
\begin{aligned}
\min_{\Delta \in [\underline{\Delta}, \overline{\Delta}]} \max_{(d,\lambda)\in\mathcal{U},\pi,q} \quad & w(\Delta) + \sum_{v \in V} \max\{0, \underline{\pi}_v - \pi_v, \pi_v - \overline{\pi}_v\} \\
\text{s.t.} \quad & Mq = d \ , \\
& (M^\top \pi)_a = \Delta_a \qquad \forall a \in \mathcal{A}_c \ , \\
& (M^\top \pi)_a = -\lambda_a q_a |q_a| \qquad \forall a \in \mathcal{A}_{pi} \ , \\
& (q, \pi) \in \mathbb{R}^{|\mathcal{A}|} \times \mathbb{R}^{|\mathcal{V}|} \ .
\end{aligned} \tag{5.5}
$$

Thereby, with the under- and overlined variables as fixed lower and upper bounds on the corresponding variables the uncertainty set $\mathcal{U}$ is defined as

$$
\mathcal{U} := \left\{ (d, \lambda) \,\middle|\, \lambda \in [\underline{\lambda}, \overline{\lambda}], d_i \in [\underline{d}, \overline{d}], \sum_{i=1}^{n} d_i = 0 \right\} \ .
$$

Fixing the uncertain parameters to some values, there is a unique physical state, i.e., unique flow and pressure variables, that fulfills the physical constraints [6, 22]. Due to this fact, the optimization problem (5.5) can be reformulated as a box-constrained optimization problem by writing the pressure as a function of the other parameters (see [71]):

$$
\min_{\Delta \in [\underline{\Delta}, \overline{\Delta}]} \max_{(d,\lambda)\in\mathcal{U}} \quad w(\Delta) + \sum_{v \in V} \max\{0, \underline{\pi}_v - \pi_v(\Delta; d, \lambda), \pi_v(\Delta; d, \lambda) - \overline{\pi}_v\} \ .
$$

To solve such a problem, Martina Kuchlbauer has developed an adaptive bundle method during her PhD phase, see [71]. Therefore, the bundle method is applied to the outer minimization problem with the optimal value function of the inner maximization problem as its objective function. As in the bundle method, an approximate function evaluation is required, the inner maximization problem has to be solved approximately in every iteration. This inner *adversarial problem* is the following nonconvex constrained optimization problem:

$$
\begin{aligned}
\max_{(d,\lambda)\in\mathcal{U},\pi,q} \quad & \sum_{v \in \mathcal{V}} \max\{0, \underline{\pi}_v - \pi_v, \pi_v - \overline{\pi}_v\} \\
\text{s.t.} \quad & Mq = d \ , \\
& (M^\top \pi)_a = \Delta_a \qquad \forall a \in \mathcal{A}_c \ , \\
& (M^\top \pi)_a = -\lambda_a q_a |q_a| \qquad \forall a \in \mathcal{A}_{pi} \ , \\
& (q, \pi) \in \mathbb{R}^{|\mathcal{A}|} \times \mathbb{R}^{|\mathcal{V}|} \ .
\end{aligned} \tag{5.6a}
$$

In [71] this adversarial problem is solved approximately via piecewise linear relaxation. The adaptive bundle method only allows for a certain error in the optimal objective value. As a relaxation that fulfills a requested error bound, for each of the pressure loss constraints (5.6a), piecewise linear relaxation via the delta method [6, 38, 83] is used.

From a set of given interpolation points x_i with corresponding interpolation values y_i, an auxiliary result can be used to convert the corresponding piecewise linear interpolation function into Abs-Linear Form. The next lemma is taken from a not yet published article, which was kindly provided by the authors [42].

Lemma 5.6. *In terms of the values y_0, y_n the outer slopes s_0, s_{n+1}, and the inner slopes $s_i = (y_i - y_{i-1})/(x_i - x_{i-1})$ for $i = 1, \ldots, n$ the interpolant $y(x)$ is given by*

$$y = \frac{1}{2}\left[y_0 + s_0(x - x_0) + \sum_{i=0}^{n}((s_{i+1} - s_i)|x - x_i|) + y_n + s_{n+1}(x - s_n)\right] .$$

Here, the two linear functions at the beginning and the end can be combined to $[y_0 - s_0 x_0 + y_n - s_{n+1}x_n + (s_0 + s_{n+1})x]/2$.

In the adaptive bundle method [71], an error bound on the optimal objective value of the adversarial problem is requested and a consequent bound for the error in the pressure loss constraints is provided. As this theoretical bound turned out to be not very tight, the strategy in [71] is to allow for large errors in the constraints and to refine in case of a too large a posteriori error in the objective (see [71, Section 5.1.1, 5.1.2]). As also described in [71], solving the adversarial problem up to the requested error has the biggest impact on the run time. To solve the piecewise linearly relaxed adversarial problem mixed-integer programming (MIP)-Solvers, e.g., Gurobi [52], are used. Therefore, a major motivation is the reduction of the runtime of an inner solver for these piecewise linear problems. In particular, a method that allows for warm start strategies has the potential to speed up computations, because of the use of the refinement strategy. Therefore, sequences of refined relaxations are solved. This sequence of solutions of refined relaxation is referred to as a *cascade*. In the bundle method as described above, the MIP is completely solved anew after each refinement, without the old solution having any influence since so far no warm start strategy is known for MIP solvers. In contrast to that, CASM perform a warm start for the inner loop. That is, if the inner problem is solved for a given discretization, a new starting point for the next model, with a finer discretization, is calculated with the help of the previous solution. For this purpose, the calculated values for demand d and pressure loss coefficient λ are taken from the solution and new starting values for pressure and flow are determined for the refined model. This step coincides with the one for the starting point, i.e., the coarsest discretization.

The application of CASM as a solver for the single piecewise linear problem is studied numerically in [67]. The results shown here, represent a further improvement compared to [67] due to additionally exploiting sparsity of the matrices and efficient updating of the system matrix, as described in Section 5.1.

5.3.1 GasLib-Instances

As a basis for data, a library of realistic gas network instances [88, 95] is used. In [67], results on instances GasLib-11, GasLib-40 and GasLib-134 have already been presented and in this thesis, these test cases and additionally the instance GasLib-582 are considered. The respective number in the name of the problem instance indicates the number of nodes in the network. Over all of them, different aspects are considered from instance to instance. On the one hand, different choices of the compressor control Δ are investigated. One option here is to use randomly chosen controls, which may be not robust feasible. Another option is to consider a robust feasible control implying that the optimal value of the adversarial problem is equal to 0. On the other hand, results for different choices of the piecewise linear relaxations will be shown. That is to impose different allowed error tolerances up to which the relaxed problem deviates from the original one in terms of the nonconvex pressure loss constraints leading to different discretizations in the piecewise linear approximation of the nonconvex term. As described above, the error bounds are possibly refined during one iteration of the outer bundle method. This results in an applicability of CASM for such a cascade of refinements. Finally, physically useful values are considered as feasible starting points for CASM for most instances, but physically infeasible starting values are also used for the penalty approaches from Section 2.3 to determine a feasible starting point.

The data sets used for the optimizations are from random iterations of the adaptive bundle method and were provided by Martina Kuchlbauer. The goal is to show the successful application of CASM for such optimization problems. A complete integration of the method as an inner solver in the adaptive bundle method is the subject of future research.

All network plots in the following subsections are given by internal data of the *Sonderforschungsbereich/Transregio 154 Mathematical Modelling, Simulation and Optimization using the Example of Gas Networks* (project ID: 239904186) [97]. There input nodes are marked in blue, output nodes red, inner nodes black, pipes and short pipes black, compressor stations red, resistors cyan, valves green, and control valves are orange.

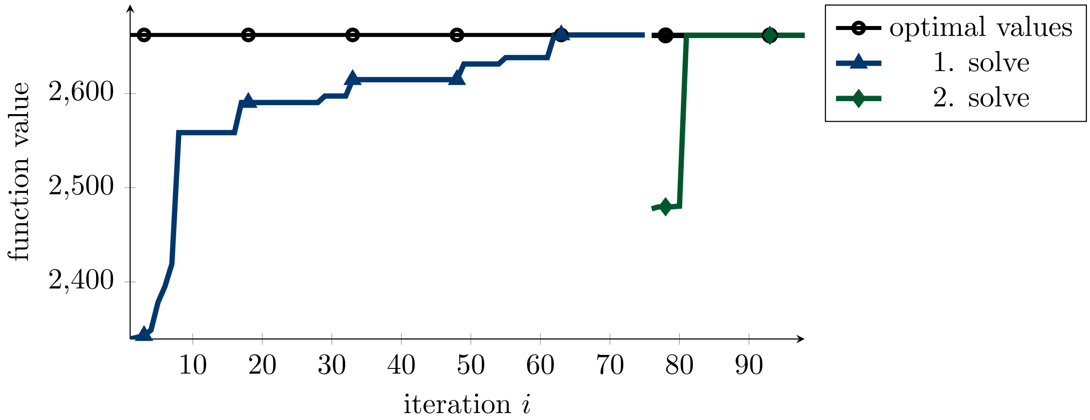

Figure 5.4: Optimization history of CASM for the GasLib-11 cascade

GasLib-11

To begin with, consider the small and academic test instance GasLib-11. The topology of it is shown in Figure 5.3. For this instance the adversarial problem (5.6) with an initial compressor control for two typical sizes of given error bounds is solved.

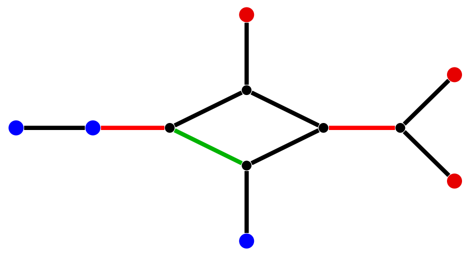

Figure 5.3: Topology of GasLib-11

Figure 5.4 shows the development of the function values during the optimization runs using CASM for two discretizations of the nonconvex function (5.6a). The blue line depicts the function values for the coarse discretization (1. solve) and the green line the function value for the fine discretization (2. solve). The black line represents the optimal value, which was determined by the solver Gurobi [52] for the respective problem. This line does not depend on the number of iterations, but only represents the target to be reached. For this example, both the first and the second discretization have the same optimal objective function value. Here, one can see that at the end of each optimization the function value does not increase further, but the algorithm still needs some iterations to terminate. The reason for this behavior is the regularization term which is used in CASM. In the graphic the function value is shown without the regularization term. If one would add this term, it can be seen that the value still increases slightly until the algorithm terminates.

Table 5.5 lists some essential parameters for the respective problems. The number of variables, the equality and inequality constraints remain the same in both optimizations, but a finer approximation of the Eq. (5.6a) increases the number of switching variables and thus the dimension of the saddle point system (4.25).

For the first optimization CASM required 75 iterations and for the second one 23 additional iterations are needed. In Figure 5.4 it can be seen from the beginning of the green line that after the refinement and thus after the warm start the optimal function value is not directly reached again. This is due to the fact that - as described at the beginning of this section - the initial values for the pressure and the flow are recalculated using the solution for the demand and the pressure loss coefficients from the first optimization and do not coincide with those of the previous solution. For this instance, the optimal values for both relaxations coincide. For some of the next instances it will be seen that this does not always need to be the case.

relaxation	1.	2.
variables n	44	
equal. const. m	19	
inequal. const. p	70	
switching variables s	175	183
rows/columns of saddle point sys.	484	500
iterations	75	23

Table 5.5: Complexity of GasLib-11 instances for their respective optimizations and iterations needed by CASM

The plots and tables in the following subchapters are structured very similarly to those here, thus the description will be limited to the essential changes.

GasLib-40

After the GasLib-11, now with the GasLib-40 (cf. Figure 5.5) a somewhat larger instance is considered in two different settings and in cascades of three.

First, the adaptive bundle method was applied with an uncertainty set for demand d and pressure loss coefficients λ that is $[0.95d, 1.05d] \times [\lambda, 1.1\lambda]$. For this case, multiple refinements of the relaxation of the adversarial problem are requested in the bundle method's last iteration, when a robust feasible compressor control is investigated.

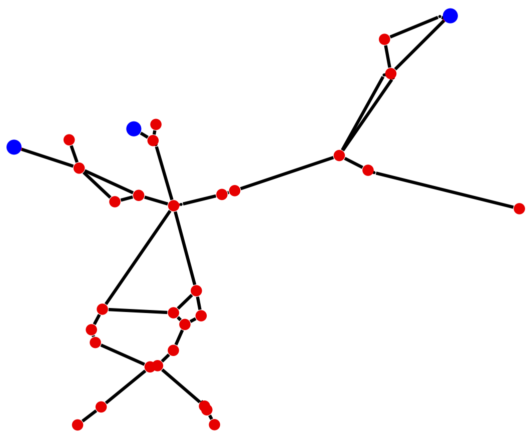

Figure 5.5: Topology of GasLib-40

	nonrobust feasible			robust feasible		
relaxation	1.	2.	3.	1.	2.	3.
variables n	170					
equal. const. m	54					
inequal. const. p	314					
switching variables s	331	341	573	315	315	319
rows/columns of saddle point sys.	1206	1226	1690	1174	1174	1182
iterations	965	620	707	621	213	213

Table 5.6: Complexity of GasLib-40 instances for their respective optimizations and iterations needed by CASM

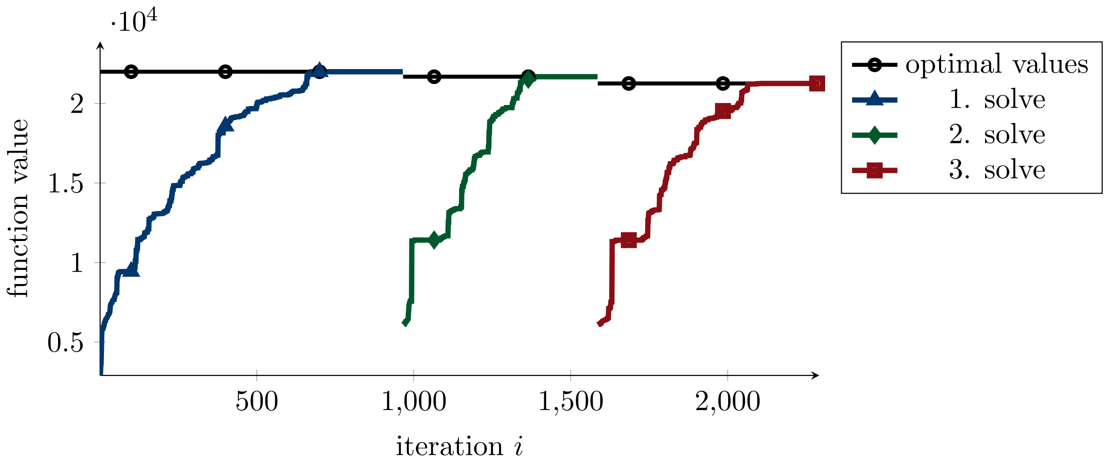

Figure 5.6: Optimization history of CASM for the GasLib-40 cascade

Second, the uncertainty set is enlarged to $[0.9d, 1.1d] \times [\lambda, 1.5\lambda]$. In this case, multiple refinements are requested in an earlier iteration of the bundle method in which the compressor control is not robust feasible. Therefore, for this instance it is distinguished between a nonrobust feasible and a robust feasible control.

Analogous to the previous subsection, the main quantities for this problem are given in Table 5.6. In addition, two different plots are shown for the nonrobust feasible problem in Figure 5.6 and 5.7. The first one shows, analogous to the plot for GasLib-11 (cf. Figure 5.4), the development of the function values over the iterations but in this case for the three individual optimizations. In contrast to the calculation for GasLib-11, however, the optimal function value decreases after the refinement.

To highlight the effect of the warm start strategy, refer to Figure 5.7. Here, on the one hand, the function values of the iterates corresponding to the third optimization from Figure 5.6 is shown in red (warm start). On the other hand, the function values for the same problem is shown in sand color (cold start) this time with the initial values from the first optimization from the three-part cascade, where 1109 iterations are needed. This shows that even the most refined problem can be solved with the original starting value, however, clearly more iterations are needed and due to the larger system matrix the cost per iteration is larger and this is less useful. The effect is even more noticeable for the GasLib-134 instance, as can be seen in the next subsection.

In addition, from a technical point of view, the considered model with the adapted finer discretization can only be generated if the solution of the previous one is known. Otherwise, a refinement that leads to the same a posteriori error would be even more complex to solve (see [67, 71]) which also supports the warm start strategy proposed here.

Despite the fact that the models become more complex with each refinement, the number of iterations does not increase in the same way.

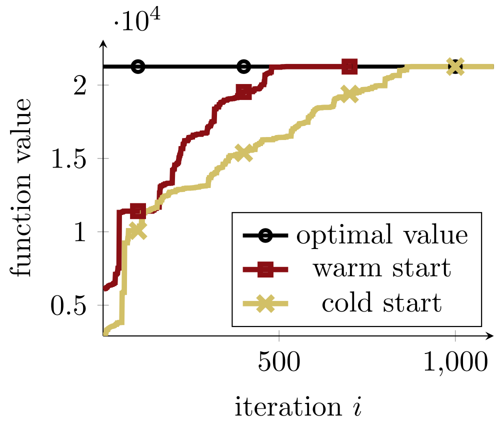

Figure 5.7: Comparison of the third optimization for nonrobust feasible GasLib-40 with and without warm start

Next, the results for the GasLib-40 instance with a robust feasible compressor control are given. For the first two relaxations of the three-part cascade the optimal function value is 0.6965 and for the finest one the optimal function value is 0 corresponding to a robust feasible compressor control. For this setting, CASM needs 621 iterations to solve the first model with the coarsest discretization to reach a local optimum that is not globally optimal. However, since CASM determines only locally optimal points this fits to the theoretical analysis of CASM as described in Section 4.2. The same behavior is also observed for the second relaxation, where the number of iterations is clearly reduced by the warm start (cf. Table 5.6). In the third optimization, however, the global optimum is found again.

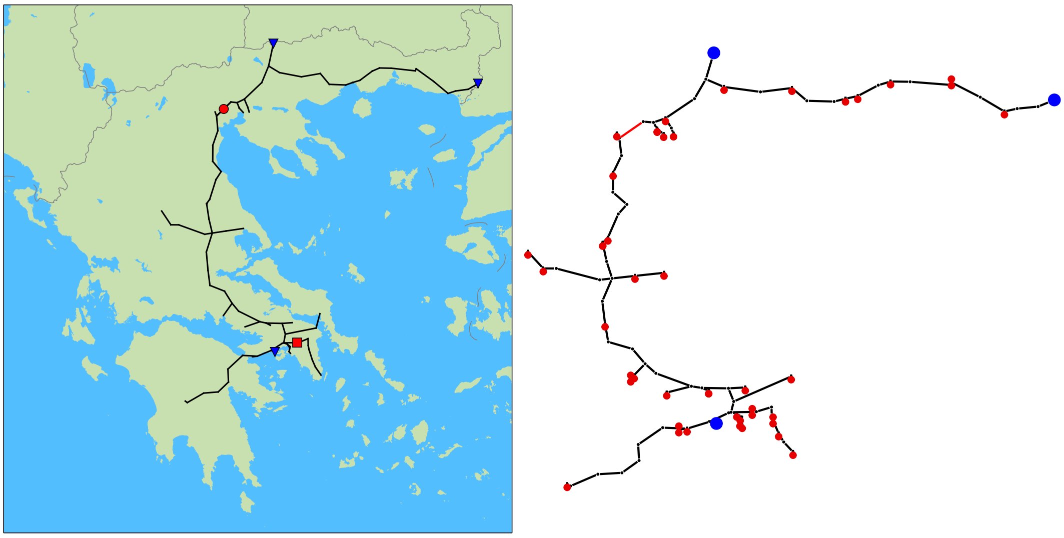

Figure 5.8: Topology of GasLib-134

GasLib-134

GasLib-134 models the gas network of Greece, which has a tree structure, see Figure 5.8. Typically, solvers handle tree structures better, since the gas can flow only in one direction and only along one path to arrive at a certain location.

relaxation	1.	2.	3.
variables n	534		
equal. const. m	230		
inequal. const. p	784		
switching variables s	737	1107	1985
rows/columns of saddle point sys.	3022	3762	5518
iterations	732	337	337

Table 5.7: Complexity of GasLib-134 instances for their respective optimizations and iterations needed by CASM

The adversarial problems (5.6) for GasLib-134 are also taken from a run of the adaptive bundle method, namely for the uncertainty set $[0.8d, 1.2d] \times [\lambda, 2\lambda]$. Again, CASM is applied to multiple refinements that are requested for a compressor control that is not robust feasible.

The progress of the function values is shown in Figure 5.9, while the essential quantities are listed in Table 5.7. In Figure 5.9 the three optimizations with the respective refinements using the warm start are given again (blue, green and red line). A noteworthy observation is that after the warm start in the second and third iteration the optimal function value is already reached, but - as already described for the GasLib-11 - because of the regularization still further iterations are needed before

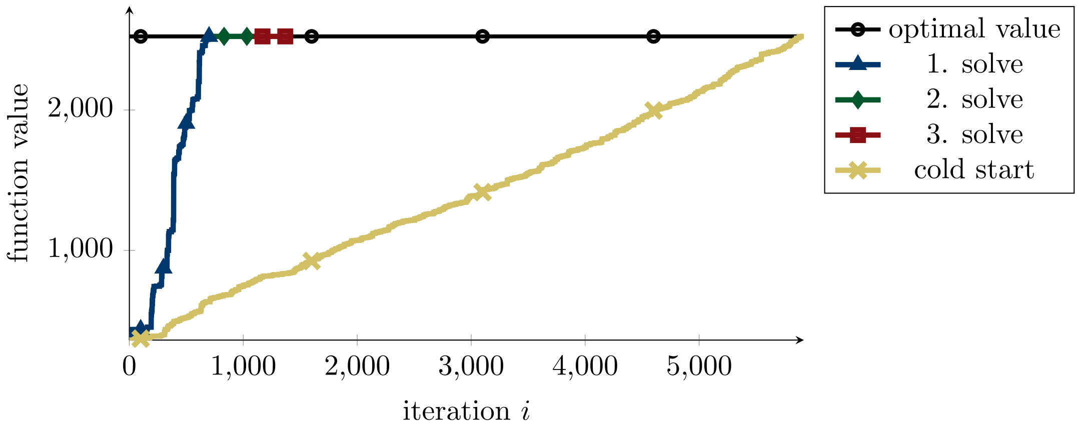

Figure 5.9: Optimization history of CASM for the GasLib-134 cascade

CASM terminates. Moreover, Figure 5.9 shows in sand color, the cold start optimization, i.e., the same optimization problem with the same refinement as in the last part of the three-part cascade was used, but with the same starting value as in the first optimization. Here it can be seen particularly well that approximately four times as many iterations are necessary, namely 5919. Note that in each iteration the largest system of equations from the cascade has to be solved. This is also reflected in the runtime: While the three-step cascade takes about 20 seconds in total, the cold start variant takes about 120 seconds.

GasLib-582

A much larger instance, namely the GasLib-582, whose topology is shown in Figure 5.10, is considered as the last instance.

Again a nonrobust feasible setting is used, so that a nonzero function value is attained in the optimum. Using this instance, the performance of CASM is now compared with a penalty approach. For this purpose, the constraints are added to the objective function via a l_1-penalty, as described in Section 2.3, and the resulting unconstrained and still piecewise linear optimization problem is solved with Algorithm 3, i.e., ASM. For each constraint another nonsmoothness, i.e., an argument in the absolute value, is added resulting in a total number of $m + p + s$ switching variables.

The essential data and the overview of the optimization process are again presented in Table 5.8 and three plots, namely Figure 5.11 for a given feasible starting point and Figures 5.12 as well as 5.13 without a given feasible starting point. Note the different scaling of the function values. In Figure 5.12 the optimal value is of course the same as in

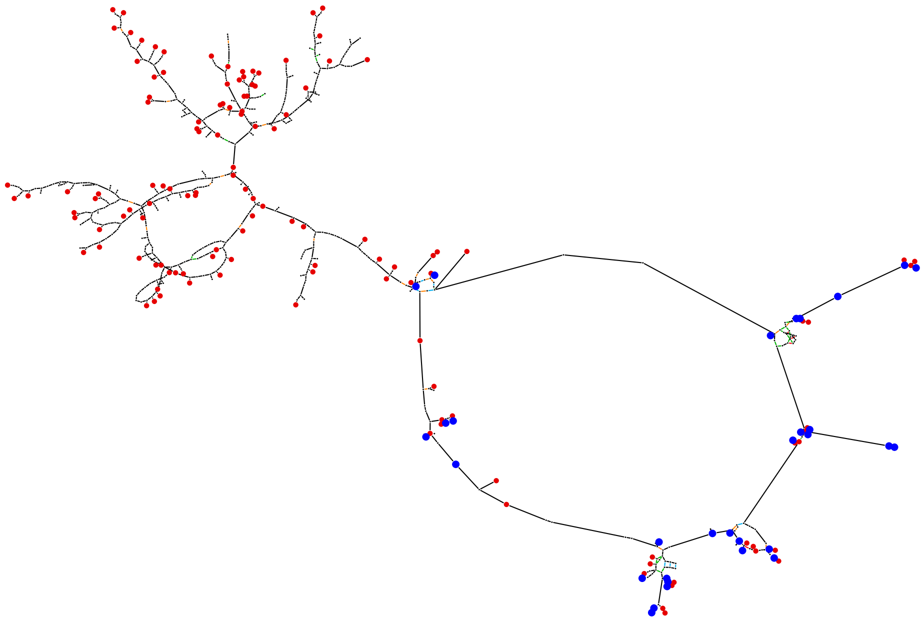

Figure 5.10: Topology of GasLib-582

Figures 5.11 and 5.13 and given by $7.2074 \cdot 10^4$ but because of the scaling it is optically almost close to zero. On the one hand, for an already known feasible starting point, the performance of CASM and ASM combined with the l_1-penalty approach is compared. On the other hand, if no feasible starting point is available, a two-stage strategy is compared with ASM together with the l_1-penalty approach. The two-stage strategy consists in the first step of a Phase I Method as in the representation (2.11), but without the part of the actual objective function to ignore its nonsmoothnesses. After obtaining a feasible starting point, in the second step CASM is used to finally solve the problem. For the comparison without a feasible starting point, the zero vector is used as an initial point. The solution of all approaches is finally compared again with the solution by the optimizer Gurobi, still marked in the figures by the blue lines.

When comparing the optimization with the given starting point (cf. Figure 5.11), it is noticeable that CASM requires significantly more iterations, but is still faster than ASM in combination with the penalty approach. This is due to the fact that the complexity of this problem has reached a size where the regularization term is often adjusted for a better condition number of the system matrix. The effect of this adjustment is to make the residual of the solution of the system of equations sufficiently small, i.e., smaller than a given ε-tolerance. As a heuristic, if the system of equations (4.25) would not be solved

	feasible		nonfeasible		
	CASM	ASM + Penalty	Phase I	CASM	ASM + Penalty
variables n	2382	2382	2383	2382	2382
equal. const. m	1246	-	-	1246	-
inequal. const. p	2832	-	5325	2832	-
switching variables s	2569	6647	2569	2569	6647
rows/columns of saddle point sys.	11598	15676	12846	11598	15676
iterations	3245	1042	5820	3180	8233
runtime (sec.)	2533	3296	30810	3591	5331

Table 5.8: Comparison of the complexity of GasLib-582 instances for their respective optimizations and iterations needed by CASM and ASM with a l_1-penalty approach

sufficiently well enough, the matrix Q is scaled by a factor of 10 and the system of equations is solved again. This results in the system having to be solved twice on average in almost every iteration. Thus, there are almost 2000 solutions of the saddle point system, which in combination with the larger dimension of the system leads to a larger runtime. Both approaches yield the same result at the end, however the explicit treatment of the constraint as in CASM lead to a better numerical stability, in particular with optimization problems which are more complex by the dimensions.

Considering the setting in which one starts with an infeasible starting point and therefore first has to perform a Phase I Method to find an feasible point, the performance comparison changes significantly. Compared to all other optimizations, the Phase I Method requires a significantly larger number of iterations, a multiple of this again in the number of solutions of the saddle point system and thus a much higher runtime by about 30 000 seconds $\approx$ 8.3 hours (cf. Table 5.8). Numerically it could be observed that a main reason for this is the upcoming numerical instability. Therefore, the system of equations has to be solved very often after adjusting the regularization parameter q. Hence, this problem is not particularly robust to implementation-typical choices of tolerance parameters. In addition, it was necessary here to shift the regularization term for reasons of numerical stability. For this the current iteration was used, as it was described at the end of Subsection 4.2.5. Without this shift, a function value in the order of 10^{-4} was already reached, but then a cycle occurred, in which a constraint was added and dropped repeatedly.

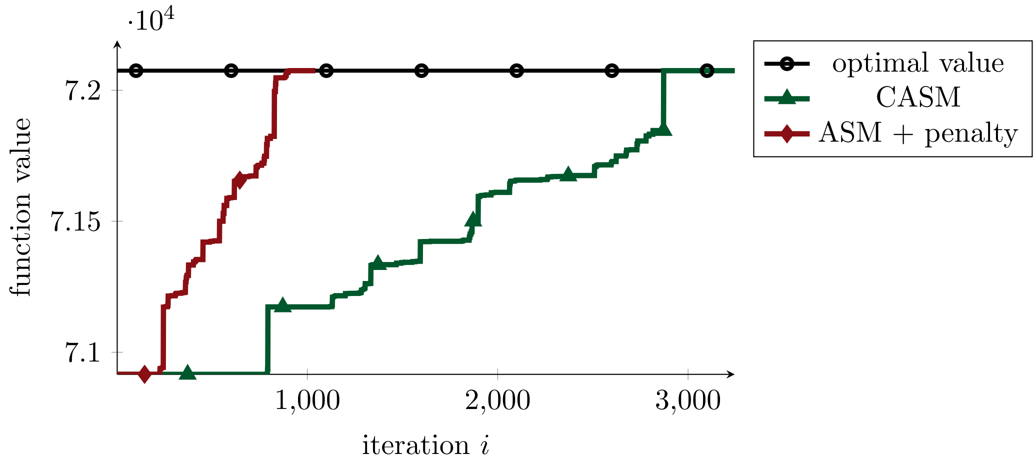

Figure 5.11: Optimization history of CASM for GasLib-582 with feasible starting point

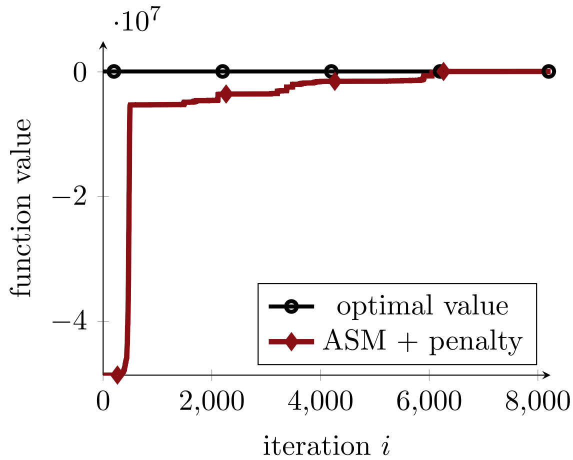

Figure 5.12: Optimization history of ASM for GasLib-582 with nonfeasible starting point

Figure 5.12 shows the history of the function values per iteration for the solution with ASM combined with the penalty approach. In Figure 5.13, on the other hand, the course of the function value of Phase I for finding a feasible starting point (in sand color) is shown on the left side. Since the target function is initially ignored here, the value of the violation of the constraint can be seen, here scaled logarithmically. In order to reach a feasible starting point, this must therefore become zero. As starting point for Phase I the zero vector is chosen. Furthermore, on the right side the subsequent history of the function values for the original optimization problem with CASM as solver using the starting value calculated from Phase I (in green) as well as the optimal value determined by Gurobi (in black) are shown. Using this combination of the Phase I and CASM, a total of about 9000 iterations are needed. In contrast, due to the Phase I Method, an even larger number of solutions of the equation system (4.25) and thus also the above mentioned relatively long runtime is observed.

In summary, this instance shows that numerically, the choice of CASM as an optimization algorithm with the explicit handling of the constraints is most stable if the user has access

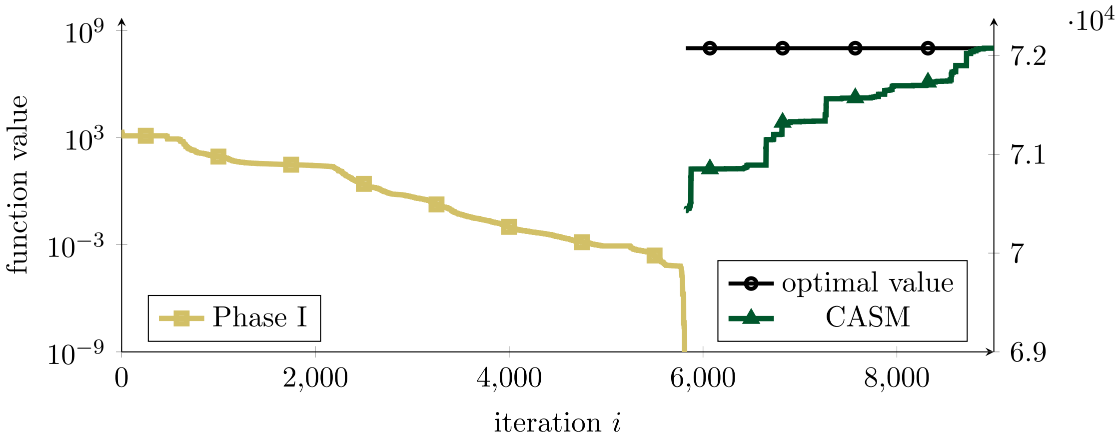

Figure 5.13: Optimization history for the Phase I and CASM of the GasLib-582 using the zero vector as nonfeasible starting point

to a feasible starting point. If the latter is not the case, a Phase I Method can be used to determine it, but this results in greater numerical instability for problems that are too complex. In this case, it is recommended to use ASM with a l_1-penalty approach to handle the constraints. In any case, the optimal solution has been found in the end, which was also found by the comparison algorithm Gurobi. The possible problems for Phase I Methods or penalty approaches as described in Section 4.3 did not occur here.

5.4 Piecewise Linear Regression in Retail

A subproblem from the retail industry marks the final example in this chapter. It arises when solving retail portfolio maximization problems.

The retail industry is governed by crucial decisions on inventory management, discount offers (promotions), and stock clearing (markdowns). These present a two-step optimization problems. The first step is an estimation problem, where the underlying objective is to predict the coefficients of demand (consumer demand/sales) elasticity with respect to product prices. This is the one that will be considered here. The second step is the dynamic revenue maximization problem, that takes in the coefficients as inputs [62, 63]. While both present nonsmooth optimization problems, the latter is a challenging nonlinear problem in high dimensions. This is further subject to constraints on inventory, inter-product relationships, and price bounds. Because of the nonlinearities, this is not a subject of this work but of future research.

Here the estimation problem is discussed. The consumer demand is usually nonlinear and nonsmooth. Therefore, the demand d as a function of price p at a time $t \in \{1, \dots, T\}$ and for a product $i \in \{1, \dots, N\}$ is modeled by a piecewise smooth function of the form

$$d_i^t(p_i^t) = \max_j (h_j(p_i^t)) \ ,$$

with finitely many smooth functions $h_j : \mathbb{R} \to \mathbb{R}$, see [62, 63]. To approximate this nonlinear demand function, there has been an extensive study in literature to multiple demand models, e.g., see [20, 74, 100]. One possibility is to represent this as a piecewise linear function of the form

$$d_i^t(p_i^t) = \max(a_i - b_i p_i^t, 0) \ , \tag{5.7}$$

where a_i and b_i are the coefficients to be determined. For this, there are anonymized, but publicly available real data from a retailer [19]. This retailer has 44 products/stock-keeping-units and over 4400 entries with individual columns for price and sales (besides others). This has been used as one of the standards for demand forecasts, e.g., in [21, 25]. The timestamps of the records cover about 100 weeks between the years 2016 to 2018.

In order to determine the coefficients for Eq. (5.7) using these data, the following piecewise linear optimization problem is considered

$$\begin{aligned} \min_{a,b \in \mathbb{R}^N} \quad & \sum_{i=1}^{N} \sum_{t=1}^{T} \left| \max(a_i - b_i p_i^t, 0) - d_i^t \right| \\ \text{s.t.} \quad & a, b \geq 0 \ . \end{aligned} \tag{5.8}$$

Since the ith coefficient depend only on the associated price p_i and demand d_i and not on the others, i.e., not on p_j and d_j for $j \neq i$, for all $i = 1, \dots, N$, this can also be viewed as N individual optimization problems of the form

$$\begin{aligned} \min_{a_i, b_i \in \mathbb{R}} \quad & \sum_{t=1}^{T} \left| \max(a_i - b_i p_i^t, 0) - d_i^t \right| \\ \text{s.t.} \quad & a, b \geq 0 \ . \end{aligned} \tag{5.9}$$

Table 5.9 shows the essential quantities for the complexity of the optimization problems (5.8) and (5.9) each with the zero vector as starting point. Note that the latter is solved 44 times, each corresponding to one of the 44 different products, one after the other. The shown

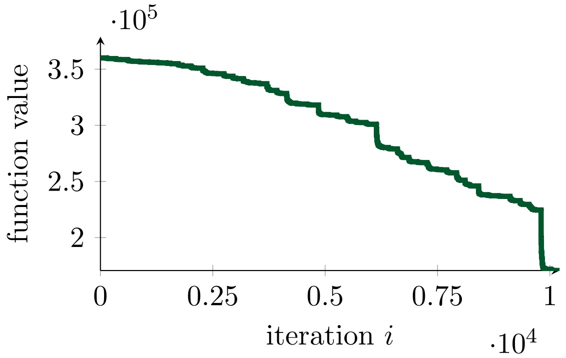

Figure 5.14: Optimization history of CASM for the optimization problem (5.8)

optimization problem	(5.8)	(5.9)
variables n	88	2
equal. const. m	0	0
inequal. const. p	88	2
switching variables s	8625	197
rows/columns of saddle point sys.	17426	398
iterations	10215	10303
runtime (sec.)	765	27

Table 5.9: Comparison of the complexity of the optimization problems (5.8) and (5.9) as well as iterations needed by CASM

running time is the total time for all 44 runs together. So it is clear to see that although the number of iterations is almost the same, the runtime is much faster. This is of course due to the much larger system matrix when all products are considered simultaneously. The history of the function values of (5.8) in each iteration is shown in Figure 5.14.

This example thus shows another field of application for CASM, namely solving piecewise linear regression problems, here using the example given by the retail industry.

6
Conclusions and Outlook

This concluding chapter will now be used to briefly summarize the most important results of this thesis once again (see Section 6.1). In addition, a selection of topics follows, which result from this work as further research fields (see Section 6.2).

6.1 Summary

In this thesis, an optimization algorithm, called Constrained Active Signature Method (CASM), for solving piecewise linear optimization problems with piecewise linear equality and inequality constraints was presented and its performance was tested using various academic as well as real-world application examples. For this purpose, it was assumed that the optimization problems are given in the so-called Abs-Linear Form, a matrix-vector representation. With the help of this form it was possible to decompose the primal image space into finitely many polyhedra. As a first step, the constraints were not considered and the Active Signature Method (ASM) from [45] was introduced and optimality conditions for the unconstrained case were given, which can be verified in polynomial time. The basic idea of ASM is to use polyhedral decomposition to solve linear subproblems on individual polyhedra and to use the optimality condition to choose the next polyhedron if optimality is not satisfied. In addition, and in comparison to the previously published literature, missing details have been added in various places. This includes in particular the finite convergence of ASM and the related assurance of a descent in the function value.

Thereafter, ASM was used as the basis for the development of CASM. For this purpose, the constraints in an active set sense were additionally taken into account, resulting in potentially more polyhedra in the decomposition, depending on the nonsmoothness in the constraints, and additional linear constraints on the polyhedra. The algorithmic procedure is then similar to ASM. First, as in Active Set Methods [87], the inequality constraints are considered as active or inactive. To switch again from polyhedron to polyhedron, the optimality conditions are used again to decide which constraints or kinks are to be activated or deactivated. It was shown, that these conditions can also be checked again in polynomial time in the constrained case. This makes it unnecessary to check whether, e.g., the zero vector lies in some kind of subdifferential, as required by other optimality conditions for nonsmooth optimization. For both algorithms, ASM as well as CASM, it was possible to show convergence in finitely many steps.

In order to test the numerical performance of CASM, the algorithm was implemented in Matlab and tested on various, both academic and real-world, application examples and compared with other optimization algorithms. Problems of smaller scale, including bi-level problems or those with linear complentarity constraints, were successfully solved. As examples with real-world background, subproblems from an adaptive bundle method [71] used to solve gas network optimization problems were studied. Therefore, different instances and scenarios were considered. Furthermore, a possible warm start strategy was shown, which can be used to solve equivalent formulations of mixed integer problems.

Here it was seen that for large optimization problems, i.e., those in which the system matrix of the system of equations to be solved in each iteration is of the order of approximately 15000x15000, first numerical inaccuracies appear. From this observation it can be concluded that, at least as a Matlab implementation, small to medium sized optimization problems can be solved. Dimensionally very large optimization problems, on the other hand, have been challenging, but can still be solved in the order of magnitude considered here. However, other programming languages or also improvements in the Matlab implementation itself could provide a more stable method.

6.2 Future Research Directions

To conclude this thesis, now a few possible further research aspects will be discussed.

Solving the saddle point system The main task, which also accounts for most of the runtime of the CASM implementation, has turned out to be solving the system of equations (4.25). In Lemma 4.9 it has already been shown that this can be reduced to a smaller system (4.26), but is more nested in the dependence of the changing matrices Σ and Ω. I.e., these matrices are usually not only multiplied from one side to others but occur more frequently in each block, so that a change of these generally changes more entries in the system matrix itself. Possibly low rank updating can be used here, as known for example from Quasi-Newton methods [37, 39, 87]. Of course, it would be ideal if the previous solution of the system of equations could be reused and the solution from the next iteration could be determined directly without solving the system itself. However, this seems highly unlikely, since to the best of the authors' knowledge no comparable methods are known, in which this could be used.

Globalization strategies Another research direction consists in globalization strategies. So far, both ASM and CASM are "only" able to find local minima. Of course, they can also be global minima, or if one adds further assumptions to the problem, this can be guaranteed (see Section 4.3). But the goal should be to find global solutions for problems without making additional assumptions and to ensure this mathematically. Due to the decomposition of the feasible set into finitely many polyhedra, strategies known from mixed-integer optimization [76] may be useful. Because of the fact that only finitely many polyhedra exist, one can, for example, number them consecutively and interpret this numbering as an integer variable. In the worst case, of course, one can also optimize on each polyhedron individually and then compare the function values belonging to the minima there. This results in the optimization problem

$$\begin{aligned}
\min_{\sigma\in\{-1,1\}^s} \quad \min_{x\in\mathbb{R}^n, z\in\mathbb{R}^s} \quad & a^\top x + b^\top z \\
\text{s.t.} \quad & 0 = g + Ax + Bz + C\Sigma z \ , \\
& 0 \geq h + Dx + Fz + G\Sigma z \ , \\
& z = c + Zx + Mz + L\Sigma z \ , \\
& 0 = (I_s - |\Sigma|)z \ , \\
& 0 \leq \Sigma z \ ,
\end{aligned}$$

where the inner problem is a linear optimization problem. This can be solved by using Simplex algorithms [85] or, in an unbounded case by an Active Set Method together with a regularization [87]. The outer minimization problem then gives a finite number of inner optimization problems. Note that in this thesis it was avoided to solve the linear optimization problems one by one which would lead to 2^s optimizations of the inner problem. Rather, the goal was to use the information given by the Lagrange multipliers to select the next polyhedron as cleverly as possible and, as seen in the examples in Chapter 5, to reduce the number of inner optimization tasks. Therefore, in principle, there are possibilities for globalization strategies, the challenge is to make the process as efficient as possible.

Solving constrained abs-smooth optimization problems In Chapter 3 the SALMIN Algorithm [28, 29] was already mentioned, which uses abs-linearization to generate an abs-linear model of an abs-smooth function and then, for example, ASM acts as a solver on the model. In a first step, CASM could also be integrated here, so that abs-smooth optimization problems with piecewise linear constraints can be solved. As a second step, however, it becomes more complicated if additional general abs-smooth constraints are to be considered. These could then also be abs-linearized with the help of the AD techniques as described at the beginning of this thesis. However, the problem arises that the linearization changes the feasible set and one does not get a consistent feasible point method anymore. In order to be able to guarantee this further, for example, projections back on the feasible set are necessary, if this set is left, or one could work with filter methods [33, 103, 105]. In comparison to the update step given by Eq. (3.41) for the unconstrained case, one option is to update in the constrained case via

$$\begin{aligned} \tilde{x}^{+} &= x + \underset{\Delta x \in \mathcal{F}_{\text{abs}}(x)}{\arg\min} \left\{ \Delta F(x; \Delta x) + \frac{\rho}{2} ||\Delta x||^2 \right\} \,, \\ x^{+} &= \operatorname{Proj}_{\mathcal{F}}(\tilde{x}^{+}) \,, \end{aligned}$$

where $\mathcal{F}$ denotes the feasible set, $\mathcal{F}_{\text{abs}}(x)$ the abs-linearization of the feasible set obtained from the base point x and $\operatorname{Proj}_{\mathcal{F}}(x^{+})$ the projection of x^{+} on the set $\mathcal{F}$.

Bagirov et al. listed a survey of various academic test problems in [7, Section 9]. One of them is the so-called MAD1 from [79, Example 1]. This is a nonsmooth linear constrained optimization problem, which will be used here to show in a simple example that the SALMIN approach with constraints can in principle be successful. Since this example has only one linear constraint, no projection is needed for this at first.

Example 6.1. *As stated in [79, Example 1], first consider the linear optimization problem*

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & \max\{x_1^2 + x_2^2 + x_1 x_2 - 1, \sin(x_1), -\cos(x_2)\} \\ s.t. \quad & -x_2 - x_1 + 0.5 \leq 0 \,. \end{aligned}$$

Second, consider the slightly modified optimization problem with a piecewise linear constraint

$$\begin{aligned} \min_{x \in \mathbb{R}^2} \quad & \max\{x_1^2 + x_2^2 + x_1 x_2 - 1, \sin(x_1), -\cos(x_2)\} \\ s.t. \quad & -x_2 + \max\{-x_1 + 0.5, -4x_1 - 1.75\} \leq 0 \,. \end{aligned}$$

This is constructed in such a way that the optimal point is the same as in the first example,

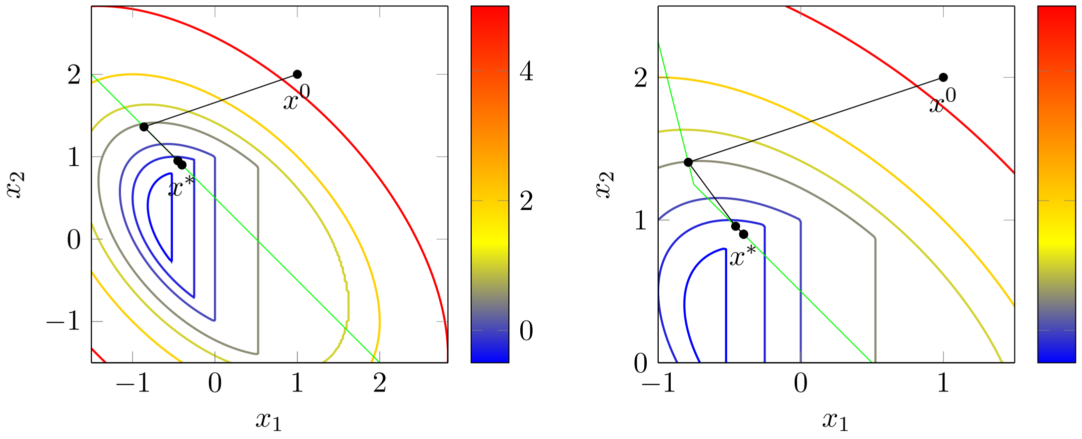

Figure 6.1: Illustration of the iteration sequences generated by SALMIN for Example 6.1

but the iteration sequence has to change, because not all iteration points from the first one are feasible anymore. Of course, the piecewise linear constraint from the second problem can also be formulated as two single linear constraints, but here a piecewise linear representation is explicitly considered. Using for both problems the starting point $x^0 = (1, 2)$ *Figure 6.1 shows two plots of the iteration sequence which are generated by SALMIN. For this purpose, the contour lines of the objective function in the* x_1*-*x_2*-plane are shown, and the iterations are marked as points. The green straight lines indicates the constraint, all points above them are feasible. In total, in both cases 6 (outer) SALMIN iterations are needed to find the optimal point* $x^* = (-0.4003, 0.9003)$ *according to the literature.*

Frank-Wolfe Algorithm for abs-smooth functions on convex and compact domains

Similar to the previous approach and as mentioned at the end of Section 3.4, the generalization to abs-smooth functions of the so-called *Frank-Wolfe* algorithm (or *conditional gradient*) [36, 75] is also the subject of current research. First results of this research, in which the author is also involved, can be found as a report in [68]. The setting is somewhat different from the one considered in this thesis. In the Frank-Wolfe algorithm for abs-smooth functions, general abs-smooth functions are allowed in the objective function and not only piecewise linear ones. On the other hand, the feasible set, denoted in the following by $C \subseteq \mathbb{R}^n$, is required to be convex and compact. These two requirements do not play any role at all in this thesis.

Classical Frank-Wolfe algorithms for smooth objective functions $f : \mathbb{R}^n \to \mathbb{R}$ compute in every iterate i a direction $v_i \in \arg\min_{v \in C} \nabla f(x_i)^\top v$, which corresponds to an optimization along the linearization of the function f. Then, with a step size $\alpha_i \in (0, 1]$, a linear combination of the previous iterate and the computed solution of the linearized problem determines the new iterated [36, 75]. Therefore, for abs-smooth objective functions, the idea is to replace the linearization by an abs-linearization, which replaces the first step by $v_i \in \arg\min_{v \in C} \Delta f(x_i; v - x_i)$. Depending on the nature of C, this interior problem can then be solved with, e.g., ASM or CASM. First statements on convergence rates, which depend among others on the choice of the step size, as well as numerical tests can be found in [68].

The possible topics listed here show that there are many directions for further research, building on CASM. There are many more, of course, but these are just a few of the ones the author is particularly interested in.

Bibliography

[1] Jean Abadie. On the Kuhn-Tucker theorem. *Nonlinear Programming*, pages 21–36, 1966.

[2] Walter Alt. *Numerische Verfahren der konvexen, nichtglatten Optimierung: Eine anwendungsorientierte Einführung*. Lehrbuch: Mathematik. Vieweg+Teubner Verlag, 2013.

[3] Tadeusz Antczak. The exact l_1 penalty function method for constrained nonsmooth invex optimization problems. In Dietmar Hömberg and Fredi Tröltzsch, editors, *System Modeling and Optimization - 25th IFIP TC 7 Conference, CSMO 2011, Berlin, Germany, September 12-16, 2011, Revised Selected Papers*, volume 391 of *IFIP Advances in Information and Communication Technology*, pages 461–470. Springer, 2011.

[4] Amir Ardestani-Jaafari and Erick Delage. Robust optimization of sums of piecewise linear functions with application to inventory problems. *Operations Research*, 64(2):474–494, 2016.

[5] Daniel Arnström, Alberto Bemporad, and Daniel Axehill. A dual active-set solver for embedded quadratic programming using recursive LDL^T updates. *IEEE Transactions on Automatic Control*, 67(8):4362–4369, 2022.

[6] Denis Aßmann, Frauke Liers, and Michael Stingl. Decomposable robust two-stage optimization: An application to gas network operations under uncertainty. *Networks*, 74(1):40–61, 2019.

[7] Adil Bagirov, Napsu Karmitsa, and Marko M Mäkelä. *Introduction to Nonsmooth Optimization: theory, practice and software*, volume 12. Springer, 2014.

[8] Adil M. Bagirov, Manlio Gaudioso, Napsu Karmitsa, Marko M. Mäkelä, and Sona Taheri. *Numerical Nonsmooth Optimization: State of the Art Algorithms*. Springer International Publishing, 2020.

[9] Jonathan F. Bard. *Practical bilevel optimization: algorithms and applications*, volume 30. Springer Science & Business Media, 2013.

[10] Richard H. Bartels and Gene H. Golub. The simplex method of linear programming using LU decomposition. *Commun. ACM*, 12(5):266–268, May 1969.

[11] Pietro Belotti, Christian Kirches, Sven Leyffer, Jeff Linderoth, James Luedtke, and

Ashutosh Mahajan. *Mixed-integer nonlinear optimization*, volume 22, pages 1–131. May 2013.

[12] Ibtihel Ben Gharbia and J. Charles Gilbert. Nonconvergence of the plain Newton-min algorithm for linear complementarity problems with a P-matrix. *Math. Program.*, 134(2, Ser. A):349–364, 2012.

[13] Ksenia Bestuzheva, Mathieu Besançon, Wei-Kun Chen, Antonia Chmiela, Tim Donkiewicz, Jasper van Doornmalen, Leon Eifler, Oliver Gaul, Gerald Gamrath, Ambros Gleixner, Leona Gottwald, Christoph Graczyk, Katrin Halbig, Alexander Hoen, Christopher Hojny, Rolf van der Hulst, Thorsten Koch, Marco Lübbecke, Stephen J. Maher, Frederic Matter, Erik Mühmer, Benjamin Müller, Marc E. Pfetsch, Daniel Rehfeldt, Steffan Schlein, Franziska Schlösser, Felipe Serrano, Yuji Shinano, Boro Sofranac, Mark Turner, Stefan Vigerske, Fabian Wegscheider, Philipp Wellner, Dieter Weninger, and Jakob Witzig. The SCIP Optimization Suite 8.0. Technical report, Optimization Online, December 2021.

[14] Christian H. Bischof, H. Martin Bücker, Bruno Lang, Arno Rasch, and Andre Vehreschild. Combining source transformation and operator overloading techniques to compute derivatives for MATLAB programs. In *Proceedings of the Second IEEE International Workshop on Source Code Analysis and Manipulation (SCAM 2002)*, pages 65–72, Los Alamitos, CA, USA, 2002. IEEE Computer Society.

[15] Richard H. Byrd, Jorge Nocedal, and Richard A. Waltz. *Knitro: An Integrated Package for Nonlinear Optimization*, pages 35–59. Springer US, Boston, MA, 2006.

[16] Martin Bücker and Paul Hovland. autodiff.org, 1999. Online; accessed 14-February-2023.

[17] Coralia Cartis, Nicholas I. M. Gould, and Philippe L. Toint. On the evaluation complexity of constrained nonlinear least-squares and general constrained nonlinear optimization using second-order methods. *SIAM Journal on Numerical Analysis*, 53(2):836–851, 2015.

[18] Frank H. Clarke. *Optimization and nonsmooth analysis.* Canadian Mathematical Society series of monographs and advanced texts, A Wiley-Interscience publication. New York: Wiley, 1983.

[19] Maxime C. Cohen, Paul-Emile Gras, Arthur Pentecoste, and Renyu Zhang. *Demand Prediction in Retail: A Practical Guide to Leverage Data and Predictive Analytics.* Springer Series in Supply Chain Management, 14. Springer Cham, Cham, 1st ed. 2022. edition, 2022.

[20] Maxime C. Cohen, Ngai-Hang Zachary Leung, Kiran Panchamgam, Georgia Perakis, and Anthony Smith. The impact of linear optimization on promotion planning. *Operations Research*, 65(2):446–468, 2017.

[21] Maxime C. Cohen, Renyu (Philip) Zhang, and Kevin Jiao. Data aggregation and demand prediction. *ERN: Statistical Decision Theory; Operations Research (Topic)*, 2019.

[22] Michael Albert Collins, Leon Cooper, Richard Helgason, Jeffery Kennington, and Larry LeBlanc. Solving the pipe network analysis problem using optimization techniques. *Management Science*, 24(7):747–760, 1978.

[23] Frank E. Curtis, Zheng Han, and Daniel P. Robinson. A globally convergent primal-dual active-set framework for large-scale convex quadratic optimization. *Computational Optimization and Applications*, 60(2):311–341, 2015.

[24] Stephan Dempe and Alain Zemkoho. *Bilevel Optimization: Theory, Algorithms, Applications and a Bibliography*, pages 581–672. Springer International Publishing, Cham, 2020.

[25] Yiting Deng, Yuexing Li, and Jing-Sheng Jeannette Song. A unified parsimonious model for structural demand estimation accounting for stockout and substitution. *SSRN*, 2022.

[26] Gabriele Eichfelder. Multiobjective bilevel optimization. *Mathematical Programming*, 123(2):419–449, 2010.

[27] Herbert Federer. *Geometric Measure Theory*. Springer, Berlin, Heidelberg, first edition, 1996.

[28] Sabrina Fiege. *Minimization of Lipschitzian piecewise smooth objective functions*. PhD thesis, University of Paderborn, 2017.

[29] Sabrina Fiege, Andrea Walther, and Andreas Griewank. An algorithm for nonsmooth optimization by successive piecewise linearization. *Mathematical Programming*, 177, 2018.

[30] Sabrina Fiege, Andrea Walther, Kshitij Kulshreshtha, and Andreas Griewank. Algorithmic differentiation for piecewise smooth functions: a case study for robust optimization. *Optimization Methods and Software*, 33(4-6):1073–1088, 2018.

[31] Frank Fischer and Christoph Helmberg. Dynamic graph generation and dynamic rolling horizon techniques in large scale train timetabling. In *ATMOS'10*, pages 45–60. 2010.

[32] Roger Fletcher. *Practical Methods of Optimization.* John Wiley & Sons, New York, NY, USA, second edition, 1987.

[33] Roger Fletcher, Sven Leyffer, and Philippe Toint. A brief history of filter methods. *Preprint ANL/MCS-P1372-0906, Argonne National Laboratory, Mathematics and Computer Science Division*, 36, October 2006.

[34] Anders Forsgren, Philip E. Gill, and Elizabeth Wong. Primal and dual active-set methods for convex quadratic programming. *Mathematical programming*, 159(1):469–508, 2016.

[35] Otto Forster. *Analysis 1.* Springer Spektrum, Wiesbaden, 2016.

[36] Marguerite Frank and Philip Wolfe. An algorithm for quadratic programming. *Naval Research Logistics Quarterly*, 3(1-2):95–110, 1956.

[37] Carl Geiger and Christian Kanzow. *Theorie und Numerik restringierter Optimierungsaufgaben.* Springer-Lehrbuch Masterclass. Springer Berlin Heidelberg, 2002.

[38] Björn Geißler, Alexander Martin, Antonio Morsi, and Lars Schewe. Using piecewise linear functions for solving MINLPs. In Jon Lee and Sven Leyffer, editors, *Mixed Integer Nonlinear Programming*, pages 287–314, New York, NY, 2012. Springer New York.

[39] Philip E. Gill, Walter Murray, and Margaret H. Wright. *Practical optimization.* Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1981.

[40] Xavier Glorot, Antoine Bordes, and Yoshua Bengio. Deep sparse rectifier neural networks. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pages 315–323. JMLR Workshop and Conference Proceedings, 2011.

[41] Andreas Griewank. On stable piecewise linearization and generalized algorithmic differentiation. *Optimization Methods and Software*, 28(6):1139–1178, 2013.

[42] Andreas Griewank, Manuel Radons, Tom Streubel, and Andrea Walther. Representation of piecewise linear functions in abs-linear form of switching depth less than $2n$. unpublished, February 2020.

[43] Andreas Griewank and Andrea Walther. *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation.* Society for Industrial and Applied Mathematics, USA, second edition, 2008.

[44] Andreas Griewank and Andrea Walther. First- and second-order optimality condi-

tions for piecewise smooth objective functions. *Optimization Methods and Software*, 31(5):904–930, 2016.

[45] Andreas Griewank and Andrea Walther. Finite convergence of an active signature method to local minima of piecewise linear functions. *Optimization Methods and Software*, 34(5):1035–1055, 2019.

[46] Andreas Griewank and Andrea Walther. Relaxing kink qualifications and proving convergence rates in piecewise smooth optimization. *SIAM J. Optim.*, 29(1):262–289, 2019.

[47] Andreas Griewank and Andrea Walther. *Beyond the Oracle: Opportunities of Piecewise Differentiation*, pages 331–361. Springer International Publishing, Cham, 2020.

[48] Andreas Griewank and Andrea Walther. Polyhedral DC decomposition and DCA optimization of piecewise linear functions. *Algorithms*, 13(7), 2020.

[49] Andreas Griewank, Andrea Walther, Sabrina Fiege, and Torsten Bosse. On lipschitz optimization based on gray-box piecewise linearization. *Mathematical Programming*, 158:383–415, 2016.

[50] Vidar Gunnerud and Bjarne Foss. Oil production optimization—a piecewise linear model, solved with two decomposition strategies. *Computers & Chemical Engineering*, 34(11):1803–1812, 2010.

[51] Mert Gürbüzbalaban and Michael L. Overton. On Nesterov's nonsmooth Chebyshev-Rosenbrock functions. *Nonlinear Anal.*, 75(3):1282–1289, 2012.

[52] Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual, 2022.

[53] Laurent Hascoët and Valérie Pascual. The Tapenade automatic differentiation tool: Principles, model, and specification. *ACM Transactions on Mathematical Software*, 39(3):20:1–20:43, 2013.

[54] Lisa C. Hegerhorst-Schultchen. *Optimality conditons for abs-normal NLPs.* PhD thesis, Gottfried Wilhelm Leibniz Universität, 2020.

[55] Lisa C. Hegerhorst-Schultchen, Christian Kirches, and Marc C. Steinbach. Relations between Abs-normal NLPs and MPCCs. Part 1: Strong constraint qualifications. *Journal of Nonsmooth Analysis and Optimization*, Volume 2, February 2021.

[56] Lisa C. Hegerhorst-Schultchen and Marc C. Steinbach. On first and second order optimality conditions for abs-normal NLP. *Optimization*, 69(12):2629–2656, 2020.

[57] Michael Hintermüller, Kazufumi Ito, and Karl Kunisch. The primal-dual active set strategy as a semismooth newton method. *SIAM Journal on Optimization*, 13(3):865–888, 2002.

[58] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Convex Functions*, pages 73–120. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001.

[59] Jean-Baptiste Hiriart-Urruty and Claude Lemaréchal. *Convex analysis and minimization algorithms I: Fundamentals*, volume 305. Springer science & business media, 2013.

[60] Xiaolin Huang, Jun Xu, and Shuning Wang. Exact penalty and optimality condition for nonseparable continuous piecewise linear programming. *Journal of Optimization Theory and Applications*, 155(1):145–164, 2012.

[61] Florian Jarre and Josef Stoer. *Optimierung*. Springer-Lehrbuch. Springer-Verlag Berlin Heidelberg, 2004.

[62] Aswin Kannan and Kiran Panchamgam. Computerized promotion and markdown price scheduling, September 2020. US Patent 10,776,803.

[63] Aswin Kannan, Kiran Panchamgam, and Su-Ming Wu. Computerized promotion and markdown price scheduling, January 2020. US Patent 10,528,903.

[64] William Karush. Minima of functions of several variables with inequalities as side constraints. *M. Sc. Dissertation. Dept. of Mathematics, Univ. of Chicago*, 1939.

[65] Ahmet B. Keha, Ismael R. de Farias, and George L. Nemhauser. A branch-and-cut algorithm without binary variables for nonconvex piecewise linear optimization. *Operations Research*, 54(5):847–858, 2006.

[66] Thorsten Koch, Benjamin Hiller, Marc E. Pfetsch, and Lars Schewe. *Evaluating Gas Network Capacities*. SIAM-MOS Series on Optimization. SIAM, March 2015.

[67] Timo Kreimeier, Martina Kuchlbauer, Frauke Liers, Michael Stingl, and Andrea Walther. Towards the solution of robust gas network optimization problems using the constrained active signature method. In Christina Büsing and Arie M. C. A. Koster, editors, *Network Optimization INOC 2022*, 2022.

[68] Timo Kreimeier, Sebastian Pokutta, Andrea Walther, and Zev Woodstock. On a Frank-Wolfe approach for abs-smooth functions. Available at https://opus4.kobv.de/opus4-trr154/frontdoor/index/index/docId/499, 2022.

[69] Timo Kreimeier, Henning Sauter, Tom Streubel, Caren Tischendorf, and Andrea

Walther. Solving least-squares collocated differential algebraic equations by successive abs-linear minimization - a case study on gas network simulation. Available at https://opus4.kobv.de/opus4-trr154/frontdoor/index/index/docId/473, 2021.

[70] Timo Kreimeier, Andrea Walther, and Andreas Griewank. An active signature method for constrained abs-linear minimization. Submitted to Computational Optimization and Applications. Available at https://opus4.kobv.de/opus4-trr154/frontdoor/index/index/docId/474, 2022.

[71] Martina Kuchlbauer, Frauke Liers, and Michael Stingl. Adaptive bundle methods for nonlinear robust optimization. *INFORMS Journal on Computing*, 34(4):2106–2124, 2022.

[72] Harold William Kuhn and Albert William Tucker. Nonlinear programming. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, Hrsg.: J. Neyman. S*, pages 481–492, 1951.

[73] Karl Kunisch and Daniel Walter. On fast convergence rates for generalized conditional gradient methods with backtracking stepsize, 2021.

[74] Timo Kunz and Sven Crone. Demand models for the static retail price optimization problem – a revenue management perspective. In *4th Student conference on operational research*, volume 37, May 2014.

[75] Evgeny S. Levitin and Boris Teodorovich Polyak. Constrained minimization methods. *USSR Computational Mathematics and Mathematical Physics*, 6(5):1–50, 1966.

[76] Frauke Liers, Alexander Martin, Maximilian Merkert, Nick Mertens, and Dennis Michaels. Solving mixed-integer nonlinear optimization problems using simultaneous convexification: a case study for gas networks. *Journal of Global Optimization*, 80:307–340, 2021.

[77] Frauke Liers and Maximilian Merkert. Structural investigation of piecewise linearized network flow problems. *SIAM J. Optim.*, 26(4):2863–2886, 2016.

[78] Phantisa Limleamthong and Gonzalo Guillén-Gosálbez. Rigorous analysis of pareto fronts in sustainability studies based on bilevel optimization: Application to the redesign of the UK electricity mix. *Journal of Cleaner Production*, 164:1602–1613, 2017.

[79] Kaj Madsen and Hans Schjær-Jacobsen. Linearly constrained minimax optimization. *Mathematical Programming*, 14(1):208–223, 1978.

[80] Christopher M Maes. *A regularized active-set method for sparse convex quadratic*

programming. PhD thesis, Stanford University, 2010.

[81] Marko M. Mäkelä, Napsu Karmitsa, and Outi Wilppu. Proximal bundle method for nonsmooth and nonconvex multiobjective optimization. In *Mathematical modeling and optimization of complex structures*, volume 40 of *Comput. Methods Appl. Sci.*, pages 191–204. Springer, Cham, 2016.

[82] Olvi Leon Mangasarian and Stanley Fromovitz. The Fritz John necessary optimality conditions in the presence of equality and inequality constraints. *Journal of Mathematical Analysis and Applications*, 17(1):37–47, 1967.

[83] Harry M. Markowitz and Alan S. Manne. On the solution of discrete programming problems. *Econometrica*, 25:84, 1956.

[84] MATLAB. *9.10.0.1739362 (R2021a) Update 5.* The MathWorks Inc., Natick, Massachusetts, 2021.

[85] John C. Nash. The (dantzig) simplex method for linear programming. *Comput. Sci. Eng.*, 2:29–31, 2000.

[86] John A. Nelder and Roger Mead. A simplex method for function minimization. *The Computer Journal*, 7(4):308–313, January 1965.

[87] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization.* Springer, New York, NY, USA, 2e edition, 2006.

[88] Marc E. Pfetsch, Armin Fügenschuh, Björn Geißler, Nina Geißler, Ralf Gollmer, Benjamin Hiller, Jesco Humpola, Thorsten Koch, Thomas Lehmann, Alexander Martin, Antonio Morsi, Jessica Rövekamp, Lars Schewe, Martin Schmidt, Rüdiger Schultz, Robert Schwarz, Jonas Schweiger, Claudia Stangl, Marc C. Steinbach, Stefan Vigerske, and Bernhard M. Willert. Validation of nominations in gas network optimization: models, methods, and solutions. *Optimization Methods and Software*, 30(1):15–53, 2015.

[89] Konstantin Pieper and Daniel Walter. Linear convergence of accelerated conditional gradient algorithms in spaces of measures, 2019.

[90] Lyudmila Polyakova and Vladimir Karelin. Exact penalty methods for nonsmooth optimization. In *2014 20th International Workshop on Beam Dynamics and Optimization (BDO)*, pages 1–2, 2014.

[91] Renan W. Prado, Sandra A. Santos, and Lucas E. A. Simões. On the convergence analysis of a penalty algorithm for nonsmooth optimization and its performance for solving hard-sphere problems. *Numerical Algorithms*, pages 1–25, 2022.

[92] Ralph Tyrrell Rockafellar. Characterization of the subdifferentials of convex functions. *Pacific Journal of Mathematics*, 17(3):497–510, 1966.

[93] Ralph Tyrrell Rockafellar and Roger Jean-Baptiste Wets. *Variational Analysis*. Springer Verlag, Heidelberg, Berlin, New York, 1998.

[94] Max Sagebaum, Tim Albring, and Nicolas R. Gauger. High-performance derivative computations using CoDiPack. *ACM Transactions on Mathematical Software*, 45(4), 2019.

[95] Martin Schmidt, Denis Aßmann, Robert Burlacu, Jesco Humpola, Imke Joormann, Nikolaos Kanelakis, Thorsten Koch, Djamal Oucherif, Marc E. Pfetsch, Lars Schewe, Robert Schwarz, and Mathias Sirvent. GasLib – A Library of Gas Network Instances. *Data*, 2(4):article 40, 2017.

[96] Stefan Scholtes. *Introduction to Piecewise Differentiable Equations*. SpringerBriefs in Optimization. New York, NY: Springer New York, 2012.

[97] SFB Transregio 154. Mathematische modellierung, simulation und optimierung am beispiel von gasnetzwerken. Online; accessed 7-October-2022.

[98] Lee Siaw Chong and Chin Jia Xin. Creating a gui solver for linear programming models in matlab. *Journal of Science and Technology*, 10(4), December 2018.

[99] Morton Slater. *Lagrange Multipliers Revisited*, pages 293–306. Springer Basel, Basel, 2014.

[100] Wanmei Soon, Gongyun Zhao, and Jieping Zhang. Complementarity demand functions and pricing models for multi-product markets. *European Journal of Applied Mathematics - EUR J APPL MATH*, 20, October 2009.

[101] Qinghua Tao, Li Li, Xiaolin Huang, Xiangming Xi, Shuning Wang, and Johan AK Suykens. Piecewise linear neural networks and deep learning. *Nature Reviews Methods Primers*, 2(1):1–17, 2022.

[102] Hoang Tuy, Athanasios Migdalas, and Peter Värbrand. A quasiconcave minimization method for solving linear two-level programs. *J. Global Optim.*, 4(3):243–263, 1994.

[103] Andreas Wächter and Lorenz T. Biegler. Line search filter methods for nonlinear programming: Motivation and global convergence. *SIAM Journal on Optimization*, 16(1):1–31, 2005.

[104] Andreas Wächter and Lorenz T. Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Program.*,

106(1):25–57, March 2006.

[105] Andrea Walther and Lorenz T. Biegler. On an inexact trust-region SQP-filter method for constrained nonlinear optimization. *Computational Optimization and Applications*, 63, April 2016.

[106] Andrea Walther and Andreas Griewank. Getting started with ADOL-C. In U. Naumann and O. Schenk, editors, *Combinatorial Scientific Computing*, chapter 7, pages 181–202. Chapman-Hall CRC Computational Science, 2012.

[107] Andrea Walther, Olga Weiß, Andreas Griewank, and Stephan Schmidt. Nonsmooth optimization by successive abs-linearization in function spaces. *Applicable Analysis*, 101(1):225–240, 2022.

[108] Douglas E. Ward. Exact penalties and sufficient conditions for optimality in nonsmooth optimization. *Journal of optimization theory and applications*, 57(3):485–499, 1988.

[109] Olga Weiß. *Non-Smooth Optimization by Abs-Linearization in Reflexive Function Spaces*. PhD thesis, Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät, 2022.

[110] Olga Weiß and Andrea Walther. A structure exploiting algorithm for non-smooth semilinear elliptic optimal control problems. *Surveys in Mathematics and its Applications*, 17:139–179, 2022.

[111] Olga Weiß, Andrea Walther, and Stephan Schmidt. *Algorithms Based on Abs-Linearization for Non-smooth Optimization with PDE Constraints*, pages 377–395. Springer International Publishing, Cham, 2022.

[112] Olga Weiß and Monika Weymuth. On solving elliptic obstacle problems by constant abs-linearization. *Humboldt-Universität zu Berlin and Universität der Bundeswehr München*, 2021.

[113] Elizabeth Wong. *Active-Set Methods for Quadratic Programming*. PhD thesis, Department of Mathematics, University of California, San Diego, June 2011.

Selbstständigkeitserklärung

Ich erkläre, dass ich die Dissertation selbständig und nur unter Verwendung der von mir gemäß § 7 Abs. 3 der Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät, veröffentlicht im Amtlichen Mitteilungsblatt der Humboldt-Universität zu Berlin Nr. 42/2018 am 11.07.2018 angegebenen Hilfsmittel angefertigt habe.

....

Ort, Datum

....

Unterschrift