Technische Universität München
Zentrum Mathematik

A New Relaxation Scheme for Mathematical Programs with Equilibrium Constraints: Theory and Numerical Experience

Sonja Veelken

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Dr. Peter Rentrop
Prüfer der Dissertation:	1.	Univ.-Prof. Dr. Michael Ulbrich
	2.	Univ.-Prof. Dr. Christian Kanzow Julius-Maximilians-Universität Würzburg

Die Dissertation wurde am 27.11.2008 bei der Technischen Universität eingereicht und durch die Fakultät für Mathematik am 20.04.2009 angenommen.

For since the fabric of the universe is most perfect and the work of a most wise creator, nothing at all takes place in the universe in which some rule of maximum or minimum does not appear.
(Leonhard Euler)

If you would be a real seeker after truth, it is necessary that at least once in your life you doubt, as far as possible, all things.
(René Descartes)

Acknowledgements

The completion of this thesis was not always an easy task and I am grateful to many people that have helped me along the way.

A special thanks goes to my supervisor Prof. Dr. Michael Ulbrich for the interesting subject of this doctoral thesis and for his continuous support during the past years. I benefitted from his valuable advice and many helpful discussions on various aspects of this work. His critical comments helped me to refine my thinking and his helpful suggestions often led to an improvement of the results.

I also want to express my gratitude to Prof. Dr. Christian Kanzow for his reading and assessment of my thesis.

Moreover, I like to thank my colleagues in Hamburg as well as in Munich for the nice and friendly atmosphere they offered to me. I enjoyed every little coffee break that I spent with them and I thank them all for their help and support during the past years.

Last but not least I want to thank all my friends and my family for their interest in the progress of my work and their understanding and encouragement over the past years. Thank you.

Abstract

In this thesis we are concerned with a specific solution approach for Mathematical Programs with Complementarity Constraints. The approach we suggest is based on the successive solution of a sequence of newly relaxed programs. The new relaxation scheme we propose differs from existing ones in the fact that it combines an exact and a relaxed reformulation of the complementarity conditions. A positive parameter determines to what extend the complementarity conditions are relaxed. In this thesis we not only study the various properties of the programs that are relaxed in this way but we also develop convergence results concerning a corresponding sequence of solutions. Moreover, we consider the new relaxation scheme in connection with Sequential Quadratic Programming and Interior Point methods. Finally we give a detailed report and an associated analysis of the numerical results that we obtained for the new solution approach.

Zusammenfassung

Die Arbeit beschäftigt sich mit einem speziellen Lösungsansatz für Optimierungsprobleme mit Komplementaritätsnebenbedingungen. Der vorgestellte Ansatz beruht auf der sukzessiven Lösung einer Folge von neuartig relaxierten Optimierungsproblemen. Die neue Relaxation unterscheidet sich von bestehenden Ansätzen durch die Kombination einer exakten und einer relaxierten Reformulierung der Komplementaritätsbedingungen. Dabei wird der Grad der Relaxation von der Größe eines positiven Parameters bestimmt. Es werden die Eigenschaften der auf diese Weise erzeugten, relaxierten Optimierungsprobleme untersucht sowie Konvergenzresultate bzgl. einer entsprechenden Folge von Lösungen entwickelt. Des Weiteren wird die neue Relaxation im Zusammenhang mit SQP- und Innere-Punkte Verfahren betrachtet. Abschließend erfolgt eine ausführliche Darstellung und Analyse der numerischen Ergebnisse die mithilfe des neuen Lösungsansatzes erzielt wurden.

Contents

Acronyms ix

Notations xi

Preface 1

1 Introduction to MPECs 5

1.1 The general MPEC . . . 5
1.2 Applications of MPECs . . . 7
1.3 MPEC Theory . . . 13
1.4 Existing Approaches . . . 24

2 New Relaxation Scheme for MPECs 33

2.1 Relaxation . . . 33
2.2 Properties . . . 40
2.3 Convergence . . . 45
2.4 Extensions . . . 57
2.5 Comparison . . . 60

3 New Relaxation in Combination with NLP Methods 69

3.1 SQP Methods . . . 69
3.2 An Interior Point Method . . . 83

4 Numerical Experience 97

4.1 Software and Test Problems . . . 97
4.2 Algorithm . . . 101
4.3 Comparison . . . 116

Summary 123

Bibliography 125

A Numerical Results 131

A.1 Simple Outer Algorithm . . . 131
A.2 Modified Outer Algorithm . . . 135
A.3 Results for M-Stationary Examples . . . 154

List of Tables

4.1 Short description of the test set . . . 100
4.2 Overview of parameter settings . . . 104
4.3 Types of Failures . . . 104
4.4 Failures for various Parameter Settings . . . 104
4.5 Failures for different Functions $\theta(z)$ . . . 106
4.6 Results for Example 4.1 . . . 110
4.7 Results for Example 4.2 . . . 111
4.8 Results for Example 4.2 . . . 112
4.9 Failures for Modified Algorithm . . . 116
4.10 Comparison of Methods: Failures . . . 118
4.11 Numerical Results for Example 4.3 . . . 120
4.12 Numerical Results for Example 4.4 . . . 120
4.13 Numerical Results for Example 2.4 . . . 120

List of Figures

1.1 Result for problem `pack-rig1c-16` of MacMPEC 9
1.2 Normalcones for C-, M- and strong stationarity 20
1.3 Example 1.2 21

2.1 Contours of $\Phi_j(x_1, x_2, t) = 0$ 36
2.2 Contour of $\Phi_2(x_1, x_2, t) = t\theta(0)$ 58
2.3 Functions $\sqrt{u^2 + 4t}$, $|u|$ and $h(u)$ for $u \in [-1, 1]$ and $t = 0.1$ 66
2.4 Functions $\theta_s(u)$, $|u|$ and $g(u)$ for $u \in [-1, 1]$ and $t = 1$ 67

3.1 Feasible Region of problem (3.12) 84

4.1 Comparison of different parameter settings for Algorithm 4.1 105
4.2 Comparison Algorithm 4.1 for different Functions $\theta(z)$ 107
4.3 Functions and $T_3\tilde{\Phi}(x; t)$ for $t = 1.0$, θ as in (2.1) and $x \in [-0.5, 2.0]$. 109
4.4 Illustration of the vertexproblem for Example 4.1 110
4.5 Illustration of the Infeasibility Problem for Example 4.2 112
4.6 Modified Algorithm for Example 4.2 113
4.7 Comparison of Algorithm 4.2 and Algorithm 4.1 115
4.8 Comparison of the new relaxation method to other solution methods 117
4.9 Comparison of complementarity multiplier for Example 2.4 121
4.10 Comparison of complementarity multiplier for Example 4.3 121
4.11 Convergence of iterates for Example 4.3, Example 4.4 and Example 2.4 . . 122

Acronyms

ACQ	Abadie Constraint Qualification	14
BP	Bilevel Program	6
CP	Complementarity Problem	6
CQ	Constraint Qualification	14
CRCQ	Constant Rank Constraint Qualification	50
IPM	Interior Point Method	83
KKT	Karush-Kuhn-Tucker	15
LICQ	Linear Independence Constraint Qualification	14
MPCC	Mathematical Program with Complementarity Constraints	6
MPEC	Mathematical Program with Equilibrium Constraints	5
MPEC-CRCQ	MPEC-Constant Rank Constraint Qualification	51
MPEC-LICQ	MPEC- Linear Independence Constraint Qualification	21
MPEC-SOSC	MPEC-Second Order Sufficient Condition	22
MFCQ	Mangasarian-Fromowitz Constraint Qualification	15
NCP	Nonlinear Complementarity Problem	6
NLP	Nonlinear Program	13
QP	Quadratic Program	69
RNLP	Relaxed Nonlinear Program	21
RNLP-SOSC	RNLP-Second Order Sufficient Condition	22
SOSC	Second Order Sufficient Condition	16
SQP	Sequential Quadratic Programming	69
SSOSC	Strong Second Order Sufficient Condition	23
VI	Variational Inequality	5

Notations

Scalars

$(a)^+$	$\max(0, a)$ with $a \in \mathbb{R}$

Vectors

$x \in \mathbb{R}^n$	a column vector in $\mathbb{R}^n$
(x, y)	$(x^T, y^T)^T \in \mathbb{R}^{n_1+n_2}$ for two vectors $x \in \mathbb{R}^{n_1}$, $y \in \mathbb{R}^{n_2}$
$(x_0, x_1, x_2) \in \mathbb{R}^{n+2p}$	the vector $x \in \mathbb{R}^{n+p+p}$ with $x_0 \in \mathbb{R}^n$, $x_1 \in \mathbb{R}^p$ and $x_2 \in \mathbb{R}^p$
x_j	the jth component of a vector $x \in \mathbb{R}^n$
x_{1j}	the $n + j$th component of a vector $x = (x_0, x_1, x_2) \in \mathbb{R}^{n+2p}$
x_{2j}	the $n + p + j$th component of a vector $x = (x_0, x_1, x_2) \in \mathbb{R}^{n+2p}$
e_j	the jth unit vector in $\mathbb{R}^n$
e_{1j}	the $n + j$th unit vector in $\mathbb{R}^{n+2p}$
e_{2j}	the $n + p + j$th unit vector in $\mathbb{R}^{n+2p}$
$\mathbf{e}$	the vector $(1, \ldots, 1) \in \mathbb{R}^n$
$\|x\|$	the Eucledian norm of a vector $x \in \mathbb{R}^n$
$\|x\|_p$	the p-norm of a vector $x \in \mathbb{R}^n$
$x \geq y$	the componentwise ordering $x_j \geq y_j$ for all $j = 1, \ldots, n$
$x > y$	the componentwise strict ordering $x_j > y_j$ for all $j = 1, \ldots, n$
$(x)_I$	the vector consisting of the components x_j, with $j \in I$
$[x]^+$	the vector $y \in (\mathbb{R}_0^+)^n$ consisting of the components $y_j := (x_j)^+$

Matrices

I	the identity matrix
$\mathbf{0}$	the null matrix, where all components are equal to 0
$diag(a_j) \in \mathbb{R}^{p \times p}$	the diagonal matrix in $\mathbb{R}^{p \times p}$ with the diagonal entries a_j

Sets

$\mathbb{N}$	the set of natural numbers
$\mathbb{R}$	the set of real numbers
$\mathbb{R}^+$	the set of strictly positive real numbers
$\mathbb{R}_0^+$	the set of nonnegative real numbers
$\mathbb{R}^n$	the n-dimensional space of real numbers
$(\mathbb{R}^+)^n$	the strictly positive orthant of $\mathbb{R}^n$
$(\mathbb{R}_0^+)^n$	the nonnegative orthant of $\mathbb{R}^n$
$B_\varepsilon(x^*)$	the open ball with radius ε around x^*
(a, b)	the open interval between a and b in $\mathbb{R}$
$[a, b]$	the closed interval between a and b in $\mathbb{R}$
$A \cup B$	the union of A and B

$A \cap B$	the intersection of A and B
$A \backslash B$	the complement of B in A
$(A \cup B)(x)$	the union of $A(x)$ and $B(x)$, that is $A(x) \cup B(x)$
$(A \cap B)(x)$	the intersection of $A(x)$ and $B(x)$, that is $A(x) \cap B(x)$
$(A \backslash B)(x)$	the complement of $B(x)$ in $A(x)$, that is $A(x) \backslash B(x)$
$A \subseteq B$	A is a subset of B
$A \subsetneq B$	A is a proper subset of B
$\|A\|$	cardinality of A
$\prod_{j=1}^{m} A_j$	Cartesian product of sets A_j
$\mathcal{Z}$	the feasible region of the MPEC (1.19)
$\mathcal{Z}(t)$	the feasible region of $R(t)$

Sequences

(x^k)	sequence with elements x^k
$(x^k)_{k \in \mathcal{K}}$	subsequence of (x^k) corresponding to the index set $\mathcal{K} \subseteq \mathbb{N}$
$x^k \to \bar{x}$	convergent sequence (x^k) with limit $\bar{x}$
$x^k \searrow \bar{x}$	convergent sequence (x^k) with limit $\bar{x}$ and $x^k > \bar{x}$ for all $k \in \mathbb{N}$

Functions

$\nabla f(x)$	the gradient of a function $f : \mathbb{R}^n \to \mathbb{R}$
$\nabla_{x_1} f(x_1, x_2)$	the partial gradient of a function $f : \mathbb{R}^{n_1+n_2} \to \mathbb{R}$ with respect to x_1
$\nabla^2_{xx} f(x, y)$	the Hessian matrix of a function $f : \mathbb{R}^{n_1+n_2} \to \mathbb{R}$ with respect to x
$\nabla g(x)$	the transposed Jacobian of a vector function $g : \mathbb{R}^{n_1+n_2} \to \mathbb{R}^m$
$\nabla_{x_1} g(x_1, x_2)$	the transposed partial Jacobian of a vector function $g : \mathbb{R}^{n_1+n_2} \to \mathbb{R}^m$ with respect to x_1
$\partial g_i / \partial x_j(x)$	the partial derivative of the component function g_i of a vector function $g : \mathbb{R}^n \to \mathbb{R}^m$ with respect to x_j

Problems

$R(t)$	by new relaxation scheme relaxed MPEC, with $t \in \mathbb{R}_0^+$
$NLP(t)$	by regularization scheme of [Sch01] relaxed MPEC, with $t \in \mathbb{R}_0^+$
$R(\delta, t)$	by modified two-sided relaxation scheme relaxed MPEC, with $t \in \mathbb{R}_0^+$ and $\delta \in (\mathbb{R}_0^+)^n$

Special Symbols

$\square$	end of proof

Preface

Mathematical Programs with Equilibrium Constraints, or MPECs, form a special class of Nonlinear Programming problems. Their characteristic feature concerns their constraint structure which contains the so-called equilibrium constraints. The terminology already reveals that this special type of constraints originally described certain system equilibria, such as Nash equilibria in game theory or equilibria of forces in structural mechanics. The term Mathematical Program with Equilibrium Constraints is presumed to have been introduced by Harker and Pang in 1988 [KOZ98],[LPR96]. In general, the equilibrium constraint corresponds to a parametric variational inequality [LPR96]. However, under some suitable conditions the variational inequality can be replaced by a complementarity problem, that consists of a system of equalities and complementarity conditions. In this case the MPEC is also referred to as a Mathematical Program with Complementarity Constraints, or MPCC. In this thesis we confine ourselves to discussing only the type of MPECs, that can be rewritten as an MPCC, though we still call them MPECs.

The general MPEC has its origin in Bilevel Programming [KOZ98], where nonlinear programs are considered that contain another (lower-level) optimization problem as constraint. These problems were introduced by Bracken and McGill in the 1970s [LPR96] and gained more importance by the various applications of the Stackelberg game in economic sciences. Further applications in the natural and engineering sciences then led to the extension of the bilivel programs to MPECs. The multitude and variety of recent research results concerning MPECs points out the significance and the current interest in this field of nonlinear programming.

The answer to the question why we have to treat MPECs separately from standard nonlinear programming concerns the failure of standard constraint qualifications. MPECs have a highly nonconvex constraint structure and a representation of the complementarity constraints by suitable, continuous functions (so-called NCP functions) results in a nonsmooth nonlinear program. Moreover, due to the complementarity conditions and their combinatorial structure, the regularity assumptions concerning the constraints of a nonlinear program are generally not satisfied. Therefore, the standard optimization theory for nonlinear programming and the numerical methods that are based on it are not directly applicable.

Methods that have recently been proposed to solve MPECs concern smoothing or regularization schemes, see for example [FJQ99], [JR00], [Sch01], [LF03a] and [RW04], penalty approaches see [LPRW96], [SS99] or [HR04] and the direct or adapted application of particular NLP methods, such as Sequential Quadratic Programming, see for example [Ani05b], [FLRS06] or [LY07], and Interior-Point Methods, see for example [LS04], [BR05], [DFNS05] and [LLCN06].

In this doctoral thesis we present a new relaxation method and discuss its theoretical properties as well as its numerical performance. This new relaxation method can be regarded as a combination of ideas from regularization schemes with the relaxation-free (or exact) approach by Fletcher et al. in [FLRS06].

The direct application of an SQP algorithm to the exactly reformulated (in other words not relaxed) program performs numerically quite promising [FL02b]. However, this approach can only be guaranteed to be successful for MPECs that have solutions of a special type, namely strongly stationary minimizers.

The regularization approaches are generally not subject to this restriction, but in general, due to a regularization parameter, multiple nonlinear programs have to be solved to find an approximate solution of the original MPEC, which cause an inferior numerical performance compared to the relaxation-free approach. Moreover, in general an infinite number of parameterized nonlinear programs have to be solved to obtain a solution that is exactly feasible for the original MPEC.

In order to avoid the restriction of the relaxation-free approach of Fletcher et al. [FLRS06], but maintain its good numerical performance for MPECs with strongly stationary solutions at the same time, we combine both methods. We relax only those parts of the original feasible region, where we suppose that the relaxation-free approach might get in trouble or fail and use an exact representation of the feasible set for the remaining part. A corresponding relaxed nonlinear program is obtained by introducing a strictly positive relaxation parameter $t > 0$. If we set $t = 0$, then the parameterized nonlinear program corresponds to the original MPEC. Therefore, we are able to show that, under reasonable conditions, by solving a sequence of such NLPs parameterized by t, we will find a solution of the original MPEC as soon as $t > 0$ is sufficiently small.

Before we present our new relaxation method, in Chapter 1 we first start with an introduction into the field of MPECs. We introduce the general MPEC as defined in [LPR96] and discuss its close connection to bilevel programming and how an MPCC can be derived from the MPEC. Then we continue giving a brief overview of engineering and economic applications of MPECs. The main part of Chapter 1 concerns a presentation and discussion of the most important theoretical properties of MPECs and of the MPEC specific terminology, that we will use in the remaining part of this thesis. We finish the first chapter by a short review of some recent solution approaches for MPECs.

Chapter 2 together with Chapter 4 forms the main part of this thesis. We first derive the relaxed nonlinear program that provides the basis of our relaxation method and present some basic properties of it. Then we relate stationary points and solutions of the original MPEC and the relaxed nonlinear program. This will be followed by the convergence analysis of a sequence of solutions of a sequence of parameterized, relaxed nonlinear programs (for a strictly positive, decreasing sequence of parameters). Finally, we compare the theoretical properties of our new relaxation scheme to those ones of the relaxation-free approach of [FLRS06] and the regularization scheme proposed by Scholtes in [Sch01], which seems to be the most appropriate regularization scheme for a comparison to our method.

Chapter 3 is devoted to an analysis of the new relaxation method in combination with two main solution methods for standard nonlinear programs, namely Sequential Quadratic Programming and Interior Point Methods. First, we discuss the convergence behaviour of a standard local SQP algorithm applied to the relaxed nonlinear program. Then we present a variant of the two-sided relaxation scheme for MPECs, presented by DeMiguel et al. in [DFNS05], that incorporates the new relaxation scheme, which we introduced in Chapter 2.

Furthermore, we demonstrate how the convergence results of [DFNS05] can be transferred to the modified two-sided relaxation scheme.

In Chapter 4 we present the numerical results we obtained with the new relaxation method. First, we introduce a simple outer algorithm that uses a standard SQP solver, as for example `filterSQP` [FL98], as a black box and discuss its numerical performance. Then we explain the two main failures that occurred for some of the test problems using this simple outer algorithm and present a modified outer algorithm that circumvents these failures. We examine its numerical performance individually, but we also compare the numerical results we obtained for the modified algorithm to those ones we obtained using the relaxation-free approach of [FLRS06] and the regularization scheme of [Sch01], respectively.

We finish this doctoral thesis with a short summary combined with a critical review of the results we presented in this thesis and an outlook on some possible subsequent research topics.

1 Introduction to MPECs

This chapter is devoted to a general introduction into the field of MPECs. First we introduce the general MPEC and discuss its various forms and its close connection to Bilevel Programming, which will be followed by a presentation of some engineering and economic applications of MPECs. We will then illustrate the main difficulties concerning MPECs and present the definitions and notations that we will need in the following chapters. Finally, we give a brief overview of some recent solution approaches for MPECs.

1.1 The general MPEC

In general, *Mathematical Programs with Equilibrium Constraints* (MPEC) are nonlinear programs (confer (1.13)) that contain a *Variational Inequality* (VI) as a constraint [FP03]. These variational inequalities describe certain system equilibria resulting from the underlying model, for example traffic equilibria, Nash equilibria, equilibria of forces and so on. For further details and general informations about MPECs we refer to the monographs [LPR96] and [KOZ98].

The general MPEC as introduced in [LPR96] has the form

$$\begin{array}{rlcl} \min & f(x_0, x_1) & & \\ \text{subject to} & (x_0, x_1) & \in & \mathcal{X} \\ & x_1 & \in & S(K(x_0), \Upsilon(x_0, \cdot)), \end{array} \tag{1.1}$$

where $x_0 \in \mathbb{R}^n$, $x_1 \in \mathbb{R}^p$ and $\mathcal{X} \subseteq \mathbb{R}^{n+p}$ denotes the joint feasible set of x_0 and x_1. Furthermore, $S(K(x_0), \Upsilon(x_0, \cdot))$ denotes the solution set of a Variational Inequality, denoted by $\mathrm{VI}(K(x_0), \Upsilon(x_0, \cdot))$, that depends on a set $K(x_0) \subseteq \mathbb{R}^p$ and a function $\Upsilon(x_0, \cdot)$, where $\Upsilon : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}^p$, which are both parameterized by x_0. A vector x_1 represents a solution of $\mathrm{VI}(K(x_0), \Upsilon(x_0, \cdot))$, if x_1 satisfies the conditions

$$\begin{array}{lll} (i) & x_1 \in K(x_0) & \\ (ii) & (v - x_1)^T\, \Upsilon(x_0, x_1) \geq 0 & \forall\, v \in K(x_0). \end{array} \tag{1.2}$$

These conditions can be transformed, if $K(x_0)$ represents a cone.

Definition 1.1. Let $\mathcal{M} \subseteq \mathbb{R}^\ell$, then $\mathcal{M}$ is said to be a *cone* if and only if

$$x \in \mathcal{M}, \quad \lambda \geq 0 \quad \Longrightarrow \lambda x \in \mathcal{M}.$$

Moreover, the *polar cone* of a set $\mathcal{M}$ is defined by

$$\mathcal{M}^\circ := \{ y \in \mathbb{R}^\ell \mid y^T x \leq 0 \ \ \forall\, x \in \mathcal{M} \}$$

and the *dual cone* of $\mathcal{M}$ is defined by

$$\mathcal{M}^* := \{ y \in \mathbb{R}^\ell \mid y^T x \geq 0 \ \ \forall\, x \in \mathcal{M} \}.$$

If $K(x_0) \subseteq \mathbb{R}^p$ represents a cone, then the conditions (1.2) are equivalent to the conditions [FP03]

$$\begin{array}{lll} (i) & x_1 \in K(x_0) & \\ (ii) & \Upsilon(x_0, x_1) \in K(x_0)^* & \\ (iii) & x_1^T \Upsilon(x_0, x_1) = 0 . & \end{array} \tag{1.3}$$

The VI$(K(x_0), \Upsilon(x_0, \cdot))$ is then often called a *Complementarity Problem* (CP) and denoted by $CP(K(x_0), \Upsilon(x_0, \cdot))$ [FP03]. Assuming further that $K(x_0) = (\mathbb{R}_0^+)^p$, then (1.2) can be simplified even more to

$$0 \leq x_1 \perp \Upsilon(x_0, x_1) \geq 0 , \tag{1.4}$$

which is a short notation of the conditions of (1.3), if $K(x_0) = (\mathbb{R}_0^+)^p$. Problems of this kind are referred to as *Nonlinear Complementarity Problems* (NCP).

If the VI of an MPEC of the form (1.1) can be replaced by the condition (1.4), we obtain

$$\begin{array}{rl} \min & f(x_0, x_1) \\ \text{subject to} & (x_0, x_1) \in \mathcal{X} \\ & 0 \leq x_1 \perp \Upsilon(x_0, x_1) \geq 0 . \end{array} \tag{1.5}$$

This problem is then also referred to as *Mathematical Program with Complementarity Constraints* (MPCC). Finally, introducing slack variables x_2 and assuming that

$$\mathcal{X} = \{ x \in \mathbb{R}^{n+2p} \mid h(x) = 0 \quad \text{and} \quad g(x) \geq 0 \} ,$$

where $h : \mathbb{R}^{n+2p} \to \mathbb{R}^q$ and $g : \mathbb{R}^{n+2p} \to \mathbb{R}^m$ are twice continuously differentiable functions, we end up with the problem

$$\begin{array}{rl} \min & f(x) \\ \text{subject to} & h(x) = 0 \\ & g(x) \geq 0 \\ & 0 \leq x_1 \perp x_2 \geq 0 , \end{array} \tag{1.6}$$

where $x = (x_0, x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^p$ and $f : \mathbb{R}^{n+2p} \to \mathbb{R}$ is a twice continuously differentiable function. MPECs of this type are the object of this thesis.

We have shown under which circumstances we can transform an MPEC of the form (1.1) into one that has the form of (1.6), other applications for instance in structural design [FTL99] and robotics [AAP04], however, result directly in a problem of the form (1.6), as we will see in the following section.

A class of optimization problems that is closely related to MPECs are *Bilevel Programs* (BP). The name reflects the feature that the constraints of a so called *upper-level problem* comprehend an additional optimization problem, called the *lower-level problem*, that is parametrized by the so-called *upper-level variables*. The general BP as formulated in [CMS07] is

$$\begin{array}{rl} \min_{x,y} & F(x, y) \\ \text{subject to} & x \in \mathcal{X} \\ & G(x, y) \leq 0 \\ & y \text{ solves } \quad \begin{array}{rl} \min_y & f(x, y) \\ \text{subject to} & g(x, y) \leq 0 , \end{array} \end{array} \tag{1.7}$$

where $x \in \mathbb{R}^n$ are the *upper-level variables* and $y \in \mathbb{R}^m$ are the *lower-level variables*. For each fixed upper-level variable $\hat{x} \in \mathcal{X}$, the lower-level variables $y \in \mathbb{R}^m$ are only feasible for (1.7), if they solve the lower-level problem

$$\begin{aligned} \min_y \quad & f(\hat{x}, y) \\ \text{subject to} \quad & g(\hat{x}, y) \leq 0\,, \end{aligned} \tag{1.8}$$

that is parametrized by $\hat{x}$. Assuming that the lower-level problem of (1.7) is convex and satisfies some regularity conditions, the KKT-conditions of (1.8) are necessary and sufficient. Hence, we can replace (1.8) by its KKT-conditions (confer (1.16)) and obtain the problem

$$\begin{aligned} \min_{x,y} \quad & F(x, y) \\ \text{subject to} \quad & x \in \mathcal{X} \\ & G(x, y) \leq 0 \\ & \nabla_y \mathcal{L}(x, y, \lambda) = 0 \\ & 0 \leq \lambda \perp -g(x, y) \geq 0\,, \end{aligned} \tag{1.9}$$

where $\mathcal{L}(x, y, \lambda) := f(x, y) + \sum_{j=1}^{p} \lambda_j g_j(x, y)$ denotes the Lagrangian function of (1.8). Hence, under these conditions the BP can be transformed into an MPEC of the form (1.5).

On the other hand, if the function $\Upsilon(x_0, \cdot)$ in (1.1) corresponds to a partial gradient map of a convex, continuously differentiable function $f(x, y)$, with $f : \mathbb{R}^{n+m} \to \mathbb{R}$, that is

$$\Upsilon(x_0, x_1) = \nabla_{x_1} f(x_0, x_1)\,,$$

then for every $\hat{x}_0 \in \mathcal{X}$, the VI

$$x_1 \in K(\hat{x}_0) \quad \text{and} \quad (v - x_1)^T \, \nabla_{x_1} f(\hat{x}_0, x_1) \geq 0 \qquad \forall\, v \in K(\hat{x}_0)$$

represents the stationarity conditions of the minimization problem

$$\begin{aligned} \min_{x_1} \quad & f(\hat{x}_0, x_1) \\ \text{subject to} \quad & x_1 \in K(\hat{x}_0)\,. \end{aligned} \tag{1.10}$$

If $K(x_0)$ is described by a finite number of inequalities, that is

$$K(x_0) = \{\, x_1 \in \mathbb{R}^m \mid g(x_0, x_1) \leq 0 \,\}\,,$$

where $g : \mathbb{R}^{n+m} \to \mathbb{R}^p$, then (1.1) can therefore be rewritten as a BP of the form (1.7). An overview of the state-of-the-art of bilevel programming can be found in the survey paper [CMS07]. For a more comprehensive treatment of the subject, we refer the reader to the monographs [Bar98] and [Dem02].

1.2 Applications of MPECs

MPECs have many applications in the field of engineering as well as in economic sciences. Often, they arise if some design parameters are to optimize subject to certain system equilibria. Economic equilibria (Nash equilibria), equilibria of forces, network equilibria and others are examples for such system equilibria. Many of these can be described by complementarity conditions. Hence, the resulting optimization problem corresponds to a Mathematical Program with Equilibrium Constraints of the form (1.6). Other applications result directly in a Mathematical Program with Complementarity Constraints.

Before we discuss the economic applications, we first describe some engineering applications. A detailed overview of MPEC applications can be found in [FP97] and the monographs [KOZ98] and [LPR96].

1.2.1 Engineering Applications

An important engineering application of MPECs arises in the area of contact problems. Consider a physical contact (with friction) between two or more solid bodies. The complementarity condition then models the fact that the contact forces that have to be involved in the model (for example friction forces) can only be strictly positive if the distance between the bodies vanishes, thus if the gap between them is equal to zero.

Applications involving frictional contact problems are discussed for example in [KM06]. A recent area in which such problems play an important role occurs in robotics [LPR96], see for example [AAP04].

Another kind of contact problems arise in the area of shape optimization. Modelling of an elastic membrane with a rigid or compliant obstacle involves complementarity conditions as we will see next. We briefly describe a shape optimization problem where we try to minimize a domain $\Omega(\alpha)$ subject to the condition that the membane is supposed to come into contact with a rigid obstacle for a predefined subset $\Omega_0 \subseteq \Omega(\alpha)$. The description here conforms with the detailed discussion of such problems in [KOZ98].

Let

- α be the design parameter that describes the part of the boundary of $\Omega(\alpha)$ that is to be optimized
- $J(\alpha)$ be a measure of the domain $\Omega(\alpha)$,
- U_{ad} be a set of admissible design parameters α,
- Ω_0 be a given minimum contact region,
- u be the deflection of the membrane,
- $S(\alpha)$ be the graph of the function that describes the surface of the rigid obstacle,
- $f(\alpha)$ be the force that acts perpendicularly on the membrane,
- $A(\alpha)$ be the corresponding stiffness matrix and
- $C(\alpha)$ be the resulting contact region.

Then the resulting discretized optimization problem is of the form

$$\begin{array}{rl} \min_{\boldsymbol{\alpha}} & J(\boldsymbol{\alpha}) \\ \text{subject to} & \boldsymbol{\alpha} \in \mathbf{U}_{ad} \\ & \boldsymbol{\Omega}_0 \subseteq \mathbf{C}(\boldsymbol{\alpha}) \\ & \mathbf{u}_j = 0 \qquad \text{for all} \qquad j \in D(\boldsymbol{\alpha}) \\ & 0 \leq \mathbf{u} - \mathbf{S}(\boldsymbol{\alpha}) \quad \perp \quad A(\boldsymbol{\alpha})\mathbf{u} - \mathbf{f}(\boldsymbol{\alpha}) \geq 0\,, \end{array}$$

where $\boldsymbol{\alpha}$, $\mathbf{u}$, $\mathbf{S}(\boldsymbol{\alpha})$ and $\mathbf{f}(\boldsymbol{\alpha})$ are the vectors of the corresponding function values at the nodes of the discretized domain $\Omega(\alpha)$. Furthermore, $\boldsymbol{\Omega}_0$ and $\mathbf{C}(\boldsymbol{\alpha})$ correspond to the sets of indices of nodes lying in Ω_0 and $C(\alpha)$, respectively. Finally, the set $D(\boldsymbol{\alpha})$ denotes the set of indices of nodes lying in $\partial\Omega(\alpha)$, hence we also assumed the discretized version of the boundary condition $u = 0$ on $\partial\Omega(\alpha)$.

Examples for such problems are the `pack-comp` (with a compliant obstacle) and the `pack-rig` (with a rigid obstacle) problems of the MacMPEC test problem set (see Section 4.1 and the Appendix). Figure 1.1 illustrates the result for problem `pack-rig1c-16`, which

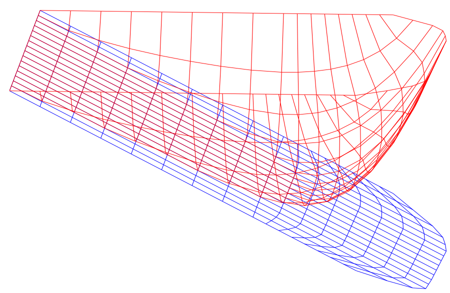


Figure 1.1: Result for problem `pack-rig1c-16` of MacMPEC

is a 2D obstacle problem on $[0,1]\times[0,1]$ and corresponds to Example 9.1 in [KOZ98], Section 9.2. Here we have $\Omega_0 = [0.25, 0.5] \times [0.25, 0.75]$, $f(\alpha)(x, y) = -1.0$, $S(\alpha)(x, y) = -0.05x$. The set $\Omega(\alpha)$ is discretized using triangular finite elements with linear basis functions and a discretization parameter $h = 1/16$. Furthermore, the set of admissible design variables $\boldsymbol{\alpha}$ is

$$\mathbf{U}_{ad} = \{\, \boldsymbol{\alpha} \in \mathbb{R}^{17} \,|\, 0.6 \leq \boldsymbol{\alpha} \leq 1.0,\ |\boldsymbol{\alpha}_{i-1} - \boldsymbol{\alpha}_i| < 3.0 \,\}\,.$$

Problems closely related to the problems we just described occur in the area of structural mechanics. Given a basic structure and an external load, the object of interest is the optimal design of this mechanical structure. The objective function might either be the total weight or volume of the structure or its compliance.

Consider a basic structure which is given by the nodes of the potential bars of a truss. Given the total volume V of the bars and the external load vector f, we are interested in a truss design, determined by the volumes v_i of the bars (hence the v_i are the design variables), which minimizes the compliance $f^T x$, where x denotes the vector of nodal displacements.

The potential energy of the truss structure is then given by [JKZ98]

$$\mathrm{E}(v,x) = \frac{1}{2}x^T A(v)x - f^T x\,,$$

where

$$A(v) = \sum_{j=1}^{n} v_j A_j$$

and A_j is the stiffness matrix corresponding to the jth bar. The matrix $A(v)$ is symmetric and positive semi-definit. (Furthermore, under reasonable assumptions it is positive definite, if $v_j > 0$ $j = 1, \dots, n$.) If x minimizes $\mathrm{E}(v,x)$, then we have a state of an equilibrium of forces [JKZ98].

Next, we involve the constraint of a rigid obstacle, thus we assume that it cannot be penetrated by the nodes of the truss and we have the additional condition

$$Cx \leq d\,,$$

where C is a kinematic transformation matrix and d consists of the distances between the obstacle and the initial nodes.

Since the resulting optimization problem

$$\begin{array}{rl} \min_x & \mathrm{E}(v,x) \\ \text{subject to} & Cx \leq d \end{array}$$

is convex, the solutions of it are characterized by the KKT-conditions (confer (1.16))

$$\begin{aligned} -f + A(v)x + C^T\lambda &= 0 \\ Cx - d &\leq 0 \\ \lambda &\geq 0 \\ \lambda^T(Cx-d) &= 0\,. \end{aligned}$$

The Lagrangian multiplier λ can be interpreted as the contact force based on the rigid obstacle. Since this is only positive, if the truss comes into contact with the obstacle, we obtain a complementarity condition.

As we are interested in a truss with minimal compliance subject to a given total volume of the truss bars, we end up with the MPEC

$$\begin{array}{rl} \min_{x,v,\lambda} & f^T x \\ \text{subject to} & \sum_{j=1}^{n} v_j = V \\ & -f + A(v)x + C^T\lambda = 0 \\ & Cx - d \leq 0 \\ & \lambda \geq 0 \\ & \lambda^T(Cx-d) = 0\,. \end{array}$$

For a brief survey of some structural design applications of MPECs see for example [AW05]. Other engineering applications of MPECs concern network design problems [LPR96], chemical engineering problems [BRB07] and further parameter identification problems [TLQ01].

1.2.2 Economic Applications

In many economic applications models serve to gain some insights into the interactions of economic systems. On the basis of the obtained information one hopes to be able to develope more adequate and sophisticated political or economic strategies. Such strategies are for example economic reforms (for example deregularization of public networks), taxation, business expansion, toll pricing and others.

Consider an economic system that deals with a finite number of commodities. With each commodity a specific price is associated. Moreover, it is assumed that there exist only two different type of agents: on the one hand the sectors (companies or producers) on the other hand the consumers.

The general equilibrium problem [FP97] is concerned with finding commodity prices and determining the behavior of the agents such that

1. each sector maximizes its profit,
2. supply exceeds demand and
3. expenditure equals income.

The third statement is called the *Walras law*. It constitutes the condition that the total expenditure of the consumers is equal to the income that originates from the trade with the commodities.

Furthermore, it is generally assumed that the consumers do not have a preferred producer but decide by price advantages. Hence, we consider the case of a perfect competition between the producers. Finally, each agent has perfect information about the prices.

The general equilibrium that is described by these conditions can be represented by complementarity conditions [FP97], such that we obtain an MPEC, if we choose some of the input data as design parameters that we want to optimize in a particular sense. Design parameters are for example production level, facility locations and distribution of goods. The described general equilibrium problem has to be specified according to the object of interest: production, consumption, taxation and subsidies and so on. Established fields of economic applications of MPECs are for example toll pricing problems [LH04], modeling of electric power markets [CHLM06],[HMP00] and transportation networks [HP06].

Most of these economic applications of MPECs are based on the concept of the *Stackelberg game*. These game theoretic problems are closely connected to MPECs, as we will explain in the following. A general introduction to game theory can be found in [HI06] or [Sch04].

The Stackelberg game is an extension of the *Nash game*. For this reason, we will first describe this basic concept, before we turn to the Stackelberg game.

Consider a finite number $M \in \mathbb{N}$ of players $i \in \{1, \dots, M\}$. Each of the players possesses a set of strategies $s_i \in S_i \subseteq \mathbb{R}^{m_i}$. The aim of each player is to minimize her cost function $\theta_i(s_i, \tilde{s}_i)$, where

$$\tilde{s}_i \in \prod_{\substack{j=1 \\ j \neq i}}^{M} S_j$$

denotes the vector of strategies of the remaining players (in other words all players except for player i). Each player observes the strategies of the other players and chooses her optimal strategy under the assumption that they will not change their chosen strategy. Furthermore, it is assumed that the players do not cooperate with each other.

A combination of strategies

$$\mathbf{s}^* \in \prod_{j=1}^{M} S_j$$

is called a *Nash equilibrium* if there exists no incentive for any player i to change her strategy $s_i^* \in S_i$. This situation can mathematically be described by

$$s_i^* \in \operatorname{argmin}\{\theta_i(s_i, \tilde{s}_i^*) \,|\, s_i \in S_i\} \qquad \text{for all} \quad i \in \{1, \ldots, M\}\,.$$

Until now, all players are equal in the sense that they have the same information and can choose freely any strategy $s_i \in S_i$ subject to this information. If we consider the Stackelberg game, we have in contrast a distinct player, called the *Stackelberg leader*, which can influence the remaining players, called the *Stackelberg followers*. The leader can anticipate the reactions of the followers and choose its optimal strategy according to his knowledge. Furthermore, the leader's choice influences the sets of strategies of the followers. In other words the sets of strategies of the followers $S_i(x) \subseteq \mathbb{R}^{m_i}$ are parameterized by the leader's strategy x. Moreover, the cost functions θ_i may also be parameterized by x, thus we have

$$\theta_i(x, \cdot) : \prod_{j=1}^{M} \mathbb{R}^{m_j} \to \mathbb{R}\,.$$

Hence, first the leader chooses its strategy $x^* \in X$, where X denotes the set of strategies of the leader, and afterwards the followers play a Nash game that is parameterized by the leader's choice x^*. The joint answer of the followers

$$\mathbf{s}^* := (s_i^*)_{i \in \{1,\ldots,M\}} \in \prod_{j=1}^{M} S_j(x^*)$$

corresponding to x^* satisfies

$$s_i^* \in \operatorname{argmin}\{\theta_i(x^*, s_i, \tilde{s}_i^*) \,|\, s_i \in S_i(x^*)\} \qquad \text{for all} \quad i \in \{1, \ldots, M\}\,.$$

Suppose $f : \mathbb{R}^m \times \prod_{j=1}^{M} \mathbb{R}^{m_j} \to \mathbb{R}$ denotes the cost function of the Stackelberg leader, then finding a solution vector (x^*, s^*) that solves the Stackelberg game corresponds to solving the BP (confer (1.7))

$$\begin{array}{lll} \min_{x,s} & f(x,s) & \\ \text{subject to} & x \in X & \\ & s_i \text{ solves} \quad \begin{array}{ll} \min_{s_i} & \theta_i(x,s) \\ \text{subject to} & s_i \in S_i(x) \end{array} & i = 1, \ldots, M \end{array} \tag{1.11}$$

Assume that each set of strategies $S_i(x)$ is a nonempty, closed convex set and

$$\theta_i(x, \cdot, \tilde{s}_i) : \mathbb{R}^{m_i} \to \mathbb{R}$$

is convex and continuously differentiable in s_i. Then we can replace the lower-level problems

$$\begin{array}{rlcl} \min_{s_i} & \theta_i(x,s) & & \\ \text{subject to} & s_i & \in & S_i(x) \end{array} \tag{1.12}$$

of problem (1.11) by the VI

$$s \ \in \ \mathbf{S}(x) \quad \text{and} \quad (v-s)^T \, \boldsymbol{D}(x,s) \geq 0 \qquad \forall \, v \ \in \mathbf{S}(x)\,,$$

where

$$\mathbf{S}(x) = \prod_{j=1}^{M} S_j(x)$$

and

$$D(x,s) := (d_i(x,s))_{i \in \{1,\ldots,M\}}$$

with

$$d_i(x,s) := \nabla_{s_i} \theta_i(x,s) \qquad \text{for all} \quad i \in \{1,\ldots,M\}$$

(for a detailed description see [FP03]).

Moreover, if the sets of strategies can be represented by a finite number of continuously differentiable functions, then we can replace the lower-level problems by the corresponding KKT-conditions and obtain an MPCC (see the foregoing Section 1.1).

1.3 MPEC Theory

Before we start our discussion of the specific theoretical characteristics and terms of MPECs, we will first briefly review some basic optimality conditions of Nonlinear Programming.

Basic Optimality Conditions for Nonlinear Programming

Consider the general Nonlinear Program (NLP)

$$\begin{array}{rl} \min & f(x) \\ \text{subject to} & h(x) = 0 \\ & g(x) \geq 0\,, \end{array} \tag{1.13}$$

where $f : \mathbb{R}^n \to \mathbb{R}$, $h : \mathbb{R}^n \to \mathbb{R}^q$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ are here assumed to be twice continuously differentiable functions. The following definitions [GK02] are essential in the theory of Nonlinear Programming.

The general stationarity condition for an ordinary NLP of the form (1.13) is

$$\nabla f(x^*)^T d \geq 0 \qquad \forall \, d \in \mathcal{T}(\mathcal{X}, x^*)\,, \tag{1.14}$$

where $\mathcal{X} := \{x \in \mathbb{R}^n : h(x) = 0,\, g(x) \geq 0\}$ denotes the feasible region and $\mathcal{T}(\mathcal{X}, x^*)$ denotes the tangent cone that is defined as follows [GK02]:

Definition 1.2. Let $\mathcal{M} \subseteq \mathbb{R}^\ell$ denote a nonempty set and let $x \in \mathcal{M}$. The *tangent cone* of $\mathcal{M}$ at x is defined by

$$\begin{aligned}\mathcal{T}(\mathcal{M}, x) \;=\; &\Big\{ d \in \mathbb{R}^\ell \;\Big|\; \exists\,(x^k) \subset \mathcal{M}, \\ &\exists\,(\eta_k) \subset \mathbb{R}, \eta_k \searrow 0 : \; x^k \to x \;\text{ and }\; (x^k - x)/\eta_k \to d \Big\}.\end{aligned}$$

Moreover the corresponding *normal cone* of $\mathcal{M}$ at x is

$$\mathcal{N}(\mathcal{M}, x) = (\mathcal{T}(\mathcal{M}, x))^\circ .$$

Condition (1.14) represents the fact that there exists no feasible descent direction at a local optimum x^* and it is equivalent to

$$-\,\nabla f(x^*) \;\in\; \mathcal{N}(\mathcal{X}, x^*)\,. \tag{1.15}$$

As these two stationarity conditions are difficult to verify they are in particular not well practicable for numerical purposes. Some constraint qualifications (CQ) are therefore typically used to guarantee that the unwieldy tangent cone $\mathcal{T}(\mathcal{X}, x^*)$ can be replaced by the linearized tangent cone

$$\mathcal{T}_{lin}(\mathcal{X}, x^*) := \{ d \in \mathbb{R}^\ell \mid \nabla h_j(x^*)^T d = 0,\; \forall\, j \in I_h(x^*),\; \nabla g(x^*)^T d \geq 0,\; \forall\, j \in I_g(x^*) \},$$

where

$$\begin{aligned} I_h(x) &= \{ i \in \{1, .., q\} \;:\; h_i(x) = 0 \}, \\ I_g(x) &= \{ i \in \{1, .., m\} \;:\; g_i(x) = 0 \}, \end{aligned}$$

denote the sets of the active constraints in x.

One of the most basic constraint qualification is the so-called *Abadie Constraint Qualification* [GK02].

Definition 1.3. Let $x^* \in \mathcal{X}$, then x^* is said to satisfy the *Abadie Constraint Qualification* (ACQ), if $\mathcal{T}(\mathcal{X}, x^*) = \mathcal{T}_{lin}(\mathcal{X}, x^*)$.

Suppose x^* satisfies the ACQ, then (1.15) can be replaced by

$$-\,\nabla f(x^*) \;\in\; (\mathcal{T}_{lin}(\mathcal{X}, x^*))^\circ$$

which can then by the Farkas Lemma (Lemma 2.27 in [GK02]) proved to be equal to the KKT-conditions (see Definition 1.6), which mostly form the basis of solution methods and software for NLPs. Since the ACQ is difficult to verify, often some stronger constraint qualifications are used that imply the ACQ and hence the admissibility of the KKT-conditions. Two basic regularity assumptions concerning the feasible region of the NLP that imply the ACQ are:

Definition 1.4. Let x^* be feasible for (1.13), then x^* is said to satisfy the *Linear Independence Constraint Qualification (LICQ)*, if the family

$$\begin{aligned} &\nabla h_i(x^*) \quad i \in \{1, \ldots, q\}, \\ &\nabla g_j(x^*) \quad j \in I_g(x^*) \end{aligned}$$

is linear independent.

Definition 1.5. Let x^* be feasible for (1.13), then x^* is said to satisfy the *Mangasarian-Fromowitz Constraint Qualification (MFCQ)*, if

1. the family $\nabla h_i(x^*)$ $i = 1, \ldots, q$ is linear independent and
2. there exists a vector $d \in \mathbb{R}^n$ that satisfies the conditions $\nabla g_j(x^*)^T d > 0$ for all $j \in I_g(x^*)$ and $\nabla h_i(x^*)^T d = 0$ for all $i \in \{1, \ldots, q\}$

It can be proved (see for example [GK02]), that these two constraint qualifications satisfy the implications LICQ $\Rightarrow$ MFCQ $\Rightarrow$ ACQ.

Next we define the *Karush-Kuhn-Tucker (KKT-) conditions* that form a necessary optimality condition [GK02] for Nonlinear Programming problems.

Definition 1.6. Let $x^* \in \mathbb{R}^n$. We call the conditions

$$\begin{array}{rcll} \nabla f(x^*) - \nabla g(x^*)\lambda^* - \nabla h(x^*)\mu^* & = & 0 & \\ h(x^*) & = & 0 & \\ g(x^*) & \geq & 0 & \\ \lambda^* & \geq & 0 & \\ g_i(x^*)\lambda_i^* & = & 0 & \quad i = 1, \ldots, m \end{array} \tag{1.16}$$

Karush-Kuhn-Tucker (KKT-) conditions. Moreover, if there exist $\lambda^* \in \mathbb{R}^m$ and $\mu^* \in \mathbb{R}^q$, such that (x^*, λ^*, μ^*) satisfies (1.16), then we call x^* a *stationary point* of (1.13) and the vectors λ^* and μ^* *Lagrange multipliers* of x^*.

Suppose that a local solution x^* of (1.13) satisfies either LICQ or MFCQ, then the existence of vectors $\lambda^* \in \mathbb{R}^m$ and $\mu^* \in \mathbb{R}^q$, such that (x^*, λ^*, μ^*) satisfies the KKT-conditions form a necessary optimality condition.

Theorem 1.1. *Let $x^* \in \mathbb{R}^n$ be a local solution of (1.13). If x^* satisfies either LICQ or MFCQ, then there exist vectors $\lambda^* \in \mathbb{R}^m$ and $\mu^* \in \mathbb{R}^q$, such that (1.16) is satisfied.*

Proof. See for example Theorem 2.39 and 2.41 in [GK02]. ☐

Define the sets

$$I_g^+(x, \lambda) = \{i \in \{1, \ldots, m\} \,:\, g_i(x) = 0,\ \lambda_i > 0\},$$
$$I_g^0(x, \lambda) = \{i \in \{1, \ldots, m\} \,:\, g_i(x) = 0,\ \lambda_i = 0\}$$

and

$$\begin{array}{rcll} \mathcal{S}(x, \lambda) & = & \{ & d \in \mathbb{R}^n \backslash \{0\} | \\ & & & \nabla h_i(x)^T d = 0, \quad i \in \{1, \ldots, q\}, \\ & & & \nabla g_j(x)^T d = 0, \quad j \in I_g^+(x, \lambda), \\ & & & \nabla g_j(x)^T d \geq 0, \quad j \in I_g^0(x, \lambda)\, \} \end{array}$$

and let

$$\mathcal{L}(x, \lambda, \mu) = f(x) - \sum_{j=1}^{m} \lambda_j g_j(x) - \sum_{i=1}^{q} \mu_i h_i(x) \tag{1.17}$$

denote the *Lagrangian function* of (1.13), then we can define a standard Second Order Sufficient Condition (SOSC) for x^* to be a local solution of (1.13).

Definition 1.7. Let x^* be a stationary point of (1.13) with multipliers λ^* and μ^* and suppose that

$$d^T \nabla^2_{xx} \mathcal{L}(x^*, \lambda^*, \mu^*)\, d > 0 \quad \forall\, d \in \mathcal{S}(x^*, \lambda^*)\,, \tag{1.18}$$

then x^* is said to satisfy the *Second Order Sufficient Condition* (SOSC) for (1.13).

Theorem 1.2. *Let (x^*, λ^*, μ^*) satisfy the KKT-conditions and the SOSC, then x^* is a strict local solution of (1.13).*

Proof. See for example Theorem 2.55 in [GK02]. □

For further details concerning Nonlinear Programming we refer the interested reader to [GK02], [Fle00] or [CGT00]. Having reviewed the basic notions of the theory of Nonlinear Programming, that we will need in the following, we now start our discussion of the theory of MPECs.

As mentioned before, in this thesis we consider MPECs of the form

$$\begin{array}{rl} \min & f(x) \\ \text{subject to} & h(x) = 0 \\ & g(x) \geq 0 \\ & 0 \leq x_1 \perp x_2 \geq 0, \end{array} \tag{1.19}$$

where $x = (x_0, x_1, x_2) \in \mathbb{R}^n \times \mathbb{R}^p \times \mathbb{R}^p$ and all (in)equalities are meant componentwise. Throughout, we will assume that $f : \mathbb{R}^{n+2p} \to \mathbb{R}$, $h : \mathbb{R}^{n+2p} \to \mathbb{R}^q$ and $g : \mathbb{R}^{n+2p} \to \mathbb{R}^m$ are twice continuously differentiable functions.

Note that MPECs formulated with the seemingly more general complementarity condition

$$0 \leq G(x) \perp H(x) \geq 0\,,$$

where G and H are twice continuously differentiable functions, mapping $\mathbb{R}^{n+2p}$ to $\mathbb{R}^p$, can be transformed to the form (1.19) by introducing slack variables:

$$\begin{array}{l} G(x) - s_1 = 0 \\ H(x) - s_2 = 0 \\ 0 \leq s_1 \perp s_2 \geq 0\,. \end{array}$$

The complementarity constraint

$$0 \leq x_1 \perp x_2 \geq 0, \tag{1.20}$$

can equivalently be replaced by one of the conditions

$$\begin{array}{rl} \text{(i)} & x_1 \geq 0, \quad x_2 \geq 0, \quad x_1^T x_2 = 0, \\ \text{(ii)} & x_{1j} \geq 0, \quad x_{2j} \geq 0, \quad x_{1j} x_{2j} = 0, \quad j = 1, \ldots, p, \\ \text{(iii)} & x_{1j} \geq 0, \quad x_{2j} \geq 0, \quad x_{1j} = 0 \ \text{ or } \ x_{2j} = 0, \quad j = 1, \ldots, p, \\ \text{(iv)} & \phi(x_{1j}, x_{2j}) = 0, \quad j = 1, \ldots, p, \end{array} \tag{1.21}$$

where ϕ denotes an NCP function (see Definition 1.13 in Section 1.4). Any reformulation of (1.20) by smooth constraints results in positively linear dependent gradients of the active constraints. Furthermore, note that at every feasible point x of (1.19) the complementarity condition does not admit strictly feasible points thus the MFCQ is also violated at every feasible point (see also [SS00]).

The combinatorial, nonconvex structure of the complementarity constraints may even prevent the ACQ (see Definition 1.3) to hold in a solution x^*. These unfavourable properties of MPECs make these problems rather difficult to solve. Moreover, due to the lack of the constraint qualifications, the standard necessary and sufficient conditions for NLPs cannot be applied straightforwardly to MPECs, such that this class of optimization problems needs to be handled with special care. In the following, we will consider these difficulties in more detail.

The stationarity condition for MPECs that corresponds to (1.14) is referred to as B-stationarity condition and is the most fundamental stationarity concept for MPECs. It is discussed in more detail in [FP99, SS00, Ye05].

Definition 1.8. Let $\mathcal{Z}$ be the feasible region of (1.19) and let $x^* \in \mathcal{Z}$. Then x^* is called *B-(Bouligand)-stationary*, if

$$\nabla f(x^*)^T d \geq 0 \qquad \forall\, d \in \mathcal{T}(\mathcal{Z}, x^*).$$

or equivalently

$$-\nabla f(x^*) \;\in\; (\mathcal{T}(\mathcal{Z}, x^*))^\circ\,.$$

Hence, B-stationarity is a necessary optimality condition for local solutions of an MPEC.

Definition 1.9. Let $x = (x_0, x_1, x_2) \in \mathbb{R}^{n+2p}$ be feasible for (1.19), then

1. The components of a pair (x_{1j}, x_{2j}) are called *degenerate*, if they do not satisfy strict complementarity, that is if $x_{1j} = 0$ and $x_{2j} = 0$.
2. If the pair satisfies strict complementarity, that is either $x_{1j} > 0$ or $x_{2j} > 0$, then x_{1j} and x_{2j} are called *nondegenerate*.
3. If all components of x_1 and x_2 are nondegenerate, then x is said to satisfy *strict complementarity*.

If a feasible point x of (1.19) possesses degenerate components x_{1j} and x_{2j} and $e_{1j}, e_{2j} \in \mathcal{T}(\mathcal{Z}, x)$, then the tangent cone $\mathcal{T}(\mathcal{Z}, x)$ is nonconvex. However, as the linearized tangent cone is always convex in this case the ACQ is inherently not satisfied. Thus, applying the KKT-condition (1.16) to a smoothly reformulated MPEC is inadequate as the following example illustrates.

Example 1.1.

$$\begin{aligned} &\textit{min}\; x_2 - x_1 \\ &\textit{subject to}\; x_1^2 + x_2^2 - 2x_2 \leq 0 && :\lambda \\ &\qquad\quad 0 \leq\; x_1 \;\perp\; x_2 \;\geq 0, && :\nu_1, \nu_2, \xi \end{aligned}$$

We reformulate the complementarity constraint by an alternative of (ii) of (1.21), such that we have two nonnegative multipliers ν_1 and ν_2 corresponding to $x_1 \geq 0$ and $x_2 \geq 0$, respectively, and a third multiplier ξ corresponding to $x_1 x_2 \leq 0$.

The solution of this MPEC is $(x_1^*, x_2^*) = (0,0)$ *and the tangent cone in* $(0,0)$ *is*

$$\mathcal{T}(\mathcal{Z},(0,0)) = \{d \in \mathbb{R}^2 \mid d_1 = 0 \ \ \textit{and} \ \ d_2 \geq 0\} = \{0\} \times \mathbb{R}_0^+$$

such that

$$\nabla f(0,0)^T d = -d_1 + d_2 = d_2 \geq 0 \quad \forall \ d \ \in \ \mathcal{T}(\mathcal{Z},(0,0))$$

and $(0,0)$ *is by Definition 1.8 B-stationary. It also satisfies (1.14) for all smooth reformulations. However, as*

$$\begin{pmatrix}-1\\1\end{pmatrix} + \lambda^* \begin{pmatrix}0\\-2\end{pmatrix} - \nu_1^* \begin{pmatrix}1\\0\end{pmatrix} - \nu_2^* \begin{pmatrix}0\\1\end{pmatrix} - \xi^* \begin{pmatrix}0\\0\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$

can only be satisfied if $\nu_1^* = -1 < 0$, $(0,0)$ *is not a KKT-point of the reformulated MPEC. This is not contradictory, since the linearized tangent cone in* $(0,0)$ *is*

$$\mathcal{T}_{lin}(\mathcal{Z},(0,0)) = \{d \in \mathbb{R}^2 \mid d_1 \geq 0 \ \ \textit{and} \ \ d_2 \geq 0\} = \mathbb{R}_0^+ \times \mathbb{R}_0^+$$

such that $\mathcal{T}(\mathcal{Z},(0,0)) \neq \mathcal{T}_{lin}(\mathcal{Z},(0,0))$ *and the ACQ does not hold. Hence, replacing the stationarity condition (1.14) by the KKT-conditions of a smoothly reformulated MPEC was not admissible.*

The problem that an MPEC cannot be treated as an ordinary NLP led to a variety of theoretical examinations of the MPEC, its tangent and normal cones and the development of some potentially more appropriate constraint qualifications and stationarity concepts. Most of them are discussed in more detail in [LPR96, KOZ98, FP99, FK05a, SS00].

Next to B-stationarity, the so-called C- and M-stationarity and in particular strong stationarity are subject of recent research [FK05b, Fle05, Ye99, Ye05]. To ease the notation of the conditions of these stationarities, we first introduce some further notations. Let

$$\mathcal{L}_{MPEC}(x,\lambda,\mu,\nu_1,\nu_2) = f(x) - \sum_{j=1}^{m} \lambda_j g_j(x) - \sum_{i=1}^{q} \mu_i h_i(x) - \nu_1^T x_1 - \nu_2^T x_2 \tag{1.22}$$

denote the Lagrangian of (1.19) and let

$$\begin{aligned} I_1(x) &= \{j \in \{1,\ldots,p\} \,:\, x_{1j} = 0\}, \\ I_2(x) &= \{j \in \{1,\ldots,p\} \,:\, x_{2j} = 0\} \end{aligned}$$

denote the index sets of the active constraints concerning the complementarity conditions for the MPEC (1.19). The definition of the mentioned stationarity concepts are then as follows.

Definition 1.10.

1. A point x^* is called *C-(Clarke)-stationary*, if there exist multipliers $\lambda^* \in \mathbb{R}^m$, $\mu^* \in \mathbb{R}^q$, $\hat{\nu}_1 \in \mathbb{R}^p$ and $\hat{\nu}_2 \in \mathbb{R}^p$, such that the following system of (in)equalities

$$\begin{array}{rcll} \nabla_x \mathcal{L}_{MPEC}(x^*, \lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2) & = & 0 & \\ h(x^*) & = & 0 & \\ g(x^*) & \geq & 0 & \\ \lambda^* & \geq & 0 & \\ g_i(x^*)\lambda_i^* & = & 0 & i = 1, \ldots, m \\ x_1^* & \geq & 0 & \\ x_2^* & \geq & 0 & \\ x_{1j}^* = 0 \text{ or } x_{2j}^* & = & 0 & j = 1, \ldots, p \\ x_{1j}^* \hat{\nu}_{1j} & = & 0 & j = 1, \ldots, p \\ x_{2j}^* \hat{\nu}_{2j} & = & 0 & j = 1, \ldots, p \end{array} \tag{1.23}$$

   is satisfied and

$$\forall\, j \in (I_1 \cap I_2)(x^*) : \ \hat{\nu}_{1j}\hat{\nu}_{2j} \geq 0$$

   holds.

2. A point x^* is called *M-(Mordukhovich)-stationary*, if there exist multipliers λ^*, μ^*, $\hat{\nu}_1$, and $\hat{\nu}_2$, such that (1.23) is satisfied and

$$\forall\, j \in (I_1 \cap I_2)(x^*) : \ \hat{\nu}_{1j}, \hat{\nu}_{2j} > 0 \quad \text{or} \quad \hat{\nu}_{1j}\hat{\nu}_{2j} = 0$$

   holds.

3. A point x^* is called *strongly stationary*, if there exist multipliers λ^*, μ^*, $\hat{\nu}_1$, and $\hat{\nu}_2$, such that (1.23) is satisfied and

$$\forall\, j \in (I_1 \cap I_2)(x^*) : \ \hat{\nu}_{1j} \geq 0 \ \text{ and } \ \hat{\nu}_{2j} \geq 0.$$

   holds.

Notice, that the stationarity concepts differ only in the additional condition on the multipliers $\hat{\nu}_{1j}$ and $\hat{\nu}_{2j}$ for indices $j \in (I_1 \cap I_2)(x^*)$, in other words for the degenerate components of x^*. Hence, if x^* satisfies strict complementarity, the stationarity conditions in Definition 1.10 are all equal. Otherwise we have the implications: *strong stationarity* $\Rightarrow$ *M-stationarity* $\Rightarrow$ *C-stationarity*.

In the case that x^* satisfies strict complementarity, the complementarity conditions locally correspond to an equality constraint, such that the combinatorial structure of the feasible set and accordingly of the tangent cone vanishes in the vicinity of x^*. Hence, replacing the complementarity conditions by such equality constraints and assuming the resulting constraints satisfy the ACQ in x^*, we then have $\mathcal{T}(\mathcal{Z}, x^*) = \mathcal{T}_{lin}(\mathcal{Z}, x^*)$ for the reformulated MPEC, such that we may apply the KKT-conditions to verify the B-stationarity of x^*.

The stationarity conditions of Definition 1.10 can also be stated in terms of the corresponding normal cone (see for example [FKO07]). To illustrate the differences of the normal cones we consider the simple MPEC

$$\begin{aligned} &\min\ f(x_1, x_2) \\ &\text{subject to } 0 \leq\ x_1\ \perp\ x_2\ \geq 0\,. \end{aligned} \tag{1.24}$$

The corresponding normal cones $\mathcal{N}_C(x_1^*, x_2^*)$, $\mathcal{N}_M(x_1^*, x_2^*)$ and $\mathcal{N}_S(x_1^*, x_2^*)$ for $(x_1^*, x_2^*) = (0, 0)$ to be C-, M- or strongly stationary if and only if $\nabla f(x_1^*, x_2^*) \in \mathcal{N}_u$, $u = C, M$ or S are demonstrated in Figure 1.2.

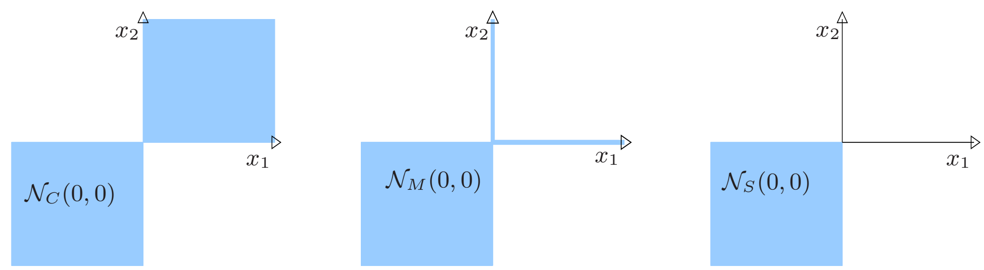


Figure 1.2: Corresponding normalcones in $(x_1, x_2) = (0, 0)$ for (1.24) for C-, M- and strong stationarity

In Example 1.1 the solution $(x_1^*, x_2^*) = (0, 0)$ was a B-stationary point that was M-stationary, though not strongly stationary as any multiplier that satisfies (1.23) is of the form $(\lambda^*, \hat{\nu}_1, \hat{\nu}_2) = (1/2(1-a), -1, a)$ with $a \in [0, 1]$. This illustrates that not every B-stationary point is strongly stationary, thus the strong stationarity is not always a necessary optimality condition. However, the next example demonstrates that the M-stationarity conditions are too weak to guarantee B-stationarity in general.

Example 1.2.

$$\begin{aligned} &\mathit{min} && f(x) = -2x_1 + (x_2 - x_1)^2 \\ &\mathit{subject\ to} && 0 \leq x_1 \quad \perp \quad x_2 \geq 0\,, \end{aligned}$$

then $(x_1^*, x_2^*) = (0, 0)$ *is M-stationary with multipliers* $(\hat{\nu}_1, \hat{\nu}_2) = (-2, 0)$*. However,* $d = (1, 0)$ *is a feasible descent direction as* $\nabla f(x^*)^T d = -2d_1 = -2 < 0$ *and*

$$\begin{aligned} d \ \in\ \mathcal{T}(\mathcal{Z}, (0,0)) &= \{d \in \mathbb{R}^2 \mid \min(d_1, d_2) = 0\} \\ &= \mathbb{R}_0^+ \times \{0\} \ \cup\ \{0\} \times \mathbb{R}_0^+\,. \end{aligned}$$

Thus, given an MPEC we need further properties to decide which stationarity condition is suitable to characterize its candidates of local solutions. As discussed in [Fle05] and [SS00] the fulfillment or non-fulfillment of certain constraint qualifications can serve as such decision support. There exists a variety of constraint qualifications for MPECs. The strongest condition that is mainly used and furthermore easiest to verify is the MPEC-LICQ.

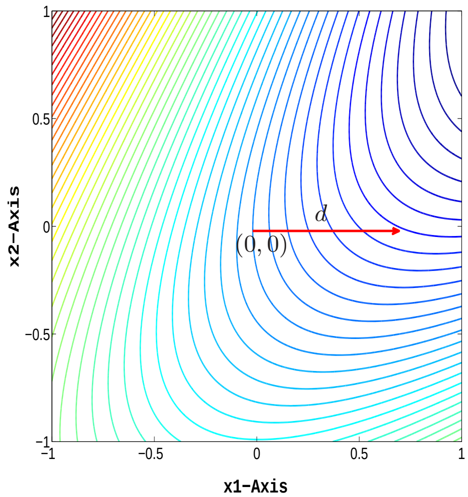


Figure 1.3: Contours of f and direction $d = -\delta\nabla f(0,0)$ ($\delta \in (0,1)$) of Example 1.2.

Definition 1.11. Let x^* be a feasible point of (1.19). Then the *MPEC-LICQ (MPEC-Linear Independence Constraint Qualification)* is said to hold at x^* if the gradients

$$\begin{array}{rl} \nabla h_i(x^*) & i \in \{1,\dots,q\}, \\ \nabla g_j(x^*) & j \in \ I_g(x^*), \\ e_{n+j} & j \in \ I_1(x^*), \\ e_{n+p+j} & j \in \ I_2(x^*), \end{array}$$

where e_j denotes the j-th unit vector in $\mathbb{R}^{n+2p}$, are linearly independent.

Notice, that the definition of the MPEC-LICQ differs from the standard LICQ: the complementarity constraints are smoothly reformultated by (ii) of (1.21) to obtain differentiable constraints, however only the gradients of the conditions $x_1 \geq 0$ and $x_2 \geq 0$ are considered and the gradients concerning the condition $x_{1j}x_{2j} = 0$ $j = 1,\dots,p$ are left out. However, the MPEC-LICQ represents the standard LICQ for the following related optimization problem known as *Relaxed Nonlinear Program* [FL04a, FP99, SS00].

$$\begin{array}{llrcll} RNLP & \min & & f(x) & & \\ & \text{subject to} & & h(x) & = & 0 \\ & & & g(x) & \geq & 0 \\ & & x_{1j} = 0, & x_{2j} & \geq & 0 \quad j \in (I_1 \backslash I_2)(x^*) \\ & & x_{2j} = 0, & x_{1j} & \geq & 0 \quad j \in (I_2 \backslash I_1)(x^*) \\ & & x_{1j} \geq 0, & x_{2j} & \geq & 0 \quad j \in (I_1 \cap I_2)(x^*) \end{array}$$

Note that the feasible set of RNLP in this case is determined by the point x^*. Sometimes, it is more useful to determine the feasible set of RNLP not by x^* but by a different point, for instance by the current iterate x^k of an iterative method. This, however, needs then to be clearly indicated to avoid ambiguity.

It can be shown that strong stationarity corresponds to the standard stationarity for NLPs (confer (1.16)) applied to the RNLP [FL04a]. Furthermore, if the MPEC-LICQ holds in x^*, B-stationarity of x^* implies strong stationarity [SS00]. Hence, under the assumption that the MPEC-LICQ holds, strong stationarity is a necessary optimality condition. More relations concerning constraint qualifications and stationarity conditions for MPECs can also be found in [FK05a, FK05b, Fle05, Ye05].

Next, before defining the second order optimality conditions, we introduce the following sets of critical directions in x^*, following the definitions in [RW04]. Let

$$\begin{aligned}\bar{\mathcal{S}}(x^*,\lambda^*,\mu^*,\hat{\nu}_1,\hat{\nu}_2) \;=\; \{\,(d_0,d_1,d_2)\in\mathbb{R}^{n+2p}\backslash\{0\}\,|\; & \nabla h_i(x^*)^T d = 0, && i\in\{1,\dots,q\}\\ & \nabla g_j(x^*)^T d = 0, && j\in I_g^+(x^*,\lambda^*)\\ & \nabla g_j(x^*)^T d \geq 0, && j\in I_g^0(x^*,\lambda^*)\\ & d_{1j} = 0, && j\in (I_1\backslash I_2)(x^*)\\ & d_{1j} = 0, && j\in (I_1\cap I_2)(x^*) \text{ and } \hat{\nu}_{1j}>0\\ & d_{1j} \geq 0, && j\in (I_1\cap I_2)(x^*) \text{ and } \hat{\nu}_{1j}=0\\ & d_{2j} = 0, && j\in (I_2\backslash I_1)(x^*)\\ & d_{2j} = 0, && j\in (I_1\cap I_2)(x^*) \text{ and } \hat{\nu}_{2j}>0\\ & d_{2j} \geq 0, && j\in (I_1\cap I_2)(x^*) \text{ and } \hat{\nu}_{2j}=0\,\}\end{aligned}$$

and

$$\begin{aligned}\mathcal{S}^*(x^*,\lambda^*,\mu^*,\hat{\nu}_1,\hat{\nu}_2) = \;& \{(d_0,d_1,d_2)\in\bar{\mathcal{S}}(x^*\lambda^*,\mu^*,\hat{\nu}_1,\hat{\nu}_2)\,|\\ & \min(d_{1j},d_{2j}) = 0,\quad j\in (I_1\cap I_2)(x^*) \text{ and } \hat{\nu}_{1j}=\hat{\nu}_{2j}=0\,\},\end{aligned}$$

as well as

$$\begin{aligned}\tilde{\mathcal{S}}(x^*,\lambda^*,\mu^*,\hat{\nu}_1,\hat{\nu}_2) = \{\,(d_0,d_1,d_2)\in\mathbb{R}^{n+2p}\backslash\{0\}\;|\; & \nabla h_i(x^*)^T d = 0, && i\in\{1,\dots,q\}\\ & \nabla g_j(x^*)^T d = 0, && j\in I_g^+(x^*,\lambda^*)\\ & d_{1j} = 0, && j\,:\,\hat{\nu}_{1j}\neq 0,\\ & d_{2j} = 0, && j\,:\,\hat{\nu}_{2j}\neq 0\,\}\end{aligned}$$

be the MPEC specific sets of critical directions. The first two sets differ only in the additional condition on the components of a direction $d\in\mathbb{R}^{n+2p}\backslash\{0\}$ corresponding to indices $j\in(I_1\cap I_2)(x^*)$, with $\hat{\nu}_{1j}=\hat{\nu}_{2j}=0$, whereas for the third set there exists no condition on directions $d\in\mathbb{R}^{n+2p}\backslash\{0\}$ for indices for which the multipliers vanish. The defined sets hence satisfy the relationship $\mathcal{S}^*(x^*,\lambda^*,\mu^*,\hat{\nu}_1,\hat{\nu}_2)\subseteq\bar{\mathcal{S}}(x^*,\lambda^*,\mu^*,\hat{\nu}_1,\hat{\nu}_2)\subseteq\tilde{\mathcal{S}}(x^*,\lambda^*,\mu^*,\hat{\nu}_1,\hat{\nu}_2)$.

Definition 1.12.

1. If x^* is a strongly stationary point of (1.19) with multipliers $(\lambda^*,\mu^*,\hat{\nu}_1,\hat{\nu}_2)$ satisfying (1.23), then x^* is said to satisfy the *MPEC-SOSC (MPEC-Second Order Sufficient Condition)*, if
$$d^T\nabla^2_{xx}\mathcal{L}_{MPEC}(x^*,\lambda^*,\mu^*,\hat{\nu}_1,\hat{\nu}_2)\,d>0 \qquad (1.25)$$
holds for all $d=(d_0,d_1,d_2)\in\mathcal{S}^*(x^*,\lambda^*,\mu^*,\hat{\nu}_1,\hat{\nu}_2)$.

2. If (1.25) holds for all $d\in\bar{\mathcal{S}}(x^*,\lambda^*,\mu^*,\hat{\nu}_1,\hat{\nu}_2)$, then x^* is said to satisfy the *RNLP-SOSC (RNLP-Second Order Sufficient Condition)*.

3. If (1.25) holds for all $d \in \tilde{\mathcal{S}}(x^*, \lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2)$, then x^* satisfies the *SSOSC (Strong Second Order Sufficient Condition).*

Considering the definitions of these sets, we obtain the relationship SSOSC $\Rightarrow$ RNLP-SOSC $\Rightarrow$ MPEC-SOSC for the second order conditions. Moreover note that the RNLP-SOSC corresponds to the standard SOSC (confer (1.18)) applied to RNLP.

Finally, we state a result concerning the second order sufficient conditions, which is a simple application of the corresponding result in [SS00].

Theorem 1.3. *If x^* is a strongly stationary point of the MPEC (1.19) that satisfies the MPEC-SOSC for any multiplier vector satisfying the strong stationarity conditions, then x^* is a strict local minimum of (1.19).*

Proof. If for any $d \neq 0$ satisfying

$$\begin{aligned} \nabla g_j(x^*)^T d &\geq 0, && j \in I_g(x^*) \\ \nabla h(x^*)^T d &= 0, && j \in \{1, \ldots, q\} \\ d_{1j} &= 0 && j \in (I_1 \backslash I_2)(x^*) \\ d_{2j} &= 0 && j \in (I_2 \backslash I_1)(x^*) \\ \min(d_{1j}, d_{2j}) &= 0 && j \in (I_1 \cap I_2)(x^*)\,, \end{aligned} \tag{1.26}$$

and additionally $\nabla f(x^*)^T d = 0$ it holds for any suitable multiplier vector $(x^*, \lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2)$

$$d^T \nabla^2_{xx} \mathcal{L}_{MPEC}(x^*, \lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2)\, d > 0\,,$$

then x^* is a strict local minimum of (1.19) by Theorem 7 of [SS00]. Hence, if we can show that any $d \neq 0$ satisfying (1.26) and $\nabla f(x^*)^T d = 0$ is an element of $\mathcal{S}^*(x^*,\, \lambda^*,\, \mu^*,\, \hat{\nu}_1,\, \hat{\nu}_2)$, then x^* is a strict local minimum of (1.19).

Therefore, let $d \neq 0$ satisfy (1.26) and additionally $\nabla f(x^*)^T d = 0$. It then follows in combination with the strong stationarity of x^* that

$$0 = \nabla f(x^*)^T d = \sum_{j \in I_g^+(x^*, \lambda^*)} \lambda_j^* \nabla g_j(x^*)^T d + \sum_{i=1}^{q} \mu_i^* \nabla h_i(x^*)^T d + \sum_{j \in (I_1 \cap I_2)(x^*)} (\, \hat{\nu}_{1j}\, d_{1j} + \hat{\nu}_{2j}\, d_{2j}\,)\,.$$

This, however, can only be satisfied if $d_{1j} = 0$ for all $j \in (I_1 \cap I_2)(x^*)$ with $\hat{\nu}_{1j} > 0$ and $d_{2j} = 0$ for all $j \in (I_1 \cap I_2)(x^*)$ with $\hat{\nu}_{2j} > 0$, respectively, as well as $\nabla g_j(x^*)^T d = 0$ for all $j \in I_g^+(x^*, \lambda^*)$. Hence we can derive the following additional conditions for the direction d

$$\begin{aligned} \nabla g_j(x^*)^T d &= 0, && j \in I_g^+(x^*, \lambda^*) \\ d_{1j} &= 0, && j \in (I_1 \cap I_2)(x^*) \quad \text{and} \quad \hat{\nu}_{1j} > 0 \\ d_{2j} &= 0, && j \in (I_1 \cap I_2)(x^*) \quad \text{and} \quad \hat{\nu}_{2j} > 0\,. \end{aligned}$$

and conclude that $d \in \mathcal{S}^*(x^*,\, \lambda^*,\, \mu^*,\, \hat{\nu}_1,\, \hat{\nu}_2)$. □

The MPEC-SOSC is thus sufficient to guarantee the local optimality of a strong stationary point x^* for (1.19) .

1.4 Existing Approaches

Recent numerical approaches to solve MPECs follow a variety of directions. A survey of the different ideas can be found in [FL04b]. The majority of them have a special treatment of the complementarity constraints in common, either by relaxation or as a penalty term or elsewise.

Regularization schemes are based on introducing a parameter $t > 0$ such that the resulting parametrized nonlinear program P_t is regular and for $t = 0$ the original MPEC is recovered. Although the different methods vary in the way of the parametrization, they share the feature that the smaller $t > 0$, the better P_t approximates the MPEC. Thus, solving a sequence of such problems P_t, a sequence of approximate solutions is obtained, which converges to a solution of the original MPEC, under suitable assumptions.

Similar ideas concern penalization techniques, where the complementarity constraints are added as a penalty term to the objective function. In some of the approaches the penalty term does not only contain the complementarity constraints but also the ordinary nonlinear constraints of the MPEC.

Finally, there are approaches that concern the direct or adapted application of particular NLP methods, as Sequential Quadratic Programming, Interior Point and Trust-Region Methods to solve MPECs.

In the sequel, we will give a brief overview of each of these directions starting with regularization methods. To keep the consistency of the notation, we sometimes slightly adapt the notation of the approaches and the corresponding results. The modifications, however, mainly concern the form of the complementarity constraints that are discussed.

Smoothing and Regularization Methods

Often smoothing methods for MPECs are based on a reformulation the complementarity constraint

$$0 \leq x_1 \;\perp\; x_2 \geq 0$$

with $(x_1, x_2) \in \mathbb{R}^{2p}$ by means of an NCP function, which is defined as follows [GK02].

Definition 1.13. A function $\phi : \mathbb{R}^2 \longrightarrow \mathbb{R}$ is called *NCP (Nonlinear Complementarity Problem) function* , if

$$\phi(a,b) = 0 \quad \Longleftrightarrow \quad a \geq 0,\; b \geq 0,\; ab = 0$$

holds.

Two main examples for NCP functions are the minimum function

$$\phi^{Min}(a,b) := 2\min(a,b)$$

and the Fischer-Burmeister function

$$\phi^{FB}(a,b) := a + b - \sqrt{a^2 + b^2}\,.$$

Reformulating accordingly the MPEC, we obtain the problem

$$\begin{array}{rrl} \min & f(x) & \\ \text{subject to} & h(x) = 0 & \\ & g(x) \geq 0 & \\ & \phi(x_{1j}, x_{2j}) = 0 & j = 1, \ldots, p\,, \end{array} \tag{1.27}$$

which has an identical feasible region to the original problem. However, since these NCP functions are not differentiable in $(a, b) = (0, 0)$, it is not suitable to apply an NLP method to (1.27). Therefore, one parametrizes (smoothes) the NCP function in a way, such that the parametrized function is differentiable for every strictly positive parameter t and coincides with the original function for $t = 0$. The smoothed minimum function is

$$\phi_t^{Min}(a, b) := a + b - \sqrt{(a-b)^2 + 4t^2}\,.$$

and the smoothed Fischer-Burmeister function has the form

$$\phi_t^{FB}(a, b) := a + b - \sqrt{a^2 + b^2 + 2t^2}\,.$$

Using a smoothed NCP function ϕ_t to solve the MPEC, we hence get a family of parameterized NLPs with differentiable constraint functions:

$$\begin{array}{lrrl} P_t & \min & f(x) & \\ & \text{subject to} & h(x) = 0 & \\ & & g(x) \geq 0 & \\ & & \phi_t(x_{1j}, x_{2j}) = 0 & j = 1, \ldots, p\,. \end{array} \tag{1.28}$$

The feasible regions of these problems are cumulatively approximating the feasible region of the MPEC for every positive, decreasing sequence of parameters $t_k \searrow 0$. Hence, to solve the MPEC, one has to solve such a sequence of smoothed problems P_{t_k}.

Facchinei et al. [FJQ99] present such a solution approach using the smoothed Minimum function and approximately solving a sequence of problems P_{t_k}.

Jiang and Ralph [JR00] distinguish two similar approaches. The first one, the explicit smooth SQP method, is based on solving a sequence of modified quadratic models of P_{t_k} at which the parameter $t_k > 0$ is updated after each solution of such Quadratic Programs (QP). The alternative, the implicit smooth SQP method, is based on solving a sequence of modified SQP subproblems of

$$\begin{array}{lrrl} P & \min_{x,t} & f(x) & \\ & \text{subject to} & h(x) = 0 & \\ & & g(x) \geq 0 & \\ & & \phi(x_{1j}, x_{2j}, t) = 0 & j = 1, \ldots, p \\ & & e^t - 1 = 0 & \end{array}$$

where t is treated as an additional variable and for each t it holds $\phi(x_{1j}, x_{2j}, t) := \phi_t(x_{1j}, x_{2j})$ and e denotes Euler's constant. The terms explicit and implicit correspond to the update of t since for the first variant one has to update the parameter explicitly, whereas for the second one the constraint $e^t - 1 = 0$ implicitly forces $t = 0$.

Zhu et al. [ZLZ07] propose to solve the MPEC using the nonsmooth Fischer-Burmeister function, however, combined with the idea of successive approximation, at which the exact nonsmooth function is used outside $B_\varepsilon(0,0)$ and a smooth approximation of it inside of $B_\varepsilon(0,0)$. The resulting function Φ_ε of this combination then replaces the original Fischer-Burmeister function and problems of the form

$$\begin{array}{lcr} P_\varepsilon & \min & f(x) \\ & \text{subject to} & h(x) = 0 \\ & & g(x) \geq 0 \\ & & \Phi_\varepsilon(x_{1j}, x_{2j}) = 0. \end{array}$$

are solved.

Zhu et al. [ZLZ07] consider global solutions of the MPEC and the reformulated problems as well as the convergence of a sequence of stationary points of P_{ε_k} for a decreasing sequence of parameters ε_k. One main assumption of their convergence result concerns the nondegeneracy of the accumulation point that is considered. Numerical results have not yet been reported.

Another solution approach is the regularization scheme proposed by Scholtes in [Sch01] and further discussed in [RW04]. Therein, the complementarity constraint (1.20) is relaxed by a strictly positive parameter t

$$x_{1j} \geq 0, \quad x_{2j} \geq 0, \quad x_{1j}x_{2j} \leq t, \quad j = 1, \ldots, p$$

such that the resulting parameterized problem is

$$\begin{array}{lcrl} NLP(t) & \min & f(x) & \\ & \text{subject to} & h(x) = 0 & \\ & & g(x) \geq 0 & \\ & & x_1 \geq 0, \; x_2 \geq 0 & \\ & & x_{1j}x_{2j} \leq t & \quad j = 1, \ldots, p\,. \end{array} \tag{1.29}$$

The bilinear function $\phi^{Bil}(a,b) = ab$ is not an NCP function, since the implication

$$\phi^{Bil}(a,b) = 0 \quad \Longrightarrow \quad a \geq 0, \; b \geq 0$$

does not hold. Therefore, these conditions have to be additionally included in the constraints of NLP(t). The main convergence results given in [Sch01] can be stated as follows.

Theorem 1.4. *Let (t_k) be a sequence of positive scalars tending to zero, let x^k be a stationary point of $NLP(t_k)$ tending to $\bar{x}$ and suppose the MPEC-LICQ holds at $\bar{x}$. Let*

$$I_0 = \{\, j \mid x_{1j}^k x_{2j}^k = t_k \quad \text{for infinitely many } k \,\}\,.$$

Then the following statements hold:

1. *For every sufficiently large k NLP(t_k) has unique Langrange multipliers λ^k, μ^k, γ^k,ν^k, δ^k at x^k.*

2. *The point $\bar{x}$ is a C-stationary point of NLP(0) with unique multipliers $\bar{\lambda}$, $\bar{\mu}$, $\hat{\nu}_1$ and $\hat{\nu}_2$ that satisfy*

$$
\begin{aligned}
\bar{\lambda}_j &= \lim_{k\to\infty} \lambda_j^k \geq 0 \\
\bar{\mu}_i &= \lim_{k\to\infty} \mu_i^k \\
\hat{\nu}_{1k} &= \lim_{k\to\infty} \nu_{1k}^k \geq 0 \qquad k \notin I_0 \\
\hat{\nu}_{2k} &= \lim_{k\to\infty} \nu_{2k}^k \geq 0 \qquad k \notin I_0 \\
\hat{\nu}_{1m} &= \lim_{k\to\infty} -\delta_{1m}^k x_{1m}^k \leq 0 \qquad m \in I_0 \\
\hat{\nu}_{2m} &= \lim_{k\to\infty} -\delta_{2m}^k x_{1m}^k \leq 0 \qquad m \in I_0
\end{aligned}
$$

3. *The point $\bar{x}$ is B-stationary if and only if $\hat{\nu}_{1m} = \hat{\nu}_{2m} = 0$ for all $m \in (I_1 \cap I_2)(\bar{x}) \cap I_0$.*

4. *If in addition the second order optimality conditions (see Definition 1.7) hold at each x^k, then $\bar{x}$ is M-stationary.*

Theorem 1.5. *Suppose $\bar{x}$ is a B-stationary point of NLP(0) and MPEC-LICQ as well as SSOSC hold at $\bar{x}$. Assume further that $\hat{\nu}_{1j} \neq 0$ if $j \in I_1(\bar{x})$ and $\hat{\nu}_{2j} \neq 0$ if $j \in I_2(\bar{x})$ for every $j = 1, \ldots, p$. Then there exists an open neighbourhood $\mathcal{U}$ of $\bar{x}$, a scalar $\bar{t} > 0$ and a piecewise smooth function $x : (-\bar{t}, \bar{t}) \longrightarrow \mathcal{U}$ such that $x(t)$ is the unique stationary point of NLP(t) for every $0 < t < \bar{t}$. Moreover, $x(t)$ satisfies second order sufficient conditions.*

The result of Theorem 1.5 has been improved and extended by Ralph and Wright in [RW04]. However, since it is not of particular importance to which of these variants of Theorem 1.5 we compare our results with, we confine ourselves to state here only the basic result of [Sch01].

A variant of this approach is discussed in [BS07], where the inequality constraints $x_{1j}x_{2j} \leq t$ are replaced by equality constraints $x_{1j}x_{2j} = t$ for $j = 1, \ldots, p$. Preliminary numerical experiments mentioned in [Sch01] reveal, however, that this solution approach is not likely to be advantageous compared with the one using $x_{1j}x_{2j} \leq t$.

In contrast to the feasible region of P_t for $t > 0$, the feasible region of $NLP(t)$ contains the feasible region of the MPEC. Thus, by solving $NLP(t)$ for a strictly positive parameter t one might already find a solution of the MPEC without having to solve further problems $NLP(t)$ for reduced parameters t.

Another modified regularization scheme is proposed by Fukushima and Lin [FL05]. They consider solving a sequence of parameterized problems

$$
\begin{array}{lrl}
\min & f(x) & \\
\text{subject to} & h(x) = 0 & \\
& g(x) \geq 0 & \\
& x_{1j}x_{2j} \leq t^2 & j = 1, \ldots, p \\
& (x_{1j} + t)(x_{2j} + t) \geq t^2 & j = 1, \ldots, p\,.
\end{array}
$$

The convergence results they state resemble the results of Theorem 1.4.

Finally, Lin and Fukushima [LF03a] have also analyzed a relaxation method that is based on solving a sequence of problems (applied to (1.19))

$$\begin{array}{rl} \min & f(x) \\ \text{subject to} & h(x) = 0 \\ & g(x) \geq 0 \\ & x_1 \geq 0 \\ & (e_j^k - x_1)\, x_2 \geq 0 \qquad j = 0, \ldots, p\,, \end{array} \tag{1.30}$$

where

$$e_j^k := (1/k)\,\mathbf{e} + k\, e_j \qquad j = 0, \ldots, p$$

with $e_0 := (0, \ldots, 0)^T$, $\mathbf{e} = (1, \ldots, 1)^T$, e_j denotes the j-th unit vector and $k \in \mathbb{N}$. They discuss the limiting behaviour of global solutions as well as of stationary points of a sequence of problems (1.30) and report some computational results of their method.

Penalty Methods

Other ideas to deal with the equilibrium or complementarity constraints use penalty approaches of various kinds. Luo et al. [LPRW96] consider MPECs of the form (1.1), which are then, however, reformulated by the KKT-conditions for the VI, such that the resulting MPEC has a similar form to (1.9). Then, a so-called *penalty equivalent* is introduced which applied to (1.9) corresponds to

$$\begin{array}{rl} \min_{x,y} & F(x,y) + \varpi\, r(x,y,\lambda)^{1/N} \\ \text{subject to} & x \quad \in \quad \mathcal{X} \\ & G(x,y) \leq 0 \\ & \lambda \geq 0 \\ & \|\lambda\| \leq c\,, \end{array} \tag{1.31}$$

with

$$r(x,y,\lambda) := \|\nabla_y \mathcal{L}(x,y,\lambda)\| + \sum ((g_i(x,y))^+ + \lambda_i |g_i(x,y)|)$$

and $N \in \mathbb{N}$. It is proved that under suitable conditions, the solution set of problem (1.31) is identical to the one of the original MPEC (1.1).

A further specification of this approach for NCP constrained programs is discussed in [LPR96]. Applied to problem (1.19), the penalty equivalent then has the form

$$\begin{array}{rl} \min & f(x) + \varpi\, x_1^T x_2 \\ \text{subject to} & h(x) = 0 \\ & g(x) \geq 0 \\ & x_1 \geq 0,\;\; x_2 \geq 0\,, \end{array}$$

which, again under suitable assumptions, is shown to have an identical solution set to (1.19). The results for the exact penalty approach of [LPRW96] have been further extended in [LF03b].

Another exact penalty approach for a similar type of MPECs is considered in [SS99]. Therein Scholtes and Stöhr prove the existence of an exact penalty function for such MPECs under suitable assumptions. Furthermore, they suggest a solution approach that is based on an extension of the $S\ell_1QP$ method of Fletcher [Fle00], which is an SQP trust-region method on the basis of an ℓ_1 penalty function. Therein, linearized constraints are directly substituted into the QP objective function by an ℓ_1 penalty term, such that the only constraints of the QP subproblems are those implied by the trust-region. In [SS99] a suitable extension of the ℓ_1 penalty term is proposed that handles the complementarity constraints specifically. Applied to (1.19) the proposed ℓ_1 penalty term of the QP subproblem in a current iterate x^k has the form:

$$\begin{aligned} p_{I_k}(d) \;\; = \;\; & \sum_{\substack{j=1 \\ j\in I_k}}^{p} \max\{|x_{1j}^k + d_{1j}|, -(x_{2j}^k + d_{2j})\} + \sum_{\substack{j=1 \\ j\notin I_k}}^{p} \max\{|x_{2j}^k + d_{2j}|, -(x_{1j}^k + d_{1j})\} \\ & + \sum_{j=1}^{q} |h_j(x^k) + \nabla h_j(x^k)^T\, d| + \sum_{j=1}^{m} (-g_j(x^k) - \nabla g_j(x^k)^T\, d)^+ \end{aligned}$$

with $I_k \subseteq \{1, \ldots, p\}$ and I_k is adapted appropriately after each solution of a QP subproblem. The QP subproblems then have the form

$$\begin{aligned} \min \quad & q_k(d) + \varpi p_{I_k}(d) \\ \text{subject to} \quad & \|d\| \leq \Delta_k\,, \end{aligned}$$

where $q_k(d) = \nabla f(x^k)^T\, d + 1/2\, d^T\, \nabla_{xx}^2 \mathcal{L}(x^k, \pi^k)\, d$ with $\mathcal{L}(x, \pi)$ being the corresponding Lagrangian function and Δ_k being the actual trust-region radius.

A different penalty approach is suggested in [HR04]. Hu and Ralph discuss the convergence properties of the penalty problem (applied to (1.19))

$$\begin{aligned} \min \quad & f(x) + \varpi p(x_1, x_2) \\ \text{subject to} \quad & h(x) = 0 \\ & g(x) \geq 0 \\ & x_1 \geq 0,\;\; x_2 \geq 0\,, \end{aligned} \tag{1.32}$$

with $p(x_1, x_2)$ satisfying appropriate conditions such that reasonable convergence properties result. Two functions satisfying these conditions are

$$p(x_1, x_2) = \sum_{j=1}^{p} x_{1j} x_{2j} \qquad \text{or} \qquad p(x_1, x_2) = \sum_{j=1}^{p} \phi^{FB}(x_{1j}, x_{2j})^3\,.$$

In contrast to the methods that are concerned with penalty equivalents of the MPEC, the penalty method of [HR04] is not exact in the sense that the solution set of problem (1.32) is not considered to be identical to the one of (1.19). However, under suitable assumptions it is shown that if an iterate x^0 is a stationary point of (1.32) that is close enough to a B-stationary point x^*, and the corresponding penalty parameter ϖ is sufficiently large, then the generated sequence of stationary points x^k satisfies $x^k \to x^*$.

SQP Methods

The approaches we will discuss next consider SQP methods for MPECs. They differ not only in the specific SQP variant that is proposed, but also in the class of MPECs the algorithms are designed for.

We begin with solution approaches for MPECs with linear constraints. Fukushima et al. [FLP98] propose an SQP algorithm that is applied to a reformulation of the MPEC using the smoothed Fischer-Burmeister function $\phi_t^{FB}(a,b)$. A line-search applied to an ℓ_1 penalty function is further used to promote global convergence. Their algorithm is shown to be globally convergent under suitable conditions including nondegeneracy of the solution. A further modification of the algorithm is discussed in [JLM06].

Another kind of SQP algorithm for this class of MPECs is suggested by Zhang et al. [ZLW04]. Their SQP method is based on an approximately active search and similar to the *Piecewise SQP Approach* explained in [LPR96]. The algorithm is globally convergent and is further extended to a more general class of MPECs in [LY07].

Finally, Júdice et al. [JSRF07] present an active set algorithm designed for MPECs with linear constraints.

For general MPECs of the form of (1.19) Fletcher et al. [FLRS06] recommend to apply directly ordinary SQP methods to (1.19) reformulated as the NLP

$$\begin{array}{rrcl} \min & f(x) & & \\ \text{subject to} & h(x) & = & 0 \\ & g(x) & \geq & 0 \\ & x_1 \geq 0,\ x_2 & \geq & 0 \\ & x_1^T x_2 & \leq & 0\,. \end{array} \tag{1.33}$$

They establish superlinear convergence near a strongly stationary point x^* under reasonable assumptions, including the MPEC-LICQ and a second order sufficient condition to hold in x^*. Convincing numerical results for this approach are presented in [FL02b].

An extension of the results of [FLRS06] is considered by Leyffer in [Ley06]. Therein, Leyffer applies an ordinary SQP method to problems that are exact reformulations of (1.19) using nonsmoothed NCP functions. The convergence results of [FLRS06] can also be established for these reformulations.

Another extension is considered by Anitescu in [Ani05b]. He discusses the usage of SQP methods with an *elastic-mode*. In [ATW07] Anitescu et al. demonstrate global convergence of an SQP method with elastic-mode to C-, M- or strongly stationary points.

Interior Point Methods

Recently, Interior Point Methods (see Section 3.2) have also been used to solve MPECs and receive an increasing interest. The methods we discuss in the following are alike, as they all originate in solving a barrier problem, where the complementarity constraints receive a special treatment. The methods differ, however, in the way they treat the complementarity constraints.
To begin with, Liu and Sun [LS04] propose a method that solves a relaxed barrier problem

of the form

$$\begin{array}{rl} \min & f(x) - \rho \sum_{j=1}^{m+3p} \ln(s_j) \\ \text{subject to} & h(x) = 0 \\ & H(x,s,t) = 0\,, \end{array} \tag{1.34}$$

where $s = (s_g, s_1, s_2, s_c) \in \mathbb{R}^{m+3p}$,

$$H(x,s,t) = \begin{pmatrix} g(x) - s_g \\ x_1 - s_1 \\ x_2 - s_2 \\ X_1 x_2 + s_c - t\mathbf{e} \end{pmatrix},$$

and $X_1 = diag(x_{1j}) \in \mathbb{R}^{p\times p}$, $\rho > 0$ is the barrier parameter and $t > 0$ denotes the relaxation parameter. Note that we obtain (1.34), if we introduce slack variables s_i to problem $NLP(t)$ and remove the remaining inequality constraints $s \geq 0$ by inserting the corresponding barrier-term $\rho \sum_{j=1}^{m+3p} \ln(s_j)$ into the objective function.

Liu and Sun further propose to link the reduction of the barrier parameter ρ with the reduction of the relaxation parameter t by choosing ρ to be a fixed fraction of t. They prove global convergence of their algorithm under suitable assumptions.

A similar interior point approach is considered by Raghunatan and Biegler [BR05]. They introduce a barrier problem akin to (1.34). However, they do not introduce the slack variables s_1 and s_2 for the inequality constraints $x_1 \geq 0$ and $x_2 \geq 0$ but remove these constraints by directly adding a corresponding part to the barrier term. Thus, their objective function has the form

$$f(x) - \rho \sum_{j=1}^{m} \ln(s_{gj}) - \rho \sum_{j=1}^{p} \ln(x_{1j}) - \rho \sum_{j=1}^{p} \ln(x_{2j}) - \rho \sum_{j=1}^{m} \ln(s_{cj})$$

and the corresponding constraint function $H(x,s,t)$ is

$$H(x,s,t) = \begin{pmatrix} g(x) - s_g \\ X_1 x_2 + s_c - t\,\mathbf{e} \end{pmatrix}.$$

They also link the updates of the barrier parameter and of the relaxation parameter. The algorithm that is proposed attains quadratic convergence to a strongly stationary point under suitable assumptions.

The interior point method of DeMiguel et al. [DFNS05] differs from the preceding two in that it uses a two-sided relaxation scheme. The scheme is obtained by not only relaxing the constraint $x_{1j}x_{2j} = 0$ but also the nonnegativity condition on x_{1j} and x_{2j}. To be more precise, the condition

$$0 \leq x_{1j} \;\perp\; x_{2j} \geq 0,$$

is replaced by

$$x_{1j} \geq -\delta_{1j}\,, \qquad x_{2j} \geq -\delta_{2j}\,, \qquad x_{1j}x_{2j} \leq t_j\,.$$

Slack variables are again introduced and a problem similar to (1.34) is solved. The relaxation parameters are then reduced in each iteration according to the sign of a corresponding multiplier estimate for a strongly stationary point. Either δ_{ij} for $i = 1, 2$ is reduced or the parameter t_j is reduced, but not both. This procedure ensures that the strictly feasible set does not correspond to the empty set even in the limit, in contrast to the other relaxation methods we described before. This property makes it attractive for interior-point algorithms. The algorithm presented in [DFNS05] is shown to converge superlinearly near a strongly stationary point x^* if x^* satisfies suitable conditions including MPEC-LICQ and SSOSC.

Finally, an interior point approach is analyzed by Leyffer et al. [LLCN06] that is based on a barrier problem, where the complementarity constraints are not relaxed by introducing a relaxation parameter t. Instead the conditions $x_{1j}x_{2j} = 0$ are removed from the constraints and a corresponding penalty term is added to the objective function. The barrier problem that is solved is

$$\begin{array}{rl} \min & f(x) + \varpi\, x_1^T x_2 - \rho \sum_{j=1}^{m} \ln(s_{gj}) - \rho \sum_{j=1}^{p} \ln(x_{1j}) - \rho \sum_{j=1}^{p} \ln(x_{2j}) \\ \text{subject to} & h(x) = 0 \\ & g(x) - s_g = 0 \end{array}$$

Global and local convergence results are discussed and numerical experiences for different penalty parameter update modes are reported.

Other Methods

Although we tried to cover a large variety of methods to solve MPECs, we clearly could not mention all approaches that have been proposed and tested. Other methods, as for example Implicit Programming approaches which are described e.g. in [LPR96] and [KOZ98] have not been discussed here.

However, the selection we made corresponds to the main purpose of this section, which is to give some brief background information about the state-of-the-art of methods for MPECs that are somehow related to the relaxation scheme we introduce and analyse in the following chapters.

2 New Relaxation Scheme for MPECs

In the following chapter we will present the new relaxation scheme for MPECs. The scheme we propose can be regarded as a combination of ideas from regularization schemes, as for example the one proposed by Scholtes in [Sch01], with the relaxation-free approach proposed by Fletcher et al. [FLRS06].

First, we derive and illustrate the new relaxation, followed by a detailed discussion of its properties. In particular, we analyse the stationary points and solutions of the relaxed problems. Then we are concerned with the convergence behaviour of sequences (x^k) that are determined by solving a sequence of relaxed problems. Afterwards, we briefly consider possible extensions of the relaxation and finally we compare our new relaxation with the regularization scheme of Scholtes and the direct approach of Fletcher et al..

2.1 Relaxation

For $t > 0$, our relaxation for each pair $(x_{1j}, x_{2j}) \in \mathbb{R}^2$ is done only on a subset of the triangle with the vertices $(0, 0)$, $(t, 0)$, and $(0, t)$. Therefore, if the relaxation parameter is sufficiently small then around a local solution x^* of (1.19) our relaxed problem only modifies the complementarity constraints that correspond to *degenerate* components of the vector x^*. Hence, we merge the overall relaxation for all components of Scholtes [Sch01] with the exactness for the strictly complementary components of the approach of Fletcher et al. [FLRS06]. Moreover, if $t = 0$, then the parametrized nonlinear program corresponds to the original MPEC.

To derive the relaxation scheme that we analyse in this thesis, we first consider the scalar complementarity condition

$$x_1 \geq 0, \qquad x_2 \geq 0 \qquad \text{and} \qquad x_1 x_2 = 0, \tag{2.1}$$

with $x_1, x_2 \in \mathbb{R}$. Now, if we introduce new Cartesian coordinates

$$y := x_1 + x_2 \qquad \text{and} \qquad z := x_1 - x_2,$$

then the constraints of (2.1) simplify to $y = |z|$. Next, to smooth out the kink of the absolute value function on the interval $[-1, 1]$, we define a function $\theta : \mathcal{D} \subseteq \mathbb{R} \to \mathbb{R}$, where $\mathcal{D}$ is supposed to be an open subset of $\mathbb{R}$ with $[-1, 1] \subseteq \mathcal{D}$, which satisfies the following conditions

Assumptions 2.1.

1. *θ is twice continuously differentiable on* $[-1, 1]$*,*

2. $\theta(1) = \theta(-1) = 1$*,*

3. $\theta'(-1) = -1$ *and* $\theta'(1) = 1$,

4. $\theta''(-1) = \theta''(1) = 0$ *and*

5. θ *is strictly convex on* $(-1, 1)$.

Throughout this thesis we will assume that Assumptions 2.1 hold. Suitable examples for θ are

$$\theta_s(z) := \frac{2}{\pi} \sin\left(z\frac{\pi}{2} + \frac{3\pi}{2} \right) + 1 \tag{2.2}$$

or

$$\theta_p(z) := \frac{1}{8}(-z^4 + 6z^2 + 3) \tag{2.3}$$

Combining θ inside $[-1, 1]$ with the absolute value function outside $[-1, 1]$ and scaling the interval leads to a $\mathcal{C}^2$-function

$$\psi(z,t) := \begin{cases} |z| & |z| \geq t \\ t\,\theta(\frac{z}{t}) & |z| < t\,, \end{cases}$$

which gives us a relaxation of $y = |z|$ of the form

$$y \geq -z, \qquad y \geq z \qquad \text{and} \qquad y \leq \psi(z,t)\,. \tag{2.4}$$

Switching back to our original coordinate system we obtain

$$x_1 + x_2 \geq x_2 - x_1, \qquad x_1 + x_2 \geq x_1 - x_2 \qquad \text{and} \qquad x_1 + x_2 \leq \varphi(x_1, x_2, t)\,,$$

with $\varphi(x_1, x_2, t) := \psi(x_1 - x_2, t)$. Hence, relaxing each of the complementarity constraints of the MPEC as described above, we get a parametric nonlinear program $R(t)$ of the form:

$$\begin{array}{llrcl} R(t) & \min & f(x) & & \\ & \text{subject to} & h(x) & = & 0 \\ & & g(x) & \geq & 0 \\ & & x_1, x_2 & \geq & 0 \\ & & \Phi(x_1, x_2, t) & \leq & 0\,, \end{array} \tag{2.5}$$

where $\Phi(x_1, x_2, t) : \mathbb{R}^p \times \mathbb{R}^p \times \mathbb{R}_0^+ \to \mathbb{R}^p$ is defined by

$$\Phi(x_1, x_2, t) := x_1 + x_2 - \varphi(x_1, x_2, t)$$

and

$$\varphi_j(x_1, x_2, t) := \begin{cases} |x_{1j} - x_{2j}| & |x_{1j} - x_{2j}| \geq t \\ t\,\theta(\frac{x_{1j} - x_{2j}}{t}) & |x_{1j} - x_{2j}| < t\,. \end{cases}$$

Since θ is assumed to be strictly convex on $(-1, 1)$ the graph of $\theta(z/t)$ lies entirely above the two linearizations of θ in $z = -t$ and $z = t$, respectively. In other words $\theta(z/t) > |z/t|$ for all $z \in (-t, t)$. Furthermore, for any $z \in \mathbb{R}\backslash(-t, t)$ we have $\psi(z, t) = |z|$.

Now, consider a point $\hat{x}$ that is feasible for (1.19). Then $\hat{x}$ is also feasible for $R(t)$ for any $t \geq 0$, as for all complementarity pairs $(\hat{x}_{1j}, \hat{x}_{2j})$ with $|\hat{x}_{1j} - \hat{x}_{2j}| \geq t$ by the definition of Φ

$$\Phi_j(\hat{x}_1, \hat{x}_2, t) = \hat{x}_{1j} + \hat{x}_{2j} - |\hat{x}_{1j} - \hat{x}_{2j}| = 2\min(\hat{x}_{1j}, \hat{x}_{2j}) = 0. \tag{2.6}$$

and for all pairs $(\hat{x}_{1j}, \hat{x}_{2j})$ with $|\hat{x}_{1j} - \hat{x}_{2j}| < t$ we have

$$\begin{aligned} \Phi_j(\hat{x}_1, \hat{x}_2, t) &= \hat{x}_{1j} + \hat{x}_{2j} - t\theta\left(\tfrac{\hat{x}_{1j} - \hat{x}_{2j}}{t}\right) \\ &< \hat{x}_{1j} + \hat{x}_{2j} - t\left|\tfrac{\hat{x}_{1j} - \hat{x}_{2j}}{t}\right| \\ &= \hat{x}_{1j} + \hat{x}_{2j} - |\hat{x}_{1j} - \hat{x}_{2j}| = 2\min(\hat{x}_{1j}, \hat{x}_{2j}) = 0. \end{aligned} \tag{2.7}$$

Hence, by the feasibility of a point $\hat{x}$ for (1.19) we do not only know that $\hat{x}$ is feasible for any $R(t)$ with $t \geq 0$, but we can also directly deduce, which indices $j \in \{1, \dots, p\}$ are contained in the active set of $R(t)$ concerning the constraints $\Phi_j(x_1, x_2, t) \leq 0$ which we will denote by

$$I_\Phi(x, t) = \{\, j \in \{1, \dots, p\} \,|\, \Phi_j(x_1, x_2, t) = 0 \,\}.$$

The next Lemma summarizes the observations we just made.

Lemma 2.1. *Let $\hat{x}$ be feasible for (1.19), then*

1. *$\hat{x}$ is also feasible for $R(t)$ for every $t \geq 0$ and*
2. *$j \in I_\Phi(\hat{x}, t)$ if and only if $\max(\hat{x}_{1j}, \hat{x}_{2j}) \geq t$.*

Proof. The problem $R(t)$ differs from (1.19) only in the constraints concerning the complementarity of $\hat{x}_1$ and $\hat{x}_2$. By (2.6) and (2.7), however, $\hat{x}$ satisfies the constraint $\Phi_j(x_1, x_2, t) \leq 0$ for all $j \in \{1, \dots, p\}$ and for all $t \geq 0$. Moreover, if $\max(\hat{x}_{1j}, \hat{x}_{2j}) < t$, thus $|\hat{x}_{1j} - \hat{x}_{2j}| < t$, then (2.7) implies that $\Phi_j(x_1, x_2, t) \leq 0$ is inactive, whereas if $\max(\hat{x}_{1j}, \hat{x}_{2j}) \geq t$, thus $|\hat{x}_{1j} - \hat{x}_{2j}| \geq t$, then $j \in I_\Phi(\hat{x}, t)$ by (2.6). □

On the other hand, the strict convexity of θ implies that a point $\hat{x}$ that is feasible for $R(t)$ and satisfies $\Phi_j(\hat{x}_1, \hat{x}_2, t) = 0$ for a pair $(\hat{x}_{1j}, \hat{x}_{2j})$ with $|\hat{x}_{1j} - \hat{x}_{2j}| < t$ is infeasible for (1.19) as

$$\begin{aligned} 0 = \Phi_j(\hat{x}_1, \hat{x}_2, t) &= \hat{x}_{1j} + \hat{x}_{2j} - t\theta\left(\tfrac{\hat{x}_{1j} - \hat{x}_{2j}}{t}\right) \\ &< \hat{x}_{1j} + \hat{x}_{2j} - t\left|\tfrac{\hat{x}_{1j} - \hat{x}_{2j}}{t}\right| \\ &= 2\min(\hat{x}_{1j}, \hat{x}_{2j}). \end{aligned} \tag{2.8}$$

This demonstrates the relaxing property of $R(t)$, since for all x with $\Phi_j(x_1, x_2, t) = 0$ and $|x_{1j} - x_{2j}| < t$, by (2.8) we have $(x_{1j}, x_{2j}) \in \mathbb{R}^+ \times \mathbb{R}^+$.

Now consider a strictly positive decreasing sequence (t_k) and a corresponding sequence $(x^k) \subseteq \mathbb{R}^{n+2p}$, where each x^k is feasible for $R(t_k)$. Assume the sequence (x^k) converges to a point $\hat{x}$ that is feasible for (1.19). Then, by view of (2.8) , (2.6) and the corresponding remarks, it becomes clear, that the relaxation concerns increasingly only the pairs (x^k_{1j}, x^k_{2j}) that converge to the degenerate complementarity pairs of $\hat{x}$.

Moreover, if $t = 0$, then the feasible region of the MPEC and of $R(t)$ are equal. Figure 2.1 illustrates the new relaxation scheme and how it combines the ideas of the regularization scheme proposed by Scholtes in [Sch01] with the ideas of the approach to solve the MPEC by using exact NCP-functions (for example the *minimum function*) analysed in [Ley06] (confer Section 1.4).

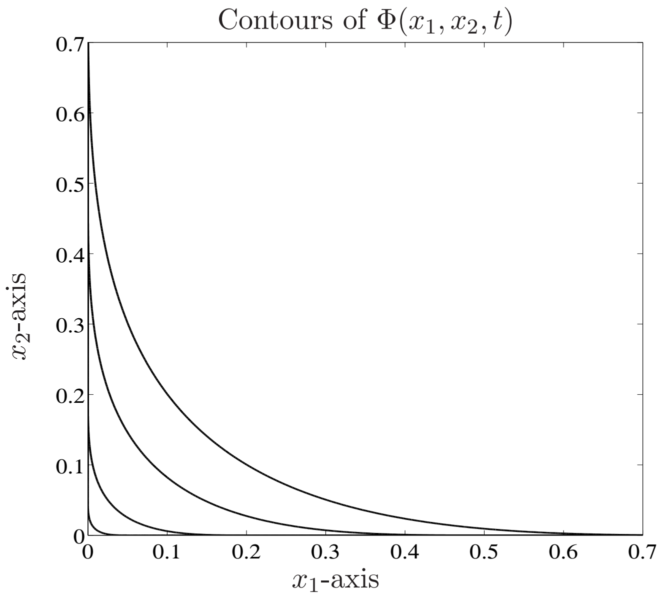


Figure 2.1: Contours of $\Phi_j(x_1, x_2, t) = 0$ with $\theta_s(z)$ as defined by (2.2) for different values of $t > 0$, namely $t \in \{0.05, 0.2, 0.5, 0.8\}$

Having discussed the relaxation property of $R(t)$ in relation to (1.19), we also need to relate the feasible regions of $R(t)$ for different values of t in order to be able to relate their solutions.

Lemma 2.2. *Let $\mathcal{Z}(t)$ denote the feasible region of $R(t)$, then for every pair of parameters $t_1, t_2 \in \mathbb{R}$, with $0 \le t_1 < t_2$:*

$$\mathcal{Z}(t_1) \subseteq \mathcal{Z}(t_2)\,. \tag{2.9}$$

Proof. Since for different parameters t_k the feasible regions of $R(t_k)$ differ only in the relaxing constraints $\Phi_j(x_1, x_2, t_k) \le 0$ for $j \in \{1, .., p\}$, we have to prove (2.9) only for these constraints. First, we show that for any $0 \le t_1 < t_2$

$$\Phi_j(x_1, x_2, t_2) \le 0 \tag{2.10}$$

is implied by $\Phi_j(x_1, x_2, t_1) \le 0$ for all $j \in \{1, .., p\}$.

Suppose $|x_{1j} - x_{2j}| \ge t_1$, then (2.10) is an immediate consequence of the definition of Φ_j. Therefore, we assume that $|x_{1j} - x_{2j}| < t_1$. Let $z := x_{1j} - x_{2j}$ then

$$\left|\frac{z}{t_1}\right|, \left|\frac{z}{t_2}\right| < 1\,.$$

It follows from the Assumptions 2.1 for θ in conjunction with Lemma 2.1 that

$$\theta\left(\frac{z}{t}\right) > \left|\frac{z}{t}\right| \qquad \text{and} \qquad \left|\theta'\left(\frac{z}{t}\right)\right| < 1\,,$$

such that

$$\theta\left(\frac{z}{t}\right) > \left|\frac{z}{t}\right| > \left|\frac{z}{t}\right|\left|\theta'\left(\frac{z}{t}\right)\right| \geq \frac{z}{t}\theta'\left(\frac{z}{t}\right)\,.$$

Hence, for pairs (x_{1j}, x_{2j}) with $|x_{1j} - x_{2j}| < t_1$, the function $\varphi_j(x_1, x_2, t)$ is strictly monotone increasing in $t \geq 0$, as we have

$$\frac{\partial \varphi_j}{\partial t}(x_1, x_2, t) = \theta\left(\frac{z}{t}\right) - \frac{z}{t}\theta'\left(\frac{z}{t}\right) > 0\,.$$

Now, the strict monotonicity of φ_j, the definition of Φ_j and the condition $\Phi_j(x_1, x_2, t_1) \leq 0$ imply (2.10). □

Consider a sequence (t_k) with $0 \leq t_{k+1} < t_k < \ldots < t_0$, then

$$\mathcal{Z}(0) \subseteq \mathcal{Z}(t_{k+1}) \subseteq \mathcal{Z}(t_k) \subseteq \ldots \subseteq \mathcal{Z}(t_0)\,.$$

This, however, implies that if we have found a local solution x^* of $R(\hat{t})$ with $\hat{t} > 0$ that is feasible for (1.19), then x^* is a local solution of $R(t)$ for all $t \in [0, \hat{t}]$.

Lemma 2.3. *Let Assumptions 2.1 hold and let x^* be a strict local minimum of $R(\hat{t})$ in an ε-neighbourhood $\mathcal{B}_\varepsilon(x^*)$ of x^*, that satisfies the complementarity constraints $0 \leq x_1 \perp x_2 \geq 0$. Then x^* is a strict local minimum of $R(t)$ for every $t \in [0, \hat{t}\,]$ in the same ε-neighbourhood $\mathcal{B}_\varepsilon(x^*)$.*

Proof. Suppose x^* is a strict local minimum of $R(\hat{t}\,)$ that satisfies the complementarity constraints, thus $x^* \in \mathcal{Z}(0)$. Then by Lemma 2.2 it follows that

$$\forall\, x \,\in\, (\mathcal{Z}(t) \cap \mathcal{B}_\varepsilon(x^*)) \subseteq \left(\mathcal{Z}(\hat{t}) \cap \mathcal{B}_\varepsilon(x^*)\right) \;:\; f(x) > f(x^*) \tag{2.11}$$

for every $t \in [0, \hat{t}\,)$. Hence, x^* is also a strict local minimum for $R(t)$ for every $t \in [0, \hat{t}\,)$ in the same ε-neighbourhood $\mathcal{B}_\varepsilon(x^*)$ of x^*. □

In Section 2.3 we will consider sequences $(x^k) \subseteq \mathbb{R}^{n+2p}$, where each x^k is feasible for $R(t_k)$ for positive, strictly decreasing parameter sequences (t_k). We are therefore further interested in the question whether a limit point $\bar{x}$ of such a sequence (x^k) is feasible for (1.19). The next Lemma addresses to this question.

Lemma 2.4. *Let $\mathcal{Z}(t)$ denote the feasible region of $R(t)$ and let $\mathcal{Z}$ be the feasible region of (1.19). Furthermore, let $(t_k) \subseteq \mathbb{R}^+$ be a sequence that satisfies $t_k \to 0$. Then*

$$\mathcal{Z} = \mathcal{Z}(0) = \bigcap_{k=1}^{\infty} \mathcal{Z}(t_k) \tag{2.12}$$

holds, such that any accumulation point $\bar{x}$ of a sequence (x^k) with $x^k \in \mathcal{Z}(t_k)$ satisfies $\bar{x} \in \mathcal{Z}$, in other words $\bar{x}$ is feasible for (1.19).

Proof. As $\Phi_j(x_1, x_2, 0) = 2\min(x_{1j}, x_{2j}) \le 0$ implies $0 \le x_1 \perp x_2 \ge 0$, the first equation of (2.12) is clearly satisfied. Furthermore, Lemma 2.2 implies $\mathcal{Z}(0) \subseteq \bigcap_{k=1}^{\infty} \mathcal{Z}(t_k)$.

To prove $\mathcal{Z}(0) \supseteq \bigcap_{k=1}^{\infty} \mathcal{Z}(t_k)$, we assume that there exists an $\bar{x} \in \bigcap_{k=1}^{\infty} \mathcal{Z}(t_k) \setminus \mathcal{Z}(0)$. Then $\bar{x}_{1j}\bar{x}_{2j} > 0$ for at least one index $j \in \{1, \ldots, p\}$, thus $\bar{x}_{1j} > 0$ and $\bar{x}_{2j} > 0$. Since

$$0 \ge \Phi_j(x_1, x_2, t_k) = x_{1j} + x_{2j} - |x_{1j} - x_{2j}| = 2\min(x_{1j}, x_{2j})$$

for all pairs (x_{1j}, x_{2j}) with $|x_{1j} - x_{2j}| \ge t_k$, by the feasibility of $\bar{x}$ for each $R(t_k)$ and as $\min(\bar{x}_{1j}, \bar{x}_{2j}) > 0$, we conclude that $|\bar{x}_{1j} - \bar{x}_{2j}| < t_k$ for all $k \in \mathbb{N}$. However, this implies

$$0 \le |\bar{x}_{1j} - \bar{x}_{2j}| \le \lim_{k\to\infty} t_k = 0\,,$$

and hence $\bar{x}_{1j} = \bar{x}_{2j}$. Define $\vartheta := \bar{x}_{1j} \ge 0$, then for all $k \in \mathbb{N}$

$$0 \ge \Phi_j(\bar{x}_1, \bar{x}_2, t_k) = 2\vartheta - \varphi_j(\vartheta, \vartheta, t_k) = 2\vartheta - t_k\,\theta(0)\,,$$

such that $2\vartheta \le t_k\theta(0)$ for all $k \in \mathbb{N}$. Therefore,

$$0 \le 2\vartheta \le \theta(0) \lim_{k\to\infty} t_k = 0\,,$$

which contradicts the assumption that $\bar{x}_{1j}\bar{x}_{2j} > 0$. Hence, $\mathcal{Z}(0) = \bigcap_{k=1}^{\infty} \mathcal{Z}(t_k)$. □

Let $\mathcal{L}_{R(t)}(x, \lambda, \mu, \nu_1, \nu_2, \xi)$ denote the corresponding Lagrangian function for $R(t)$, with

$$\begin{aligned} \mathcal{L}_{R(t)}(x, \lambda, \mu, \nu_1, \nu_2, \xi) &= f(x) - \textstyle\sum_{j=1}^{m} \lambda_j g_j(x) - \sum_{i=1}^{q} \mu_i h_i(x) \\ &\quad - \nu_1^T x_1 - \nu_2^T x_2 + \textstyle\sum_{j=1}^{p} \xi_j \Phi_j(x_1, x_2, t)\,. \end{aligned} \tag{2.13}$$

We can then state the first order stationarity conditions for $R(t)$ as follows (confer (1.16)): a point x is said to be a stationary point or KKT-point for $R(t)$ if it is feasible for $R(t)$ and there exist multipliers $\lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*$ such that the system

$$\begin{array}{rcll} \nabla_x \mathcal{L}_{R(t)}(x, \lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*) &=& 0 & \\ h(x) &=& 0 & \\ g(x) &\ge& 0 & \\ \lambda^* &\ge& 0 & \\ g_j(x)\lambda_j^* &=& 0 & j = 1, \ldots, m \\ x_1 \ \ge\ 0, \quad x_2 &\ge& 0 & \\ \nu_1^* \ \ge\ 0, \quad \nu_2^* &\ge& 0 & \\ x_{1j}\nu_{1j}^* \ =\ 0, \quad x_{2j}\nu_{2j}^* &=& 0 & j = 1, \ldots, p \\ \Phi(x_1, x_2, t) &\le& 0 & \\ \xi^* &\ge& 0 & \\ \xi_j^*\Phi_j(x_1, x_2, t) &=& 0 & j = 1, \ldots, p\,. \end{array} \tag{2.14}$$

is satisfied. In view of the definition of $\Phi_i(x_1, x_2, t)$, the matrix $\nabla_x \Phi(x_1, x_2, t)$ is very sparse. Its entries $\partial\Phi_i(x_1, x_2, t)/\partial x_{1j}$, $\partial\Phi_i(x_1, x_2, t)/\partial x_{2j}$ all vanish except for the two diagonals

$$\alpha_j := \frac{\partial \Phi_j(x_1, x_2, t)}{\partial x_{1j}} = \begin{cases} 0 & x_{1j} \ge x_{2j} + t \\ 2 & x_{1j} \le x_{2j} - t \\ 1 - \theta'\left(\frac{x_{1j} - x_{2j}}{t}\right) & \text{otherwise} \end{cases}$$

$$\beta_j := \frac{\partial \Phi_j(x_1, x_2, t)}{\partial x_{2j}} = \begin{cases} 0 & x_{2j} \geq x_{1j} + t \\ 2 & x_{2j} \leq x_{1j} - t \\ 1 + \theta'(\frac{x_{1j} - x_{2j}}{t}) & \text{otherwise}. \end{cases}$$

for all $t > 0$. It follows that

$$\nabla_x \Phi(x_1, x_2, t) = \begin{pmatrix} 0 \\ D_1 \\ D_2 \end{pmatrix}$$

with $D_1 = \text{diag}(\alpha_j) \in \mathbb{R}^{p \times p}$ and $D_2 = \text{diag}(\beta_j) \in \mathbb{R}^{p \times p}$.

Remark 2.1. Note that due to the conditions on $\theta(z)$ for all feasible points $x \in \mathbb{R}^{n+2p}$, we obtain $0 \leq \alpha_j \leq 2$ and $0 \leq \beta_j \leq 2$. Moreover, $\beta_j = 2 - \alpha_j$ for all $j \in \{1, .., p\}$ and the values of α_j are strictly monotonically decreasing for $z := (x_{1j} - x_{2j})/t$ in $(-1, 1)$.

As we will make use of this conclusion later on in Section 2.3, we state it more precisely in the following lemma.

Lemma 2.5. *Suppose $\theta(z)$ satisfies Assumptions 2.1 and let $\alpha(z) := 1 - \theta'(z)$ and $\beta(z) := 1 + \theta'(z)$. Then $\alpha(z)$ is strictly monotonically decreasing and $\beta(z)$ is strictly monotonically increasing for $z \in (-1, 1)$.*

Proof. If $\theta(z)$ satisfies the Assumptions 2.1, then for all $z \in (-1, 1)$

$$\frac{\partial \alpha}{\partial z} = -\theta''(z) < 0 \qquad \text{and} \qquad \frac{\partial \beta}{\partial z} = \theta''(z) > 0.$$

□

The following lemma relates the active index sets $I_1(x)$ and $I_2(x)$ to $I_\Phi(x, t)$. These relations will later be used in the analysis of the stationary points and the local minima of the MPEC in comparison to those of $R(t)$.

Lemma 2.6. *Let $t > 0$ and let x be a feasible point of $R(t)$, then it follows that*

1. *if $j \in I_1(x)$ and $x_{2j} \geq t$ or $j \in I_2(x)$ and $x_{1j} \geq t$, then $j \in I_\Phi(x, t)$ and the corresponding gradients of the active constraints are positive linear dependent.*
2. *If $x_{1j} < t$ and $x_{2j} < t$, then at most either $j \in I_\Phi(x, t)$ or $j \in (I_1 \cup I_2)(x)$ but not both.*

Proof. The first part of the lemma is due to the definition of $\Phi_j(x_1, x_2, t)$, the special structure of $\nabla_x \Phi_j(x_1, x_2, t)$ and the values of α_j and β_j, respectively. The second part follows by Lemma 2.1. □

Combining the two statements of Lemma 2.6 leads to the fact that either the constraints $x_{1j} \geq 0$ or $x_{2j} \geq 0$ and $\Phi_j(x_1, x_2, t) \leq 0$ are both active and the corresponding gradients are positive linear dependent or at most one of them appears in the KKT-system with a possibly non-vanishing multiplier.

2.2 Properties

In this section we will first relate the stationary points of the MPEC (1.19) to the stationary points of $R(t)$. This will be done by comparing the corresponding KKT-systems in order to derive relations between suitable Lagrange multipliers of both problems. We then correlate the local minima of (1.19) and of $R(t)$ using the informations about the stationary points and comparing the second order sufficient conditions of both problems.

A crucial assumption that we will need to relate a strongly stationary point x^* of (1.19) to a stationary point of $R(t)$, concerns the admissible size of the parameter t. In order to prove that x^* is stationary for $R(t)$, we have to assume that $t \geq 0$ is small enough. The admissible size of t thereby depends on the nonvanishing components of the complementarity pairs of x^*. To simplify the notation of the corresponding results and their proofs, we therefore introduce the following notation.

Definition 2.1. Let x^* be a strongly stationary point of (1.19), then we define

$$\tau(x^*) := \min\{x^*_{ij} \mid i \in \{1,2\} \ , \ j \in \{1,\dots,p\} \quad \text{and} \quad x^*_{ij} > 0 \, \} \, .$$

If $x^*_{ij} = 0$ for all $j \in \{1,\dots,p\}$ and $i \in \{1,2\}$, then we set $\tau(x^*) := +\infty$.

The relation between strongly stationary points of (1.19) and stationary points of $R(t)$ can now be stated as follows.

Lemma 2.7. *Suppose that x^* is a strongly stationary point of (1.19) with multipliers λ^*,μ^*,$\hat{\nu}_1$,$\hat{\nu}_2$. Then for every $t \in (0, \tau(x^*)]$ there exist multipliers ν_1^*, ν_2^*, ξ^* such that the vector (x^*, λ^*, μ^*, ν_1^*, ν_2^*, ξ^*) satisfies the stationarity conditions (2.14) of $R(t)$.*

Proof. By Lemma 2.1, the feasibility of x^* for $R(t)$ follows directly from the feasibility of x^* for (1.19). Now, let $t \in (0, \tau(x^*)]$, then $x^*_{1j} \geq t$ for all $j \notin I_1(x^*)$ and $x^*_{2j} \geq t$ for all $j \notin I_2(x^*)$. Hence,

$$\begin{array}{llll} \alpha_j = 0 & \text{and} \quad \beta_j = 2 & \text{for} & j \notin I_1(x^*) \\ \alpha_j = 2 & \text{and} \quad \beta_j = 0 & \text{for} & j \notin I_2(x^*) \, . \end{array} \tag{2.15}$$

As x^* is feasible for (1.19), we have that $I_1(x^*)^\perp \cap I_2(x^*)^\perp = \emptyset$ (with $I_i(x^*)^\perp$ denoting the complement of $I_i(x^*)$ in $\{1,..,p\}$ $(i = 1,2)$). Considering the special structure of $\nabla_x \Phi(x_1, x_2, t)$, the values of α_j and β_j of (2.15) and comparing (2.14) with the conditions for strong stationarity yields

$$\begin{array}{llll} \hat{\nu}_{1j} = \nu^*_{1j} - 2\xi^*_j & \text{and} \quad 0 = \hat{\nu}_{2j} = \nu^*_{2j} & \text{for} & j \notin I_2(x^*) \\ \hat{\nu}_{2j} = \nu^*_{2j} - 2\xi^*_j & \text{and} \quad 0 = \hat{\nu}_{1j} = \nu^*_{1j} & \text{for} & j \notin I_1(x^*) \, . \end{array} \tag{2.16}$$

If $\hat{\nu}_{ij} < 0$ $(i = 1,2)$, then we choose $\xi^*_j > 0$ such that $\nu^*_{ij} = \hat{\nu}_{ij} + 2\xi^*_j \geq 0$.

As a consequence of Lemma 2.1, ξ^*_j has to vanish for all $j \in (I_1 \cap I_2)(x^*)$ if $t > 0$. Hence, we have

$$0 \leq \hat{\nu}_{1j} = \nu^*_{1j} \quad \text{and} \quad 0 \leq \hat{\nu}_{2j} = \nu^*_{2j} \qquad \text{for all} \quad j \in (I_1 \cap I_2)(x^*) \, .$$

□

Remark 2.2. Note that examining the proof of Lemma 2.7 we get the answer to the question why t has to be smaller then the parameter $\tau(x^*)$. If $t > \tau(x^*)$, then $\Phi_j(x_1^*, x_2^*, t) \leq 0$ might not be active for all $j \in \{1, \dots, p\} \backslash (I_1 \cap I_2)(x^*)$ and we might not be able to increase ξ_j^* in order to guarantee $\nu_{1j}^* \geq 0$ or $\nu_{2j}^* \geq 0$, respectively.

Lemma 2.7 does not give any information about the uniqueness of the multipliers of the stationary point of $R(t)$. In fact, due to the positive linear dependence of active constraint gradients, as mentioned in Lemma 2.6, it cannot be guaranteed that the chosen multipliers are unique, even if the multipliers for the strongly stationary point x^* are unique. However, assuming the MPEC-LICQ holds in x^*, we can prove the uniqueness for a special choice of multipliers, which we call *basic multipliers* and define as follows (similar to the basic multipliers defined in [Ley06]).

Definition 2.2. Let x^* be a strongly stationary point of (1.19) with multipliers λ^*, μ^*, $\hat{\nu}_1$, and $\hat{\nu}_2$ then we call the multipliers defined by

$$\begin{aligned}
\lambda^b &:= \lambda^* \\
\mu^b &:= \mu^* \\
\nu_{1j}^b &:= (\hat{\nu}_{1j})^+ \qquad j = 1, \dots, p \\
\nu_{2j}^b &:= (\hat{\nu}_{2j})^+ \qquad j = 1, \dots, p \\
\xi_j^b &:= \begin{cases} \left(-\frac{\hat{\nu}_{1j}}{2}\right)^+ & j \in (I_1 \backslash I_2)(x^*) \\ \left(-\frac{\hat{\nu}_{2j}}{2}\right)^+ & j \in (I_2 \backslash I_1)(x^*) \\ 0 & j \in (I_1 \cap I_2)(x^*) \end{cases}
\end{aligned}$$

the *basic multipliers* of x^* for $R(t)$.

Suppose x^* is a strongly stationary point of (1.19) and let $t \in (0, \tau(x^*)]$. Furthermore, let $\Lambda(x^*, t)$ denote the set of all feasible multipliers of x^* for $R(t)$ (thus multipliers that satisfy (2.14)), which is not empty by Lemma 2.7. It turns out, that assuming the MPEC-LICQ holds in x^*, the basic multipliers are not only feasible and unique, but also solve the minimization problem

$$\begin{aligned}
\min \quad & \|(\lambda, \mu, \nu_1, \nu_2, \xi)\|_1 \\
\text{subject to} \quad & (\lambda, \mu, \nu_1, \nu_2, \xi) \in \Lambda(x^*)\,.
\end{aligned}$$

Lemma 2.8. *Let x^* be a strongly stationary point of (1.19), with multipliers λ^*, μ^*, $\hat{\nu}_1$ and $\hat{\nu}_2$ that satisfies the MPEC-LICQ and let $t \in (0, \tau(x^*)]$. Then,*

1. *the basic multipliers (λ^b,μ^b, ν_1^b, ν_2^b, ξ^b) are unique and feasible for $R(t)$, thus*

$$(\lambda^b, \mu^b, \nu_1^b, \nu_2^b, \xi^b) \in \Lambda(x^*, t)$$

and

2. *they form the unique solution of*

$$\begin{aligned}
\textit{min} \quad & \|(\lambda, \mu, \nu_1, \nu_2, \xi)\|_1 \\
\textit{subject to} \quad & (\lambda, \mu, \nu_1, \nu_2, \xi) \in \Lambda(x^*, t)\,.
\end{aligned} \tag{2.17}$$

Proof. First, the MPEC-LICQ in x^* implies the uniqueness of the multipliers λ^*, μ^*, $\hat{\nu}_1$ and $\hat{\nu}_2$. The uniqueness of the multipliers $(\lambda^b, \mu^b, \nu_1^b, \nu_2^b, \xi^b)$ hence directly follows by Definition 2.2.

Next, note that $(\lambda^b, \mu^b, \nu_1^b, \nu_2^b, \xi^b)$ satisfies all the conditions on $(\lambda, \mu, \nu_1, \nu_2, \xi)$ made in the proof of Lemma 2.7. Hence, the basic multipliers are feasible for $R(t)$.

Since λ^*, μ^*, $\hat{\nu}_1$ and $\hat{\nu}_2$ are unique by the MPEC-LICQ and the corresponding feasible multipliers $(\lambda, \mu, \nu_1, \nu_2, \xi)$ of $R(t)$ therefore must satisfy $\lambda = \lambda^*$ and $\mu = \mu^*$, we can reduce (2.17) to

$$\begin{aligned} \min \quad & \|(\nu_1, \nu_2, \xi)\|_1 \\ \text{subject to} \quad & (\lambda, \mu, \nu_1, \nu_2, \xi) \in \Lambda(x^*, t)\,. \end{aligned} \tag{2.18}$$

Now consider (2.14) and notice that the feasible multipliers of x^* for $R(t)$ with $t \in (0, \tau(x^*)]$ must additionally satisfy the conditions

$$\begin{array}{llll} \nu_{1j} = \hat{\nu}_{1j} + 2\xi_j & & \text{for} & j \in (I_1 \backslash I_2)(x^*) \\ \nu_{2j} = \hat{\nu}_{2j} + 2\xi_j & & \text{for} & j \in (I_2 \backslash I_1)(x^*) \\ \nu_{1j} = \hat{\nu}_{1j}, \quad \nu_{2j} = \hat{\nu}_{2j} & \text{and} \quad \xi_j = 0 & \text{for} & j \in (I_1 \cap I_2)(x^*) \\ \nu_{1j} \geq 0, \quad \nu_{2j} \geq 0 & \text{and} \quad \xi_j \geq 0 & \text{for} & j \in \{1, \ldots, p\}\,. \end{array}$$

Substituting ν_{ij} with $j \in \{1, \ldots, p\}$ and $i = 1, 2$, in the objective function of (2.18) and into the inequality conditions, we obtain the problem

$$\begin{array}{lll} \min & \sum_{j=1}^p (\hat{\nu}_{1j} + \hat{\nu}_{2j}) + 3\sum_{j=1}^p \xi_j & \\ \text{subject to} & \xi_j \geq \left(-\frac{\hat{\nu}_{1j}}{2}\right)^+ & j \in (I_1 \backslash I_2)(x^*) \\ & \xi_j \geq \left(-\frac{\hat{\nu}_{2j}}{2}\right)^+ & j \in (I_2 \backslash I_1)(x^*) \\ & \xi_j = 0 & j \in (I_1 \cap I_2)(x^*)\,, \end{array}$$

which is clearly solved by ξ^b. Finally, the corresponding values of ν_1 and ν_2 that are uniquely derived by the values of $\hat{\nu}_1$, $\hat{\nu}_2$ and ξ^b coincide with the values of the definition of ν_1^b and ν_2^b. □

Next, we consider the inverse direction of Lemma 2.7. We obviously need to require the feasibility of x^* for (1.19), since it cannot be guaranteed by the feasibility of x^* for $R(t)$. However, provided x^* is feasible for (1.19), the stationarity of x^* for $R(t)$ implies that x^* is strongly stationary.

Lemma 2.9. *Suppose $t > 0$ and $(x^*(t), \lambda^*(t), \mu^*(t), \nu_1^*(t), \nu_2^*(t), \xi^*(t))$ satisfies (2.14). Moreover, let $x^*(t)$ be feasible for (1.19). Then, $x^*(t)$ is strongly stationary with multipliers $\lambda^* = \lambda^*(t), \mu^* = \mu^*(t)$ and*

$$\begin{aligned} \hat{\nu}_{1j} &= \nu_{1j}^*(t) - \alpha_j\, \xi_j^*(t) \\ \hat{\nu}_{2j} &= \nu_{2j}^*(t) - (2 - \alpha_j)\, \xi_j^*(t)\,. \end{aligned}$$

Proof. First suppose $t > 0$. Considering the feasibility assumption on $x^*(t)$, the choice of the multipliers and the values of α_j and $\xi_j^*(t)$, respectively, the conditions of (1.23) are a direct consequence of the conditions of (2.14). Furthermore, the nonnegativity of $\hat{\nu}_{ij}$ ($i = 1,2$) for $j \in (I_1 \cap I_2)(x^*(t))$ is implied by $\xi_j^*(t) = 0$ (since by Lemma 2.1 $\Phi_j(x_1^*(t), x_2^*(t), t) < 0$) and the nonnegativity of $\nu_{ij}^*(t)$ ($i = 1,2$). □

Having related the first order conditions of (1.19) and $R(t)$, it remains to relate the second order sufficient conditions in order to describe the relation between the strict local minima of both problems. We therefore compare the sets of critical directions of both problems and the corresponding Hessians of the Lagrangian functions with respect to x. It turns out, that the sets of critical directions and the corresponding Hessians of the Lagrangian functions are identical. Hence, the second oder sufficient conditions are identical and can be replaced by each other. Let

$$\begin{aligned}
\mathcal{S}_{R(t)}(x^*, \lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*, t) = \{\, d \in \mathbb{R}^{n+2p} \backslash \{0\} \quad | & \\
\nabla h_i(x^*)^T d = 0, &\quad i \in \{1, \ldots, q\} \\
\nabla g_j(x^*)^T d = 0, &\quad j \in I_g^+(x^*, \lambda^*) \\
\nabla g_j(x^*)^T d \geq 0, &\quad j \in I_g^0(x^*, \lambda^*) \\
\nabla \Phi_j(x_1^*, x_2^*, t)^T d = 0, &\quad j \in I_\Phi^+(x^*, t, \xi^*) \\
\nabla \Phi_j(x_1^*, x_2^*, t)^T d \leq 0, &\quad j \in I_\Phi^0(x^*, t, \xi^*) \\
d_{1j} = 0, &\quad j \in I_1^+(x^*, \nu_1^*) \\
d_{1j} \geq 0, &\quad j \in I_1^0(x^*, \nu_1^*) \\
d_{2j} = 0, &\quad j \in I_2^+(x^*, \nu_2^*) \\
d_{2j} \geq 0, &\quad j \in I_2^0(x^*, \nu_2^*) \} \,.
\end{aligned}$$

denote the set of critical directions of $R(t)$ in x^*. Suppose x^* is a strongly stationary point of (1.19) or a stationary point of $R(t)$ with $t \in (0, \tau(x^*)]$, and we choose the multipliers as described in Lemma 2.7 or Lemma 2.9, respectively. Then we obtain the identity

$$\bar{\mathcal{S}}(x^*, \lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2) = \mathcal{S}_{R(t)}(x^*, \lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*, t)$$

as well as $\nabla_{xx} \mathcal{L}_{R(t)}(x^*, \lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*) = \nabla_{xx} \mathcal{L}_{MPEC}(x^*, \lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2)$. Hence, in this case the RNLP-SOSC for (1.19) and the SOSC for $R(t)$ are identical conditions.

Theorem 2.1. *Let x^* be feasible for (1.19). Then*

1. *if x^* is a strongly stationary point of the MPEC (1.19) that satisfies the RNLP-SOSC, then x^* is a stationary point of $R(t)$ for every $t \in (0, \tau(x^*)]$ that satisfies the SOSC for $R(t)$. Thus, x^* is a strict local minimum of $R(t)$.*
2. *On the other hand, if x^* is a stationary point for $R(t)$ that satisfies the SOSC, then x^* is a strongly stationary point of the MPEC (1.19) that satisfies the RNLP-SOSC. Therefore, it is also a strict local minimum of (1.19) .*

Proof. First the stationarity of x^* for $R(t)$ or (1.19), respectively, results from Lemma 2.7 or from Lemma 2.9, respectively.

Hence, it remains to show, that the second order conditions in x^* imply each other. We therefore compare the second derivatives with respect to x^* of the Lagrangian functions $\mathcal{L}_{MPEC}(x^*,\lambda^*, \mu^*,\hat{\nu}_1,\hat{\nu}_2)$ and $\mathcal{L}_{R(t)}(x^*, \lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*)$:

$$\nabla^2_{xx}\mathcal{L}_{MPEC}(x^*,\lambda^*,\mu^*,\hat{\nu}_1,\hat{\nu}_2) = \nabla^2_{xx}f(x^*) - \sum_{j\in I_g}\nabla^2_{xx}g_j(x^*)\lambda_j^* - \sum_{i=1}^{q}\nabla^2_{xx}h_i(x^*)\mu_i^*$$

$$\begin{aligned}\nabla^2_{xx}\mathcal{L}_{R(t)}(x^*,\lambda^*,\mu^*,\nu_1^*,\nu_2^*,\xi^*) = & \nabla^2_{xx}f(x^*) - \sum_{j\in I_g}\nabla^2_{xx}g_j(x^*)\lambda_j^* - \sum_{i=1}^{q}\nabla^2_{xx}h_i(x^*)\mu_i^* \\ & + \sum_{j=1}^{p}\begin{pmatrix} 0 & 0 \\ 0 & \nabla^2_{yy}\Phi_j(x_1^*,x_2^*,t)\end{pmatrix}\xi_j^*\,,\end{aligned}$$

where $y := (x_1, x_2)$. We observe that they only differ in the additional term

$$M(x^*,t) = \sum_{j=1}^{p}\begin{pmatrix} 0 & 0 \\ 0 & \nabla^2_{yy}\Phi_j(x_1^*,x_2^*,t)\end{pmatrix}\xi_j^*\,.$$

Now, consider Lemma 2.1 and the special structure of $\nabla_x\Phi(x_1,x_2,t)$ and notice that either $\xi_j^* = 0$ or $\nabla^2_{yy}\Phi_j(x_1^*,x_2^*,t) = \mathbf{0}$, such that $M(x^*,t)$ vanishes completely. Hence,

$$\nabla^2_{xx}\mathcal{L}_{MPEC}(x^*,\lambda^*,\mu^*,\hat{\nu}_1,\hat{\nu}_2) = \nabla^2_{xx}\mathcal{L}_{R(t)}(x^*,\lambda^*,\mu^*,\nu_1^*,\nu_2^*,\xi^*)\,.$$

Next, we prove that $\bar{\mathcal{S}}(x^*, \lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2) = \mathcal{S}_{R(t)}(x^*, \lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*, t)$. First, the conditions on the directions $d \in \mathbb{R}^{n+2p}\backslash\{0\}$ corresponding to the constraints $g_j(x) \geq 0$ with $j \in \{1,\dots,m\}$ and $h_i(x) = 0$ with $i \in \{1,\dots,q\}$ are completely identical. Moreover, the remaining conditions can be transformed into each other as follows:

First assume $d \in \mathcal{S}_{R(t)}(x^*,\lambda^*,\mu^*,\nu_1^*,\nu_2^*,\xi^*,t)$. If $j \in (I_1\backslash I_2)(x^*)$, then provided $t < \tau(x^*)$,

$$\alpha_j\, d_{1j} + (2-\alpha_j)\, d_{2j} = 2\, d_{1j} \leq 0 \qquad \text{and} \qquad d_{1j} \geq 0$$

which implies $d_{1j} = 0$. The same argument implies that $d_{2j} = 0$ for $j \in (I_2\backslash I_1)(x^*)$. If $j \in (I_1 \cap I_2)(x^*)$, then the conditions on the corresponding multipliers $\xi_j^* = 0$ and $\hat{\nu}_{1j} > 0$ lead to $\nu_{1j}^* > 0$, such that $d_{1j} = 0$. Likewise, it follows that $d_{2j} = 0$ for indices $j \in (I_1 \cap I_2)(x^*)$ with $\hat{\nu}_{2j} > 0$. Furthermore, if $j \in (I_1 \cap I_2)(x^*)$ then the nonnegativity of d_{1j} and d_{2j} follows immediately. Hence, $d \in \bar{\mathcal{S}}(x^*, \lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2)$ and we obtain

$$\mathcal{S}_{R(t)}(x^*,\lambda^*,\mu^*,\nu_1^*,\nu_2^*,\xi^*,t) \;\subseteq\; \bar{\mathcal{S}}(x^*,\lambda^*,\mu^*,\hat{\nu}_1,\hat{\nu}_2).$$

Now let $d \in \bar{\mathcal{S}}(x^*, \lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2)$, then $d_{1j} = 0$ for $j \in (I_1\backslash I_2)(x^*)$ or $d_{2j} = 0$ for $j \in (I_2\backslash I_1)(x^*)$, respectively . Since $I_\Phi(x^*,t) \cap (I_1 \cap I_2)(x^*) = \emptyset$ and

$$\begin{aligned} j \in (I_1\backslash I_2)(x^*) \quad &\Rightarrow \quad \nabla_x\Phi_j(x_1^*,x_2^*,t)^T d = 2\, d_{1j} = 0 \\ j \in (I_2\backslash I_1)(x^*) \quad &\Rightarrow \quad \nabla_x\Phi_j(x_1^*,x_2^*,t)^T d = 2\, d_{2j} = 0\,,\end{aligned}$$

the conditions concerning to the constraints $\Phi_j(x_1, x_2, t)$ are satisfied. Moreover,

$$j \in \{I_1(x^*)|\nu_{1j}^* > 0\} \quad \Rightarrow \quad j \in (I_1\backslash I_2)(x^*) \quad \text{or} \quad j \in (I_1 \cap I_2)(x^*)\,.$$

If $j \in (I_1 \backslash I_2)(x^*)$, then $d_{1j} = 0$ follows directly by the definition of $\bar{\mathcal{S}}(x^*, \lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2)$. If on the other hand $j \in (I_1 \cap I_2)(x^*)$, then $\nu^*_{1j} = \hat{\nu}_{1j}$, such that $\nu^*_{1j} > 0$ implies $\hat{\nu}_{1j} > 0$ and hence $d_{1j} = 0$. Accordingly, $d_{2j} = 0$ for $j \in \{I_2(x^*) | \nu^*_{2j} > 0\}$. Finally, with $d_{1j} \geq 0$ for $j \in \{I_1(x^*) | \nu^*_{1j} = 0\}$ and $d_{2j} \geq 0$ for $j \in \{I_2(x^*) | \nu^*_{2j} = 0\}$, respectively, we obtain

$$\bar{\mathcal{S}}(x^*, \lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2) = \mathcal{S}_{R(t)}(x^*, \lambda^*, \mu^*, \nu^*_1, \nu^*_2, \xi^*, t)$$

for the related multipliers, introduced in Lemma 2.7 and Lemma 2.9. □

Having answered the question whether a local minimum of (1.19) is also a local minimum of $R(t)$, we can extend Theorem 2.1 by combining it with Lemma 2.3.

Corollary 2.1. *Let x^* be a strongly stationary point of (1.19) that satisfies the RNLP-SOSC. Then there exists an $\varepsilon > 0$, such that for every $t \in [0, \tau(x^*)]$, x^* is a strict local minimum in $\mathcal{B}_\varepsilon(x^*)$ of $R(t)$ which satisfies the SOSC.*

Proof. In view of Theorem 2.1, it follows that x^* is a strict local minimum of $R(t)$ for every $t \in (0, \tau(x^*)]$. Hence, if we choose $\hat{t} = \tau(x^*)$, then there exists an $\varepsilon > 0$, such that x^* is a strict local minimum in $\mathcal{B}_\varepsilon(x^*)$ of $R(\hat{t})$. Next, if we apply Lemma 2.3, then x^* is a strict local minimum of $R(t)$ in $\mathcal{B}_\varepsilon(x^*)$ for every $t \in [0, \tau(x^*)]$. □

2.3 Convergence

In the preceding section, we related the stationary points and local solutions of (1.19) and of $R(t)$. In particular, we proved that a local solution of (1.19) that satisfies the strong stationarity conditions and the RNLP-SOSC, is a local minimizer of $R(t)$, provided $t \geq 0$ is small enough.

However, as our solution approach is based on solving a sequence of problems $R(t_k)$, where (t_k) is a positive, decreasing sequence, we are in addition interested in the convergence behaviour of thus determined sequences of solutions (x^k).

At first we consider a sequence of global solutions of problems $R(t_k)$. The starting point of the following theorem is therefore a positive, strictly decreasing sequence of parameters t_k and a convergent sequence of global solutions x^k of the corresponding problems $R(t_k)$.

Theorem 2.2. *Let $(t_k) \subseteq \mathbb{R}^+$ be a sequence that satisfies $t_k \to 0$ and let (x^k) be a sequence of global solutions of the corresponding problems $R(t_k)$. Furthermore, let $\bar{x}$ be an accumulation point of (x^k). Then, $\bar{x}$ is a global solution of (1.19).*

Proof. As $\bar{x}$ is an accumulation point of (x^k), there exist a subsequence $(x^k)_{k\in\mathcal{K}}$, such that

$$\lim_{\substack{k\to\infty \\ k\in\mathcal{K}}} x^k = \bar{x}\,.$$

It follows by Lemma 2.4, that $\bar{x}$ is feasible for (1.19). Now, suppose $\bar{x}$ is not a global minimum of (1.19), then there exists an $\hat{x} \in \mathcal{Z}$ with $f(\bar{x}) > f(\hat{x})$. This, however, implies, by the continuity of f and the convergence of the subsequence $(x^k)_{k\in\mathcal{K}}$, that for all $k \in \mathcal{K}$ sufficiently large $f(x^k) > f(\hat{x})$ which contradicts the assumption that x^k is a global minimum of $R(t_k)$. □

Although interesting from a theoretical viewpoint, the convergence of a sequence of global solutions is not very serviceable for practical purposes, as NLP solvers are often not designed to find the global solutions, but rather the stationary points of an NLP. In the remaining part of this section, we will therefore focus on sequences of stationary points.

The starting point of the next theorem is again a positive sequence of parameters $t_k \to 0$. Though now, we consider a convergent sequence of stationary points x^k of the problems $R(t_k)$. If the limit point $\bar{x}$ satisfies the MPEC-LICQ, then we can prove that it is a C-stationary point of (1.19). Moreover, if we tie together the multipliers of the two positive linear dependent gradients, the sequence of multiplier vectors of the stationary points x^k converges to the unique multiplier vector of $\bar{x}$. If in addition each x^k satisfies the SOSC for $R(t_k)$, then we can proof that $\bar{x}$ is even M-stationary.

Theorem 2.3. *Let $(t_k) \subseteq \mathbb{R}^+$ be a sequence that satisfies $t_k \to 0$. Furthermore, let $(x^k) \subseteq \mathbb{R}^{n+2p}$ be a sequence of stationary points of $R(t_k)$ that converges to a limit point $\bar{x}$ and suppose the MPEC-LICQ holds in $\bar{x}$.*

1. *Then $\bar{x}$ is a C-stationary point of the MPEC with unique multipliers $\bar{\lambda}, \bar{\mu}, \bar{\nu}_1$ and $\bar{\nu}_2$ that satisfy*

$$\begin{aligned}
\bar{\lambda}_j &= \lim_{k\to\infty} \lambda_j^k \geq 0 && j \in I_g(\bar{x}) \\
\bar{\mu}_i &= \lim_{k\to\infty} \mu_i^k && i \in \{1,\ldots,q\} \\
\bar{\nu}_{1m} &= \lim_{k\to\infty} (\nu_{1m}^k - \xi_m^k \alpha_m^k) && m \in I_1(\bar{x}) \\
\bar{\nu}_{2m} &= \lim_{k\to\infty} (\nu_{2m}^k - \xi_m^k (2-\alpha_m^k)) && m \in I_2(\bar{x})
\end{aligned}$$

   and

$$\bar{\lambda}_j = 0 \quad j \notin I_g(\bar{x}), \qquad \bar{\nu}_{1m} = 0 \quad m \notin I_1(\bar{x}), \qquad \bar{\nu}_{2m} = 0 \quad m \notin I_2(\bar{x})\,.$$

2. *If in addition the SOSC holds in each stationary point x^k of $R(t^k)$, then $\bar{x}$ is M-stationary.*

3. *Finally, $\bar{x}$ is strongly stationary, if and only if $\bar{\nu}_{1m} \geq 0$ and $\bar{\nu}_{1m} \geq 0$ for all indices $m \in (I_1 \cap I_2)(\bar{x}) \cap I_\Phi^{inf}$ with*

$$I_\Phi^{inf} := \{j \in \{1,\ldots,p\} \,|\, j \in I_\Phi(x^k, t_k) \quad \textit{for infinitely many} \quad k \in \mathbb{N}\}.$$

Proof. Since $g(x)$ is continuous $I_g(x^k) \subseteq I_g(\bar{x})$ for sufficiently large $k \in \mathbb{N}$ and as each x^k is a stationary point of $R(t_k)$, it satisfies the KKT conditions (2.14). Hence,

$$\begin{aligned}
\nabla f(x^k) \;=\; & \sum_{j\in I_g(\bar{x})} \lambda_j^k \nabla g_j(x^k) \;+\; \sum_{i=1}^{q} \mu_i^k \nabla h_i(x^k) \\
& + \sum_{m=1}^{p} (\nu_{1m}^k - \xi_m^k \alpha_m^k)\, e_{1m} \\
& + \sum_{m=1}^{p} (\nu_{2m}^k - \xi_m^k (2-\alpha_m^k))\, e_{2m}
\end{aligned} \tag{2.19}$$

As $x^k \to \bar{x}$ and $t_k \to 0$, then for sufficiently large $k \in \mathbb{N}$ it holds $x^k_{1m} \geq t_k$ for all $m \notin I_1(\bar{x})$. Hence, by the feasibility of x^k and the first part of Lemma 2.6, we have $m \in I_2(x^k) \cap I_\Phi(x^k, t_k)$ for all $m \notin I_1(\bar{x})$. Accordingly, we get $m \in I_1(x^k) \cap I_\Phi(x^k, t_k)$ for all $m \notin I_2(\bar{x})$. Hence, by the second part of Lemma 2.6 and the values of α_j for sufficiently large $k \in \mathbb{N}$,

$$\begin{array}{lclcl} \nu^k_{1m} = 0 & \text{and} & \alpha^k_m = 0 & \text{for all} & m \notin I_1(\bar{x}) \\ \nu^k_{2m} = 0 & \text{and} & (2 - \alpha^k_m) = 0 & \text{for all} & m \notin I_2(\bar{x}) \,, \end{array}$$

such that (2.19) can be rewritten as $\mathbf{A}_k^T \pi_k = \nabla f(x^k)$, where $\pi_k = ((\lambda^k)_{I_g(\bar{x})}, \mu^k, \gamma^k_1, \gamma^k_2)$ with $\gamma^k_{1m} := (\nu^k_{1m} - \xi^k_m \alpha^k_m)_{I_1(\bar{x})}$ and $\gamma^k_{2m} := (\nu^k_{2m} - \xi^k_m (2 - \alpha^k_m))_{I_2(\bar{x})}$ and $\mathbf{A}_k$ denotes the matrix consisting of the row vectors

$$\begin{array}{rl} \nabla g_j(x^k)^T & j \in I_g(\bar{x}) \\ \nabla h_i(x^k)^T & i \in \{1, \ldots, q\} \\ e^T_{1m} & m \in I_1(\bar{x}) \\ e^T_{2m} & m \in I_2(\bar{x}) \,. \end{array}$$

Due to the continuous differentiability of g and h the row vectors $\nabla g_j(x^k)^T$ and $\nabla h_i(x^k)^T$ converge to $\nabla g_j(\bar{x})^T$ and $\nabla h_i(\bar{x})^T$, respectively. Hence, the sequence $(\mathbf{A}_k)$ converges to the matrix $\mathbf{A}$ consisting of the row vectors

$$\begin{array}{rl} \nabla g_j(\bar{x})^T & j \in I_g(\bar{x}) \\ \nabla h_i(\bar{x})^T & i \in \{1, \ldots, q\} \\ e^T_{1m} & m \in I_1(\bar{x}) \\ e^T_{2m} & m \in I_2(\bar{x}) \,. \end{array}$$

Since the MPEC-LICQ is assumed to hold in $\bar{x}$, these vectors are linear independent, such that $\mathbf{A}$ has full row rank. Hence, there exists a unique solution vector π solving $\mathbf{A}^T \pi = \nabla f(\bar{x})$. The full row rank of A implies that AA^T is invertible, such that by the convergence of (A_k) and the perturbation lemma (see for example Lemma 5.23 in [GK02]) $A_k A_k^T$ is invertible for sufficiently large $k \in \mathbb{N}$. Hence there exists a unique solution vector $\pi_k = (A_k A_k^T)^{-1}(A_k \nabla f(x^k))$. Furthermore, since $\nabla f(x^k)$ converges to $\nabla f(\bar{x})$

$$\pi_k = (A_k A_k^T)^{-1}(A_k \nabla f(x^k)) \longrightarrow (AA^T)^{-1}(A \nabla f(\bar{x})) = \pi \,.$$

Therefore, if we define

$$\bar{\lambda}_j := 0 \quad j \notin I_g(\bar{x}), \qquad \bar{\nu}_{1m} := 0 \quad m \notin I_1(\bar{x}), \qquad \bar{\nu}_{2m} := 0 \quad m \notin I_2(\bar{x}) \,,$$

then $(\lambda^k, \mu^k, (\nu^k_{1m} - \xi^k_m \alpha^k_m), (\nu^k_{2m} - \xi^k_m (2 - \alpha^k_m)))$ converges to the unique multiplier vector $(\bar{\lambda}, \bar{\mu}, \bar{\nu}_1, \bar{\nu}_2)$ of $\bar{x}$ satisfying (1.23).

Since the feasibility of $\bar{x}$ follows by Lemma 2.4, to prove that $\bar{x}$ is C-stationary it remains to show that $\bar{\nu}_{1m}\bar{\nu}_{2m} \geq 0$ holds for all $m \in (I_1 \cap I_2)(\bar{x})$. Suppose without loss of generality there exists an index $m \in (I_1 \cap I_2)(\bar{x})$ with $\bar{\nu}_{1m} < 0$ and $\bar{\nu}_{2m} > 0$. It follows from the convergence $(\nu_{1m}^k - \xi_m^k \alpha_m^k) \to \bar{\nu}_{1m}$ and $(\nu_{2m}^k - \xi_m^k(2 - \alpha_m^k)) \to \bar{\nu}_{2m}$ that

$$\nu_{1m}^k - \xi_m^k \alpha_m^k < 0 \tag{2.20}$$

and

$$\nu_{2m}^k - \xi_m^k(2 - \alpha_m^k) > 0 \tag{2.21}$$

for sufficiently large $k \in \mathbb{N}$. As $\xi_m^k(2 - \alpha_m^k) \geq 0$ for all $k \in \mathbb{N}$ condition (2.21) implies that $\nu_{2m}^k > 0$ for sufficiently large $k \in \mathbb{N}$. Therefore, $m \in I_2(x^k)$ must hold. Hence, either $m \in I_2(x^k) \backslash I_\Phi(x^k, t_k)$ or $m \in I_2(x^k) \cap I_\Phi(x^k, t_k)$. If $m \in I_2(x^k) \backslash I_\Phi(x^k, t_k)$, then $\xi_m^k = 0$. This however implies that $\nu_{1m}^k - \xi_m^k \alpha_m^k = \nu_{1m}^k \geq 0$ which contradicts (2.20). If on the other hand $m \in I_2(x^k) \cap I_\Phi(x^k, t_k)$, then by the second part of Lemma 2.6 $\nu_{1m}^k = 0$ and $\alpha_m^k = 0$ for sufficiently large $k \in \mathbb{N}$. Therefore, $\nu_{1m}^k - \xi_m^k \alpha_m^k = 0$ which again contradicts (2.20).

To prove the second part of the theorem, assume that the SOSC holds in each x^k and that $\bar{x}$ is not M-stationary. Then, there exists at least one index $m_0 \in (I_1 \cap I_2)(\bar{x})$ with $\bar{\nu}_{1m_0} < 0$ and $\bar{\nu}_{2m_0} < 0$. By the convergence of the multipliers, we thus have

$$\nu_{1m_0}^k - \xi_{m_0}^k \alpha_{m_0}^k < 0 \qquad \text{and} \qquad \nu_{2m_0}^k - \xi_{m_0}^k(2 - \alpha_{m_0}^k) < 0\,, \tag{2.22}$$

for sufficiently large $k \in \mathbb{N}$. However, since $\nu_{1m_0}^k \geq 0$ and $\nu_{2m_0}^k \geq 0$, by (2.22) it follows that

$$\xi_{m_0}^k > 0 \qquad \text{and} \qquad 2 > \alpha_{m_0}^k > 0$$

for every $k \in \mathbb{N}$ that is large enough. Since Lemma 2.6 implies $I_\Phi(x^k, t_k) \cap (I_1 \cap I_2)(x^k) = \emptyset$ and considering the values of α_j for $j \in I_\Phi(x^k, t_k) \cap (I_1 \backslash I_2)(x^k)$ or $j \in I_\Phi(x^k, t_k) \cap (I_2 \backslash I_1)(x^k)$ respectively, for k large enough it follows that $m_0 \in I_\Phi(x^k, t_k) \backslash (I_1 \cup I_2)(x^k)$. Hence, $\nu_{1m_0}^k = 0$ and $\nu_{2m_0}^k = 0$ for all $k \in \mathbb{N}$ being sufficiently large. Therefore,

$$0 > \bar{\nu}_{1m_0} = -\lim_{k \to \infty} (\xi_{m_0}^k \alpha_{m_0}^k) \quad \text{and} \quad 0 > \bar{\nu}_{2m_0} = -\lim_{k \to \infty} (\xi_{m_0}^k (2 - \alpha_{m_0}^k))\,. \tag{2.23}$$

As $0 \leq \alpha_{m_0}^k \leq 2$, there cannot exist a subsequence $(\xi_{m_0}^k)_{k \in \mathcal{K}}$ of $(\xi_{m_0}^k)$ that converges to zero.

Suppose there exist a subsequence $(\alpha_{m_0}^k)_{k \in \mathcal{K}}$ that converges to zero. Then $((2 - \alpha_{m_0}^k))_{k \in \mathcal{K}} \to 2$ as $k \to \infty$. This, however, contradicts (2.23), since then either $\xi_{m_0}^k(2 - \alpha_{m_0}^k) \to \infty$ or $\xi_{m_0}^k \alpha_{m_0}^k \to 0$ for $k \in \mathcal{K}$ and $k \to \infty$. By Lemma 2.5 it therefore follows in addition that there exists no subsequence $(x^k)_{k \in \mathcal{K}}$ of (x^k) that satisfies $(x_{1m_0}^k - x_{2m_0}^k)/t_k \to 1$ for $k \in \mathcal{K}$ and $k \to \infty$. The same arguments imply that there exists no subsequence $\alpha_{m_0}^k \to 2$ with $k \in \mathcal{K} \subseteq \mathbb{N}$, in other words, there exists no subsequence $(x_k)_{k \in \mathcal{K}}$ with $(x_{1m_0}^k - x_{2m_0}^k)/t_k \to -1$ as $k \to \infty$ and for $k \in \mathcal{K}$. Furthermore, it follows that $(\xi_{m_0}^k)$ has to be bounded.

Because the MPEC-LICQ holds in $\bar{x}$, for sufficiently large $k \in \mathbb{N}$ we can construct a sequence $(d^k) \subseteq \mathbb{R}^{n+2p}$ such that

$$\begin{array}{rcll} \nabla h_i(x^k)^T d^k & = & 0, & i \in \{1,\dots,q\} \\ \nabla g_j(x^k)^T d^k & = & 0, & j \in I_g(\bar{x}) \\ d^k_{1j} & = & 0, & j \in I_1(\bar{x})\backslash\{m_0\} \\ d^k_{2j} & = & 0, & j \in I_2(\bar{x})\backslash\{m_0\} \\ d^k_{1m_0} & = & 1, & \\ d^k_{2m_0} & = & -\frac{\alpha^k_{m_0}}{(2-\alpha^k_{m_0})}\,. & \end{array}$$

These d^k are well defined and bounded, because there exists an $\varepsilon > 0$ such that $\varepsilon < \alpha^k_{m_0} < 2-\varepsilon$. Moreover, these directions are contained in the corresponding sets of critical directions $\mathcal{S}_{R(t)}(x^k, \lambda^k, \mu^k, \nu^k_1, \nu^k_2, \xi^k, t_k)$, as

$$\nabla\Phi_{m_0}(x^k_1, x^k_2, t_k)^T d^k = \alpha^k_{m_0} d_{1m_0} + (2-\alpha^k_{m_0}) d_{2m_0} = 0\,.$$

The twice continuous differentiability of f, g and h and the convergence of λ^k and μ^k imply that the first three parts (2.24), (2.25) and (2.26) of the right-hand side of

$$\begin{aligned} {d^k}^T\, \nabla^2_{xx}\mathcal{L}_{R(t_k)}(x^k, \lambda^k, \mu^k, \nu^k_1, \nu^k_2, \xi^k)\, d^k = \quad & {d^k}^T \nabla^2_{xx} f(x^k) d^k && (2.24) \\ & - \sum_{j\in I_g(\bar{x})} {d^k}^T \nabla^2_{xx} g_j(x^k) \lambda^k_j d^k && (2.25) \\ & - \sum_{i=1}^{q} {d^k}^T \nabla^2_{xx} h_i(x^k) \mu^k_i d^k && (2.26) \\ & + \sum_{j=1}^{p} {d^k}^T \begin{pmatrix} 0 & 0 \\ 0 & \nabla^2_{yy}\Phi_j(x^k_1, x^k_2, t_k) \end{pmatrix} \xi^k_j d^k\,, && (2.27) \end{aligned}$$

where $y := (x_1, x_2)$, are bounded for $k \to \infty$. Furthermore, for (2.27), we have

$$\begin{array}{rcl} \sum_{j=1}^{p} {d^k}^T \begin{pmatrix} 0 & 0 \\ 0 & \nabla^2_{yy}\Phi_j(x^k_1, x^k_2, t_k) \end{pmatrix} \xi^k_j d^k & = & \sum_{j\in I^k_\Phi\backslash(I^k_1\cup I^k_2)} c_j(x^k, t_k)\, \xi^k_j (d^k_{1j} - d^k_{2j})^2 \\ & = & c_{m_0}(x^k, t_k)\, \xi^k_{m_0} (d^k_{1m_0} - d^k_{2m_0})^2 \\ & = & c_{m_0}(x^k, t_k)\, \xi^k_{m_0} (1 + \frac{\alpha^k_{m_0}}{(2-\alpha^k_{m_0})})^2 \\ & < & c_{m_0}(x^k, t_k)\, \xi^k_{m_0}\,, \end{array}$$

where

$$c_j(x,t) = \begin{cases} 0 & |x_{1j} - x_{2j}| \geq t \\ -\frac{1}{t}\theta''(\frac{x_{1j}-x_{2j}}{t}) & |x_{1j} - x_{2j}| < t \end{cases}$$

and $I^k_\Phi = I_\Phi(x^k, t_k)$ and $I^k_1 \cup I^k_2 = (I_1 \cup I_2)(x^k)$. Since we have shown that (x^k) has no subsequence with $|x^k_{1m_0} - x^k_{2m_0}|/t_k \to 1$ for $k \in \mathcal{K} \subseteq \mathbb{N}$ and $k \to \infty$ and θ'' is continuous and strictly positive on $(-1, 1)$, it follows that we can find a $\delta > 0$ such that

$$\theta''\left(\frac{x^k_{1m_0} - x^k_{2m_0}}{t_k}\right) > \delta$$

for all $k \in \mathbb{N}$ sufficiently large. Hence,

$$\lim_{k\to\infty} c_{m_0}(x^k, t_k) = -\lim_{k\to\infty} \frac{1}{t_k}\theta''\left(\frac{x^k_{1m_0} - x^k_{2m_0}}{t_k}\right) < -\lim_{k\to\infty} \frac{\delta}{t_k} = -\infty\,.$$

However, as $(\xi^k_{m_0}) > \varepsilon$ for some strictly positive ε, this implies

$$\sum_{j=1}^{p} d^{k^T} \begin{pmatrix} 0 & 0 \\ 0 & \nabla^2\Phi_j(x_1^k, x_2^k, t_k) \end{pmatrix} \xi_j^k d^k < c_{m_0}(x^k, t_k)\xi^k_{m_0} \to -\infty$$

for the last term (2.27) and hence

$$d^{k^T} \nabla^2_{xx}\mathcal{L}_{R(t_k)}(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)d^k \to -\infty\,,$$

for $k \to \infty$ which contradicts the assumption that the SOSC holds in each x^k.
Finally, if $m \notin I_\Phi^{inf}$, then there exists a $k_0 \in \mathbb{N}$, such that $\xi^k_m = 0$ for all $k > k_0$. Hence

$$\bar{\nu}_{1m} = \lim_{k\to\infty} (\nu^k_{1m} - \xi^k_m \alpha^k_m) = \lim_{k\to\infty} \nu^k_{1m} \geq 0$$

and

$$\bar{\nu}_{2m} = \lim_{k\to\infty} (\nu^k_{2m} - \xi^k_m (2 - \alpha^k_m)) = \lim_{k\to\infty} \nu^k_{2m} \geq 0\,,$$

such that if $\bar{\nu}_{1m} \geq 0$ and $\bar{\nu}_{1m} \geq 0$ for all $m \in (I_1 \cap I_2)(\bar{x}) \cap I_\Phi^{inf}$, then $\bar{x}$ is strongly stationary. The inverse direktion clearly holds, as $\bar{x}$ is not strongly stationary if either $\bar{\nu}_{1m} < 0$ or $\bar{\nu}_{1m} < 0$ for at least one $m \in (I_1 \cap I_2)(\bar{x}) \cap I_\Phi^{inf} \subseteq (I_1 \cap I_2)(\bar{x})$. □

As B-stationary points that satisfy the MPEC-LICQ are strongly stationary, it is of particular interest, what can be proven, if the assumptions of Theorem 2.3 were relaxed. In particular, we are interested in a relaxation of the constraint qualification. As we will see, assuming that a weaker condition, namely the MPEC-CRCQ, does hold we can still prove convergence to a C-stationary point. Although replacing the MPEC-LICQ by the MPEC-CRCQ the convergence of the sequence of multipliers cannot be established anymore.

The original Constant Rank Constraint Qualification for NLPs as introduced in [Jan84] can be defined as follows.

Definition 2.3. Let $\bar{x}$ be feasible for (1.13). Then the *CRCQ (Constant Rank Constraint Qualification)* is said to hold in $\bar{x}$, if for any subsets $\mathcal{K}_g \subseteq I_g(\bar{x})$ and $\mathcal{K}_h \subseteq I_h(\bar{x})$ the family

$$\{\nabla g_j(y)\,|\, j \in \mathcal{K}_g\} \cup \{\nabla h_j(y)\,|\, j \in \mathcal{K}_h\}$$

has the same rank as the family

$$\{\nabla g_j(\bar{x})\,|\, j \in \mathcal{K}_g\} \cup \{\nabla h_j(\bar{x})\,|\, j \in \mathcal{K}_h\}$$

for any y in a neighbourhood of $\bar{x}$.

The following definition of the MPEC-CRCQ corresponds to the original CRCQ applied to RNLP (estimated in $\hat{x}$).

Definition 2.4. Let $\hat{x}$ be feasible for the MPEC. Then the *MPEC-CRCQ (MPEC-Constant Rank Constraint Qualification)* is said to hold in $\hat{x}$, if for any $\mathcal{K}_g \subseteq I_g(\hat{x})$, $\mathcal{K}_1 \subseteq I_1(\hat{x})$, $\mathcal{K}_2 \subseteq I_2(\hat{x})$ and $\mathcal{K}_h \subseteq \{1, \ldots, q\}$ there exists a neighbourhood $\mathcal{U}(\hat{x})$ such that for any $y \in \mathcal{U}(\hat{x})$ the family of gradient vectors

$$\{\nabla g_j(y) \,|\, j \in \mathcal{K}_g\} \cup \{\nabla h_j(y) \,|\, j \in \mathcal{K}_h\} \cup \{e_{1j} \,|\, j \in \mathcal{K}_1\} \cup \{e_{2j} \,|\, j \in \mathcal{K}_2\}\,,$$

has the same rank as the family

$$\{\nabla g_j(\hat{x}) \,|\, j \in \mathcal{K}_g\} \cup \{\nabla h_j(\hat{x}) \,|\, j \in \mathcal{K}_h\} \cup \{e_{1j} \,|\, j \in \mathcal{K}_1\} \cup \{e_{2j} \,|\, j \in \mathcal{K}_2\}\,.$$

Next, we further need an auxiliary lemma for the proof of the next theorem.

Lemma 2.10. *Let x be a stationary point of the NLP*

$$\begin{array}{rl} \text{min} & f(x) \\ \text{subject to} & h(x) = 0 \\ & g(x) \geq 0. \end{array}$$

Then there exist feasible multipliers $(\bar{\lambda}, \bar{\mu})$, such that the family

$$\begin{array}{ll} \nabla g_j(x) & j \in I_g^+(x, \bar{\lambda}) \\ \nabla h_j(x) & j \in \{i \in I_h \,|\, \bar{\mu} \neq 0\} \end{array}$$

is linearly independent.

Proof. As x is assumed to be stationary, there exist multipliers (λ, μ), such that

$$\nabla f(x) - \sum_{j \in I_g(x)} \lambda_j \nabla g_j(x) - \sum_{j \in I_h(x)} \mu_j \nabla h_j(x) = 0\,, \qquad \lambda_j \geq 0 \tag{2.28}$$

If we substitute $\mu_j := \mu_j^+ - \mu_j^-$ in (2.28) and claim that $\mu_j^+, \mu_j^- \geq 0$, then finding multipliers satisfying (2.28) corresponds to finding a solution to

$$Az = b, \qquad z \geq 0$$

where the columns of A correspond to the gradient vectors ∇g_j with $j \in I_g(x)$, ∇h_j and $-\nabla h_j$, $b = \nabla f(x)$ and $z = (\lambda, \mu^+, \mu^-)$. Applying a main result of Linear Optimization (see for example Chapter 3.1 in [Pad99]) we obtain the conclusion. □

Furthermore, we will make use of another definition to abbreviate the notation of the proof of the next theorem.

Definition 2.5. Let $\lambda \in \mathbb{R}^m$ then the *support* of λ is defined as

$$\mathit{supp}(\lambda) := \{\, j \in \{1, \ldots, m\} \,|\, \lambda_j \neq 0\,\}\,.$$

Now we have provided all necessary information that we need to prove the following convergence result.

Theorem 2.4. *Let $(t_k) \subseteq \mathbb{R}^+$ be a sequence that satisfies $t_k \to 0$, further let $(x^k) \subseteq \mathbb{R}^{n+2p}$ be a sequence of stationary points of $R(t_k)$ that satisfies $x^k \to \bar{x}$ and suppose the MPEC-CRCQ holds in $\bar{x}$. Then $\bar{x}$ is a C-stationary point of (1.19).*

Proof. Let x^k be a stationary point of $R(t_k)$. Since $g(x)$ is continuous, $I_g(x^k) \subseteq I_g(\bar{x})$ for sufficiently large k, as well as $I_1(x^k) \subseteq I_1(\bar{x})$ and $I_2(x^k) \subseteq I_2(\bar{x})$. Hence, as x^k is stationary for $R(t_k)$, there exist λ^k, μ^k, ν_1^k, ν_2^k

$$\begin{aligned} -\nabla f(x^k) + \sum_{i \in I_g(\bar{x})} \lambda_i^k \nabla g_i(x^k) + \sum_{j=1}^{q} \mu_j^k \nabla h_j(x^k) \\ + \sum_{m \in I_1(\bar{x})} \nu_{1m}^k e_{1m} + \sum_{m \in I_2(\bar{x})} \nu_{2m}^k e_{2m} \\ - \sum_{m \in I_\Phi(x^k, t_k)} \xi_m^k (\alpha_m^k e_{1m} + (2 - \alpha_m^k) e_{2m}) = 0\,. \end{aligned} \tag{2.29}$$

If we define

$$\gamma_{1j}^k := \alpha_j^k \, \xi_j^k \qquad \text{and} \qquad \gamma_{2j}^k := (2 - \alpha_j^k)\, \xi_j^k$$

for all $k \in \mathbb{N}$ and for all $j \in \{1, \ldots, p\}$, then $\gamma_{1j}^k \geq 0$ and $\gamma_{2j}^k \geq 0$. Moreover,

$$\begin{aligned} \gamma_{1j}^k > 0 \quad &\Longleftrightarrow \quad \alpha_j^k > 0 \quad \text{and} \quad \xi_j^k > 0 \quad \Longrightarrow \quad j \in I_\Phi(x^k, t_k) \backslash I_2(x^k) \\ \gamma_{2j}^k > 0 \quad &\Longleftrightarrow \quad \alpha_j^k < 2 \quad \text{and} \quad \xi_j^k > 0 \quad \Longrightarrow \quad j \in I_\Phi(x^k, t_k) \backslash I_1(x^k) \end{aligned} \tag{2.30}$$

Hence, for sufficiently large $k \in \mathbb{N}$

$$\begin{aligned} supp(\gamma_1^k) \subseteq I_\Phi(x^k, t_k) \backslash I_2(x^k) \subseteq I_1(\bar{x}) \\ supp(\gamma_2^k) \subseteq I_\Phi(x^k, t_k) \backslash I_1(x^k) \subseteq I_2(\bar{x})\,. \end{aligned} \tag{2.31}$$

Since

$$\begin{aligned} supp(\lambda^k) \subseteq I_g(\bar{x}), \quad supp(\mu^k) \subseteq I_h(\bar{x}) \\ supp(\nu_1^k) \subseteq I_1(\bar{x}), \quad supp(\nu_2^k) \subseteq I_2(\bar{x})\,, \end{aligned} \tag{2.32}$$

we can then write (2.29) as

$$\begin{aligned} -\nabla f(x^k) + \sum_{j \in supp(\lambda^k)} \lambda_j^k \nabla g_j(x^k) + \sum_{j \in supp(\mu^k)} \mu_j^k \nabla h_j(x^k) \\ + \sum_{j \in supp(\nu_1^k)} \nu_{1j}^k e_{1j} + \sum_{j \in supp(\nu_2^k)} \nu_{2j}^k e_{2j} \\ + \sum_{j \in supp(\gamma_1^k)} \gamma_{1j}^k (-e_{1j}) + \sum_{j \in supp(\gamma_2^k)} \gamma_{2j}^k (-e_{2j}) = 0\,. \end{aligned} \tag{2.33}$$

As a consequence of Lemma 2.10, we can find a multiplier vector $(\bar{\lambda}^k, \bar{\mu}^k, \bar{\nu}_1^k, \bar{\nu}_2^k, \bar{\gamma}_1^k, \bar{\gamma}_2^k)$ for every x^k, such that (2.33) is satisfied and the family

$$\begin{aligned} \{\nabla g_j(x^k) \,|\, j \in supp(\bar{\lambda}^k)\} \cup \{\nabla h_j(x^k) \,|\, supp(\bar{\mu}^k)\} \cup \{e_{1j} \,|\, supp(\bar{\nu}_1^k)\} \\ \cup \{e_{2j} \,|\, supp(\bar{\nu}_2^k)\} \cup \{-e_{1j} \,|\, supp(\bar{\gamma}_1^k)\} \cup \{-e_{2j} \,|\, supp(\bar{\gamma}_2^k)\} \end{aligned} \tag{2.34}$$

is linear independent. Furthermore,

$$\begin{aligned} &supp(\bar{\lambda}^k) \subseteq supp(\lambda^k), && supp(\bar{\mu}^k) \subseteq supp(\mu^k) \\ &supp(\bar{\nu}_1^k) \subseteq supp(\nu_1^k), && supp(\bar{\nu}_2^k) \subseteq supp(\nu_2^k), \\ &supp(\bar{\gamma}_1^k) \subseteq supp(\gamma_1^k), && supp(\bar{\gamma}_2^k) \subseteq supp(\gamma_2^k)\,. \end{aligned} \tag{2.35}$$

The linear independence of the family implies that

$$supp(\bar{\nu}_1^k) \cap supp(\bar{\gamma}_1^k) = \emptyset \qquad \text{and} \qquad supp(\bar{\nu}_2^k) \cap supp(\bar{\gamma}_2^k) = \emptyset\,. \tag{2.36}$$

By (2.31), (2.32) and (2.35) we have that

$$supp(\bar{\nu}_1^k) \cup supp(\bar{\gamma}_1^k) \subseteq I_1(\bar{x}) \qquad \text{and} \qquad supp(\bar{\nu}_2^k) \cup supp(\bar{\gamma}_2^k) \subseteq I_2(\bar{x})\,. \tag{2.37}$$

Next, we distinguish two cases: either the sequence $((\bar{\lambda}^k,\, \bar{\mu}^k,\, \bar{\nu}_1^k,\, \bar{\nu}_2^k,\, \bar{\gamma}_1^k,\, \bar{\gamma}_2^k))$ does have a bounded subsequence or it does not have one.

Let us first consider the case that it does not have a bounded subsequence. This implies that the sequence $((\bar{\lambda}^k,\, \bar{\mu}^k,\, \bar{\nu}_1^k,\, \bar{\nu}_2^k,\, \bar{\gamma}_1^k,\, \bar{\gamma}_2^k))$ is not bounded either. Hence, we can find an index $k_0 \in \mathbb{N}$, such that $\| (\bar{\lambda}^k,\, \bar{\mu}^k,\, \bar{\nu}_1^k,\, \bar{\nu}_2^k,\, \bar{\gamma}_1^k,\, \bar{\gamma}_2^k) \| > 0$ for all $k \geq k_0$. We therefore consider the normalized (thus bounded) sequence of multipliers (starting with index $k = k_0$)

$$(\tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\gamma}_1^k, \tilde{\gamma}_2^k) = \frac{(\bar{\lambda}^k, \bar{\mu}^k, \bar{\nu}_1^k, \bar{\nu}_2^k, \bar{\gamma}_1^k, \bar{\gamma}_2^k)}{\| (\bar{\lambda}^k, \bar{\mu}^k, \bar{\nu}_1^k, \bar{\nu}_2^k, \bar{\gamma}_1^k, \bar{\gamma}_2^k) \|}\,.$$

As this sequence is bounded, it has a convergent subsequence

$$(\tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\gamma}_1^k, \tilde{\gamma}_2^k)_{\kappa \in \mathcal{K}} \longrightarrow (\tilde{\lambda}, \tilde{\mu}, \tilde{\nu}_1, \tilde{\nu}_2, \tilde{\gamma}_1, \tilde{\gamma}_2)\,.$$

Moreover,

$$\begin{array}{rcccl} supp(\tilde{\lambda}) & \subseteq & supp(\tilde{\lambda}^k) & = & supp(\bar{\lambda}^k), \\ supp(\tilde{\mu}) & \subseteq & supp(\tilde{\mu}^k) & = & supp(\bar{\mu}^k), \\ supp(\tilde{\nu}_1) & \subseteq & supp(\tilde{\nu}_1^k) & = & supp(\bar{\nu}_1^k), \\ supp(\tilde{\nu}_2) & \subseteq & supp(\tilde{\nu}_2^k) & = & supp(\bar{\nu}_2^k), \\ supp(\tilde{\gamma}_1) & \subseteq & supp(\tilde{\gamma}_1^k) & = & supp(\bar{\gamma}_1^k), \\ supp(\tilde{\gamma}_2) & \subseteq & supp(\tilde{\gamma}_2^k) & = & supp(\bar{\gamma}_2^k), \end{array} \tag{2.38}$$

holds for all $k \in \mathcal{K}$ that are large enough. Therefore, for all $k \in \mathcal{K}$ sufficiently large, by

(2.33) and the continuity of $\nabla f(x)$, $\nabla h(x)$ and $\nabla g(x)$ we get

$$\begin{aligned} 0 &= \lim_{\substack{\kappa\in\mathcal{K}\\ \kappa\to\infty}} \Bigg(-\frac{\nabla f(x^\kappa)}{\omega_\kappa} + \sum_{i\in I_g(\bar{x})} \frac{\bar{\lambda}_i^\kappa}{\omega_\kappa}\nabla g_i(x^\kappa) + \sum_{j\in I_h(\bar{x})} \frac{\bar{\mu}_j^\kappa}{\omega_\kappa}\nabla h_j(x^\kappa) \\ &\quad + \sum_{j\in I_1(\bar{x})} \frac{\bar{\nu}_{1j}^\kappa}{\omega_\kappa} e_{1j} + \sum_{j\in I_2(\bar{x})} \frac{\bar{\nu}_{2j}^\kappa}{\omega_\kappa} e_{2j} \\ &\quad + \sum_{j\in I_1(\bar{x})} \frac{\bar{\gamma}_{1j}^\kappa}{\omega_\kappa}(-e_{1j}) + \sum_{j\in I_2(\bar{x})} \frac{\bar{\nu}_{2j}^\kappa}{\omega_\kappa}(-e_{2j}) \Bigg) \\ &= \sum_{j\in supp(\tilde{\lambda})} \tilde{\lambda}_j \nabla g_j(\bar{x}) + \sum_{j\in supp(\tilde{\mu})} \tilde{\mu}_j \nabla h_j(\bar{x}) + \\ &\quad \sum_{j\in supp(\tilde{\nu}_1)} \tilde{\nu}_{1j} e_{1j} + \sum_{j\in supp(\tilde{\nu}_2)} \tilde{\nu}_{2j} e_{2j} + \\ &\quad \sum_{j\in supp(\tilde{\gamma}_1)} \tilde{\gamma}_{1j}(-e_{1j}) + \sum_{j\in supp(\tilde{\gamma}_2)} \tilde{\gamma}_{2j}(-e_{2j}) \end{aligned} \tag{2.39}$$

where $\omega_\kappa = \|(\bar{\lambda}^\kappa, \bar{\mu}^\kappa, \bar{\nu}_1^\kappa, \bar{\nu}_2^\kappa, \bar{\gamma}_1^\kappa, \bar{\gamma}_2^\kappa)\|$. As $\|(\tilde{\lambda}, \tilde{\mu}, \tilde{\nu}_1, \tilde{\nu}_2, \tilde{\gamma}_1, \tilde{\gamma}_2)\| = 1$, there have to exist some entries of $(\tilde{\lambda}, \tilde{\mu}, \tilde{\nu}_1, \tilde{\nu}_2, \tilde{\gamma}_1, \tilde{\gamma}_2)$ that do not vanish. Hence,

$$\begin{gathered} \{\nabla g_j(\bar{x}) \,|\, j \in supp(\tilde{\lambda})\} \cup \{\nabla h_j(\bar{x}) \,|\, supp(\tilde{\mu})\} \cup \{e_{1j} \,|\, supp(\tilde{\nu}_1)\} \\ \cup\{e_{2j} \,|\, supp(\tilde{\nu}_2)\} \cup \{-e_{1j} \,|\, supp(\tilde{\gamma}_1)\} \cup \{-e_{2j} \,|\, supp(\tilde{\gamma}_2)\} \end{gathered}$$

is linear dependent. Then it follows by (2.38) that the family

$$\begin{gathered} \{\nabla g_j(\bar{x}) \,|\, j \in supp(\bar{\lambda}^k)\} \cup \{\nabla h_j(\bar{x}) \,|\, supp(\bar{\mu}^k)\} \cup \{e_{1j} \,|\, supp(\bar{\nu}_1^k)\} \\ \cup\{e_{2j} \,|\, supp(\bar{\nu}_2^k)\} \cup \{-e_{1j} \,|\, supp(\bar{\gamma}_1^k)\} \cup \{-e_{2j} \,|\, supp(\bar{\gamma}_2^k)\} \end{gathered} \tag{2.40}$$

is linear dependent for all $k \in \mathcal{K}$ that are sufficiently large. Now as $\bar{x}$ satisfies the MPEC-CRCQ, there exists a neighbourhood $\mathcal{U}(\bar{x})$, such that for any $y \in \mathcal{U}(\bar{x})$ it follows by (2.40), (2.32), (2.35), (2.36) and (2.37) that the family of (2.40) evaluated in y has the same rank. Therefore it is also linear dependent.
Furthermore, as x^k converges to $\bar{x}$, there exists a $k_1 \in \mathbb{N}$, such that x^k lies within $\mathcal{U}(\bar{x})$ for all $k > k_1$. This however implies that

$$\begin{gathered} \{\nabla g_j(x^k) \,|\, j \in supp(\bar{\lambda}^k)\} \cup \{\nabla h_j(x^k) \,|\, supp(\bar{\mu}^k)\} \cup \{e_{1j} \,|\, supp(\bar{\nu}_1^k)\} \\ \cup\{e_{2j} \,|\, supp(\bar{\nu}_2^k)\} \cup \{-e_{1j} \,|\, supp(\bar{\gamma}_1^k)\} \cup \{-e_{2j} \,|\, supp(\bar{\gamma}_2^k)\} \end{gathered}$$

is linear dependent for all $k \in \mathbb{N}$ sufficiently large which contradicts the linear independence of the family of gradients of (2.34).

We therefore may assume that the sequence of multipliers $(\bar{\lambda}^k, \bar{\mu}^k, \bar{\nu}_1^k, \bar{\nu}_2^k, \bar{\gamma}_1^k, \bar{\gamma}_2^k)$ has a bounded subsequence which implies that it has also a convergent subsequence

$$(\bar{\lambda}^k, \bar{\mu}^k, \bar{\nu}_1^k, \bar{\nu}_2^k, \bar{\gamma}_1^k, \bar{\gamma}_2^k)_{k\in\mathcal{K}} \longrightarrow (\lambda^*, \mu^*, \nu_1^*, \nu_2^*, \gamma_1^*, \gamma_2^*)\,.$$

By the continuity of $\nabla f(x)$, $\nabla h(x)$ and $\nabla g(x)$ and (2.33) these limit multipliers satisfy

$$\begin{aligned} -\nabla f(\bar{x}) + \sum_{i \in I_g(\bar{x})} \lambda_i^* \nabla g_i(\bar{x}) + \sum_{j=1}^{q} \mu_j^* \nabla h_j(\bar{x}) \\ + \sum_{j \in I_1(\bar{x})} \hat{\nu}_{1j} e_{1j} + \sum_{j \in I_2(\bar{x})} \hat{\nu}_{2j} e_{2j} = 0 \end{aligned} \tag{2.41}$$

with

$$\hat{\nu}_{1j} := \begin{cases} \nu_1^* & j \in supp(\nu_1^*) \\ -\gamma_1^* & j \in supp(\gamma_1^*) \\ 0 & \text{elsewise} \end{cases} \qquad \hat{\nu}_{2j} := \begin{cases} \nu_2^* & j \in supp(\nu_2^*) \\ -\gamma_2^* & j \in supp(\gamma_2^*) \\ 0 & \text{elsewise} . \end{cases} \tag{2.42}$$

Because of the convergence of the subsequence

$$supp(\nu_1^*) \subseteq supp(\bar{\nu}_1^k) \qquad \text{and} \qquad supp(\gamma_1^*) \subseteq supp(\bar{\gamma}_1^k),$$

as well as

$$supp(\nu_2^*) \subseteq supp(\bar{\nu}_2^k) \qquad \text{and} \qquad supp(\gamma_2^*) \subseteq supp(\bar{\gamma}_2^k)$$

for all $k \in \mathcal{K}$ that are sufficiently large. Hence, by (2.36)

$$supp(\nu_1^*) \cap supp(\gamma_1^*) = \emptyset \qquad \text{and} \qquad supp(\nu_2^*) \cap supp(\gamma_2^*) = \emptyset\,,$$

which together with (2.37) implies that the multipliers $\hat{\nu}_1$ and $\hat{\nu}_2$ are well-defined. Thus by Lemma 2.4 $\bar{x}$ is weakly stationary with multipliers $(\lambda^*,\, \mu^*,\, \hat{\nu}_1,\, \hat{\nu}_2)$. Suppose it is not C-stationary, then there exists at least one index $j_0 \in (I_1 \cap I_2)(\bar{x})$ with either $\hat{\nu}_{1j_0} < 0$ and $\hat{\nu}_{2j_0} > 0$ or $\hat{\nu}_{1j_0} > 0$ and $\hat{\nu}_{2j_0} < 0$. Suppose without loss of generality that $\hat{\nu}_{1j_0} < 0$ and $\hat{\nu}_{2j_0} > 0$, then by the convergence of

$$(\bar{\lambda}^k, \bar{\mu}^k, \bar{\nu}_1^k, \bar{\nu}_2^k, \bar{\gamma}_1^k, \bar{\gamma}_2^k)_{k \in \mathcal{K}}\,,$$

it follows that

$$j_0 \in supp(\bar{\gamma}_1^k) \qquad \text{and} \qquad j_0 \in supp(\bar{\nu}_2^k)$$

for $k \in \mathcal{K}$ sufficiently large. Then, however, by (2.35)

$$j_0 \in supp(\gamma_1^k) \qquad \text{and} \qquad j_0 \in supp(\nu_2^k),$$

for all $k \in \mathcal{K}$ being large enough. It follows together with (2.31) and (2.32) that

$$j_0 \in I_\Phi(x^k, t_k) \backslash I_2(x^k) \qquad \text{and} \qquad j_0 \in I_2(x^k)$$

for all $k \in \mathcal{K}$ sufficiently large. This obviously constitutes a contradiction. Hence, $\bar{x}$ has to be C-stationary.

□

Next, we describe an example that demonstrates that, in order to proof M-stationarity of a limit point x^*, we would need further conditions.

Example 2.1.

$$\begin{aligned} &\mathit{min}\ \tfrac{1}{2}((x_1-1)^2+(x_2-1)^2) \\ &\mathit{subject\ to}\ 0 \le\ x_1\ \perp\ x_2\ \ge 0, \qquad : \nu_1, \nu_2, \xi \end{aligned}$$

Consider the sequence $(x_1^k, x_2^k) = (\vartheta_k, \vartheta_k)$ *with* $\vartheta_k \searrow 0$*, which is a sequence of KKT-points for* $R(t_k)$ *for a sequence* (t_k) *with* $t_k \searrow 0$ *(determined such that* $\Phi(\vartheta_k, \vartheta_k, t_k) = 0$ *for all* k*), as*

$$\begin{pmatrix}\vartheta_k-1\\ \vartheta_k-1\end{pmatrix} + \xi^k \begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$

with $\xi^k = 1-\vartheta_k$*,* $\nu_1^k = 0$ *and* $\nu_2^k = 0$*. The vector* (x_1^k, x_2^k) *converges to* $(\bar{x}_1, \bar{x}_2) = (0,0)$ *which clearly satisfies the MPEC-CRCQ. The MPEC multipliers* $(\hat{\nu}_1, \hat{\nu}_2)$ *that we can construct accordingly to Theorem 2.4 from the limits of* ν_1^k*,* ν_2^k*,* ξ^k *and* α^k *are* $(\hat{\nu}_1, \hat{\nu}_2) = (-1,-1)$*. In contrast to the next example, for this example we cannot construct multipliers* $(\hat{\nu}_1, \hat{\nu}_2)$ *that satisfy the M-stationarity condition* $\hat{\nu}_1\hat{\nu}_2 = 0$ *or* $\hat{\nu}_1 > 0$ *and* $\hat{\nu}_2 > 0$*. Thus* $(0,0)$ *is C-stationary.*

However, the next example demonstrates that it would be a valuable task to proof M-stationarity without the need of the MPEC-LICQ, as the limit point $\bar{x}$ in the next example is M-stationary though it does not satisfy the MPEC-LICQ and Theorem 2.3 cannot be applied.

Example 2.2.

$$\begin{aligned} &\mathit{min}\ x_2 - 2x_1 \\ &\mathit{subject\ to}\ x_2 - x_1 \ge 0 \qquad : \lambda \\ &\qquad 0 \le\ x_1\ \perp\ x_2\ \ge 0 \qquad : \nu_1, \nu_2, \xi \end{aligned}$$

The solution of this MPEC is $(\bar{x}_1, \bar{x}_2) = (0,0)$ *and the MPEC-CRCQ is satisfied in* $(\bar{x}_1, \bar{x}_2)$*, though not the MPEC-LICQ.*
Consider the sequence $(x_1^k, x_2^k) = (\vartheta_k, \vartheta_k)$ *with* $\vartheta_k \searrow 0$*, which is a sequence of KKT-points for* $R(t_k)$ *for a sequence* (t_k) *with* $t_k \searrow 0$ *(such that* $\Phi(\vartheta_k, \vartheta_k, t_k) = 0$ *for all* k*), as*

$$\begin{pmatrix}-2\\1\end{pmatrix} - \lambda^k \begin{pmatrix}-1\\1\end{pmatrix} + \xi^k \begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix},$$

with $\lambda^k = 1.5$ *and* $\xi^k = 0.5$*. In the limit these multipliers lead to the MPEC multipliers* $\lambda^* = 1.5$*,* $\hat{\nu}_1 = -\alpha^*\xi^* = -0.5$ *and* $\hat{\nu}_2 = -(2-\alpha^*)\xi^* = -0.5$*. Hence* $(0,0)$ *is at least C-stationary.*

However, for $(\bar{x}_1, \bar{x}_2) = (0,0)$ *the active gradients* $(-1,1)^T$*,* $(1,0)^T$ *and* $(0,1)^T$ *satisfy a linear dependency, such that by substituting either* $(1,0)^T$ *or* $(0,1)^T$ *we can construct MPEC multipliers with either* $\lambda^* = 2$*,* $\hat{\nu}_1 = 0$ *and* $\hat{\nu}_2 = -1$ *or* $\lambda^* = 1$*,* $\hat{\nu}_1 = -1$ *and* $\hat{\nu}_2 = 0$ *both satisfying the M-stationarity condition.*

Remark 2.3. Notice that Theorem 2.3 as well as Theorem 2.4 can be extended to the convergence of a sequence (x^k) that satisfies (2.14) only approximately, that is

$$\nabla_x \mathcal{L}_{R(t)}(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k) = \varepsilon_k$$

for a sequence (ε_k) with $\varepsilon_k \searrow 0$.

2.4 Extensions

In this section we consider two possible alternatives of the relaxation, we have discussed in the previous sections. The idea of the first alternative is to merge the "relaxing constraints" $\Phi_j(x_1, x_2, t) \leq 0$ with $j \in \{1, \ldots p\}$ into one single constraint $\Phi(x_1, x_2, t)^T \mathbf{e} \leq 0$, such that $R(t)$ becomes

$$\begin{array}{rrcl} \min & f(x) & & \\ \text{subject to} & h(x) & = & 0 \\ & g(x) & \geq & 0 \\ & x_1, x_2 & \geq & 0 \\ & \Phi(x_1, x_2, t)^T \mathbf{e} & \leq & 0\,. \end{array}$$

where $\Phi(x_1, x_2, t) : \mathbb{R}^p \times \mathbb{R}^p \times \mathbb{R}_0^+ \to \mathbb{R}^p$ and $\mathbf{e} \in \mathbb{R}^p$ is the vector of all ones. The single constraint $\Phi(x_1, x_2, t)^T \mathbf{e} \leq 0$ hence conforms to the condition

$$\sum_{j=1}^{p} \Phi_j(x_1, x_2, t) \leq 0.$$

This idea has been discussed for example by Fletcher et al. in [FLRS06] for the exact bilinear reformulation of the complementarity constraint (see Section 1.4) and by Scholtes in [Sch01] for the relaxed bilinear reformulation. This extension, however, does not maintain the characteristics of our relaxation method, as we will explain now.

In the context of a reformulation of the complementarity constraints in terms of $x_{1j}x_{2j}$ it is reasonable to consider this extension. The conditions $x_1 \geq 0$ and $x_2 \geq 0$ imply that each product $x_{1j}x_{2j}$ is nonnegative. Hence, due to the fact that the summands cannot even out each other $x_1^T x_2 \leq 0$ implies that $x_{1j}x_{2j} \leq 0$ for all $j \in \{1, \ldots, p\}$. Accordingly $x_1^T x_2 - t \leq 0$ implies that $x_{1j}x_{2j} - t \leq 0$ for all $j \in \{1, \ldots, p\}$. For the bilinear reformulations we can therefore replace the p constraints $x_{1j}x_{2j} \leq 0$ by one single constraint $x_1^T x_2 \leq 0$ or $x_{1j}x_{2j} - t \leq 0$ by $x_1^T x_2 - t \leq 0$, respectively, and maintain (or tighten) the feasibility properties of the reformulated problem.

Applying this idea to our relaxation method we obtain the difficulty that the conditions $x_1 \geq 0$ and $x_2 \geq 0$ do not imply either sign of $\Phi_j(x_1, x_2, t)$. Hence, the summands might even out each other and $\Phi_j(x_1, x_2, t) \leq 0$ cannot be guaranteed by $\Phi(x_1, x_2, t)^T \mathbf{e} \leq 0$ for all $j \in \{1, \ldots, p\}$.

Consider for example the case $p = 2$ and let $(x_{11}, x_{21}) = (0, 0)$. Then $\Phi_1(x_1, x_2, t) = \Phi_1(0, 0, t) = -t\theta(0) < 0$ and the constraint $\Phi(x_1, x_2, t)^T \mathbf{e} \leq 0$ is also satisfied if $0 < \Phi_2(x_1, x_2, t) \leq t\theta(0)$. Figure 2.4 illustrates the set of points (x_{12}, x_{22}) that satisfy $(x_{12}, x_{22}) \geq 0$ and $\Phi_2(x_1, x_2, t) \leq t\theta(0)$.

Hence, for all $t > 0$, we loose the guarantee of the exact feasibility of points that satisfy $|x_{1j} - x_{2j}| \geq t$. Since this is one of the main properties of the relaxed problem $R(t)$ to distinguish it from $NLP(t)$, we do not further investigate this extension of our approach here.

The second idea that comes into ones mind concerns the relaxation parameter. So far we considered a relaxed nonlinear program $R(t)$ that is parameterized by a scalar parameter $t \in \mathbb{R}$. However, as we will see later in Chapter 4, it is sometimes more favourable to use a parameter vector $t = (t_1, \ldots, t_p) \in (\mathbb{R}^+)^p$ instead of a scalar one. In this case we

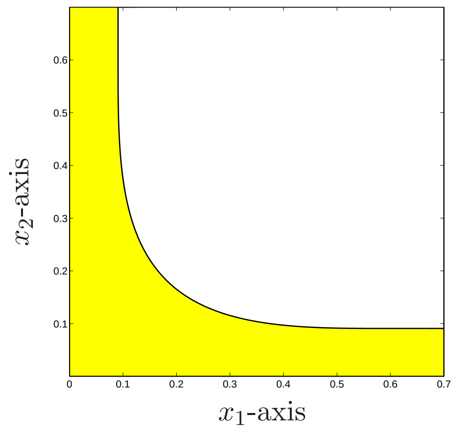


Figure 2.2: Contour of $\Phi_2(x_1, x_2, t) = t\theta_s(0)$ with $\theta_s(z)$ as defined by (2.2) and feasible region (concerning Φ_2) for $t = 0.5$.

use an independent parameter $t_j \in \mathbb{R}^+$ for each constraint $\Phi_j(x_1, x_2, t) \leq 0$. The relaxed, parametrized problem then has the form

$$\begin{array}{rrcl} \min & f(x) & & \\ \text{subject to} & h(x) & = & 0 \\ & g(x) & \geq & 0 \\ & x_1, x_2 & \geq & 0 \\ & \Phi(x_1, x_2, t) & \leq & 0\,, \end{array} \tag{2.43}$$

where $\Phi(x_1, x_2, t) : \mathbb{R}^p \times \mathbb{R}^p \times (\mathbb{R}_0^+)^p \to \mathbb{R}^p$,

$$\Phi_j(x_1, x_2, t) := x_{1j} + x_{2j} - \varphi_j(x_{1j}, x_{2j}, t_j)$$

and $\varphi_j(x_{1j}, x_{2j}, t_j)$ as defined in Section 2.1. We then have

$$\nabla_x \Phi_j(x_1, x_2, t) = \alpha_j\, e_{1j} + (2 - \alpha_j)\, e_{2j} \tag{2.44}$$

with

$$\alpha_j := \frac{\partial \Phi_j(x_1, x_2, t)}{\partial x_{1j}} = \begin{cases} 0 & x_{1j} \geq x_{2j} + t_j \\ 2 & x_{1j} \leq x_{2j} - t_j \\ 1 - \theta'(\frac{x_{1j} - x_{2j}}{t_j}) & \text{otherwise} \end{cases}$$

Although we have to save and update a possibly large parameter vector $t \in \mathbb{R}^p$, the advantage of this alternative is that we can adapt the parameter values individually for each constraint $\Phi_j(x_1, x_2, t) \leq 0$. Moreover, if we undertake some slight modifications, the theoretical results of the foregoing two sections do also hold for these alternatively relaxed programs $R(t)$.

Next, we briefly discuss the modifications of the assumptions and the adjusted results for most of the results of the previous sections, before we state the alternative variants of Theorem 2.1 and Corollary 2.1 for (2.43).

Considering Lemma 2.1, we only have to replace the conditions $\max(\hat{x}_{1j}, \hat{x}_{2j}) < t$ and $\max(\hat{x}_{1j}, \hat{x}_{2j}) > t$ by $\max(\hat{x}_{1j}, \hat{x}_{2j}) < t_j$ or $\max(\hat{x}_{1j}, \hat{x}_{2j}) > t_j$, respectively.

Replacing the condition $0 \leq t_1 < t_2$ in Lemma 2.2 by the conditions $0 \leq t_{1j} \leq t_{2j}$ for all $j \in \{1, \dots, p\}$, the conclusion $\mathcal{Z}(t_1) \subseteq \mathcal{Z}(t_2)$ still holds.

Accordingly, the statement of Lemma 2.3 can be transferred to problem (2.43), if we replace the condition $t \in [0, \hat{t}\,]$ by $0 \leq t_j \leq \hat{t}_j$ for all $j \in \{1, \dots, p\}$.

Furthermore, Lemma 2.4 can easily be adjusted by replacing $t_k \to 0$ by $\|t^k\| \to 0$ and the conclusion $\mathcal{Z} = \bigcap_{k=1}^{\infty} \mathcal{Z}(t^k)$ holds.

In Lemma 2.6 we need to adapt the bounds on the variables x_{ij} with $i = 1, 2$, thus $x_{ij} \geq t$ by $x_{ij} \geq t_j$ and $x_{ij} < t$ by $x_{ij} < t_j$ and the conclusion holds.

Now, considering the results of Section 2.2, we generally replace the condition $t \in (0, \tau(x^*)]$ by $0 < t_j \leq \max(x^*_{1j}, x^*_{2j})$ for all $j \in \{1, \dots, p\} \backslash (I_1 \cap I_2)(x^*)$ and $t_j > 0$ for all $j \in (I_1 \cap I_2)(x^*)$ and the corresponding conclusions hold.

The corresponding variant of Theorem 2.1 for problem (2.43) can be stated as follows. As the proof follows the same ideas of the proof of Theorem 2.1, we will only give a brief sketch of it.

Theorem 2.5. *Let x^* be feasible for (1.19), then*

1. *if x^* is a strongly stationary point of the MPEC (1.19) that satisfies the RNLP-SOSC, then x^* is a stationary point of $R(t)$ for every $t \in \mathbb{R}^p$ that satisfies $0 < t_j \leq \max(x^*_{1j}, x^*_{2j})$ for all $j \in \{1, \dots, p\} \backslash (I_1 \cap I_2)(x^*)$ and $t_j > 0$ for all $j \in (I_1 \cap I_2)(x^*)$, which satisfies the SOSC for $R(t)$. Thus, it is a strict local minimum of $R(t)$.*

2. *if x^* is a stationary point for $R(t)$ that satisfies the SOSC, then x^* is also a strongly stationary point of the MPEC (1.19) that satisfies the RNLP-SOSC. Therefore, it is also a strict local minimum of (1.19) .*

Proof. First the stationarity of x^* results from corresponding variants of Lemma 2.7 or respectively of Lemma 2.9 as explained above.

Hence, it remains to show, that the second order conditions in x^* imply each other. Therefore, we again compare the second derivative with respect to x^* of the Lagrangian functions of the MPEC (1.22) with the suitable variant for problem (2.43). If the parameter t satisfies $0 < t_j \leq \max(x^*_{1j}, x^*_{2j})$ for all $j \in \{1, \dots, p\} \backslash (I_1 \cap I_2)(x^*)$, then by view of (2.44) the second derivatives $\nabla^2_{xx} \Phi_j(x^*_1, x^*_2, t)$ vanish. On the other hand, as x^* is feasible for the MPEC, by $t_j > 0$ for all $j \in (I_1 \cap I_2)(x^*)$ it follows that $\Phi_j(x_1, x_2, t) < 0$ (confer (2.7)), such that for $j \in (I_1 \cap I_2)(x^*)$, the corresponding multiplier satisfies $\xi^*_j = 0$. Thus, as for the original problem, again both Lagrangian functions are equal.

Notice then, that by (2.44) the proof that $\bar{\mathcal{S}}(x^*, \lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2)$ is equal to the set of critical directions in x^* of (2.43) can be done in the same way as in the proof of Theorem 2.1. Hence we obtain the result. □

The corresponding Corollary has the form:

Corollary 2.2. *Let x^* be a strongly stationary point of (1.19) that satisfies the RNLP-SOSC. Then there exists an $\varepsilon > 0$, such that for every $t \in \mathbb{R}^p$ that satisfies $0 < t_j \leq \max(x^*_{1j}, x^*_{2j})$ for all $j \in \{1, \dots, p\} \backslash (I_1 \cap I_2)(x^*)$ and $t_j > 0$ for all $j \in (I_1 \cap I_2)(x^*)$, x^* is a strict local minimum in $\mathcal{B}_\varepsilon(x^*)$ of $R(t)$, which even satisfies the SOSC.*

We finish this section by considering the convergence results of Section 2.3. Replacing the scalar sequence $(t_k) \subseteq \mathbb{R}_0^+$ by a convergent sequence $(t^k) \subseteq (\mathbb{R}_0^+)^p$, the statements of Theorem 2.3 can directly be applied to problem (2.43). The only changes that have to be made in the proof of the theorem concerns the sequence (x^k). Considering (2.44) we have to show that there exists no subsequence with $|x_{1m_0}^k - x_{2m_0}^k|/t_{m_0}^k \to 1$ as $k \to \infty$ with $k \in \mathcal{K} \subseteq \mathbb{N}$. This, however, can be done in the same way as in the proof of Theorem 2.3.

Finally, the convergence result of Theorem 2.4 also still holds, if the scalar parameter t_k is replaced by a parameter vector t^k.

2.5 Comparison

As we mentioned in the beginning of this chapter, our new relaxation scheme can be interpreted as a combination of the exact bilinear reformulation of Fletcher et al. [FLRS06] and the relaxed bilinear reformulation proposed by Scholtes [Sch01]. We chose this regularization scheme as we think the relaxed bilinear reformulation is the representative of the variety of regularization and smoothing methods that is closest to our approach, because it relaxes the feasible set of original MPEC on the one hand, but on the other hand the original feasible set is still completely contained in the feasible set of the corresponding $NLP(t)$ (confer Section 1.4).

In this section, we therefore compare the theoretical results and properties we presented in the previous sections with the results of these two approaches. We begin with a comparison of our scheme to the approach using the exact bilinear reformulation (again confer Section 1.4).

The advantage of that approach is the equality of the set of the strongly stationary points of the original MPEC (1.19) and the set of the stationary points of the reformulated problem (1.33). This result is due to the fact that the feasible sets of both problems are equal and that the multipliers of strongly stationary points are nonnegative for the degenerate components.

It is therefore clear, that this result cannot exactly be established for our relaxation, since for any strictly positive parameter t the feasible set of (1.19) is in general a proper subset of the feasible set of $R(t)$. However, Lemma 2.7 and Lemma 2.9 provide comparable results.

The main differences to the corresponding result in [FLRS06] concerns either the admissible size of the parameter t or the feasibility assumption concerning a stationary point $x^*(t)$ of $R(t)$.

One main disadvantage of the direct approach of Fletcher et al. is the fact that the method can only be guaranteed to work close to strongly stationary points. This is due to the fact that B-stationary points, that are not strongly stationary, do not correspond to a stationary point of the reformulated problem (1.33). However, as we have seen in Section 1.3, if the MPEC-LICQ fails to hold, strong stationarity is not a necessary condition for a point x^* to be B-stationary. Hence, the exact bilinear reformulation is not a suitable approach to solve MPECs that do not have strongly stationary but B-stationary points. One such problem is Example 1.1 (here the complementarity constraint is reformulated using the exact bilinear reformulation):

$$\begin{array}{rll} \min & x_2 - x_1 & \\ \text{subject to} & x_1^2 + x_2^2 - 2x_2 \leq 0 & : \lambda \\ & 0 \leq \ x_1 \ \perp \ x_2 \ \geq 0, & : \nu_1, \nu_2, \xi \end{array}$$

The solution of this MPEC is $(x_1^*, x_2^*) = (0,0)$, which is a B-stationary point that is M-stationary, but not strongly stationary. First note, that $(0,0)$ is not a stationary point of the reformulated problem as

$$\begin{pmatrix} -1 \\ 1 \end{pmatrix} + \lambda^* \begin{pmatrix} 0 \\ -2 \end{pmatrix} - \begin{pmatrix} \nu_1^* \\ \nu_2^* \end{pmatrix} + \xi^* \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

cannot be satisfied by a multiplier vector $(\lambda^*,\ \nu_1^*,\ \nu_2^*,\ \xi^*) \geq 0$.

Moreover, even if the method produces feasible iterates that converge to the solution x^*, the corresponding KKT-residuals will not fall below a given tolerance unless the multiplier estimates of ξ^* are becoming very large. This demonstrates that the method might either fail or might become numerically unstable for problems such as Example 1.1 (compare also the numerical results at the end of Section 4.3).

The other method we want to compare our relaxation scheme with, is the relaxed bilinear reformulation. Comparing the main convergence results given in [Sch01] (see also Section 1.4) with the ones we proved in the previous section it becomes clear, that Theorem 1.4 and Theorem 2.3 are very similar. However, comparing Theorem 1.5 with Theorem 2.1 or further Corollary 2.1, it turns out that these two results given in Section 2.2 are stronger than Theorem 1.5. For our relaxation scheme the assumptions are weaker and the piecewise smooth function $x(t)$, for t being small enough, is identical to the strongly stationary point $\bar{x}$, that is $x(t) = \bar{x}$ for $t \in [0, \tau(\bar{x})]$. We exemplify this difference of the results by the following example (see also [Sch01]):

Example 2.3.

$$\begin{array}{rll} \textit{min} & \frac{1}{2}\left((x_1 - 1)^2 + (x_2 - 1)^2\right) & \\ \textit{subject to} & 0 \leq \ x_1 \ \perp \ x_2 \ \geq 0 & : \nu_1, \nu_2, \xi \end{array}$$

There exist two strongly stationary points $(1,0)$ *and* $(0,1)$ *with MPEC multipliers* $(0,-1)$ *and* $(-1,0)$*, respectively.*

We consider the strongly stationary point $(1,0)$*. If we reformulate the problem using the relaxed bilinear reformulation, then the stationary points* $x(t)$ *must satisfy*

$$\begin{pmatrix} x_1(t) - 1 \\ x_2(t) - 1 \end{pmatrix} - \begin{pmatrix} \nu_1(t) \\ \nu_2(t) \end{pmatrix} + \xi(t) \begin{pmatrix} x_2(t) \\ x_1(t) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{2.45}$$

where $(\nu_1(t), \nu_2(t), \xi(t))$ *are the corresponding multipliers for* $NLP(t)$ *that must additionally satisfy* $(\nu_1(t), \nu_2(t), \xi(t)) \geq 0$*. If we insert the strongly stationary point* $(1,0)$ *into (2.45), then we obtain*

$$\begin{pmatrix} 0 \\ -1 \end{pmatrix} - \begin{pmatrix} \nu_1(t) \\ \nu_2(t) \end{pmatrix} + \xi(t) \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{2.46}$$

Since for $t > 0$ the constraint $x_1(t)x_2(t) - t \leq 0$ is inactive for $(x_1(t), x_2(t)) = (1, 0)$, we have that $\xi(t) = 0$. Thus $(1, 0)$ is not a stationary point of the corresponding $NLP(t)$ for any $t > 0$, as (2.46) cannot be satisfied.

By condition (2.45) for $t < 1/4$ we can derive the smooth function $x : \mathbb{R}^+ \to \mathbb{R}^2$ of stationary points $x(t) = 1/2(1 + \sqrt{1-4t}, 1 - \sqrt{1-4t})$ that satisfy (2.45). Clearly $x(t)$ converges to $(1, 0)$ as t approaches zero.

In contrast to that $(1, 0)$ is a stationary point of the corresponding $R(t)$ for any $0 \leq t \leq 1$ as

$$\begin{pmatrix} 0 \\ -1 \end{pmatrix} - \begin{pmatrix} \nu_1(t) \\ \nu_2(t) \end{pmatrix} + \xi(t) \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

is satisfied by $\nu_1(t) = 0$, $\nu_2(t) = 0$ and $\xi(t) = 1/2$. Here $\xi(t) > 0$ is admissible, as $x_1(t) = 1 \geq t$ and hence $\Phi(x_1(t), x_2(t), t) \leq 0$ is active.

The illustrated example demonstrates that although the MPEC does have strongly stationary points that in addition satisfy strict complementarity, in contrast to our relaxation scheme, the regularization scheme discussed in [Sch01] and [RW04] might yield the solution only in the limit $t \to 0$.

Next, we want to consider MPECs that do not have B-stationary points that are also strongly stationary. A suitable example is

Example 2.4. *(*`scholtes4`*)*

$$\begin{array}{rlr} \textit{min} & x_1 + x_2 - x_0 & \\ \textit{subject to} & -4x_1 + x_0 \leq 0 & : \lambda_1 \\ & -4x_2 + x_0 \leq 0 & : \lambda_2 \\ & 0 \leq \ x_1 \ \perp \ x_2 \ \geq 0, & : \nu_1, \nu_2, \xi \end{array}$$

The solution of this MPEC is $(x_0^*, x_1^*, x_2^*) = (0, 0, 0)$ with MPEC multipliers either (λ_1^*, λ_2^*, $\hat{\nu}_1$, $\hat{\nu}_2$) $=(3/4, 1/4, -2, 0)$ or (λ_1^*, λ_2^*, $\hat{\nu}_1$, $\hat{\nu}_2$) $=(1/4, 3/4, 0, -2)$. The first equation of the KKT-conditions (see (1.16)) for the corresponding $NLP(t)$ corresponds to

$$\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} + \lambda_1(t) \begin{pmatrix} 1 \\ -4 \\ 0 \end{pmatrix} + \lambda_2(t) \begin{pmatrix} 1 \\ 0 \\ -4 \end{pmatrix} - \begin{pmatrix} 0 \\ \nu_1(t) \\ \nu_2(t) \end{pmatrix} + \xi(t) \begin{pmatrix} 0 \\ x_2(t) \\ x_1(t) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

which is satisfied by the smooth function $x(t) = \sqrt{t}(4, 1, 1)$. Furthermore, as t approaches zero, $x(t)$ converges to the M-stationary point $(0, 0, 0)$. However, if we consider the corresponding multiplier vector ($\lambda_1(t)$, $\lambda_2(t)$, $\nu_1(t)$, $\nu_2(t)$, $\xi(t)$) $= (1/2, 1/2, 0, 0, 1/\sqrt{t})$ then we notice that at the same time as t converges to zero and $x(t)$ approaches the solution, the multiplier $\xi(t)$ diverges.

The unboundedness of the corresponding multiplier vector, as illustrated by the previous example, might cause numerical difficulties. Furthermore, it is not a special feature of this example but an inevitable consequence, if the relaxed bilinear reformulation produces a sequence x^k that converges to a B-stationary point that is not strongly stationary.

Lemma 2.11. *Let* $(t_k) \subseteq \mathbb{R}^k$ *be a sequence that satisfies* $t_k \to 0$ *and let* $(x_k) \subseteq \mathbb{R}^{n+2p}$ *be a sequence of stationary points of* $NLP(t_k)$ *that converges to a point* $\bar{x}$ *that is not strongly stationary. Then any sequence of corresponding multipliers* $(\lambda^k,\ \mu^k,\ \nu_1^k,\ \nu_2^k,\ \xi^k)$ *is unbounded.*

Proof. Assume there exists a bounded sequence of multipliers $(\lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$ associated with the sequence of stationary points x^k. Then, $(\lambda^k,\ \mu^k,\ \nu_1^k,\ \nu_2^k,\ \xi^k)$ has a convergent subsequence

$$(\lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)_{\kappa \in \mathcal{K}} \longrightarrow (\tilde{\lambda}, \tilde{\mu}, \tilde{\nu}_1, \tilde{\nu}_2, \tilde{\xi})$$

with $k \in \mathcal{K} \subseteq \mathbb{N}$. As x^k is stationary and converges to $\bar{x}$ and, since f, g and h are assumed to be twice continuously differentiable,

$$\begin{aligned} 0 \;&=\; \lim_{\substack{\kappa \in \mathcal{K} \\ \kappa \to \infty}} \Bigg(f(x^k) - \sum_{j=1}^{m} \lambda_j^k \nabla g_j(x^k) \;-\; \sum_{i=1}^{q} \mu_i^k \nabla h_i(x^k) \\ &\qquad - \sum_{j=1}^{p} \nu_{1j}^k\, e_{1j} - \sum_{j=1}^{p} \nu_{2j}^k\, e_{2j} \\ &\qquad + \sum_{j=1}^{p} \xi_j^k x_{2j}^k\, e_{1j} + \sum_{j=1}^{p} \xi_j^k x_{1j}^k\, e_{2j} \Bigg) \\ &=\; f(\bar{x}) - \sum_{j=1}^{m} \tilde{\lambda}_j \nabla g_j(\bar{x}) \;-\; \sum_{i=1}^{q} \tilde{\mu}_i \nabla h_i(\bar{x}) \\ &\qquad - \sum_{j=1}^{p} (\tilde{\nu}_{1j} - \tilde{\xi}_j \bar{x}_{2j})\, e_{1j} - \sum_{j=1}^{p} (\tilde{\nu}_{2j} - \tilde{\xi}_j \bar{x}_{1j})\, e_{2j}\,. \end{aligned}$$

Hence, we can choose $(\tilde{\lambda},\ \tilde{\mu},\ \tilde{\gamma}_1,\ \tilde{\gamma}_2)$ with $\tilde{\gamma}_{1j} = (\tilde{\nu}_{1j} - \tilde{\xi}_j \bar{x}_{2j})$ and $\tilde{\gamma}_{2j} = (\tilde{\nu}_{2j} - \tilde{\xi}_j \bar{x}_{1j})$ for $j \in \{1, \ldots, p\}$ as a multiplier vector for $\bar{x}$. That $(\tilde{\lambda},\ \tilde{\mu},\ \tilde{\gamma}_1, \tilde{\gamma}_2)$ satisfies the complementarity condition and $\tilde{\lambda}$ satisfies the nonnegativity conditions follows by the corresponding conditions for the multipliers $(\lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$ and the relation of the active sets in x^k and in $\bar{x}$. Finally, consider the limit multipliers $\tilde{\gamma}_{1j}$ and $\tilde{\gamma}_{2j}$ for $j \in (I_1 \cap I_2)(\bar{x})$ and note that as for all these indices

$$\tilde{\gamma}_{1j} = \lim_{\substack{\kappa \in \mathcal{K} \\ \kappa \to \infty}} \nu_{1j}^k - \xi_j^k x_{2j}^k = \tilde{\nu}_{1j} \geq 0 \quad \text{and} \quad \tilde{\gamma}_{2j} = \lim_{\substack{\kappa \in \mathcal{K} \\ \kappa \to \infty}} \nu_{2j}^k - \xi_j^k x_{1j}^k = \tilde{\nu}_{2j} \geq 0\,.$$

Hence, we have found a multiplier vector such that $\bar{x}$ is a strongly stationary point in contradiction to our assumption. □

In contrast to the previous result, we can prove that for the new relaxation scheme the sequence of multipliers $(\lambda^k,\ \mu^k,\ \nu_1^k,\ \nu_2^k,\ \xi^k)$ of a sequence of stationary points of $R(t_k)$ is bounded, provided the family of constraint gradients corresponding to the strictly positive entries of $(\nu_1^k, \nu_2^k, \xi^k)$ are linearly independent and the sequences (λ^k) and (μ^k) are bounded.

Lemma 2.12. *Let $(t_k) \subseteq \mathbb{R}^+$ be a sequence that satisfies $t_k \to 0$ and let $(x^k) \subseteq \mathbb{R}^{n+2p}$ be a sequence of stationary points of $R(t_k)$ that converges to $\bar{x}$. Furthermore, let $(\lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$ be a corresponding sequence of multipliers of x^k and assume that the family*

$$\{e_{1j} \mid j \in supp(\nu_1^k)\} \cup \{e_{2j} \mid j \in supp(\nu_2^k)\} \cup \{-\alpha_j^k\, e_{1j} - (2-\alpha_j^k)\, e_{2j} \mid j \in supp(\xi^k)\} \tag{2.47}$$

is linear independent for each $k \in \mathbb{N}$. Finally, suppose that the the sequences (λ^k) and (μ^k) are bounded. Then the sequence $(\lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$ is also bounded.

Proof. As $x^k \to \bar{x}$ and f is assumed to be continuously differentiable, it holds

$$\lim_{k\to\infty} \frac{\partial f}{\partial x_{ij}}(x^k) = \frac{\partial f}{\partial x_{ij}}(\bar{x}) \tag{2.48}$$

for $i = 1, 2$ and for all $j \in \{1, \dots, p\}$. Furthermore, since each x^k is a stationary point of $R(t_k)$, we can represent $\nabla f(x^k)$ by the active constraint gradients and the corresponding multipliers $(\lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$. Hence,

$$\left|\frac{\partial f}{\partial x_{1j}}(x^k)\right| = \left|\left(\sum_{i=1}^{m} \lambda_i^k \nabla g_i(x^k) + \sum_{i=1}^{q} \mu_i^k \nabla h_i(x^k)\right)_{n+j} + (\nu_{1j}^k - \xi_j^k \alpha_j^k)\right|$$

for each $j \in \{1, \dots, p\}$. By the boundedness of (λ^k) and (μ^k) and the differentiability assumptions for g and h there exists an $M \in \mathbb{R}^+$ such that

$$\left|\frac{\partial f}{\partial x_{1j}}(x^k)\right| \geq -M + |\, \nu_{1j}^k - \xi_j^k \alpha_j^k \,| \tag{2.49}$$

and accordingly

$$\left|\frac{\partial f}{\partial x_{2j}}(x^k)\right| \geq -M + |(\nu_{2j}^k - \xi_j^k (2 - \alpha_j^k))| \tag{2.50}$$

for all $j \in \{1, \dots, p\}$ and all $k \in \mathbb{N}$.

Next, we prove that either $\nu_{1j}^k = 0$ and $\nu_{2j}^k = 0$ or $\xi_j^k = 0$ for every $j \in \{1, \dots, p\}$ and for every $k \in \mathbb{N}$.

First, suppose that $x_{1j}^k \geq t_k$, then it holds that $\nu_{1j}^k = 0$ and $j \in I_2(x^k) \cap I_\Phi(x^k, t_k)$, which implies $\alpha_j^k = 0$. Therefore, $(-\alpha_j^k\, e_{1j} - (2 - \alpha_j^k)\, e_{2j}) = -2e_{2j}$, such that by the linear independence of (2.47), it follows that either $\nu_{2j}^k = 0$ or $\xi_j^k = 0$.

Accordingly, if $x_{2j}^k \geq t_k$, then $\nu_{2j}^k = 0$ and, since in this case $\alpha_j^k = 2$, $(-\alpha_j^k\, e_{1j} - (2 - \alpha_j^k)\, e_{2j}) = -2e_{1j}$. Thus, by the linear independence of (2.47), it follows that either $\nu_{1j}^k = 0$ or $\xi_j^k = 0$.

Hence, in both cases ($x_{1j}^k \geq t_k$ or $x_{2j}^k \geq t_k$) either $\nu_{1j}^k = 0$ and $\nu_{2j}^k = 0$ or $\xi_j^k = 0$.

If $x_{1j}^k < t_k$ and $x_{2j}^k < t_k$, then by Lemma 2.6 it follows that at most either $j \in I_\Phi(x^k, t_k)$ or $j \in (I_1 \cup I_2)(x^k)$ but not both. Hence, again either $\nu_{1j}^k = 0$ and $\nu_{2j}^k = 0$ or $\xi_j^k = 0$.

Using this information we will now prove the boundedness of (ν_{1j}^k), (ν_{2j}^k) and (ξ_j^k).

Assume that (ν_{1j}^k) is not bounded. Then there exists a subsequence $(\nu_{1j}^k)_{k\in\mathcal{K}}$ with $\nu_{1j}^k > 0$ for all $k \in \mathcal{K}$ and $\nu_{1j}^k \to \infty$. As $\nu_{1j}^k > 0$ and by the previous considerations $\xi_j^k = 0$ for all $k \in \mathcal{K}$. However, by (2.49) it then follows

$$\lim_{\substack{k\in\mathcal{K}\\k\to\infty}} \left|\frac{\partial f}{\partial x_{1j}}(x^k)\right| \geq -M + \lim_{\substack{k\in\mathcal{K}\\k\to\infty}} |\nu_{1j}^k| = \infty\,,$$

which contradicts (2.48). Analogously we obtain a contradiction if (ν_{2j}^k) is assumed to be unbounded.

Now assume that (ξ_j^k) is unbounded. Then there exists again a subsequence $(\xi_j^k)_{k\in\mathcal{K}}$ satisfying $\xi_j^k > 0$ for all $k \in \mathcal{K}$ (such that $\nu_{1j}^k = 0$ and $\nu_{2j}^k = 0$ for all $k \in \mathcal{K}$) and $\xi_j^k \to \infty$. If there exists a $\delta > 0$, such that $\alpha_j^k > \delta$ holds for all $k \in \mathcal{K}$, then

$$\lim_{\substack{k\in\mathcal{K}\\k\to\infty}} \left|\frac{\partial f}{\partial x_{1j}}(x^k)\right| \geq -M + \lim_{\substack{k\in\mathcal{K}\\k\to\infty}} |-\xi_j^k\alpha_j^k\,| > -M + \delta \lim_{\substack{k\in\mathcal{K}\\k\to\infty}} \xi_j^k = \infty\,,$$

which again contradicts (2.48).

Finally, assume there does not exist a $\delta > 0$ such that $\alpha_j^k > \delta$ holds for all $k \in \mathcal{K}$. Then there exists, however, a subsequence $(\alpha_j^k)_{k\in\mathcal{L}}$ with $\alpha_j^k \to 0$ and $k \in \mathcal{L} \subseteq \mathcal{K}$. Since then $(2 - \alpha_j^k) \to 2$ for $k \in \mathcal{L} \subseteq \mathcal{K}$, it follows that

$$\lim_{\substack{k\in\mathcal{L}\\k\to\infty}} \left|\frac{\partial f}{\partial x_{2j}}(x^k)\right| \geq -M + \lim_{\substack{k\in\mathcal{L}\\k\to\infty}} |-\xi_j^k(2-\alpha_j^k)\,| = -M + 2 \lim_{\substack{k\in\mathcal{L}\\k\to\infty}} \xi_j^k = \infty\,,$$

which also contradicts (2.48), such that $(\lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$ has to be bounded. □

Next, consider Example 2.3 and Example 2.4 and note, that in both cases a sequence of stationary points of a sequence of problems $NLP(t_k)$ satisfies $x_1(t_k)x_2(t_k) - t_k = 0$ for all $k \in \mathbb{N}$. Hence, it is of interest to have an estimate of the maximum distance of points satisfying $x_1(t_k)x_2(t_k) - t_k = 0$ to points satisfying the complementarity condition.

To derive such an estimate, we first transform the original coordinate system by rotating the coordinate system by 45° and obtain new coordinates u and v with:

$$u(x_1, x_2) = x_1 - x_2 \qquad \text{and} \qquad v(x_1, x_2) = x_1 + x_2\,.$$

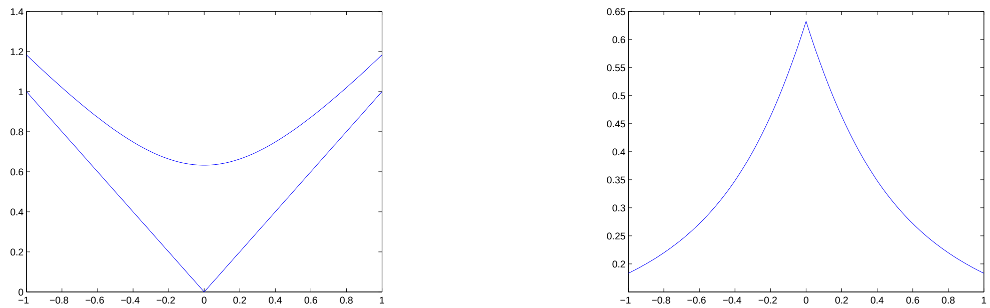
The inverse transformation is given by

$$x_1(u, v) = \frac{1}{2}(u + v) \qquad \text{and} \qquad x_2(u, v) = \frac{1}{2}(v - u)\,.$$

Next, we describe the curve $s(x_1, x_2) := x_1x_2 - t = 0$ in terms of u and v

$$\tilde{s}(u, v) := s(x_1(u, v), x_2(u, v)) = \frac{1}{4}(u + v)(v - u) - t = \frac{1}{4}(v^2 - u^2) - t = 0\,.$$

As implicit function $\tilde{v}(u)$, such that $\tilde{s}(u, \tilde{v}(u)) = 0$ holds, we get $\tilde{v}(u) = \sqrt{u^2 - 4t}$. To evaluate an upper bound for the maximum distance of points (x_1, x_2) with $x_1x_2 - t = 0$ to



(a) $\sqrt{u^2+4t}$ and $|u|$

(b) $h(u) = \sqrt{u^2+4t} - |u|$

Figure 2.3: Functions $\sqrt{u^2+4t}$, $|u|$ and $h(u)$ for $u \in [-1, 1]$ and $t = 0.1$

points that satisfy $x_1 x_2 = 0$, we consider the distance of points $(u, |u|)$ to points $(u, v) = \Pi(u, |u|)$, where $\Pi(u, |u|)$ denotes the projection of $(u, |u|)$ onto the curve $(u, \tilde{v}(u))$. This distance is bounded by $|h(u)|$ for all $u \in \mathbb{R}$, where $h(u) := \tilde{v}(u) - |u|$. Figure 2.3 illustrates the curves $(u, |u|)$, $(u, \tilde{v}(u))$ and $(u, h(u))$, respectively for $u \in [-1, 1]$ and $t = 0.1$.

Since

$$h(u) = \tilde{v}(u) - |u| = \sqrt{u^2 - 4t} - |u| < 2\sqrt{t} \qquad \text{for all} \quad u \in \mathbb{R}\backslash\{0\}$$

and $h(0) = 2\sqrt{t}$, it follows that the maximum distance of points $(u, |u|)$ to points $(u, v) = \Pi(u, |u|)$ is bounded above by $2\sqrt{t}$. Moreover, because this upper bound is reached for $(u, |u|) = (0, 0)$, it cannot be tightened any further.

To obtain a corresponding estimate for our relaxation scheme, we consider the curve

$$\hat{s}(u, v) := \Phi(x_1(u, v), x_2(u, v), t) = v - t\theta\left(\frac{u}{t}\right) = 0\,.$$

for $u \in [-t, t]$. Hence, $\hat{s}(u, \hat{v}(u)) = 0$ holds for $\hat{v}(u) = t\theta\left(\frac{u}{t}\right)$ and $u \in [-t, t]$.

To evaluate the maximum distance of points with $\Phi(x_1, x_2, t) = 0$ to points that satisfy $x_1 x_2 = 0$, we only have to consider the distance of points $(u, |u|)$ to points $(u, v) = \Pi(u, |u|)$, where $\Pi(u, |u|)$ now denotes the projection of $(u, |u|)$ onto the curve $(u, \hat{v}(u))$ with $u \in [-t, t]$. Again similar as for the relaxed bilinear reformulation this distance is bounded by $|g(u)|$ for all $u \in [-t, t]$, where $g(u) := \hat{v}(u) - |u|$. Figure 2.4 illustrates the curves $(u, |u|)$, $(u, \hat{v}(u))$ and $(u, h(u))$, respectively for $u \in [-1, 1]$ and $t = 1$.

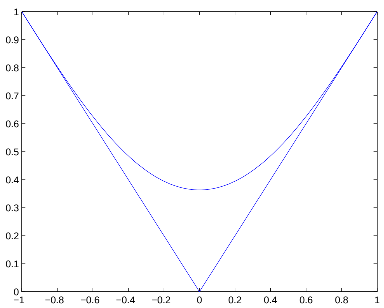
Applying Assumptions 2.1 we get that $g(u)$ must also have its maximum in $u = 0$. As $g(0) = t\theta(0)$ the distance directly depends on the function θ. For the two examples of θ we proposed in Section 2.1, we get

$$t\theta_s(0) = \frac{2t}{\pi}\sin\left(\frac{3\pi}{2}\right) + t = \left(-\frac{2}{\pi} + 1\right)t < \frac{1}{2}t$$

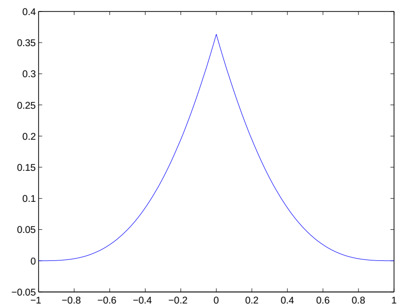
and

$$t\theta_p(0) = \frac{3}{8}t\,.$$

To obtain the maximum distance in the original coordinate system, we have to scale these values by $1/2\sqrt{2}$. However, this scaling does not alter the order. Hence, the maximum



(a) $\theta_s(u)$ and $|u|$



(b) $\theta_s(u) - |u|$

Figure 2.4: Functions $\theta_s(u)$, $|u|$ and $g(u)$ for $u \in [-1, 1]$ and $t = 1$

distance estimate of the relaxed bilinear approach is of order $O(\sqrt{t})$, whereas for our relaxation scheme we obtain an estimate of order $O(t)$. Again, consider Example 2.4 and note that for the stationary points $x^*(t)$ of $NLP(t)$ we get the rate of convergence $\|x^*(t) - x^*\| = \sqrt{2}\sqrt{t} = O(\sqrt{t})$, whereas for the stationary points $x^*(t)$ of $R(t)$ we have $\|x^*(t) - x^*\| = 1/2\sqrt{2}\,\theta(0)\,t = O(t)$.

Finally, we consider the order of convergence depending on t of the product $x_1(t)x_2(t)$. Since $x_1(t)x_2(t) \leq t$ for the relaxed bilinear approach, the convergence of $x_1(t)x_2(t)$ is of order $O(t)$.

For our relaxation scheme we have by contrast for pairs $(x_1(t), x_2(t))$ with $|x_1(t) - x_2(t)| < t$, that is $u(x_1(t), x_2(t)) \in (-t, t)$,

$$\begin{aligned} x_1(t)x_2(t) &\leq x_1(u, \tilde{v}(u))x_2(u, \tilde{v}(u)) = \frac{1}{4}\left(t^2\theta^2\left(\frac{u}{t}\right) - u^2\right) \\ &= \frac{1}{4}t^2\left(\theta^2\left(\frac{u}{t}\right) - \left(\frac{u}{t}\right)^2\right) \leq \frac{1}{4}\,t^2\,. \end{aligned}$$

Hence, we get the order $O(t^2)$.

3 New Relaxation in Combination with NLP Methods

3.1 SQP Methods

In this chapter we consider the behaviour of two important methods to solve nonlinear programs, namely Sequential Quadratic Programming (SQP) and Interior Point Methods (IPM), if we combine them with the new relaxation scheme we introduced in the previous chapter.

In the first section we are concerned with the local convergence properties of a general SQP method applied to problem $R(t)$. We will show that a general local SQP method applied to the relaxed problem $R(t)$ will generate a quadratically convergent sequence of iterates that converges to a strongly stationary point x^* of the original MPEC under suitable assumptions.

Afterwards, in the second section, we consider a modified two-sided relaxation scheme, which forms a combination of the relaxation we discussed in Chapter 2 with the two-sided relaxation scheme proposed by DeMiguel et al. in [DFNS05] (see also Section 1.4).

3.1.1 A Brief Introduction to SQP Methods

Sequential Quadratic Programming methods belong to the most important algorithms to solve NLPs. As the name already reveals, these methods are based on the successive solution of quadratic programs (QP). These QPs form an approximation of the NLP in a current iterate x^k. In this subsection we give a brief sketch of a general SQP algorithm, explain some of the main ideas behind it and finally mention some convergence results, which we will need in the subsequent convergence analysis of the SQP method applied to our problem $R(t)$.

Since we want to introduce a general SQP method, in this section we consider a standard NLP of the form (1.13). The quadratic programs that are successively solved by an SQP algorithm for (1.13) are of the form

$$\begin{aligned} \min_d \quad & q(d) = \nabla f(x^k)^T d + \frac{1}{2} d^T H^k d \\ \text{subject to} \quad & h(x^k) + \nabla h(x^k)^T d = 0 \\ & g(x^k) + \nabla g(x^k)^T d \geq 0 \,, \end{aligned} \tag{3.1}$$

where H^k denotes the Hessian matrix $\nabla^2_{xx}\mathcal{L}(x^k, \lambda^k, \mu^k)$ (confer (1.17)), x^k denotes the current iterate and λ^k and μ^k denote the current multiplier estimates. The constraints of this QP are the linearizations of the (nonlinear) constraints of (1.13). The objective function nearly corresponds to a second order Taylor approximation of the objective function of (1.13) in x^k. However, the constant term $f(x^k)$ is left out and instead of the Hessian matrix of f in x^k the Hessian matrix of $\mathcal{L}(x^k, \lambda^k, \mu^k)$ with respect to x is used. The application of this alternative approximation is motivated by the Lagrange-Newton iteration for nonlinear programs with equality constraints (see Section 5.5.2 in [GK02] for further details). Similar to Quasi-Newton methods, the exact Hessian matrix $\nabla^2_{xx}\mathcal{L}(x^k, \lambda^k, \mu^k)$ is often substituted by a suitable symmetric approximation of it.

In each outer iteration of the SQP method the solution of such a QP yield a new search direction d and associated multipliers λ_{qp} and μ_{qp}. A new solution estimate is then obtained by setting $x^{k+1} = x^k + d$ and taking λ_{qp} and μ_{qp} as new multiplier estimates λ^{k+1} and μ^{k+1}, respectively. Hence, a general local SQP algorithm is of the form:

Algorithm 3.1: Local SQP Algorithm

Choose initial values for $x^0 \in \mathbb{R}^n$, $\lambda^0 \in \mathbb{R}^m$ *and* $\mu^0 \in \mathbb{R}^q$.

repeat

- *Compute local minimizer* d^k *of (3.1) that is closest to the origin and compute associated multipliers* λ^k_{qp} *and* μ^k_{qp}.
- *Update the iterate and the multipliers:*

$$x^{k+1} \leftarrow x^k + d^k$$
$$\lambda^{k+1} \leftarrow \lambda^k_{qp}$$
$$\mu^{k+1} \leftarrow \mu^k_{qp}$$

- $k \leftarrow k + 1$

until (x^k, λ^k, μ^k) satisfies the KKT conditions.

In general, the quadratic subprograms are solved either by an active set strategy or by Interior Point Methods. Since for the local SQP algorithm it can only be guaranteed to yield a convergent sequence of iterates, if the initial point (x^0, λ^0, μ^0) is close enough to a solution of the NLP (see Theorem 3.1), most SQP algorithms incorporate a globalization strategy. The most popular approaches to promote global convergence of an SQP algorithm concern linesearch methods applied to a suitable penalty or merit function, trust-region approaches or most recently filter methods (see chapter 15 in [CGT00]).

A main local convergence result for SQP methods (Theorem 15.2.2 in [CGT00]) is:

Theorem 3.1. *Suppose that the second derivatives of f, g and h exist and are Lipschitz continuous in some neighbourhood Ω of a stationary point x^* of (1.13) with multipliers λ^* and μ^*. Assume that the LICQ and the SOSC hold in (x^*, λ^*, μ^*) and furthermore that $\lambda^*_j \neq 0$ for all $j \in I_g(x^*)$. Then the following holds:*

1. *Consider any sequence (λ^k, μ^k) converging to (λ^*, μ^*). Then there exists a neighbourhood $\mathcal{X} \subset \Omega$ of x^* for which the sequence (x^k) generated by Algorithm 3.1 converges q-superlinearly to x^* from any starting point $x_0 \in \mathcal{X}$. Furthermore, if*

$$\|(\lambda^k, \mu^k) - (\lambda^*, \mu^*)\| = O(\|x^k - x^*\|),$$

then the convergence is q-quadratic.

2. *Let (x^k) and (d^k) be the sequences generated by Algorithm 3.1 and let $(\lambda^{k+1}, \mu^{k+1})$ be the Lagrange multipliers associated with d^k. Then there is a neighbourhood $\mathcal{X} \subset \Omega$ of x^* and another neighbourhood $\mathcal{Y}$ of (λ^*, μ^*) for which the sequence $((x^k, \lambda^k, \mu^k))$ converges q-quadratically to (x^*, λ^*, μ^*) from any starting point $((x^0, \lambda^0, \mu^0)) \in \mathcal{X} \times \mathcal{Y}$.*
3. *In either case, the set of constraints that are active at x^* are precisely those that are active for the quadratic subproblem (3.1) at d^k for large enough k.*

3.1.2 Local Convergence for R(t)

In this section we will prove that Algorithm 3.1 applied to $R(t)$ will generate a sequence of iterates x^k that converges to a strongly stationary point x^*, if the starting point is close to x^*.

However, we know by Lemma 2.6 that the LICQ in general fails to hold near x^* for $t \in (0, \tau(x^*))$. Hence, although we have proven in the previous chapter that a strongly stationary point x^* is also a strict local solution of $R(t)$, provided the MPEC-LICQ and the RNLP-SOSC hold in x^*, Theorem 3.1 is not directly applicable to $R(t)$.

To prove the convergence behaviour of the SQP algorithm applied to $R(t)$, we will therefore make use of an auxiliary problem $P(t)$. This problem nearly corresponds to $R(t)$, though in contrast to $R(t)$, it satisfies the assumptions of Theorem 3.1 under suitable conditions. For this reason, the local SQP algorithm applied to $P(t)$ generates a convergent sequence of iterates. We will show that Algorithm 3.1 applied to $R(t)$ generates the same sequence of solutions d^k of the corresponding QPs. This implies that the sequences of iterates for $P(t)$ and $R(t)$ must coincide, thus the convergence behaviour can be transferred to $R(t)$.

A helpful Lemma concerning the representation of the active sets of a current iterate x^k that is close to a strongly stationary point x^* is:

Lemma 3.1. *Let x^* be a strongly stationary point of (1.19) and assume that $t \in (0, \tau(x^*))$ ($\tau(x^*)$ defined as in Definition 2.1). Then there exists an $\varepsilon > 0$ such that for any $x \in \mathcal{Z}(t) \cap \mathcal{B}_\varepsilon(x^*)$ it holds*

$$(I_1 \backslash I_2)(x^*) = I_1(x) \cap I_\Phi(x, t)$$

and

$$(I_2 \backslash I_1)(x^*) = I_2(x) \cap I_\Phi(x, t)\,.$$

Proof. Consider $j \in (I_1 \backslash I_2)(x^*)$. Then, since $t \in (0, \tau(x^*))$, we have $t < x^*_{2j}$. We can therefore find an $\varepsilon > 0$ such that if $x \in \mathcal{B}_\varepsilon(x^*)$, then $t < x_{2j}$. By the feasibility of x for $R(t)$ it follows that $j \in I_1(x) \cap I_\Phi(x, t)$. Hence,

$$(I_1 \backslash I_2)(x^*) \subseteq I_1(x) \cap I_\Phi(x, t)\,. \tag{3.2}$$

Analogously it can be proved that

$$(I_2 \backslash I_1)(x^*) \subseteq I_2(x) \cap I_\Phi(x, t)\,. \tag{3.3}$$

For the inverse direction we assume that there exists an index $j_0 \in I_1(x) \cap I_\Phi(x, t)$ with $j_0 \notin (I_1 \backslash I_2)(x^*)$. Then either $j_0 \in (I_1 \cap I_2)(x^*)$ or $j_0 \in (I_2 \backslash I_1)(x^*)$. If $j_0 \in (I_1 \cap I_2)(x^*)$, then by $x \in \mathcal{Z}(t) \cap \mathcal{B}_\varepsilon(x^*)$ it follows that $x_{1j_0} < t$ and $x_{2j_0} < t$, which by Lemma 2.6 (2) contradicts $j_0 \in I_1(x) \cap I_\Phi(x,t)$. If $j_0 \in (I_2 \backslash I_1)(x^*)$ then by (3.3) it follows that $j_0 \in (I_2(x) \cap I_\Phi(x,t)) \cap (I_1(x) \cap I_\Phi(x,t))$. This however contradicts Lemma 2.6 (2). Thus, if x is feasible for $R(t)$ and close enough to x^* then $(I_1 \backslash I_2)(x^*) = I_1(x) \cap I_\Phi(x,t)$. Accordingly we can prove that $(I_2 \backslash I_1)(x^*) = I_2(x) \cap I_\Phi(x,t)$. □

The next Lemma illustrates and extends Lemma 2.6. It states that the LICQ is inherently not satisfied for $R(t)$ close to a strongly stationary point x^*, if there exists at least one nondegenerate complementarity pair and explains the reason for it.

Lemma 3.2. *Let x^* be a strongly stationary point of (1.19) and let $t \in (0, \tau(x^*))$. Then there exists an $\varepsilon > 0$ such that for any $x \in \mathcal{Z}(t) \cap \mathcal{B}_\varepsilon(x^*)$*

1. *if $(I_1 \cap I_2)(x^*) \neq \{1, \dots, p\}$ then the family of active constraint gradients is linearly dependent.*
2. *If x^* satisfies the MPEC-LICQ, then the linear dependence of the family of active constraint gradients is based on the existence of a pair of positive linear dependent vectors for each $j \in \{1, \dots, p\} \backslash (I_1 \cap I_2)(x^*)$:*
$$\begin{array}{llll} e_{1j} & \textit{and} & -2e_{1j} & \textit{for} \quad j \in (I_1 \backslash I_2)(x^*) \\ e_{2j} & \textit{and} & -2e_{2j} & \textit{for} \quad j \in (I_2 \backslash I_1)(x^*). \end{array}$$
Thus, if only one representative of each such pair is chosen, then the resulting family is linear independent.

Proof. First by Lemma 3.1 we can find an $\varepsilon > 0$, such that $(I_1 \backslash I_2)(x^*) = I_1(x) \cap I_\Phi(x,t)$ and $(I_2 \backslash I_1)(x^*) = I_2(x) \cap I_\Phi(x,t)$ holds for all $x \in \mathcal{Z}(t) \cap \mathcal{B}_\varepsilon(x^*)$. Furthermore, taking into account the special form of $\nabla_x \Phi(x_1, x_2, t)$, we obtain as family of active constraint gradients for such x

$$\begin{array}{rl} \nabla h_j(x) & j \in \{1, \dots, q\} \\ \nabla g_j(x) & j \in I_g(x) \\ e_{1j} \quad \text{and} \quad -2e_{1j} & j \in (I_1 \backslash I_2)(x^*) \\ e_{2j} \quad \text{and} \quad -2e_{2j} & j \in (I_2 \backslash I_1)(x^*) \\ e_{1j} & j \in (I_1 \cap I_2)(x^*) \cap I_1(x) \\ e_{2j} & j \in (I_1 \cap I_2)(x^*) \cap I_2(x) \\ -\alpha_j\, e_{1j} - (2-\alpha_j)\, e_{2j} & j \in (I_1 \cap I_2)(x^*) \cap I_\Phi(x,t)\,, \end{array} \tag{3.4}$$

which is clearly linear dependent if $(I_1 \cap I_2)(x^*) \neq \{1, \dots, p\}$. By the continuity of g we have $I_g(x) \subseteq I_g(x^*)$ for all x that are close enough to x^*. Thus, if the MPEC-LICQ is satisfied in x^*, then the family

$$\begin{array}{rl} \nabla h_j(x^*) & j \in \{1, \dots, q\} \\ \nabla g_j(x^*) & j \in I_g(x) \\ e_{1j} & j \in (I_1 \backslash I_2)(x^*) \\ e_{2j} & j \in (I_2 \backslash I_1)(x^*) \\ e_{1j} & j \in (I_1 \cap I_2)(x^*) \\ e_{2j} & j \in (I_1 \cap I_2)(x^*)\,. \end{array} \tag{3.5}$$

is linear independent.

Consider a matrix A whose column vectors are the gradient vectors of (3.5). Then A has full column rank and $A^T A$ is invertible. Replacing the columns $\nabla h_j(x^*)$ by $\nabla h_j(x)$ for all $j \in \{1, \ldots, q\}$ and $\nabla g_j(x^*)$ by $\nabla g_j(x)$ for all $j \in I_g(x)$, we obtain a slightly perturbed matrix $\bar{A}$, which is close to A. Therefore, by the perturbation lemma (see Lemma 5.23 in [GK02]), we deduce that $\bar{A}^T \bar{A}$ is also invertible and it follows that $\bar{A}$ must have full column rank.

Finally, consider again (3.4) and note that by Lemma 2.6 at most either $j \in I_\Phi(x,t)$ or $j \in (I_1 \cup I_2)(x)$ but not both, if $x_{1j} < t$ and $x_{2j} < t$. Hence, if we choose ε small enough, then for all $j \in (I_2 \cap I_1)(x^*)$ it holds $I_\Phi(x,t) \cap (I_1 \cup I_2)(x) = \emptyset$. Therefore, if one chooses only one representative of each pair

$$\begin{array}{llll} e_{1j} & \text{and} & -2e_{1j} & \text{for} \quad j \in (I_1 \backslash I_2)(x^*) \\ e_{2j} & \text{and} & -2e_{2j} & \text{for} \quad j \in (I_2 \backslash I_1)(x^*), \end{array}$$

then the resulting family

$$\begin{array}{rrrrl} & & & \nabla h_j(x) & j \in \{1, \ldots, q\} \\ & & & \nabla g_j(x) & j \in I_g(x) \\ \text{either} & e_{1j} & \text{or} & -2e_{1j} & j \in (I_1 \backslash I_2)(x^*) \\ \text{either} & e_{2j} & \text{or} & -2e_{2j} & j \in (I_2 \backslash I_1)(x^*) \\ & & & e_{1j} & j \in (I_1 \cap I_2)(x^*) \cap I_1(x) \\ & & & e_{2j} & j \in (I_1 \cap I_2)(x^*) \cap I_2(x) \\ & & & -\alpha_j e_{1j} - (2-\alpha_j) e_{2j} & j \in (I_1 \cap I_2)(x^*) \cap I_\Phi(x,t) \end{array} \tag{3.6}$$

partly corresponds to a subset of the columns of $\bar{A}$. Only for indices $j \in (I_2 \cap I_1)(x^*) \cap I_\Phi(x,t)$ the two linear independent vectors e_{1j} and e_{2j} have been replaced by a linear combination of both of them, namely $-\alpha_j e_{1j} - (2-\alpha_j) e_{2j}$. However, as this does not disturb the linear independence, the resulting family of active constraint gradients is linear independent. □

The results of Lemma 3.2 imply that, although we have shown that a strongly stationary point x^* is also a stationary point of $R(t)$ if $t \in (0, \tau(x^*))$ (see Section 2.2), we cannot directly apply the Convergence Theorem 3.1 to problem $R(t)$ in general. However, it gives us a hint how to handle the absence of the LICQ in x^* for $R(t)$.

Consider the auxiliary problem

$$\begin{array}{lrrcll} P(t) & \min & f(x) & & & \\ & \text{subject to} & h(x) & = & 0 & \\ & & g(x) & \geq & 0 & \\ & & x_{1j} = 0,\ x_{2j} & \geq & 0 & j \in (I_1 \backslash I_2)(x^*) \\ & & x_{2j} = 0,\ x_{1j} & \geq & 0 & j \in (I_2 \backslash I_1)(x^*) \\ & & x_{1j} \geq 0,\ x_{2j} & \geq & 0 & j \in (I_1 \cap I_2)(x^*) \\ & & \Phi_j(x_1, x_2, t) & \leq & 0 & j \in (I_1 \cap I_2)(x^*). \end{array}$$

In the following, we will prove that if the MPEC-LICQ and the RNLP-SOSC are satisfied in a strongly stationary point x^*, then the SQP algorithm 3.1 applied to $P(t)$ will yield a locally convergent sequence (x^k).

First, we prove that each strongly stationary point x^* of (1.19) is a stationary point of $P(t)$ (determined for this strongly stationary point x^*).

Lemma 3.3. *Let x^* be a strongly stationary point of (1.19) with multipliers $\lambda^*, \mu^*, \hat{\nu}_1$ and $\hat{\nu}_2$. Assume further that $t \in (0, \tau(x^*))$, then x^* is a stationary point for $P(t)$ (associated with x^*) with multipliers λ^*, μ^*, $\tilde{\nu}_1 := \hat{\nu}_1$, $\tilde{\nu}_2 := \hat{\nu}_2$, and $\tilde{\xi} := 0$.*

Proof. As x^* is feasible for (1.19) it is clearly also feasible for $P(t)$. For $P(t)$ the multipliers $\tilde{\nu}_{1j}$ for $j \in (I_1 \backslash I_2)(x^*)$ as well as the multipliers $\tilde{\nu}_{2j}$ for $j \in (I_2 \backslash I_1)(x^*)$ are allowed to be of either sign. Moreover, $\hat{\nu}_{1j} \geq 0$ as well as $\hat{\nu}_{2j} \geq 0$ for $j \in (I_1 \cap I_2)(x^*)$. Hence, the multipliers λ^*, μ^*, $\tilde{\nu}_1 = \hat{\nu}_1$ and $\tilde{\nu}_2 = \hat{\nu}_2$ and $\tilde{\xi} = 0$ satisfy the KKT-conditions for $P(t)$ and x^* is a stationary point of $P(t)$. □

Next, we show, that the LICQ as well as the second order sufficient condition can be transferred to $P(t)$.

Lemma 3.4. *Suppose x^* is a strongly stationary point of (1.19) and assume the MPEC-LICQ holds in x^*. Furthermore, assume that $t \in (0, \tau(x^*))$. Then there exists an $\varepsilon > 0$ such that the LICQ holds in x for all x that are feasible for $P(t)$ and satisfy $x \in \mathcal{B}_\varepsilon(x^*)$.*

Proof. If $t \in (0, \tau(x^*))$ and x is close enough to x^*, then the family of active constraint gradients for $P(t)$ in x is

$$\begin{array}{rl} \nabla h_j(x) & j \in \{1, \ldots, q\} \\ \nabla g_j(x) & j \in I_g(x) \\ e_{1j} & j \in (I_1 \backslash I_2)(x^*) \\ e_{2j} & j \in (I_2 \backslash I_1)(x^*) \\ e_{1j} & j \in (I_1 \cap I_2)(x^*) \cap I_1(x) \\ e_{2j} & j \in (I_1 \cap I_2)(x^*) \cap I_2(x) \\ -\alpha_j \, e_{1j} - (2 - \alpha_j) \, e_{2j} & j \in (I_1 \cap I_2)(x^*) \cap I_\Phi(x, t) \, . \end{array} \tag{3.7}$$

The linear independence of (3.7) now follows by Lemma 3.2 (confer 3.5). □

Lemma 3.5. *Let x^* be a strongly stationary point of (1.19) and let $t \in (0, \tau(x^*))$. Furthermore, assume that the RNLP-SOSC holds in x^*. Then the SOSC for $P(t)$ is also satisfied in x^*.*

Proof. Since x^* is a strongly stationary point of (1.19), by Lemma 3.3 it is a stationary point of $P(t)$ with multipliers that $\tilde{\nu}_1 := \hat{\nu}_1$, $\tilde{\nu}_2 := \hat{\nu}_2$, and $\tilde{\xi} := 0$. Moreover, all second partial derivatives of $\Phi_j(x_1, x_2, t) \leq 0$ with respect to x vanish except for the indices $j \in \{1, \ldots, p\}$ with $|x_{1j} - x_{2j}| < t$, thus except for indices $j \in (I_1 \cap I_2)(x^*)$. Hence,

$$\nabla^2_{xx} \mathcal{L}_{P(t)}(x^*, \lambda^*, \mu^*, \tilde{\nu}_1, \tilde{\nu}_2, \tilde{\xi}) = \nabla^2_{xx} \mathcal{L}_{MPEC}(x^*, \lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2) \, . \tag{3.8}$$

Comparing the set of critical directions $\mathcal{S}_{P(t)}(x^*,\ \lambda^*,\ \mu^*,\ \tilde{\nu}_1,\ \tilde{\nu}_2,\ \tilde{\xi}, t)$ of $P(t)$ with the corresponding set $\mathcal{S}_{R(t)}(x^*,\ \lambda^*,\ \mu^*,\ \nu_1^*,\ \nu_2^*,\ \xi^*,\ t)$ of $R(t)$, we note that the two sets differ only in the conditions for $j \in (I_1 \backslash I_2)(x^*)$ and $j \in (I_2 \backslash I_1)(x^*)$, respectively. However, for indices $j \in (I_1 \backslash I_2)(x^*)$ we have $s_{1j} = 0$ for all $s \in \mathcal{S}_{P(t)}(x^*,\ \lambda^*,\ \mu^*,\ \tilde{\nu}_1,\ \tilde{\nu}_2,\ \tilde{\xi}, t)$, as well as we have $s_{2j} = 0$ for all $s \in \mathcal{S}_{P(t)}(x^*,\ \lambda^*,\ \mu^*,\ \tilde{\nu}_1,\ \tilde{\nu}_2,\ \tilde{\xi}, t)$ for indices $j \in (I_2 \backslash I_1)(x^*)$. Hence, $\mathcal{S}_{P(t)}(x^*,\ \lambda^*,\ \mu^*,\ \tilde{\nu}_1,\ \tilde{\nu}_2,\ \tilde{\xi}, t) \subseteq \mathcal{S}_{R(t)}(x^*,\ \lambda^*,\ \mu^*,\ \nu_1^*,\ \nu_2^*,\ \xi^*,\ t) = \bar{\mathcal{S}}(x^*, \lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2)$ (confer Theorem 2.1). The RNLP-SOSC in x^* therefore implies that the SOSC does also hold in x^* for problem $P(t)$. □

Now, if we assume that x^* is a strongly stationary point of (1.19) at which the MPEC-LICQ and the RNLP-SOSC are satisfied and if $t \in (0, \tau(x^*))$, then problem $P(t)$ satisfies almost all the conditions that are necessary to guarantee the fast local convergence of Algorithm 3.1, if applied to $P(t)$. However, to prove the local convergence we first need the following result concerning the Lipschitz-continuity of the second derivative of Φ_j.

Lemma 3.6. *Let $\Phi_j(x_1, x_2, t)$ be defined as in Section 2.1 and let $t > 0$. Furthermore, suppose θ satisfies Assumptions 2.1 and let θ'' be Lipschitz-continuous on $\mathcal{D} \supsetneq [-1, 1]$. Then $\nabla^2_{xx}\Phi_j(x_1, x_2, t)$ is Lipschitz-continuous on $\mathbb{R}^{n+2p}$.*

Proof. First we proof that if each partial derivative $\partial^2\Phi_\ell(x_1, x_2, t)/\partial x_i \partial x_j$ is Lipschitz-continuous on $\mathbb{R}^{n+2p}$, then this is also true for $\nabla^2_{xx}\Phi_\ell(x_1, x_2, t)$.

Consider a map $F : \mathbb{R}^n \to \mathbb{R}^{m \times n}$, with component functions $f_{ij}(x)$ then

$$F(x) = \sum_{i=1}^{n} \sum_{j=1}^{m} f_{ij}(x) E_{ij}\,,$$

where E_{ij} denotes the matrix that has all zero entries except for the ijth entry which is set equal to one. Now, if we assume that $f_{ij}(x)$ is Lipschitz-continuous on $\mathbb{R}^n$ with Lipschitz constant L_{ij} for all $i \in \{1, \ldots, n\}$ and all $j \in \{1, \ldots, m\}$, then there exists a constant L, such that for all $x, y \in \mathbb{R}^n$

$$\begin{aligned}
\|F(x) - F(y)\| &= \left\| \sum_{i=1}^{n} \sum_{j=1}^{m} \left(f_{ij}(x) - f_{ij}(y) \right) E_{ij} \right\| \\
&\leq \sum_{i=1}^{n} \sum_{j=1}^{m} |\, f_{ij}(x) - f_{ij}(y)\, |\, \|E_{ij}\| \\
&\leq \sum_{i=1}^{n} \sum_{j=1}^{m} L_{ij}\, \|x - y\|\, \|E_{ij}\| \\
&= \|x - y\| \sum_{i=1}^{n} \sum_{j=1}^{m} L_{ij} \|E_{ij}\| = L\|x - y\|\,,
\end{aligned}$$

where different norms can be handled by the equivalence of norms of $\mathbb{R}^n$ (see for example [Kos93]). Hence, F is Lipschitz-continuous.

Next we consider the second partial derivatives $\partial^2\Phi_\ell(x_1,x_2,t)/\partial x_i\partial x_j$ and prove, that they are Lipschitz-continuous for all $x, y \in \mathbb{R}^{n+2p}$. Consider the definition of $\Phi_\ell(x_1,x_2,t)$ and note that the second derivatives of $\Phi_\ell(x_1,x_2,t)$ with respect to x all vanish except for the case that $|x_{1\ell}-x_{2\ell}|<t$. Hence, assume that at least either $|x_{1\ell}-x_{2\ell}|<t$ or $|y_{1\ell}-y_{2\ell}|<t$. Furthermore, the only nonvanishing components of $\partial^2\Phi_\ell(x_1,x_2,t)/\partial x_i\partial x_j$ for such x with $|x_{1\ell}-x_{2\ell}|<t$ are

$$\left|\frac{\partial^2\Phi_\ell}{\partial x_{1\ell}^2}\right| = \left|\frac{\partial^2\Phi_\ell}{\partial x_{2\ell}^2}\right| = \left|\frac{\partial^2\Phi_\ell}{\partial x_{1\ell}\partial x_{2\ell}}\right| = \frac{1}{t}\,\theta''\left(\frac{x_{1\ell}-x_{2\ell}}{t}\right).$$

Next, we distinguish two cases: either $|x_{1\ell}-x_{2\ell}|<t$ as well as $|y_{1\ell}-y_{2\ell}|<t$ holds or only one of these two conditions holds.

First assume $|x_{1\ell}-x_{2\ell}|<t$ as well as $|y_{1\ell}-y_{2\ell}|<t$ holds. Then, as θ'' is supposed to be Lipschitz-continuous on $\mathcal{D} \supsetneq [-1,1]$, there exists an $L \in \mathbb{R}^+$ such that

$$\begin{aligned}
\left|\frac{\partial^2\Phi_\ell}{\partial x_{1\ell}^2}(x) - \frac{\partial^2\Phi_\ell}{\partial x_{1\ell}^2}(y)\right| &= \left|\frac{1}{t}\,\theta''\left(\frac{x_{1\ell}-x_{2\ell}}{t}\right) - \frac{1}{t}\,\theta''\left(\frac{y_{1\ell}-y_{2\ell}}{t}\right)\right| \\
&\le \frac{L_\theta}{t}\left|\left(\frac{x_{1\ell}-x_{2\ell}}{t}\right) - \left(\frac{y_{1\ell}-y_{2\ell}}{t}\right)\right| \\
&\le \frac{L_\theta}{t^2}(|x_{1\ell}-y_{1\ell}| + |x_{2\ell}-y_{2\ell}|) \\
&\le L\,\|x-y\|\,.
\end{aligned}$$

Next, suppose $|x_{1\ell}-x_{2\ell}|<t$ and $(y_{1\ell}-y_{2\ell}) \ge t$, then again there exists an $L \in \mathbb{R}^+$, such that

$$\begin{aligned}
\left|\frac{\partial^2\Phi_\ell}{\partial x_{1\ell}^2}(x) - \frac{\partial^2\Phi_\ell}{\partial x_{1\ell}^2}(y)\right| &= \left|\frac{1}{t}\,\theta''\left(\frac{x_{1\ell}-x_{2\ell}}{t}\right) - 0\right| \\
&= \frac{1}{t}\left|\theta''\left(\frac{x_{1\ell}-x_{2\ell}}{t}\right) - \theta''\left(\frac{t}{t}\right)\right| \\
&\le \frac{L_\theta}{t}\left|\left(\frac{x_{1\ell}-x_{2\ell}}{t}\right) - \left(\frac{t}{t}\right)\right| \\
&\le \frac{L_\theta}{t^2}\left((y_{1\ell}-y_{2\ell}) - (x_{1\ell}-x_{2\ell})\right) \\
&\le \frac{L_\theta}{t^2}(|x_{1\ell}-y_{1\ell}| + |x_{2\ell}-y_{2\ell}|) \\
&\le L\,\|x-y\|\,.
\end{aligned}$$

The same can be proved similarly for $(y_{1\ell}-y_{2\ell}) \le -t$ (then we choose $\theta''(-t/t)$) and accordingly for $\partial^2\Phi_\ell/\partial x_{2\ell}^2$ as well as for $\partial^2\Phi_\ell/\partial x_{1\ell}\partial x_{2\ell}$. Hence, applying the first part of the proof , we obtain the Lipschitz-continuity of $\nabla^2_{xx}\Phi_l$. □

Now we can apply the Convergence Theorem 3.1 to Problem $P(t)$.

Theorem 3.2. *Let x^* be a strongly stationary point of (1.19) and assume that the MPEC-LICQ and the RNLP-SOSC hold in x^*. Furthermore, assume that there exists a neighbourhood of x^* in which the second derivatives of f, g and h are Lipschitz-continuous and suppose θ'' is Lipschitz-continuous on an open set $\mathcal{D} \supseteq [-1,1]$. Moreover, assume that $t \in (0, \tau(x^*))$ and*

$$\begin{aligned} \lambda_j^* &> 0 \qquad j \in I_g(x^*) \\ \hat{\nu}_{1j} &> 0 \qquad j \in (I_1 \cap I_2)(x^*) \\ \hat{\nu}_{2j} &> 0 \qquad j \in (I_1 \cap I_2)(x^*)\,. \end{aligned}$$

Then there exists an $\varepsilon > 0$ such that for any starting point $(x^0,\, \lambda^0,\, \mu^0,\, \nu_1^0,\, \nu_2^0,\, \xi^0) \in \mathcal{B}_\varepsilon((x^*,\, \lambda^*,\, \mu^*,\, \hat{\nu}_1,\, \hat{\nu}_2, 0))$ Algorithm 3.1 applied to $P(t)$ generates a sequence of iterates $((x^k,\, \lambda^k,\, \mu^k,\, \nu_1^k,\, \nu_2^k,\, \xi^k))$ that converges q-quadratically to $(x^*,\, \lambda^*,\, \mu^*,\, \hat{\nu}_1,\, \hat{\nu}_2, 0)$. Moreover, the sequence (x^k) converges q-superlinearly to x^*.

Proof. By Lemma 3.3, x^* is a stationary point of $P(t)$ with multipliers $(x^*,\, \lambda^*,\, \mu^*,\, \tilde{\nu}_1,\, \tilde{\nu}_2, \tilde{\xi})$ that satisfy $\tilde{\nu}_1 := \hat{\nu}_1$, $\tilde{\nu}_2 := \hat{\nu}_2$, and $\tilde{\xi} := 0$. Moreover, by Lemma 2.6 and the condition $t > 0$, any index $j \in (I_1 \cap I_2)(x^*)$ satisfies $j \notin I_\Phi(x^*, t)$. Hence, the multipliers $(\lambda^*,\, \mu^*,\, \tilde{\nu}_1,\, \tilde{\nu}_2, 0)$ of x^* for $P(t)$ satisfy strict complementarity. By Lemma 3.6, Lemma 3.4 and Lemma 3.5 it follows that the assumptions for Theorem 3.1 are satisfied. Thus, by the second part of it, the generated sequence $((x^k,\, \lambda^k,\, \mu^k,\, \nu_1^k,\, \nu_2^k,\, \xi^k))$ converges q-quadratically to $(x^*,\, \lambda^*,\, \mu^*,\, \tilde{\nu}_1,\, \tilde{\nu}_2, \tilde{\xi})$ and by the first part of Theorem 3.1 it follows that (x^k) converges q-superlinearly to x^*. □

Knowing that Algorithm 3.1 applied to $P(t)$ yields a sequence of iterates x^k that satisfy fast local convergence in the vicinity of a local minimizer x^*, we are now interested in the question whether we can transfer this result to our problem $R(t)$. We therefore compare the corresponding QPs in a current iterate x^k. If we can deduce, that solving the QP of $R(t)$ determined in x^k yields the same step d as solving the corresponding QP of $P(t)$, then we can transfer the convergence result of Theorem 3.2 and obtain fast local convergence to x^* for $R(t)$.

The QP of $R(t)$ determined in a current iterate x^k with multipliers λ^k, μ^k, ν_1^k, ν_2^k and ξ^k has the form

$$QPR(t)(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$$

$$\begin{aligned} \min_d \quad & q_{R(t)}(d) = \nabla f(x^k)^T d + \frac{1}{2} d^T H^k d \\ \text{subject to} \quad & h(x^k) + \nabla h(x^k)^T d = 0 \\ & g(x^k) + \nabla g(x^k)^T d \geq 0 \\ & x_{1j}^k + d_{1j} \geq 0 \qquad j \in \{1, \dots, p\} \\ & x_{2j}^k + d_{2j} \geq 0 \qquad j \in \{1, \dots, p\} \\ & \Phi_j(x_1^k, x_2^k, t) + \nabla_x \Phi_j(x_1^k, x_2^k, t)^T d \leq 0 \quad j \in \{1, \dots, p\}\,, \end{aligned}$$

where H^k denotes the matrix $\nabla_{xx}^2 \mathcal{L}_{R(t)}(x^k,\, \lambda^k,\, \mu^k,\, \nu_1^k,\, \nu_2^k,\, \xi^k)$. The corresponding quadratic approximation of $P(t)$ in the same iterate x^k, which is close to the strongly stationary point

x^*, has the form

$QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$

$$\begin{array}{rl} \min_d & q_{P(t)}(d) = \nabla f(x^k)^T d + \frac{1}{2} d^T \hat{H}^k d \\ \text{subject to} & h(x^k) + \nabla h(x^k)^T d = 0 \\ & g(x^k) + \nabla g(x^k)^T d \geq 0 \\ & d_{1j} = 0, \quad x_{2j}^k + d_{2j} \geq 0 \qquad j \in (I_1 \backslash I_2)(x^*) \\ & d_{2j} = 0, \quad x_{1j}^k + d_{1j} \geq 0 \qquad j \in (I_2 \backslash I_1)(x^*) \\ & x_{1j}^k + d_{1j} \geq 0, \quad x_{2j}^k + d_{2j} \geq 0, \qquad j \in (I_1 \cap I_2)(x^*) \\ & \Phi_j(x_1^k, x_2^k, t) + \nabla_x \Phi_j(x_1^k, x_2^k, t)^T d \leq 0 \quad j \in (I_1 \cap I_2)(x^*)\,, \end{array}$$

where $\hat{H}^k$ denotes the matrix $\nabla^2_{xx} \mathcal{L}_{P(t)}(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$ with

$$\begin{aligned} \mathcal{L}_{P(t)}(x, \lambda, \mu, \nu_1, \nu_2, \xi) \;=\; & f(x) - \textstyle\sum_{j=1}^m \lambda_j g_j(x) - \sum_{i=1}^q \mu_i h_i(x) \\ & -\nu_1^T x_1 - \nu_2^T x_2 + \textstyle\sum_{j \in (I_1 \cap I_2)(x^*)} \xi_j \Phi_j(x_1, x_2, t)\,. \end{aligned} \tag{3.9}$$

Consider a strongly stationary point x^* of (1.19) and assume that $t \in (0, \tau(x^*))$. Then by Lemma 3.1 $(I_1 \backslash I_2)(x^*) \subseteq I_1(x) \cap I_\Phi(x, t)$, if x is sufficiently close to x^*.

Now compare problem $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$ with $QPR(t)(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$ and note that if x^k is close to x^*, then we have only exchanged the two inequalities $d_{1j} \geq 0$ and $\nabla_x \Phi_j(x_1^k, x_2^k, t)^T d = 2 d_{1j} \leq 0$ by one inequality, namely $d_{1j} = 0$ for all $j \in (I_1 \backslash I_2)(x^*)$, since $x_{1j}^k = 0$ and $\Phi_j(x_1^k, x_2^k, t) = 0$ for all $j \in (I_1 \backslash I_2)(x^*)$. The same applies accordingly to indices $j \in (I_2 \backslash I_1)(x^*)$.

Furthermore, compare the feasible regions of $P(t)$ and of $R(t)$ in a small neighbourhood of x^* and note that they are identical.

Suppose that we determine $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$ close to a strongly stationary point x^*, where the MPEC-LICQ is satisfied, then $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$ is consistent.

Lemma 3.7. *Let x^* be a strongly stationary point of (1.19), where the MPEC-LICQ is satisfied. Furthermore, let $t \in (0, \tau(x^*))$. Then there exists an $\varepsilon > 0$, so that $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$ is consistent for any $x^k \in \mathcal{Z}(t) \cap B_\varepsilon(x^*)$.*

Proof. Let $x^k \in \mathcal{Z}(t) \cap B_\varepsilon(x^*)$, then Lemma 3.4 implies that the LICQ holds in x for $P(t)$. We can therefore find a vector $d \in \mathbb{R}^{n+2p}$ that satisfies

$$\begin{array}{rl} \nabla h_j(x)^T d = 0 & j \in \{1, \ldots, q\} \\ \nabla g_j(x)^T d = 0 & j \in I_g(x) \\ d_{1j} = 0 & j \in (I_1 \backslash I_2)(x^*) \\ d_{2j} = 0 & j \in (I_2 \backslash I_1)(x^*) \\ d_{1j} = 0 & j \in (I_1 \cap I_2)(x^*) \cap I_1(x) \\ d_{2j} = 0 & j \in (I_1 \cap I_2)(x^*) \cap I_2(x) \\ -\alpha_j\, d_{1j} - (2 - \alpha_j) d_{2j} = 0 & j \in (I_1 \cap I_2)(x^*) \cap I_\Phi(x, t)\,. \end{array} \tag{3.10}$$

Hence, d satisfies the constraints of $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$ that correspond to the active constraints in x of $P(t)$.

Concerning the remaining inequality constraints of $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$, corresponding to the inactive inequality constraints of $P(t)$, we can find a $\delta > 0$ such that for any $d \in \mathbb{R}^{n+2p}$ the vector δd satisfies these inequality constraints of $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$. Since this holds in particular for $d \in \mathbb{R}^{n+2p}$ that satisfies (3.10), it follows that $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$ is consistent. □

Remark 3.1. By Lemma 3.7 it clearly follows that for all x^k close to a strongly stationary point of (1.19), where the MPEC-LICQ is satisfied, the problems $QPR(t)(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$ are also consistent.

Next we prove that the feasible regions and the objective functions of the problems $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$ and $QPR(t)(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$ are equal under suitable conditions. These results will then be used to prove that the two QPs have the same solutions.

Lemma 3.8. *Let x^* be a strongly stationary point of (1.19) and $t \in (0, \tau(x^*))$. Then there exists an $\varepsilon > 0$ such that if $x^k \in \mathcal{Z}(t) \cap B_\varepsilon(x^*)$, then the feasible regions of $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$ and of $QPR(t)(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$ are identical.*

Proof. Let first $d \in \mathbb{R}^{n+2p}$ be feasible for $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$. Then, for $x^k \in \mathcal{Z}(t) \cap B_\varepsilon(x^*)$, the inequalities $x_{1j}^k + d_{1j} \geq 0$ and $x_{2j}^k + d_{2j} \geq 0$ are clearly satisfied for all $j \in \{1, \ldots, p\}$. Furthermore, $x_{1j}^k > t$, $x_{2j}^k = 0$ and $d_{2j} = 0$ for all $j \in (I_2 \backslash I_1)(x^*)$. Thus,

$$\Phi_j(x_1^k, x_2^k, t) = 0 \quad \text{and} \quad \nabla \Phi_j(x_1^k, x_2^k, t)^T d = 2d_{2j} = 0$$

for all $j \in (I_2 \backslash I_1)(x^*)$. Hence $\Phi_j(x_1^k, x_2^k, t) + \nabla \Phi_j(x_1^k, x_2^k, t)^T d = 0$ for all $j \in (I_2 \backslash I_1)(x^*)$. Accordingly it can be proved that $\Phi_j(x_1^k, x_2^k, t) + \nabla_x \Phi_j(x_1^k, x_2^k, t)^T d = 0$ for all $j \in (I_1 \backslash I_2)(x^*)$. For all $j \in (I_1 \cap I_2)(x^*)$, these constraints are directly satisfied, since d is feasible for $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$. Hence, d is feasible for $QPR(t)(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$.

Now suppose $d \in \mathbb{R}^{n+2p}$ be feasible for $QPR(t)(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$. Then, in order to prove that d is feasible for $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$, we have to show that $d_{1j} = 0$ for all $j \in (I_1 \backslash I_2)(x^*)$ and $d_{2j} = 0$ for all $j \in (I_2 \backslash I_1)(x^*)$.

Since $t < \tau(x^*)$ and $x^k \in \mathcal{Z}(t)$ we can find an $\varepsilon > 0$, such that if $x^k \in B_\varepsilon(x^*)$, then $x_{1j}^k = 0$ for all $j \in (I_1 \backslash I_2)(x^*)$ (confer Lemma 3.1). Hence, by the feasibility of d for $QPR(t)(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$, it follows that $d_{1j} \geq 0$ for all $j \in (I_1 \backslash I_2)(x^*)$. Moreover,

$$0 \geq \Phi_j(x_1, x_2, t) + \nabla \Phi_j(x_1, x_2, t)^T d = 2d_{1j}$$

for all $j \in (I_1 \backslash I_2)(x^*)$. Therefore, $d_{1j} = 0$ for all $j \in (I_1 \backslash I_2)(x^*)$. Accordingly we can prove that $d_{2j} = 0$ for all $j \in (I_2 \backslash I_1)(x^*)$. Thus d is feasible for $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$. □

Lemma 3.9. *Let x^* be a strongly stationary point of (1.19) and $t \in (0, \tau(x^*))$. Then there exists an $\varepsilon > 0$, such that if $x^k \in \mathcal{Z}(t) \cap B_\varepsilon(x^*)$, then the corresponding problems $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$ and $QPR(t)(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$ have the same objective function, provided that $\tilde{\lambda}^k = \lambda^k$ and $\tilde{\mu}^k = \mu^k$, as well as $\tilde{\xi}_j^k = \xi_j^k$ for $j \in (I_1 \cap I_2)(x^*)$.*

Proof. The objective functions of $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$ and $QPR(t)(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$ differ only in their quadratic term, thus in H^k and $\hat{H}^k$. Comparing these two matrices for $x^k \in \mathcal{Z}(t) \cap B_\varepsilon(x^*)$ and keeping in mind, that for $t \in (0, \tau(x^*))$ the second derivatives of $\Phi_j(x_1, x_2, t)$ vanish for all $j \in \{1, \ldots, p\} \backslash (I_1 \cap I_2)(x^*)$, it follows that $H^k = \hat{H}^k$, provided that $\tilde{\lambda}^k = \lambda^k$ and $\tilde{\mu}^k = \mu^k$, as well as $\tilde{\xi}_j^k = \xi_j^k$ for all $j \in (I_1 \cap I_2)(x^*)$. □

Having proved that the feasible regions and the objective functions of both problems $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$ and $QPR(t)(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$ are equal in the vicinity of a strongly stationary point, we can now use this information to show that the solutions d of both problems, that are closest to the origin, are also equal.
Furthermore, we will make use of the following sensitivity result, that corresponds to Theorem 3.2.7 of [CGT00]:

Theorem 3.3. *Consider the problem*

$$\begin{aligned} \min \quad & f(x,p) \\ \text{subject to} \quad & h(x,p) = 0 \\ & g(x,p) \geq 0 \,, \end{aligned} \tag{3.11}$$

where p is a set of parameters, and (3.11) is (1.13) when $p = 0$. Suppose that the second derivatives of f, g and h are jointly continuous functions of x and p and that the LICQ and the SOSC hold as well as $\lambda_j^* \neq 0$ for all $j \in I_g(x^*)$ holds at (x^*, λ^*, μ^*). Then there is some open set $\mathcal{P}$ containing the origin and some open neighbourhood $\mathcal{X} \times \mathcal{Y}$ of (x^*, λ^*, μ^*) for which the continuous function $(x^*(p), \lambda^*(p), \mu^*(p))$, with $(x^*(0), \lambda^*(0), \mu^*(0)) = (x^*, \lambda^*, \mu^*)$, also satisfies the LICQ, the SOSC and strict complementarity of the multipliers $\lambda^*(p)$ for all $p \in \mathcal{P}$. Furthermore, $x^*(p)$ is the only strict local minimizer of (3.11) in $\mathcal{X}$, $(\lambda^*(p), \mu^*(p))$ is the unique vector of Lagrange multipliers at this point and $(I_h \cup I_g)(x^*(p)) = (I_h \cup I_g)(x^*)$.

Lemma 3.10. *Let x^* be a strongly stationary point of (1.19) with multipliers $(\lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2)$ and let $t \in (0, \tau(x^*))$. Furthermore, suppose the MPEC-LICQ and the RNLP-SOSC are satisfied in x^*. Then there exists an $\varepsilon > 0$, such that if $x^k \in \mathcal{Z}(t) \cap B_\varepsilon(x^*)$ and the corresponding multiplier estimates for $P(t)$ satisfy $(\tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k) \in B_\varepsilon((\lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2, 0))$, where $(\lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2, 0)$ denotes the multiplier vector of x^* for $P(t)$, and it holds*

$$\begin{aligned} \lambda_j^* > 0 \qquad & j \in I_g(x^*) \\ \hat{\nu}_{1j} > 0 \qquad & j \in (I_1 \cap I_2)(x^*) \\ \hat{\nu}_{2j} > 0 \qquad & j \in (I_1 \cap I_2)(x^*) \,, \end{aligned}$$

then

1. *the solution d^k of $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$ is the only strict local solution in the vicinity of $d = 0$. Moreover, the corresponding multipliers are unique.*
2. *The same d^k is also the only strict local solution for $QPR(t)(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$ in the vicinity of $d = 0$ provided that $\tilde{\lambda}^k = \lambda^k$ and $\tilde{\mu}^k = \mu^k$, as well as $\tilde{\xi}_j^k = \xi_j^k$ for the indices $j \in (I_1 \cap I_2)(x^*)$.*

Proof. Consider a current iterate x^k and its multiplier estimates $(\tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$. Then $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$ can be interpreted as a perturbation of $QPP(t)(x^*, \lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2, 0)$. Define the perturbation parameters $p_x := x^k - x^*$ and $p_\lambda := (\tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k) - (\lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2, 0)$. Then by Lemma 3.3 and Lemma 3.5, the vector $d = 0$ solves $QPP(t)(x^* + p_x, (\lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2, 0) + p_\lambda)$ for the perturbation parameters $(p_x, p_\lambda) = (0, 0)$ with corresponding multipliers $(\lambda^*, \mu^*, \tilde{\nu}_1, \tilde{\nu}_2, 0)$. By Lemma 3.4 and Lemma 3.5 it follows moreover, that Theorem 3.3 can be applied. We therefore conclude that the solution $d^k = d(p_x, p_\lambda)$ is the only strict local minimizer of $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$ in the vicinity of $d(0,0) = 0$. Furthermore, it follows by Theorem 3.3 that the LICQ and the SOSC hold in d^k for $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$ and its multipliers are accordingly unique.

By Lemma 3.8 the vector d^k is also feasible for $QPR(t)(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$. Moreover since d^k is a strict local minimum of $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$, there exists an $\bar{\epsilon}$ such that by Lemma 3.9 we have

$$q_{R(t)}(d) = q_{P(t)}(d) > q_{P(t)}(d^k) = q_{R(t)}(d^k)$$

for all feasible $d \in B_{\bar{\epsilon}}(d^k)$. Thus d^k is also a strict local minimum of $QPR(t)(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$.

Now suppose d^k is not unique for $QPR(t)(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$ in the vicinity of $d = 0$. Then there exists another vector $\bar{d}$ in the vicinity of $d = 0$ that is feasible for $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$ by Lemma 3.8 and minimizes the objective function of $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$ by Lemma 3.9. This, however, contradicts the uniqueness of the solution d^k for $QPP(t)(x^k, \tilde{\lambda}^k, \tilde{\mu}^k, \tilde{\nu}_1^k, \tilde{\nu}_2^k, \tilde{\xi}^k)$. □

Finally, we conclude that the local SQP algorithm 3.1 generates a convergent sequence, if applied to $R(t)$ in the vicinity of a strongly stationary point x^*.

Theorem 3.4. *Let x^* be a strongly stationary point of (1.19) and assume that the MPEC-LICQ and the RNLP-SOSC hold in x^*. Furthermore, assume that there exists some neighbourhood of x^* in which the second derivatives of f, g and h are Lipschitz-continuous and suppose θ'' is Lipschitz-continuous on an open set $\mathcal{D} \supseteq [-1, 1]$. Moreover, assume that $t \in (0, \tau(x^*))$ and*

$$\begin{aligned} \lambda_j^* &> 0 \qquad j \in I_g(x^*) \\ \hat{\nu}_{1j} &> 0 \qquad j \in (I_1 \cap I_2)(x^*) \\ \hat{\nu}_{2j} &> 0 \qquad j \in (I_1 \cap I_2)(x^*)\,. \end{aligned}$$

Then it holds that

1. *there exists a neighbourhood $\varepsilon > 0$ such that for any starting point $(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k) \in \mathcal{B}_\varepsilon((x^*, \lambda^*, \mu^*, \nu_1^b, \nu_2^b, \xi^b))$ Algorithm 3.1 applied to $R(t)$ generates a sequence of iterates (x^k) that converges q-superlinearly to x^*.*
2. *If the family*

$$\{\nabla g_j(x^k) \,|\, j \in supp(\lambda^k)\} \cup \{\nabla h_j(x^k) \,|\, j \in supp(\mu^k)\} \cup \{e_{1j} \,|\, j \in supp(\nu_1^k)\}$$
$$\cup \{e_{2j} \,|\, j \in supp(\nu_2^k)\} \cup \{-\alpha_j^k\, e_{1j} - (2 - \alpha_j^k)\, e_{2j} \,|\, j \in supp(\xi^k)\}$$

of the associated multiplier estimates for the local solutions d^k for $QPR(t)$ is linear independent for each subsequent $k \in \mathbb{N}$, then the sequence $((x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k))$ converges q-quadratically to $(x^*, \lambda^*, \mu^*, \nu_1^b, \nu_2^b, \xi^b)$, respectively, where the multipliers $\lambda^*, \mu^*, \nu_1^b, \nu_2^b$, and ξ^b are defined as in Definition 2.2.

Proof. Starting with $(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$ we can construct multipliers $\tilde\lambda^k, \tilde\mu^k, \tilde\nu_1^k, \tilde\nu_2^k$ and $\tilde\xi^k$ for the associated problem $P(t)$, such that the corresponding $QPR(t)(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$ and $QPP(t)(x^k, \tilde\lambda^k, \tilde\mu^k, \tilde\nu_1^k, \tilde\nu_2^k, \tilde\xi^k)$ satisfy the assumptions of Lemma 3.10. Hence, the solution d^k of $QPR(t)(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$ corresponds to the solution of $QPP(t)(x^k, \tilde\lambda^k, \tilde\mu^k, \tilde\nu_1^k, \tilde\nu_2^k, \tilde\xi^k)$. Furthermore, since the associated problem $QPP(t)$ satisfies the LICQ, it follows that the new multiplier estimates $\lambda^{k+1}, \mu^{k+1}, \nu_1^{k+1}, \nu_2^{k+1}$ and ξ^{k+1} and $\tilde\lambda^{k+1}, \tilde\mu^{k+1}, \tilde\nu_1^{k+1}, \tilde\nu_2^{k+1}$, and $\tilde\xi^{k+1}$ satisfy $\tilde\lambda^{k+1} = \lambda^{k+1}$ and $\tilde\mu^{k+1} = \mu^{k+1}$, as well as $\tilde\xi_j^{k+1} = \xi_j^{k+1}$ for all $j \in (I_1 \cap I_2)(x^*)$ (suppose these conditions are not satisfied, then we obtain a nontrivial linear combination of zero of the active constraint gradients, which contradicts the LICQ for $QPP(t)(x^k, \tilde\lambda^k, \tilde\mu^k, \tilde\nu_1^k, \tilde\nu_2^k, \tilde\xi^k)$). Hence we can apply Lemma 3.10 to the next pair of QPs, which are $QPR(t)(x^{k+1}, \lambda^{k+1}, \mu^{k+1}, \nu_1^{k+1}, \nu_2^{k+1}, \xi^{k+1})$ and $QPP(t)(x^{k+1}, \tilde\lambda^{k+1}, \tilde\mu^{k+1}, \tilde\nu_1^{k+1}, \tilde\nu_2^{k+1}, \tilde\xi^{k+1})$. Thus, the sequence (x^k) that is generated applying Algorithm 3.1 to $R(t)$ close to x^* corresponds to the sequence that would be generated by applying Algorithm 3.1 to the associated $P(t)$ close to x^*. Therefore we obtain the same convergence behaviour for the sequence (x^k) for $R(t)$ as we would obtain for $P(t)$.

If the family

$$\begin{aligned} &\{\nabla g_j(x^k)\,|\,j \in supp(\lambda^k)\} \cup \{\nabla h_j(x^k)\,|\,j \in supp(\mu^k)\} \cup \{e_{1j}\,|\,j \in supp(\nu_1^k)\} \\ &\cup \{e_{2j}\,|\,j \in supp(\nu_2^k)\} \cup \{-\alpha_j^k\, e_{1j} - (2-\alpha_j^k)\, e_{2j}\,|\,j \in supp(\xi^k)\} \end{aligned}$$

is linear independent for each subsequent $k \in \mathbb{N}$, then by Lemma 3.2 the associated multipliers of $QPR(t)(x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k)$ satisfy $\lambda^{k+1} = \tilde\lambda^{k+1}$, $\mu^{k+1} = \tilde\mu^{k+1}$ and

$$\begin{aligned} \nu_{1j}^{k+1} &= (\tilde\nu_{1j}^{k+1})^+ \quad \text{for} \quad j \in \{1,\dots,p\} \\ \nu_{2j}^{k+1} &= (\tilde\nu_{2j}^{k+1})^+ \quad \text{for} \quad j \in \{1,\dots,p\} \\ \xi_j^{k+1} &= \begin{cases} (-\frac{\tilde\nu_{1j}^{k+1}}{2})^+ & j \in (I_1 \backslash I_2)(x^*) \\ (-\frac{\tilde\nu_{2j}^{k+1}}{2})^+ & j \in (I_2 \backslash I_1)(x^*) \\ \tilde\xi_j^{k+1} & j \in (I_1 \cap I_2)(x^*)\,. \end{cases} \end{aligned}$$

By the q-quadratic convergence of $(x^k, \tilde\lambda^{k+1}, \tilde\mu^{k+1}, \tilde\nu_1^{k+1}, \tilde\nu_2^{k+1}, \tilde\xi^{k+1})$, these multipliers converge to the basic multipliers as defined in Definition 2.2. Furthermore, the sequence $((x^k, \lambda^k, \mu^k, \nu_1^k, \nu_2^k, \xi^k))$ converges q-quadratically to $(x^*, \lambda^*, \mu^*, \nu_1^b, \nu_2^b, \xi^b)$, respectively. □

Having discussed some convergence properties of a local SQP algorithm applied to the relaxed problem $R(t)$, we are now interested in the application of an Interior Point algorithm to problem $R(t)$, as these algorithms form another important class of optimization methods for NLPs.

3.2 An Interior Point Method

In the following, we want to apply an Interior Point Method to solve (1.19). Therefore, we demonstrate how the two-sided relaxation scheme proposed by DeMiguel et al. [DFNS05] (confer Section 1.4) can be combined with the new relaxation scheme we discussed in Chapter 2. Combining these two relaxation schemes, we can then apply an Interior Point Method, which produces a sequence of iterates that converge superlinearly in the vicinity of a strongly stationary point.

The two-sided relaxation scheme of DeMiguel et al. differs from the one-sided scheme of Scholtes [Sch01] in the fact that not only the condition $x_1x_2 = 0$ is relaxed by $x_1x_2 \leq \delta_c$ for δ_c, being a positive parameter, but also the two bound constraints $x_1 \geq 0$ and $x_2 \geq 0$ are relaxed by $x_1 \geq -\delta_1$ and $x_2 \geq -\delta_2$ with δ_1 and δ_2 being again two positive parameters. Depending on the sign of corresponding multiplier estimates, either δ_1 or δ_2, respectively, is approaching zero or δ_c approaches zero, but not both. This procedure ensures, that the strictly feasible region of the resulting problem is not empty even in the limit (in contrast to the one-sided regularization scheme by Scholtes). This property of the two-sided relaxation scheme is in particular valuable, since we want to apply an Interior Point Method, where we need a strict interior of the feasible region.

3.2.1 Modified Strictly Feasible Relaxation Scheme

If we apply the two-sided relaxation scheme of [DFNS05] in combination with the relaxation scheme discussed in Chapter 2 to (1.19), then the resulting relaxed nonlinear problem has the form

$$\begin{array}{rrcl} \text{min} & f(x) & & \\ \text{subject to} & h(x) & = & 0 \\ & g(x) & \geq & 0 \\ & x_1 & \geq & -\delta_1 \\ & x_2 & \geq & -\delta_2 \\ & \Phi(x_1, x_2, t) & \leq & \delta_c\,, \end{array} \tag{3.12}$$

where $\Phi(x_1, x_2, t) : \mathbb{R}^{2p} \times (\mathbb{R}_0^+)^p \to \mathbb{R}^p$, $\Phi_j(x_1, x_2, t) := x_{1j} + x_{2j} - \varphi_j(x_{1j}, x_{2j}, t_j)$ and $\varphi_j(x_{1j}, x_{2j}, t_j)$ as defined in Section 2.1. Note, that for $t_j = 0$ we have

$$\varphi_j(x_{1j}, x_{2j}, 0) = |x_{1j} - x_{2j}|\,.$$

Figure 3.2.1 illustrates the feasible region of a complementarity pair $(x_1, x_2) \in \mathbb{R}^2$ for problem (3.12).

Next, we introduce slack variables $s = (s_g, s_1, s_2, s_c) \in \mathbb{R}^{m+3p}$ to obtain the nonlinear program

$$\begin{array}{lrrcl} R(\delta, t) & \text{min}_{x,s} & f(x) & & \\ & \text{subject to} & h(x) & = & 0 \\ & & s_g - g(x) & = & 0 \\ & & s_1 - x_1 & = & \delta_1 \\ & & s_2 - x_2 & = & \delta_2 \\ & & s_c + \Phi(x_1, x_2, t) & = & \delta_c \\ & & s & \geq & 0\,. \end{array} \tag{3.13}$$

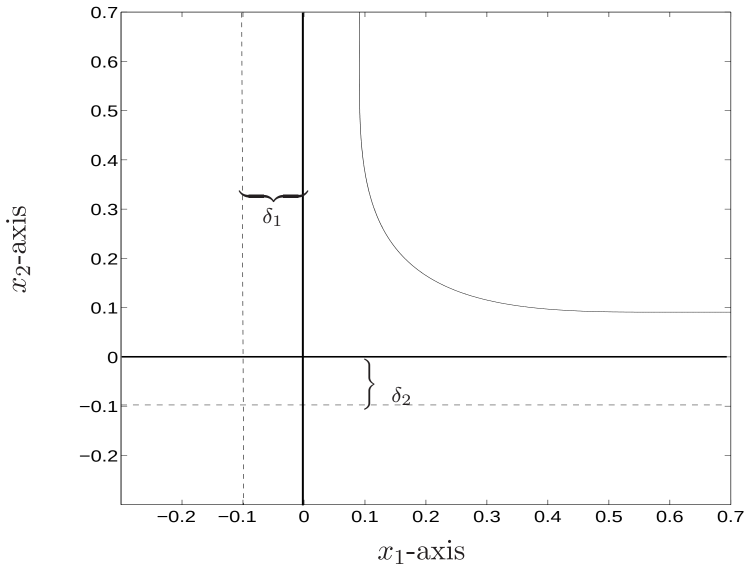


Figure 3.1: Feasible Region of the relaxed problem (3.12) for a complementarity pair $(x_1, x_2) \in \mathbb{R}^2$.

The KKT-conditions for $R(\delta, t)$ can be written as the system of nonlinear equations

$$\nabla_x \mathcal{L}_{R(\delta,t)}(x, s, \mu, v) = \nabla f(x) - \nabla h(x)\mu - \nabla_x C(x, s, t)v = 0 \tag{3.14}$$

$$\min(s_i, v_i) = 0 \tag{3.15}$$

$$h(x) = 0 \tag{3.16}$$

$$C(x, s, t) + (0, \delta_1, \delta_2, \delta_c)^T = 0\,, \tag{3.17}$$

with $v := (\lambda, \nu_1, \nu_2, \xi)$, $\mathcal{L}_{R(\delta,t)}(x, s, \mu, v) = f(x) - \sum_{i=1}^{q} \mu_i h_i(x) - \sum_{j=1}^{m+3p} v_j C_j(x, s, t)$,

$$C(x, s, t) = -\begin{pmatrix} s_g - g(x) \\ s_1 - x_1 \\ s_2 - x_2 \\ s_c + \Phi(x_1, x_2, t) \end{pmatrix}$$

and $\nabla_x C(x, s, t)$ denotes the transposed Jacobian of $C(x, s, t)$ with respect to x, thus

$$\nabla_x C(x, s, t) = \begin{pmatrix} \nabla_{x_0} g(x) & 0 & 0 & 0 \\ \nabla_{x_1} g(x) & I & 0 & -D_1 \\ \nabla_{x_2} g(x) & 0 & I & -D_2 \end{pmatrix}$$

with I denoting the unit matrix and $D_1 = \nabla_{x_1}\Phi(x_1, x_2, t)$ and $D_2 = \nabla_{x_2}\Phi(x_1, x_2, t)$ as in Section 2.1.

Remark 3.2. Note that

$$\nabla^2_{xx} \mathcal{L}_{R(\delta,t)}(x, s, \mu, v) = \nabla^2_{xx} \mathcal{L}_{R(t)}(x, \lambda, \mu, \nu_1, \nu_2, \xi). \tag{3.18}$$

Since the method is based on solving $R(\delta, t)$ for suitable values of the parameters δ and t in order to solve (1.19), we first need to relate the stationary points of $R(\delta, t)$ to the strongly stationary points of (1.19).

Theorem 3.5. *Let x^* be a strongly stationary point of (1.19) with multipliers λ^*, μ^*, $\hat{\nu}_1$ and $\hat{\nu}_2$ and let (δ^*, t^*) satisfy the following conditions*

$$\delta^*_{ij} = 0 \quad \text{if} \quad \hat{\nu}_{ij} > 0, \tag{3.19}$$
$$\delta^*_{ij} > 0 \quad \text{if} \quad \hat{\nu}_{ij} \leq 0, \tag{3.20}$$
$$\delta^*_{cj} = 0 \quad \text{if} \quad \hat{\nu}_{1j} < 0 \quad \text{or} \quad \hat{\nu}_{2j} < 0, \tag{3.21}$$
$$\delta^*_{cj} > 0 \quad \text{if} \quad \hat{\nu}_{1j} \geq 0 \quad \text{and} \quad \hat{\nu}_{2j} \geq 0, \tag{3.22}$$
$$0 \leq t^*_j \leq |x^*_{1j} - x^*_{2j}| \quad \text{if} \quad \hat{\nu}_{1j} < 0 \quad \text{or} \quad \hat{\nu}_{2j} < 0, \tag{3.23}$$
$$t^*_j > 0 \quad \text{if} \quad \hat{\nu}_{1j} \geq 0 \quad \text{and} \quad \hat{\nu}_{2j} \geq 0, \tag{3.24}$$

for $i = 1, 2$ and $j = 1, \ldots, p$. Then,

$$\delta^*_{1j} + \delta^*_{cj} > 0 \quad \text{and} \quad \delta^*_{2j} + \delta^*_{cj} > 0\,. \tag{3.25}$$

Moreover, if we define

$$s^* := (g(x^*), x^*_1 + \delta^*_1, x^*_2 + \delta^*_2, \delta^*_c - \Phi(x^*_1, x^*_2, t^*)) \tag{3.26}$$
$$(\nu^*_1, \nu^*_2) := ([\hat{\nu}_1]^+, [\hat{\nu}_2]^+) \tag{3.27}$$
$$\xi^*_j := \begin{cases} \left(-\frac{\hat{\nu}_{1j}}{2}\right)^+ & j \in (I_1 \backslash I_2)(x^*) \\ \left(-\frac{\hat{\nu}_{2j}}{2}\right)^+ & j \in (I_2 \backslash I_1)(x^*) \\ 0 & j \in (I_1 \cap I_2)(x^*)\,, \end{cases} \tag{3.28}$$

then (x^*, s^*) is a stationary point of $R(\delta^*, t^*)$ with the corresponding multiplier (μ^*, λ^*, ν^*_1, ν^*_2, ξ^*). Furthermore, if x^* satisfies the MPEC-LICQ and the SSOSC, then (x^*, s^*) satisfies the LICQ and the SOSC for $R(\delta^*, t^*)$. Moreover, if $g_j(x^*) + \lambda^*_j > 0$, then s^* and $v^* = (\lambda^*, \nu^*_1, \nu^*_2, \xi^*)$, satisfy strict complementarity.

Proof. First we prove condition (3.25). As x^* is assumed to be strongly stationary, $\hat{\nu}_{1j} > 0$ and $\hat{\nu}_{2j} < 0$ or $\hat{\nu}_{2j} > 0$ and $\hat{\nu}_{1j} < 0$, respectively, cannot hold. Hence, by view of (3.19)-(3.22), condition (3.25) holds.

Next, we show that $(x^*, s^*, \mu^*, \lambda^*, \nu^*_1, \nu^*_2, \xi^*)$ with s^*, ν^*_1, ν^*_2 and ξ^* defined by (3.26)-(3.28) is a stationary point of $R(\delta^*, t^*)$, thus satisfies (3.14)-(3.17). First, (3.16) is satisfied by the feasibility of x^* for (1.19). Furthermore, the definition of s^* implies $C(x^*, s^*, t^*) = 0$. Condition (3.14) results from the corresponding equation of (1.23), the conditions (3.19)-(3.24) and the definitions (3.26)-(3.28).

The definitions (3.27) and (3.28) imply the nonnegativity of ν^*_1, ν^*_2 and ξ^*. For s^*_1 and s^*_2 we further have $s^*_{ij} = x^*_{ij} + \delta^*_{ij} \geq 0$ for $i = 1, 2$ and for all $j \in \{1, \ldots, p\}$. As $\Phi_j(x^*_1, x^*_2, t^*) \leq 0$ for all x^* that are feasible for (1.19) and $t^* \geq 0$, it holds $s^*_{cj} = \delta^*_{cj} - \Phi_j(x^*_1, x^*_2, t^*) \geq 0$.

It remains to show that s^* and $v^* = (\lambda^*, \nu^*_1, \nu^*_2, \xi^*)$ are complementary. The complementarity of $s^*_{gj} = g_j(x^*)$ and λ^*_j follows directly by the stationarity conditions (1.23).

Next, suppose $\nu_{ij}^* > 0$, then by (3.27), it follows that $\hat{\nu}_{ij} > 0$, such that $x_{ij}^* = 0$ by (1.23). Therefore, by (3.19) we have $s_{ij}^* = x_{ij}^* + \delta_{ij}^* = 0$ for $i = 1, 2$, whereas if $\nu_{ij}^* = 0$, then by (3.20) $s_{ij}^* = x_{ij}^* + \delta_{ij}^* \geq \delta_{ij}^* > 0$.

If $\xi_j^* > 0$, then by (3.28) and (3.21) it follows that $\delta_{cj}^* = 0$ and $t_j^* \leq |x_{1j}^* - x_{2j}^*|$, thus by a suitable variant of Lemma 2.1 (confer Section 2.4) $\Phi_j(x_1^*, x_2^*, t^*) = 0$ and therefore $s_{cj}^* = \delta_{cj}^* - \Phi_j(x_1^*, x_2^*, t^*) = 0$. Suppose $\xi_j^* = 0$, then by (3.28) and (3.22) it follows that $\delta_{cj}^* > 0$ and again, as $\Phi(x_1^*, x_2^*, t^*) \leq 0$ for all x^* that are feasible for (1.19) and $t^* \geq 0$, $s_{cj}^* = \delta_{cj}^* - \Phi_j(x_1^*, x_2^*, t^*) \geq \delta_{cj}^* > 0$.

As we have just shown that s_i^* and ν_i^* are strictly complementary for $i = 1, 2$ and s_c^* and ξ^* are strictly complementary, condition (3.15) is also satisfied and hence $(x^*, s^*, \lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*)$ is a stationary point of $R(\delta^*, t^*)$. Moreover, if $g_j(x^*) + \lambda_j^* > 0$ for $j \in \{1, \dots, p\}$ then s^* and v^* satisfy strict complementarity.

Next, we proof that if the MPEC-LICQ holds in x^*, then the LICQ holds in (x^*, s^*) for $R(\delta^*, t^*)$. Therefore, assume the MPEC-LICQ holds in x^* for (1.19). Then, by the definition of the MPEC-LICQ, the family of active constraint gradients

$$\begin{array}{rl} \nabla h_i(x^*) & i \in \{1, \dots, q\}, \\ \nabla g_j(x^*) & j \in I_g(x^*), \\ e_{n+j} & j \in I_1(x^*), \\ e_{n+p+j} & j \in I_2(x^*), \end{array} \tag{3.29}$$

is linearly independent. If we can prove that the LICQ holds in x^* for

$$\begin{array}{lrcl} \min & f(x) & & \\ \text{subject to} & h(x) & = & 0 \\ & g(x) & = & 0 \\ & x_1 & \geq & -\delta_1^* \\ & x_2 & \geq & -\delta_2^* \\ & \Phi(x_1, x_2, t^*) & \leq & \delta_c^*, \end{array} \tag{3.30}$$

then it follows that the LICQ holds in (x^*, s^*) for $R(\delta^*, t^*)$. (The introduction of slack variables does not alter the property that the LICQ is satisfied.)

First, note that if $\delta_{cj}^* > 0$, then by $\Phi_j(x_1^*, x_2^*, t^*) \leq 0$ for all feasible x^* and $t^* \geq 0$, the corresponding inequality of (3.30) is strictly satisfied. Thus, if $\delta_{cj}^* > 0$ holds for all $j \in \{1, \dots, p\}$, then the LICQ for (3.30) directly follows by the MPEC-LICQ. Hence, assume $\delta_{cj}^* = 0$ for at least one $j \in \{1, \dots, p\}$. Then by (3.21) either $\hat{\nu}_{1j} < 0$ or $\hat{\nu}_{2j} < 0$, such that $j \in (I_1 \backslash I_2)(x^*)$ or $j \in (I_2 \backslash I_1)(x^*)$ and in both cases $\delta_{ij}^* > 0$ for $i = 1, 2$. Therefore, $x_{ij}^* \geq 0 > -\delta_{ij}^*$. Furthermore, we have $\nabla \Phi_j(x_1^*, x_2^*, t^*) = 2\, e_{1j}$ for $j \in (I_1 \backslash I_2)(x^*)$ and $\nabla \Phi_j(x_1^*, x_2^*, t^*) = 2\, e_{2j}$ for $j \in (I_2 \backslash I_1)(x^*)$, such that the family of active constraint gradients of (3.30) is

$$\begin{array}{rl} \nabla h_i(x^*) & i \in \{1, \dots, q\}, \\ \nabla g_j(x^*) & j \in I_g(x^*), \\ e_{n+j} & j \in I_1(x^*)^+, \\ 2e_{n+j} & j \in I_1(x^*)^-, \\ e_{n+p+j} & j \in I_2(x^*)^+, \\ 2e_{n+p+j} & j \in I_2(x^*)^-, \end{array}$$

which is linearly independent by the linear independence of the family of gradient vectors of (3.29).

Finally, we show that the SSOSC in x^* for (1.19) implies the SOSC in (x^*, s^*) for $R(\delta^*, t^*)$. Again it is sufficient to proof the SOSC for (3.30). Note that x^* is a stationary point of (3.30) with multipliers μ^*, λ^*, ν_1^*,ν_2^*, and ξ^*. Moreover, as a result of (3.28), (3.23) and (3.24) and the definition of $\Phi(x_1, x_2, t)$, either $\xi_j^* = 0$ or all second derivatives with respect to x of $\Phi_j(x_1^*, x_2^*, t^*)$ vanish. Hence, the part of $\nabla^2_{xx}\mathcal{L}_{R(\delta,t)}(x^*, s^*, \lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*)$ corresponding to $\Phi(x_1, x_2, t)$ vanishes completely (see also the proof of Theorem 2.1). Thus, it remains to show that the set of critical directions is a subset of $\tilde{\mathcal{S}}(x^*, \lambda^*,\mu^*, \hat{\nu}_1, \hat{\nu}_2)$. The conditions $\nabla h_i(x^*)^T d = 0$ for $i \in \{1, \ldots, q\}$ and $\nabla g_j(x^*)^T d = 0$ for $j \in I_g^+(x^*, \lambda^*)$ are directly satisfied. Moreover, if $\hat{\nu}_{1j} \neq 0$, then by (3.19) and (3.21)-(3.24) and the definition of ν_1^* and ξ^*, it follows that either $j \in I_1(x^*, \nu_1^*)^+$ and $\delta_{1j}^* = 0$ or $j \in I_\Phi(x^*, t^*, \xi^*)^+$ and $\delta_{cj}^* = 0$. This, however, implies that either $x_{1j}^* = 0 = -\delta_{1j}^*$ and $\nu_{1j}^* > 0$ or $\Phi_j(x_1^*, x_2^*, t^*) = 0 = -\delta_{cj}^*$ and $\xi_j^* > 0$. Thus for all critical directions d of (3.30) in x^*, in both cases $d_{1j} = 0$ for all $j \in \{1, \ldots, p\}$ with $\hat{\nu}_{1j} \neq 0$. The same implications show that $d_{2j} = 0$ for all $j \in \{1, \ldots, p\}$ with $\hat{\nu}_{2j} \neq 0$ for all critical directions d of (3.30) in x^*. Thus, by view of (3.18) and by Theorem 2.1

$$d^T \nabla_{xx}\mathcal{L}_{R(\delta,t)}(x^*, s^*, \lambda^*, \mu^*, \nu_1^*, \nu_2^*, \xi^*)\, d > 0$$

for all critical directions of (3.30). □

Remark 3.3. Note that the multipliers λ^*, μ^*, ν_1^*, ν_2^*, and ξ^* of Theorem 3.5 correspond to the basic multipliers of Definition 2.2.

Next, we prove a result concerning the inverse direction of Theorem 3.5.

Corollary 3.1. *Suppose $\delta^* \in (\mathbb{R}^+)^{3p}$ satisfies condition (3.25) and $t^* \in (\mathbb{R}^+)^p$. Moreover, assume (x^*, s^*) is a stationary point of $R(\delta^*, t^*)$ with multipliers $v^* = (\lambda^*,\, \mu^*,\, \nu_1^*,\, \nu_2^*,\, \xi^*)$ that satisfies $\min(x_{1j}^*, x_{2j}^*) = 0$ for all $j \in \{1, \ldots, p\}$. Then, x^* is a strongly stationary point of (1.19) with multipliers λ^*, μ^* and*

$$\begin{aligned} \hat{\nu}_{1j} &:= \nu_{1j}^* - \alpha_j^*\, \xi_j^* \\ \hat{\nu}_{2j} &:= \nu_{2j}^* - (2 - \alpha_j^*)\, \xi_j^*\,, \end{aligned} \tag{3.31}$$

with

$$\alpha_j^* := \frac{\partial \Phi_j(x_1^*, x_2^*, t^*)}{\partial x_{1j}}\,.$$

Proof. The feasibility of x^* for (1.19) follows by the feasibility for $R(\delta^*, t^*)$ and the assumption that $\min(x_{1j}^*, x_{2j}^*) = 0$ for all $j \in \{1, \ldots, p\}$. It remains to show that $\nabla_x\mathcal{L}_{MPEC}(x^*, \lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2) = 0$ and to check the conditions on the multipliers $\hat{\nu}_1$ and $\hat{\nu}_2$.
Comparing $\nabla_x\mathcal{L}_{MPEC}(x^*,\, \lambda^*,\, \mu^*,\, \hat{\nu}_1,\, \hat{\nu}_2) = 0$ with (3.14) and considering the values of $\hat{\nu}_1$ and $\hat{\nu}_2$ as defined by (3.31), it follows that $\nabla_x\mathcal{L}_{MPEC}(x^*,\, \lambda^*,\, \mu^*,\, \hat{\nu}_1,\, \hat{\nu}_2) = 0$ is satisfied.

The complementarity conditions concerning the multipliers λ^*, $\hat{\nu}_1$ and $\hat{\nu}_2$ are satisfied by view of (3.15). The nonnegativity of $\hat{\nu}_{1j}$ and $\hat{\nu}_{2j}$ for $j \in (I_1 \cap I_2)(x^*)$ follows by the nonnegativity of ν_1^*, ν_2^* and $\xi^* = 0$ for $j \in (I_1 \cap I_2)(x^*)$, since by $t \in (\mathbb{R}^+)^p$ and Lemma 2.1 it follows that $\Phi_j(x_1^*, x_2^*, t^*) < 0$. □

Our aim is to provide a suitable choice of δ^* and t^*, such that computing a stationary point of $R(\delta^*, t^*)$, we obtain a strongly stationary point of (1.19). At the same time, since we want to apply an Interior Point Method to solve $R(\delta^*, t^*)$, we are interested in a strict interior of the feasible region of $R(\delta^*, t^*)$.

From Corollary 3.1 we can deduce that if δ^* satisfies (3.25) and $t^* \in (\mathbb{R}^+)^p$, then a stationary point x^* of $R(\delta^*, t^*)$ is strongly stationary for (1.19), provided x^* satisfies $\min(x^*_{1j}, x^*_{2j}) = 0$ for all $j \in \{1, \dots, p\}$.

On the one hand, if we choose $\delta^* > 0$ and $t^* > 0$, then (3.25) and $t \in (\mathbb{R}^+)^p$ is satisfied and the feasible region of $R(\delta^*, t^*)$ has a strict interior. By this choice, though, we might not be able to guarantee that a stationary point x^* of $R(\delta^*, t^*)$ satisfies $\min(x^*_{1j}, x^*_{2j}) = 0$ for all $j \in \{1, \dots, p\}$.

On the other hand, if we choose $\delta^* = 0$ and $t^* = 0$, then the feasibility of a stationary point x^* for $R(\delta^*, t^*)$ implies that $\min(x^*_{1j}, x^*_{2j}) = 0$ for all $j \in \{1, \dots, p\}$. However, in this case the feasible region of $R(\delta^*, t^*)$ does not have a strict interior. Hence, we have to decide appropriately which entries of δ^* should be set equal to zero to obtain $\min(x^*_{1j}, x^*_{2j}) = 0$ for all $j \in \{1, \dots, p\}$ and which ones should remain strictly positive to maintain a strict interior.

Theorem 3.5 provides an indication how δ^* should possibly be chosen. Here we take the sign of the multipliers of a strongly stationary point x^* as an indicator which components of δ^* should be equal to zero. However, as we do not know the sign of the multipliers in advance, we start with a strictly positive parameter $\delta^k > 0$ compute a stationary point x^k and decide by the sign of the associated multiplier estimates, which parameter entries we have to reduce.

Accordingly, as we do not know the appropriate values of t^* to guarantee $\min(x_{1j}, x_{2j}) = 0$ for all $j \in \{1, \dots, p\}$ in advance, we start with a strictly positive t^k and update t^k according to the values of the multiplier and the solution estimates we obtained by solving $R(\delta^k, t^k)$.

Next, we describe how to update δ^k and t^k. As mentioned before, we start with a parameter vector (δ^k, t^k) consisting of strictly positive entries. Then we compute a corresponding new stationary point $y_{k+1} := (x^{k+1}, s^{k+1}, \mu^{k+1}, v^{k+1})$, thus it satisfies (3.14)-(3.17) for (δ^k, t^k), and corresponding multiplier estimates of $\hat{\nu}_1^{k+1}$ and $\hat{\nu}_2^{k+1}$ for the strongly stationary point by (confer Corollary 3.1)

$$
\begin{aligned}
\hat{\nu}_{1j}^{k+1} &:= \nu_{1j}^{k+1} - \alpha_j^{k+1}\, \xi_j^{k+1} \\
\hat{\nu}_{2j}^{k+1} &:= \nu_{2j}^{k+1} - (2 - \alpha_j^{k+1})\, \xi_j^{k+1}
\end{aligned} \quad , \tag{3.32}
$$

with $\alpha_j^{k+1} := \partial \Phi_j / \partial x_{1j}\, (x_1^{k+1}, x_2^{k+1}, t^k)$. By these new estimates, we then decide which parameter entries should be reduced:

Let the residual function $r(y, \delta, t)$ of the KKT-conditions be defined as

$$
r(y, \delta, t) := \begin{pmatrix} \nabla_x \mathcal{L}_{R(\delta,t)}(y) \\ Sv \\ h(x) \\ C(x, s, t) + (0, \delta_1, \delta_2, \delta_c)^T \end{pmatrix}
$$

with $y := (x, s, \mu, v)$ and $S := diag(s_i) \in \mathbb{R}^{(m+3p)\times(m+3p)}$. Then the conditions (3.14)-(3.17) are satisfied, if $r(y, \delta, t) = 0$ and $(s, v) \geq 0$. Since the limit should satisfy $r(y^*, \delta^*, t^*) = 0$ with $y^* := (x^*, s^*, \mu^*, v^*)$, we test the optimality of y_{k+1} not by the residual $r(y_{k+1}, \delta^k, t^k)$ but by $r(y_{k+1}, \delta^{k^*}, t^k)$, which is the KKT-residual for $R(\delta^{k^*}, t^k)$ where δ^{k^*} is an auxiliary parameter vector. This parameter vector is assumed to be a better approximation of δ^*, as we set those entries equal to zero that are presumed to converge to zero. Hence,

$$\delta_{ij}^{(k+1)^*} = 0 \quad \text{if} \quad \delta_{ij}^{k+1} < \delta_{ij}^k, \tag{3.33}$$

$$\delta_{ij}^{(k+1)^*} = \delta_{ij}^{k+1} \quad \text{if} \quad \delta_{ij}^{k+1} = \delta_{ij}^k, \tag{3.34}$$

$$\delta_{cj}^{(k+1)^*} = 0 \quad \text{if} \quad \delta_{cj}^{k+1} < \delta_{cj}^k, \tag{3.35}$$

$$\delta_{cj}^{(k+1)^*} = \delta_{cj}^{k+1} \quad \text{if} \quad \delta_{cj}^{k+1} = \delta_{cj}^k. \tag{3.36}$$

As we will see by Theorem 3.6, if the reduction of the corresponding parameter entries is sufficient, then the new iterates y_{k+1} are improved solution estimates, and if we start with an iterate y_k that is close enough to a solution y^* satisfying $r(y^*, \delta^{k^*}, t^k) = 0$, then the auxiliary sequence (δ^{k^*}) remains constant, that is $\delta^{(k+1)^*} = \delta^{k^*}$ for all subsequent $k \in \mathbb{N}$.

In our update algorithm, we will further make use of an upper and lower bound of the KKT-residual $r(y_{k+1}, \delta^{k^*}, t^k)$:

$$\bar{r}_{k+1}^* := \|r(y_{k+1}, \delta^{k^*}, t^k)\|^{1-\zeta} \quad \text{and} \quad \underline{r}_{k+1}^* := \|r(y_{k+1}, \delta^{k^*}, t^k)\|^{1+\zeta} \tag{3.37}$$

for a fixed parameter $0 < \zeta < 1$. The upper bound $\bar{r}_{k+1}^*$ is used to decide which parameter values of (δ^k, t^k) are to reduce and the lower bound $\underline{r}_{k+1}^*$ is used to guarantee sufficient reduction of the corresponding entries. Algorithm 3.2 represents the resulting parameter update-rule as described above.

3.2.2 Local Convergence Analysis

As we want to consider an Interior Point algorithm applied to (3.12), instead of solving the exact Newton equation

$$J(y_k, t^k)\Delta y_k = -r(y_k, \delta^k, t^k)\,,$$

where

$$J(y, t) := \nabla_y r(y, \delta, t)^T = \begin{pmatrix} H(x, s, \mu, v) & & -\nabla h(x) & -\nabla_x C(x, s, t) \\ & V & & S \\ \nabla h(x)^T & & & \\ \nabla_x C(x, s, t)^T & -I & & \end{pmatrix},$$

with

$$H(x, s, \mu, v) := \nabla_{xx}^2 \mathcal{L}_{R(\delta,t)}(x, s, \mu, v)$$

and $V := diag(v_i) \in \mathbb{R}^{(m+3p)\times(m+3p)}$, we consider the solution of the system

$$J(y_k, t^k)\Delta y_k = -r(y_k, \delta^k, t^k, \pi) := -\begin{pmatrix} \nabla_x \mathcal{L}_{R(\delta,t)}(y) \\ Sv - \pi \mathbf{e} \\ h(x) \\ C(x, s, t^k) + (0, \delta^{k^T})^T \end{pmatrix}, \tag{3.38}$$

Algorithm 3.2: Update Rule

Input: δ^k, δ^{k^*}, y_{k+1}

Set fixed parameters $\kappa, \zeta \in (0,1)$.

Compute $\bar{r}^*_{k+1}$ *and* $\underline{r}^*_{k+1}$.

Compute $\hat{\nu}_1^{k+1}$ *and* $\hat{\nu}_2^{k+1}$.

Update δ^k, δ^{k^*} *and* t^k*:*

for $j = 1, \ldots, p$ **do**
    for $i = 1, 2$ **do**
        if $\hat{\nu}_{ij}^{k+1} > \bar{r}^*_{k+1}$ **then**
            $\delta_{ij}^{k+1} \leftarrow \min(\kappa\, \delta_{ij}^k\, , \underline{r}^*_{k+1})$
            $\delta_{ij}^{(k+1)^*} \leftarrow 0$
        else
            $\delta_{ij}^{k+1} \leftarrow \delta_{ij}^k$
            $\delta_{ij}^{(k+1)^*} \leftarrow \delta_{ij}^k$
    if $\hat{\nu}_{1j}^{k+1} < -\bar{r}^*_{k+1}$ ***or*** $\hat{\nu}_{2j}^{k+1} < -\bar{r}^*_{k+1}$ **then**
        $\delta_{cj}^{k+1} \leftarrow \min(\kappa\, \delta_{cj}^k\, , \underline{r}^*_{k+1})$
        $\delta_{cj}^{(k+1)^*} \leftarrow 0$
        if $t_j^k > |x_{1j}^{k+1} - x_{2j}^{k+1}|$ **then**
            $t_j^{k+1} \leftarrow 0.5\, |x_{1j}^{k+1} - x_{2j}^{k+1}|$
    else
        $\delta_{cj}^{k+1} \leftarrow \delta_{cj}^k$
        $\delta_{cj}^{(k+1)^*} \leftarrow \delta_{cj}^k$
        $t_j^{k+1} \leftarrow t_j^k$

where the $r(y_k, \delta^k, t^k, \pi)$ corresponds to the KKT-residual of the barrier problem

$$\begin{aligned} \min \quad & f(x) - \pi \textstyle\sum_{j=1}^{m+3p} \ln(s_j) \\ \text{subject to} \quad & h(x) = 0 \\ & C(x, s, t^k) + (0, \delta^{k^T})^T = 0\,. \end{aligned} \tag{3.39}$$

Furthermore, as we presumably have to truncate the computed Newtonstep Δy_k to maintain the strict positivity of s and v, we introduce a steplength parameter σ. Motivated by DeMiguel et al. (or further by Yamashita and Yabe [YY96]), in our algorithm we use the steplength

$$\sigma = \min\left(\min\left(1, \gamma \min_{\Delta s_j < 0} \frac{-s_j}{\Delta s_j}\right), \min\left(1, \gamma \min_{\Delta v_j < 0} \frac{-v_j}{\Delta v_j}\right)\right)\,. \tag{3.40}$$

Algorithm 3.3: Interior Point Algorithm for Two-Sided Relaxed MPEC

Initialize variables (x_0, s_0, μ_0, v_0) *with* $s_0, v_0 > 0$ *and relaxation parameters* δ_0, t_0*; set* $\delta^{0^*} = \delta^0$*; choose barrier parameter* $\pi_0 > 0$*, fixed parameters* $0 < \kappa, \zeta, \bar{\gamma} < 1$*, a starting steplength parameter* $\bar{\gamma} \leq \gamma_0 < 1$ *and a convergence tolerance* ϵ*.*

repeat

- *Solve (3.38) for* Δy_k*.*
- *Determine steplength* σ_k *by (3.40)*
- *Compute the new iterate* y_{k+1}*:*

  $$y_{k+1} \leftarrow y_k + \sigma_k \Delta y_k.$$
- *Update relaxation parameters* δ_k, t_k *and* δ^{k^*} *by Algorithm 3.2.*
- *Update barrier and steplength parameter by*

  $$\pi_{k+1} \leftarrow \min(\kappa\pi_k, \underline{r}^*_{k+1})$$

  $$\gamma_{k+1} \leftarrow \max(\bar{\gamma}, 1 - \pi_{k+1})$$
- $k \leftarrow k + 1$

until $\|r(y_{k+1}, \delta^{(k+1)^*}, t^{k+1}, 0)\| < \epsilon$

For the local convergence analysis we will make use of the following general assumptions. Furthermore, x^* is supposed to be a strongly stationary point of (1.19).

Assumptions 3.1.

(A1) The second derivatives of the functions f, h *and* g *are Lipschitz continuous.*

(A2) θ *satisfies Assumptions 2.1 and* θ'' *is Lipschitz-continuous on an open set* $\mathcal{D} \supseteq [-1, 1]$

(A3) The point x^* *satisfies the MPEC-LICQ for (1.19).*

(A4) It holds $g_i(x^*) + \lambda_i^* > 0$ *for all* $i \in I_g(x^*)$*, with* $(\lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2)$ *being the multipliers of* x^**.*

(A5) The point $(x^*, \lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2)$ *satisfies the SSOSC for (1.19).*

Before we start with the convergence analysis for Algorithm 3.3, we first need two more auxiliary results.

Lemma 3.11. *Let* x^* *be a strongly stationary point of (1.19) and let* (δ^*, t^*) *satisfy (3.19)-(3.24). Let* $y^* := (x^*, s^*, \mu^*, v^*)$ *with* s^* *and* v^* *satisfying the conditions of Theorem 3.5. Furthermore, assume (A1)-(A5) hold in* x^**. Then there exists an* $\epsilon > 0$ *and an* $L > 0$ *such that*

1. $\|r(y, \delta^*, t^*) - r(y^*, \delta^*, t^*)\| \leq L\,\|y - y^*\|$ *for all* $y \in B_\epsilon(y^*)$ *and*
2. $\|J(y, t^*) - J(y^*, t^*)\| \leq L\,\|y - y^*\|$ *for all* $y \in B_\epsilon(y^*)$*.*
3. $J(y, t^*)$ *is nonsingular for all* $y \in B_\epsilon(y^*)$*.*

Proof. By Assumption (A1) the second derivatives with respect to x and s of $h_j(x)$ are Lipschitz continuous for all $j \in \{1,\dots,q\}$ as well as of $C_j(x,s,t^*)$ for $j \in \{1,\dots,m+2p\}$. Hence, it remains to show that $C_j(x,s,t^*)$ is locally twice continuously differentiable with respect to x and s for all $j \in \{m+2p+1,\dots,m+3p\}$ and moreover that the Lipschitz continuity of θ'' on $\mathcal{D} \supseteq [-1,1]$ implies the Lipschitz continuity of the second derivative of $C_j(x,s,t^*)$. Since $C_j(x,s,t^*)$ is linear in s we only have to consider the Lipschitz-continuity of $\nabla^2_{xx} C_j(x,s,t^*)$ for all $j \in \{m+2p+1,\dots,m+3p\}$.

We consider two cases: first assume that $t^*_j = 0$, then $\Phi_j(x_1,x_2,t^*) = x_{1j} + x_{2j} - |x_{1j} - x_{2j}|$. Thus all second partial derivatives of $\Phi_j(x_1,x_2,t^*)$ vanish (hence the second derivative of $\Phi_j(x_1,x_2,t^*)$ is Lipschitz continuous) for all x that satisfy $|x_{1j} - x_{2j}| \neq 0$. However, as t^* satisfies (3.23) and (3.24) and x^* is a strongly stationary point, there exists an ϵ such that $|x_{1j} - x_{2j}| \neq 0$ for all $x \in B_\epsilon(x^*)$ and for all $j \in \{1,\dots,p\}$ with $t^*_j = 0$.

Next, suppose $t^*_j > 0$, then (A2) and Lemma 3.6 imply that the second derivative of $\Phi_j(x_1,x_2,t^*)$ is Lipschitz continuous for all $x \in B_\epsilon(x^*)$. Thus, all $C_j(x,s,t^*)$ are twice Lipschitz-continuously differentiable with respect to x and s and the first two statements follow directly.

By (A1)-(A5) and Theorem 3.5 it follows that the LICQ, the SOSC and strict complementarity are satisfied in (x^*, s^*, μ^*, v^*). Hence (see for example Theorem 14 of [FC68]), $J(y^*,t^*)$ is nonsingular. The perturbation lemma (see for example Lemma 5.23 in [GK02]) then implies that there exists an $\epsilon > 0$ such that $J(y,t^*)$ is nonsingular for all $y \in B_\epsilon(y^*)$. □

Lemma 3.12. *Let x^* be a strongly stationary point of (1.19) with multipliers (λ^*, μ^*, $\hat\nu_1$, $\hat\nu_2$). Then there exists an $\epsilon > 0$ such that if $x^k \in B_\epsilon(x^*)$, $t^k_j > 0$ for all $j \in \{1,\dots,p\}$ and $t^k_j \le 0.5\,|x^k_{1j} - x^k_{2j}|$ for all $j \in \{1,\dots p\}$ with either $\hat\nu_{1j} < 0$ or $\hat\nu_{2j} < 0$, then t^k satisfies (3.23) and (3.24).*

Proof. To prove that t^k satisfies (3.23) and (3.24), we have to show that $|x^*_{1j} - x^*_{2j}| \ge t^k_j$ for all $j \in \{1,\dots p\}$ with either $\hat\nu_{1j} < 0$ or $\hat\nu_{2j} < 0$. Since x^* is supposed to be strongly stationary, the condition $\hat\nu_{1j} < 0$ or $\hat\nu_{2j} < 0$ implies that either $j \in (I_1 \backslash I_2)(x^*)$ or $j \in (I_2 \backslash I_1)(x^*)$. By $x^k \in B_\epsilon(x^*)$, we therefore have

$$|x^k_{1j} - x^k_{2j}| \ge |x^*_{1j} - x^*_{2j}| - 2\epsilon > 0\,,$$

if $0 < \epsilon$ is small enough. Thus,

$$|x^*_{1j} - x^*_{2j}| \ge |x^k_{1j} - x^k_{2j}| - 2\epsilon \ge 0.5|x^k_{1j} - x^k_{2j}| \ge t^k_j$$

holds, if $\epsilon \le |x^*_{1j} - x^*_{2j}|/6$. Hence, if ϵ is small enough, then t^k satisfies (3.23) and (3.24). □

Now we can prove the following theorem concerning the values of δ^{k^*}, determined by the Update Rule, in the vicinity of a strongly stationary point x^*.

Theorem 3.6. *Let x^* be a strongly stationary point of (1.19) with multipliers (λ^*, μ^*, $\hat\nu_1$, $\hat\nu_2$) and suppose (A1)-(A5) hold. Let further δ^{k^*} satisfy (3.19)-(3.22) and assume that $\delta^{k^*}_j = \delta^k_j > 0$ for all $j \in \{1,\dots,p\}$ with $\delta^{k^*}_j \neq 0$. Moreover, assume that $t^k_j > 0$ for all $j \in \{1,\dots,p\}$ and $t^k_j \le 0.5\,|x^k_{1j} - x^k_{2j}|$ for all $j \in \{1,\dots,p\}$ with either $\hat\nu_{1j} < 0$ or $\hat\nu_{2j} < 0$.*

Let $y_k^* = (x^*, s^{k^*}, \mu^*, v^*)$ be the corresponding stationary point of $R(\delta^{k^*}, t^k)$, and suppose that $\|y_{k+1} - y_k^*\| < \|y_k - y_k^*\|$. Then there exists an $\epsilon > 0$ such that if $y_k \in B_\epsilon(y_k^*)$, then Algorithm 3.2 yields $\delta^{(k+1)^*} = \delta^{k^*}$ and $t^{k+1} = t^k$.

Proof. First note that $r(y_k^*, \delta^{k^*}, t^k) = 0$ by the definition of y_k^*. Next, we choose $\epsilon > 0$ small enough such that Lemma 3.12 implies that t^k satisfies (3.23) and (3.24).

Furthermore, by the assumptions of the theorem and by Lemma 3.11, it follows that there exists an $\epsilon > 0$, such that we can find an $L > 0$ with

$$\|r(y_{k+1}, \delta^{k^*}, t^k)\| = \|r(y_{k+1}, \delta^{k^*}, t^k) - r(y_k^*, \delta^{k^*}, t^k)\| \leq L\,\|y_{k+1} - y_k^*\|$$

for all $y_{k+1} \in B_\epsilon(y_k^*)$. Defining $\bar{c} := L^{1-\zeta}$, we obtain

$$\bar{r}_{k+1}^* = \|r(y_{k+1}, \delta^{k^*}, t^k)\|^{1-\zeta} \leq \bar{c}\|y_{k+1} - y_k^*\|^{1-\zeta}\,. \tag{3.41}$$

By Lemma 3.11 we also know that $r(y_{k+1}, \delta^{k^*}, t^k)$ is continuously differentiable with respect to y in the vicinity of y_k^* and that $J(y_k^*, t^k)$ is regular, hence $\|J(y_k^*, t^k)^{-1}\| \leq \tilde{c}$. Then, by Taylor approximation we get

$$r(y_{k+1}, \delta^{k^*}, t^k) - r(y_k^*, \delta^{k^*}, t^k) = J(y_k^*, t^k)(y_{k+1} - y_k^*) + \mathcal{O}(\|y_{k+1} - y_k^*\|^2)\,.$$

Hence, there exists an $M > 0$ so that

$$\|y_{k+1} - y_k^*\| \leq \|J(y_k^*, t^k)^{-1}\|\|r(y_{k+1}, \delta^{k^*}, t^k)\| + M\|y_{k+1} - y_k^*\|^2\,.$$

This implies that if $\|y_k - y_k^*\|$ is small enough, then there exists a constant $\underline{c} > 0$ with

$$\underline{c}\|y_{k+1} - y_k^*\|^{1-\zeta} \leq \|r(y_{k+1}, \delta^{k^*}, t^k)\|^{1-\zeta} = \bar{r}_{k+1}^*. \tag{3.42}$$

Let $\hat{\nu}_{ij}^{k+1}$ and $\hat{\nu}_{ij}$ for $i = 1, 2$ and $j \in \{1, \ldots, p\}$ be defined by (3.32). Then,

$$\begin{aligned} |\hat{\nu}_{1j}^{k+1} - \hat{\nu}_{1j}| &= |(\nu_{1j}^{k+1} - \alpha_j^{k+1}\xi_j^{k+1}) - (\nu_{1j}^* - \alpha_j^*\xi_j^*)| \\ &\leq |\nu_{1j}^{k+1} - \nu_{1j}^*| + |\alpha_j^{k+1}\xi_j^{k+1} - \alpha_j^*\xi_j^*| \\ &\leq |\nu_{1j}^{k+1} - \nu_{1j}^*| + |\alpha_j^{k+1}||\xi_j^{k+1} - \xi_j^*| + |\alpha_j^{k+1} - \alpha_j^*||\xi_j^*| \\ &\leq 3\,\|y_{k+1} - y_k^*\| + |\alpha_j^{k+1} - \alpha_j^*||\xi_j^*|\,, \end{aligned} \tag{3.43}$$

since $|\alpha_j^{k+1}| \leq 2$. Accordingly,

$$|\hat{\nu}_{2j}^{k+1} - \hat{\nu}_{2j}| \leq 3\,\|y_{k+1} - y_k^*\| + |\alpha_j^{k+1} - \alpha_j^*||\xi_j^*|\,.$$

To prove, that Algorithm 3.2 yields $\delta^{(k+1)^*} = \delta^{k^*}$, we consider the three cases $\hat{\nu}_{ij} > 0$, $\hat{\nu}_{ij} < 0$ and $\hat{\nu}_{ij} = 0$ $(i = 1, 2)$ individually. As the arguments for $\hat{\nu}_{1j}$ and $\hat{\nu}_{2j}$ are similar, we will consider these cases only for $\hat{\nu}_1$.
First, assume $\hat{\nu}_{1j} > 0$, then $\xi^* = 0$, because δ^{k^*} satisfies (3.19)-(3.22). Hence, by (3.43)

$$|\hat{\nu}_{1j}^{k+1} - \hat{\nu}_{1j}| \leq 3\,\|y_{k+1} - y_k^*\| \tag{3.44}$$

and therefore by (3.41)

$$\hat{\nu}_{1j}^{k+1} = \hat{\nu}_{1j} + (\hat{\nu}_{1j}^{k+1} - \hat{\nu}_{1j}) \geq \hat{\nu}_{1j} - 3\,\|y_{k+1} - y_k^*\| > \bar{c}\|y_{k+1} - y_k^*\|^{1-\zeta} \geq \bar{r}_{k+1}^* \tag{3.45}$$

if $\|y_{k+1} - y_k^*\|$ is small enough.

Next, suppose $\hat{\nu}_{1j} = 0$. Then (3.44) holds, since $\xi_j^* = 0$, and we get

$$|\hat{\nu}_{1j}^{k+1}| = |\hat{\nu}_{1j}^{k+1} - \hat{\nu}_{1j}| \leq 3\,\|y_{k+1} - y_k^*\| < \underline{c}\|y_{k+1} - y_k^*\|^{1-\zeta} \leq \bar{r}_{k+1}^* \tag{3.46}$$

by (3.42) and again for $\|y_{k+1} - y_k^*\|$ being small enough.

Furthermore, notice that by (3.45) and (3.46) t_j^k might only be changed by Algorithm 3.2, if either $\hat{\nu}_{1j} < 0$ or $\hat{\nu}_{2j} < 0$, thus either $j \in (I_1 \backslash I_2)(x^*)$ or $j \in (I_2 \backslash I_1)(x^*)$. Therefore, if ϵ is small enough, then (confer the proof of Lemma 3.12)

$$|x_{1j}^{k+1} - x_{2j}^{k+1}| \geq |x_{1j}^* - x_{2j}^*| - 2\epsilon \geq |x_{1j}^k - x_{2j}^k| - 4\epsilon > 0.5|x_{1j}^k - x_{2j}^k| \geq t_j^k\,.$$

Hence, for all $j \in \{1, \ldots, p\}$ with either $\hat{\nu}_{1j} < 0$ or $\hat{\nu}_{2j} < 0$

$$|x_{1j}^{k+1} - x_{2j}^{k+1}| > t_j^k\,, \tag{3.47}$$

such that Algorithm 3.2 yields $t_j^{k+1} = t_j^k$ for all $j \in \{1, \ldots, p\}$.

Now concerning the update of $\delta_j^{k^*}$ for $j \in \{1, \ldots, p\}$ with $\hat{\nu}_{1j} < 0$, note that by the assumptions of the theorem and by (3.47) we have $\alpha_j^{k+1} = 2 = \alpha_j^*$. Therefore, inequality (3.44) holds by (3.43). Hence, again for $\|y_{k+1} - y_k^*\|$ being small enough and by (3.41)

$$\begin{aligned} \hat{\nu}_{1j}^{k+1} &\leq \hat{\nu}_{1j}^{k+1} - (\hat{\nu}_{1j}^{k+1} - \hat{\nu}_{1j}) + 3\,\|y_{k+1} - y_k^*\| = \hat{\nu}_{1j} + 3\,\|y_{k+1} - y_k^*\| \\ &< -\bar{c}\|y_{k+1} - y_k^*\|^{1-\zeta} \leq -\bar{r}_{k+1}^* \end{aligned} \tag{3.48}$$

Then (3.45), (3.46) and (3.48) imply $\delta_j^{(k+1)^*} = 0$ if and only if $\delta_j^{k^*} = 0$ which further implies that $\delta_j^{(k+1)^*} = \delta_j^{k^*}$ for all $j \in \{1, \ldots, 3p\}$. □

Remark 3.4. Note, that by Theorem 3.6 it also follows that $y_{k+1}^* = y_k^*$.

The following theorem now shows that, although we slightly change the relaxed problem $R(\delta^k, t^k)$ in each iteration by updating the relaxation parameters δ^k and t^k, we can still achieve locally superlinear convergence.

Theorem 3.7. *Let $(x^*, \lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2)$ be a strongly stationary point of (1.19) and suppose (A1)-(A5) hold. Assume that δ^{k^*} satisfies (3.19)-(3.22) and $\delta_j^{k^*} = \delta_j^k > 0$ for all $j \in \{1, \ldots, p\}$ with $\delta_j^{k^*} > 0$ and let t^k be as in Theorem 3.6. Furthermore, let $y_k^* = (x^*, s_k^*, \mu^*, v^*)$ be the stationary point of $R(\delta^{k^*}, t^k)$. Then, there exists an $\epsilon > 0$ and a $C > 0$, such that if Algorithm 3.3 is started with $y_0 \in B_\epsilon(y_k^*)$ and*

$$\begin{aligned} \|\delta^k - \delta^{k^*}\| &< C\,\|y_k - y_k^*\|^{1+\zeta} && (3.49) \\ \pi_k &< C\,\|y_k - y_k^*\|^{1+\zeta} && (3.50) \\ 1 - \gamma_k &< C\,\|y_k - y_k^*\|^{1+\zeta} && (3.51) \end{aligned}$$

then $y^*_{k+\ell} = y^*_k$ *for all* $\ell \in \mathbb{N}$ *and*

$$\|y_{k+1} - y^*_k\| = o\left(\|y_k - y^*_k\|\right)$$

for all subsequent $k \in \mathbb{N}$.

Proof. The main objective of this proof is to show, that $\|y_{k+1} - y^*_k\| = o(\|y_k - y^*_k\|)$, as we can then deduce that the theorem can be applied for each subsequent iterate $y_{k+\ell}$ with $\ell \in \mathbb{N}$ and we hence obtain the superlinear convergence for the complete sequence $(y_{k+\ell})$. By (A1)-(A5), Theorem 3.5 and Lemma 3.11 we conclude that $r(y, \delta^k, t^k)$ is differentiable and $J(y, t^k)$ is nonsingular for $\|y - y^*_k\|$ sufficiently small. Furthermore, for such y there exists a constant $\tilde{c} > 0$ with $\|J(y, t^k)^{-1}\| \leq \tilde{c}$.

Let $r^*_k := r(y_k, \delta^{k^*}, t^k, 0)$ and define

$$\varrho^*_k := (0, \pi_k\,\mathbf{e}, 0, \delta^{k^*} - \delta^k)\,,$$

then $r(y_k, \delta^k, t^k, \pi_k) = r^*_k - \varrho^*_k$ and

$$\begin{aligned} y_{k+1} - y^*_k &= y_k - y^*_k - \sigma_k\, J(y_k, t^k)^{-1} r(y_k, \delta^k, t^k, \pi_k) \\ &= (1 - \sigma_k)(y_k - y^*_k) \\ &\quad + \sigma_k J(y_k, t^k)^{-1} \left(J(y_k, t^k)(y_k - y^*_k) - r^*_k + \varrho^*_k\right) \\ &= (1 - \sigma_k)(y_k - y^*_k) + \sigma_k J(y_k, t^k)^{-1} \varrho^*_k \\ &\quad + \sigma_k J(y_k, t^k)^{-1} \left(J(y_k, t^k)(y_k - y^*_k) - r^*_k\right)\,. \end{aligned} \tag{3.52}$$

In the following, we seek bounds depending on $\|y_k - y^*_k\|$ for each term of the right-hand side of (3.52). By the assumptions, Theorem 3.5 and Lemma 5 of [YY96] it follows that for our choice of the steplength σ_k there exists a positive constant c_1 such that $0 \leq 1 - \sigma_k \leq 1 - \gamma_k + c_1\|\Delta y_k\|$. Hence,

$$\|(1 - \sigma_k)(y_k - y^*_k)\| \leq (1 - \gamma_k + c_1\|\Delta y_k\|)\|y_k - y^*_k\| \tag{3.53}$$

As $\|J(y, t^k)^{-1}\| \leq \tilde{c}$, we also have

$$\|\Delta y_k\| \leq \|J(y, t^k)^{-1}\|\|r(y_k, \delta^k, t^k, \pi_k)\| \leq \tilde{c}\left(\|r^*_k\| + \|\varrho^*_k\|\right) \tag{3.54}$$

Furthermore, by $r(y^*_k, \delta^{k^*}, t^k, 0) = 0$ and by Lemma 3.11 it follows that

$$\|r^*_k\| \leq \|r^*_k - r(y^*_k, \delta^{k^*}, t^k, 0)\| \leq c_2\|y_k - y^*_k\| \tag{3.55}$$

and by (3.49) and (3.50)

$$\|\varrho^*_k\| \leq \|\delta^k - \delta^{k^*}\| + \|\pi_k\,\mathbf{e}\| \leq c_3\|y_k - y^*_k\|^{1+\zeta}\,. \tag{3.56}$$

Then, substituting (3.54)-(3.56) in (3.53) we get an upper bound for the first term of the right-hand side of (3.52)

$$\|(1 - \sigma_k)(y_k - y^*_k)\| \leq (C + \tilde{c}\,c_1 c_3)\|y_k - y^*_k\|^{2+\zeta} + \tilde{c}\,c_1 c_2\|y_k - y^*_k\|^2\,. \tag{3.57}$$

The second term of the right-hand side of (3.52) is bounded by $\|J(y_k, t^k)^{-1}\| \le \tilde{c}$, the boundedness of σ_k and (3.56), thus

$$\|\sigma_k J(y_k, t^k)^{-1} \varrho_k^*\| \le c_4 \|y_k - y_k^*\|^{1+\zeta} . \tag{3.58}$$

The last term of the right-hand side of (3.52) can be bounded using Taylor approximation and again $\|J(y_k, t^k)^{-1}\| \le \tilde{c}$

$$\left\| \sigma_k J(y_k, t^k)^{-1} (J(y_k, t^k)\, (y_k - y_k^*) - r_k^*) \right\| \le c_5 \|y_k - y_k^*\|^2 . \tag{3.59}$$

Finally, substituting (3.57), (3.58) and (3.59) into (3.52) we obtain the upper bound for $\|y_{k+1} - y_k^*\|$:

$$\|y_{k+1} - y_k^*\| \le \bar{c}\, \|y_k - y_k^*\|^{1+\zeta} \tag{3.60}$$

This, however implies that $\|y_{k+1} - y_k^*\| < \|y_k - y_k^*\|$, provided $\|y_k - y_k^*\|$ is small enough and by Theorem 3.6 we get $\delta^{(k+1)^*} = \delta^{k^*}$ as well as $t^{k+1} = t^k$, thus $y_{k+1}^* = y_k^*$.

Furthermore, if $\|y_k - y_k^*\| < \epsilon$ and ϵ is small enough, then

$$\|y_{k+1} - y_k^*\| \le \bar{c}\, \|y_k - y_k^*\|^{1+\zeta} \le \bar{c}\, \epsilon^{1+\zeta} \le \epsilon$$

and, by Algorithm 3.2 and Step 6 of Algorithm 3.3, it follows with Lemma 3.11 that there exists a constant $L > 0$ such that

$$\pi_{k+1} \;\le\; \underline{r}_k^* \;\le\; L\, \|y_{k+1} - y_k^*\|^{1+\zeta}$$

and

$$\|\delta^{k+1} - \delta^{(k+1)^*}\| \le L\, \|y_{k+1} - y_k^*\|^{1+\zeta} .$$

Hence, we can apply Theorem 3.7 iteratively to each new $y_{k+\ell}$ with $\ell \in \mathbb{N}$, which completes the proof. □

Since the method we presented here forms a variant of the method proposed by DeMiguel et al. the algorithms and in parts the theorems (and their proofs) of this section resemble those ones presented in [DFNS05].

Remark 3.5. The numerical results DeMiguel et al. report in [DFNS05] for the problems *ex9.2.2*, *ralph1* and *scholtes4* (confer Section 4.3) give reason to the conjecture that, since they use

$$x_{1j} \ge -\delta_{1j} , \qquad x_{2j} \ge -\delta_{2j} , \qquad x_{1j} x_{2j} \le \delta_{cj} .$$

as relaxed reformulation of the complementarity condition, they are confronted with the same problems concerning MPECs that do not have strongly stationary solutions (as are for example *ex9.2.2*, *ralph1* and *scholtes4*) that we discussed before in Section 2.5 for the relaxed and the exact bilinear reformulation. In view of our discussion in Section 2.5, we further conjecture that we will not face these difficulties if the new relaxation scheme is used instead of the bilinear reformulation, thus if we solve (3.12).

4 Numerical Experience

In this chapter we present the numerical results we obtained using the new relaxation scheme introduced and analysed in Chapter 2. We have tested the relaxation scheme on a test set taken from the collection of test problems called MacMPEC, which is maintained by Leyffer and freely available [Ley00]. We used the modeling language `AMPL` and the NLP solver `filterSQP` to solve the problems. In Section 4.1 we will briefly introduce both software packages followed by some information about the test problems. In Section 4.2 we first analyse a simple outer algorithm, before we explain two failures this simple algorithm produces for some of the test problems. We then present a modified algorithm in which we try to avoid or handle these failures. The second part of Section 4.2 gives then information about the numerical results of the modified algorithm in particular in comparison to the simple outer algorithm. Finally, in the last section we compare the results of the modified algorithm to results we obtained for the relaxed bilinear approach by Scholtes [Sch01] and for the exact bilinear approach by Fletcher et al. [FLRS06], both described in Section 1.4.

4.1 Software and Test Problems

In the following, we will briefly introduce the software and the set of test problems we used for our numerical experiments.

FilterSQP

To test the proposed relaxation scheme, we used the SQP solver `filterSQP`, which was developed by Fletcher and Leyffer [FL98]. Other SQP solvers namely `SNOPT` [GMS02, GMS05] and `DONLP2` [Spe98], were also tested on a preliminary subset of test problems, though as the outcomes were approximately the same we concentrated on `filterSQP`. In `filterSQP` the SQP method is combined with a trust-region and a filter approach. Filter methods provide an alternative to penalty function methods to promote global convergence as they allow the full Newton step and one does not need to find a suitable penalty parameter [FLT07]. The difference of a filter method compared to a penalty function method can briefly be explained as follows:

Solving a Nonlinear Programming problem of the form

$$\begin{array}{rl} \text{minimize} & f(x) \\ \text{subject to} & c(x) \leq 0 \end{array}$$

comprises two targets: the minimization of the objective function $f(x)$ and of the constraint violation, which could be measured for example by $h(x) := \| [c(x)]^+ \|$. Using a penalty function these two objectives are combined into one single function and the second one is weighted with an increasing penalty parameter as feasibility has to be achieved in the solution. Instead of combining it, filter methods treat the NLP as a *biobjective optimization problem* [FLT07], where either the objective function value $f(x)$ or some measure of the infeasibility of x has to be decreased sufficiently, compared to a test set of previously determined iterates called the *filter*.

Next, we explain the software package `filterSQP` and the filter method used therein more explicitly. `filterSQP` solves NLPs of the form [FL98]

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & lb_x \le x \le ub_x \\ & lb_c \le c(x) \le ub_c \end{aligned} \tag{4.1}$$

by solving a sequence of quadratic approximations of (4.1) in the current iterate x^k within a trust-region that is determined by the condition $\|d\|_\infty \le \rho$, with ρ denoting the trust-region radius. The QPs thus have the form

$$\begin{aligned} \text{minimize} \quad & q_k(d) \\ \text{subject to} \quad & lb_x \le x^k + d \le ub_x \\ & lb_c \le c(x^k) + \nabla c(x^k)^T d \le ub_c \\ & \|d\|_\infty \le \rho\,, \end{aligned} \tag{4.2}$$

where $q_k(d) := \nabla f(x^k)^T d + \frac{1}{2} d^T \nabla^2_{xx} \mathcal{L}(x^k, \lambda^k) d$ corresponds to the quadratic approximation of the Lagrangian function $\mathcal{L}(x, \lambda)$ of (4.1). In contrast to other solver `filterSQP` uses the exact Hessian $\nabla^2_{xx}\mathcal{L}(x^k, \lambda^k)$. The QPs (4.2) are solved by `bqpd`, which is a robust QP solver that is based on a null-space active set method (for more information see [Fle00]).

The solution d^k of (4.2) gives a next trial iterate $x^{k+1} = x^k + d^k$ and it is tested, if x^{k+1} can be accepted by the filter. The filter consists of a list of pairs $(f(x^\ell), h(x^\ell))$ of previous iterates x^ℓ, that are not dominated by any other pair. The concept of domination was adopted from multiobjective optimization and is defined in [FL02a] as follows:

Definition 4.1. A pair $(f(x^\ell), h(x^\ell))$ is said to dominate another pair $(f(x^k), h(x^k))$ if and only if both $f(x^\ell) \le f(x^k)$ and $h(x^\ell) \le h(x^k)$.

Concerning the basic filter SQP algorithm (Algorithm 1 in [FL02a]), the trial iterate x^{k+1} will be accepted by the filter, if the pair $(f(x^{k+1}),\ h(x^{k+1}))$ is not dominated by any other pair of the current filter. Algorithmic extensions of the basic filter SQP algorithm that are incorporated in `filterSQP` concern a *Second Order Correction (SOC) step*, an upper bound on the constraint violation, the elemination of *blocking entries* from the filter, a *sufficient reduction test* and a *North-West* and *South-East* corner rule. However, we will not further discuss these extensions here but refer the interested reader to [FL02a].

If the trial point x^{k+1} is accepted by the filter, it is chosen to be the new iterate. The pair $(f(x^{k+1}),\ h(x^{k+1}))$ is then added to the filter and pairs $(f(x^\ell),\ h(x^\ell))$ that are dominated by $(f(x^{k+1}), h(x^{k+1}))$ are removed from the filter. If x^{k+1} is rejected by the filter, then the trail step d^k is discarded, the trust-region radius is reduced and the QP (4.2) is solved again.

Reducing the trust-region radius, though, might cause an infeasible QP, if the current iterate is not feasible for (4.1). Therefore, `filterSQP` incorporates a *feasibility restoration phase*, which aims to minimize the constraint violation by applying a trust region SQP method to solve the problem

$$\begin{aligned} &\text{minimize} && \textstyle\sum_{j\in\mathcal{J}}(c_j(x))^+ \\ &\text{subject to} && c_j(x) \le 0 \qquad j \in \mathcal{J}^\perp\,, \end{aligned} \tag{4.3}$$

where the sets $\mathcal{J}$ and $\mathcal{J}^\perp$ partition the nonlinear constraints into those ones that cannot be satisfied for the current QP (that is $c_j(x^k) + \nabla c_j(x^k)^T d > 0, \quad j \in \mathcal{J}$) and those that can be satisfied. For more details about the restoration phase and `filterSQP` in general, we refer to [FL98] and [FL02a].

The MacMPEC collection of MPECs [Ley00], we used for the testings of our algorithms, is a collection of MPECs formulated in `AMPL` (*A Mathematical Programming Language*) [FGK03]. Therefore, we retained this algebraic modeling language to reformulate and solve the test problems.

AMPL

Using a modeling language like `AMPL` has the advantage that we can formulate the problems in an algebraic form close to the mathematical notation, instead of the computational form that is expected by the specific solver we intend to apply to our test problems. The translation into the solver's specific form is done by the modeling language. This is not only in particular advantageous if there is a large list of problems that have to be changed, but also if one wants to apply different solvers to the same problem.

`AMPL` translates the problem into a representation that suits as input for many solvers. It communicates the problem's data to a solver by writing suitable files, which are then read in by the solver. The chosen solver is run as a separate program and is invoked by `AMPL`'s `solve` command. Finally, `AMPL` expects the solver to write a solution file with a termination message, which is then read in by `AMPL` and used for the output of the problem's solution. For further information on `AMPL` we refer the reader to [FGK90], [Gay97] and the book [FGK03].

We solved the test problems on a computer with 4 Dual Core AMD Opteron processors with 2.2 GHz and 32 GB RAM running a 64 bit version of Red Hat Linux (release 5.1 (Tikanga)). Furthermore, we used the GNU C compiler `gcc` (gcc version 4.1) for the `AMPL` Library and the `AMPL` solver interface for `filterSQP` with the option `-O` and we used the GNU Fortran 77 compiler `g77` (gcc version 3.4) for `filterSQP` again with the option `-O`.

Test Problem Set MacMPEC

In order to test our new relaxation scheme, we solved 133 MPECs, that we have taken from MacMPEC, which is a collection of test problems maintained by S. Leyffer [Ley00]. The test problems are taken from a variety of sources as for example [Sch01], [KOZ98],[FJQ99] or [JR99]. The problems vary in their size, type and origin, in other words academic or real life problem. Table 4.1 displays a summary of some main features of the problems contained in our set of test problems.

	nr. of variables			total nr. of constr.			nr. of compl. constr.		
type	0-9	10-99	≥ 100	0-9	10-99	≥ 100	0-9	10-99	≥ 100
LL	4	32	16	3	29	20	16	22	14
LQ	–	1	–	–	1	–	1	–	–
QL	15	14	14	15	14	14	22	10	11
QQ	3	–	–	3	–	–	3	–	–
L\ Q U	8	–	1	8	–	1	8	–	1
O	1	14	10	1	13	11	12	9	4
sum	31	61	41	30	57	46	62	41	30

Table 4.1: Short description of the test set

The abbreviations of the first column have the following meanings:

LL :	Linear objective function and linear constraint functions
LQ :	Linear objective function and quadratic constraint functions
QL :	Quadratic objective function and linear constraint functions
QQ :	Quadratic objective function and quadratic constraint functions
L\QU :	Linear or quadratic objective function and no general constraints
O :	Other types of Programs.

The classification has been taken from [Ley00], where one can find it in more detail. The final numbers of variables and constraints vary from the numbers given in Table 4.1, because the number of constraints for $R(t)$ include three more constraints for each complementarity condition (two bound constraints and one nonlinear constraint). However, the exact numbers can be found in the Table A.2 - Table A.4 in the Appendix. Moreover, in some cases we had to add slack variables or additional constraints, in order to obtain an MPEC of the form (1.19). For example, if a problem contains a complementarity constraint of the form

$$0 \quad \leq \quad G(x) \quad \perp \quad H(x) \quad \geq 0\,, \tag{4.4}$$

where $G(x)$ and $H(x)$ are two scalar functions, then we introduced slack variables, such that (4.4) becomes

$$\begin{aligned} &G(x) - s_1 = 0 \\ &H(x) - s_2 = 0 \\ &0 \;\leq\; s_1 \perp s_2 \;\geq\; 0\,. \end{aligned}$$

We also added variables and constraints in the case that the original MPEC contains a *mixed complementarity constraint* which can be described by the variational inequality

$$l \;\leq\; G(x) \;\leq\; u \qquad \text{and} \qquad H(x)(y - G(x)) \geq 0 \quad \forall\, y \in [l, u] \tag{4.5}$$

and corresponds to the conditions

$$\begin{array}{llll} & G(x) = l & \text{and} & H(x) \geq 0 \\ \text{or} & G(x) = u & \text{and} & H(x) \leq 0 \\ \text{or} & l < G(x) < u & \text{and} & H(x) = 0 \end{array}$$

(see also [FFG99]). We restated (4.5) by adding auxiliary variables s_{ij} $(i, j = 1, 2)$ and two complementarity constraints of the simple form as in (1.19). Hence (4.5) becomes

$$\begin{aligned} G(x) - l - s_{11} &= 0 \\ u - G(x) - s_{12} &= 0 \\ H(x) - (s_{21} - s_{22}) &= 0 \\ 0 \ \leq \ s_{11} \perp s_{21} \ &\geq \ 0 \\ 0 \ \leq \ s_{12} \perp s_{22} \ &\geq \ 0. \end{aligned}$$

Examples of the MacMPEC test set, that contain such constraints are each of the `gnash` problems (see [Ley00] and the Appendix).

4.2 Algorithm

The theoretical results of the previous chapters motivate the approach to solve an MPEC of the form (1.19), by solving a sequence of nonlinear programs $R(t_k)$ for a positive, descending sequence (t_k) until the complementarity condition is sufficiently satisfied.

To derive the outer algorithm that we will compare with the relaxed and the exact bilinear approach, we start off with a simple outer algorithm, discuss its performance for some varying parameter choices and two possible functions θ (which are used to describe problem $R(t)$, see Section 2.1). Then we explain the two main failures that occurred for some of the problems for these runs in more detail and give suitable modifications that might prevent these failures. Thus, we end up with the modified algorithm, which we will then analyse.

4.2.1 Outer Algorithm

The simplest outer algorithm representing our relaxation method solves a sequence of

$$\begin{array}{lllll} R(t_k) & \min & f(x) & & \\ & \text{subject to} & h(x) & = & 0 \\ & & g(x) & \geq & 0 \\ & & x_1, x_2 & \geq & 0 \\ & & \Phi(x_1, x_2, t_k) & \leq & 0\,, \end{array}$$

(see Chapter 2) using the NLP solver as a black box. Hence, we start with an initial NLP $R(t_0)$, let `filterSQP` solve this problem up to a given accuracy and check whether the solution $x^*(t_k)$ satisfies the complementarity condition. If this is the case, we stop, as according to Lemma 2.9 we have found a (approximately) strongly stationary point of the MPEC. If the complementarity condition is not sufficiently satisfied, we update the

Algorithm 4.1: Simple Outer Algorithm

Choose an initial vector $(x^0, \lambda^0, \mu^0, \nu_1^0, \nu_2^0, \xi^0)$*, an initial parameter* $t_0 > 0$*, an update parameter* $\sigma \in (0,1)$*, tolerance parameters* $\varepsilon_C > 0$ *and* $\varepsilon_{SQP} > 0$ *and a minimal parameter value* $\delta_{\min}$

repeat

    Solve $R(t_k)$ *up to the given accuracy determined by* ε_{SQP}.

    Update the relaxation parameter t_k *:*
    $t_{k+1} \leftarrow \max(\sigma t_k, \delta_{\min})$

    Set $k \leftarrow k+1$

until $\operatorname{compl}(x^*(t_k)) \leq \varepsilon_C$

parameter and solve $R(t_{k+1})$. We continue this iterative process until we have found a solution vector $x^*(t_k)$ that lies sufficiently close to the feasible region of the MPEC.

For the numerical tests, we present as next, we implemented Algorithm 4.1 as an `AMPL` script starting `filterSQP`. We used

$$\operatorname{compl}(x) = \sqrt{\sum_{j=1}^{p} \min(x_{1j}, x_{2j})^2}\,,$$

as a measure for the feasibility of an iterate x concerning the complementarity constraints. Next, before we begin with the discussion of this algorithm, we briefly introduce the performance figures we will use later on to illustrate and analyse our results.

Performance Figures

As our test set contains a large number of problems, we will use some figures in addition to summary tables to illustrate and analyse the results we obtained for the different methods and algorithm variants we tested. These figures are based on the *performance profile* that was introduced by Dolan and Moré [DM02].

Let $\mathcal{S}$ be the set of methods for (1.19) or variants of Algorithm 4.1 that we wish to compare. Moreover, let $\mathcal{P}$ denote our test set. To compare the different variants or methods of $\mathcal{S}$, we first compare them for each single problem $p \in \mathcal{P}$. Therefore, we choose a performance measure and compute the ratios of the performance of each particular variant or method s for each problem p compared to the best performance result obtained for the problem p by any variant or method $s \in \mathcal{S}$.

For our comparisons, we chose the number of SQP iterations as performance measure. For each problem p we compute the ratio of the number of iterations $i_{p,s}$ required by s to solve problem p compared to the smallest number of iterations required by any $s \in \mathcal{S}$ to solve p, thus we compute

$$r_{p,s} := \frac{i_{p,s}}{\min\{\, i_{p,s} \,|\, s \in \mathcal{S}\}}\,.$$

If a problem p was not solved by variant or method s, then we set $r_{p,s} = r_M$, where r_M is a parameter that satisfies $r_M \geq r_{p,s}$ for all $p \in \mathcal{P}$ and $s \in \mathcal{S}$ and $r_{p,s} = r_M$ if and only if p

was not solved by s. For our performance figures we used $r_M = 1.00E + 06$.

To get an impression of the relative performance of each $s \in \mathcal{S}$, we plot the graph of the performance profile

$$\varrho_s(\eta) := \frac{1}{|\mathcal{P}|} \left|\{ p \in \mathcal{P} \mid \log_2(r_{p,s}) \leq \eta \}\right|$$

for all $s \in \mathcal{S}$ in one figure. The performance profile corresponds to the probability that, given a variant or method s and a problem p, the performance of s for p is at most 2^η times the best performance, according to the chosen performance measure. Because $r_{p,s} = r_M$ if and only if problem p was not solved by s, we have that $\varrho_s(\log_2(r_M)) = 1$ and

$$\varrho_s^* := \lim_{\eta \nearrow \log_2(r_M)} \varrho_s(\eta)$$

denotes the probability that s solves a given problem. Hence, to compare the robustness of the variants or methods $s \in \mathcal{S}$, one needs to compare the values of ϱ_s^*, that is the values $\varrho_s(\eta)$ with η being large, which corresponds to the value of $\varrho_s(\eta)$ in the right-most position of the figure.

Another significant feature for each solver s is the value $\varrho_s(0)$, which corresponds to the probability that for a given problem p the variant or method s performs best.

Results for Algorithm 4.1

In this subsection we discuss the results of some numerical experiments that we carried out with Algorithm 4.1. We start the discussion by a comparison of the performance of the algorithm for a selection of parameter combinations which will be followed by comparing possible choices for the function θ (see Chapter 2).

For all experiments we set the tolerance parameter `eps` of `filterSQP` (that corresponds to ε_{SQP}) equal to 1.00E-08, as well as the complementarity tolerance ε_C. We set the maximum iteration number for `filterSQP` `maxiter=1000` and we chose the output level `iprint=1`, to get the informations we need for our discussion. Furthermore, we set the `AMPL` options `substout 0` and `presolve 0`, which turns off the the automatic substitution (of variables and constraints) and the presolve phase of `AMPL`. These options have been set, because it did not seem to benefit the performance in general (and to keep the `AMPL` options consistent, as we had to turn them off for the modified algorithm that we will present in Section 4.2.3). Finally, we set an outer iteration limit equal to 20.

To analyse the Algorithm 4.1, we begin with a comparison of its performance for a range of parameter combinations consisting of an initial parameter t_0 and an update parameter σ. The combinations we tested are listed in Table 4.2. For these experiments, we set

$$\theta(z) = s(z) := \frac{2}{\pi} \sin\left(z\frac{\pi}{2} + \frac{3\pi}{2}\right) + 1\,.$$

Remark 4.1. Since $\Phi_j(x_1, x_2, 0) = x_{1j} + x_{2j} - |x_{1j} - x_{2j}|$ is not differentiable in x with $(x_{1j}, x_{2j}) = (0, 0)$, for these x we set $\nabla_x \Phi_j(x_1, x_2, 0) := 0$.

The results are summarized in Figure 4.1 and a list of the iteration numbers can be found in Table A.1 in the Appendix, where we listed the name of the problem, the sum of inner

	data 1	data 2	data 3	data 4	data 5	data 6	data 7
t_0	0.00	1.00	1.00	1.00	10.0	10.0	10.0
σ	-	0.100	0.010	0.001	0.100	0.010	0.001

Table 4.2: Overview of parameter settings

iterations, that is the sum of SQP iterations, and in braces the number k of outer iterations, thus the number of nonlinear programs $R(t_k)$ that needed to be solved to (approximately) find a solution to (1.19). We evaluate a problem to be solved correctly if the constraint violation and the KKT-residual for the last solved problem $R(t_k)$ are sufficiently small, and if the complementarity condition is sufficiently satisfied for the solution $x^*(t_k)$ of the last solved problem $R(t_k)$, hence if $\text{compl}(x^*(t_k)) \leq \varepsilon_C$. We group the failures of Algorithm 4.1, that we document in Table A.1, into three different types: The first type concerns the

type	notation	description
1	-1	Local infeasibility of the nonlinear constraints
2	-2	KKT-residual > eps
3	-3	other failures : e.g. maximum nr. of iterations, complementarity not sufficiently satisfied, ...

Table 4.3: Description of Types of Failures and their Notation

case in which `filterSQP` is not able to find a feasible point, in other words the restoration phase fails to find a feasible solution. The second type corresponds to the case, where the solver terminates the run because either the trust-region or the step length became too small but the current iterate is not stationary, in other words the KKT-residual does not satisfy the required accuracy. In the third category we collected the remaining failures. Considering Table 4.4, where we summarized the failures of the algorithm for the different parameter combinations, we detect that the main part of the failures are either of the first or the second type, which justifies the partition we made. We summarized the failures in

type	data 1	data 2	data 3	data 4	data 5	data 6	data7
1	4	10	13	13	9	13	13
2	20	22	20	24	10	15	14
3	5	1	0	0	1	0	0
sum	29	33	33	37	20	28	27
%	22	25	25	28	15	21	20

Table 4.4: Number, type and percentage of failures for the parameter combinations of Table 4.2

Table 4.4, which gives us some useful information about the robustness of the different data variants.

It becomes clear that solving $R(t)$ with an initial parameter $t_0 = 10.0$ is more robust than solving $R(t)$ with an initial parameter $t_0 = 1.00$. Moreover, solving a sequence of $R(t_k)$ starting with $R(10.0)$ is also more robust than solving $R(0)$, since for data 1 the Algorithm fails for 29 problems, that corresponds to 22% of the test problems, whereas for data 5 - data 7 Algorithm 4.1 fails only for $15 - 21\%$ of the test problems, depending on the chosen update parameter σ. The parameter combination data 5 is the most robust, as Algorithm 4.1 solves 85% of the 133 test problems successfully for this combination.

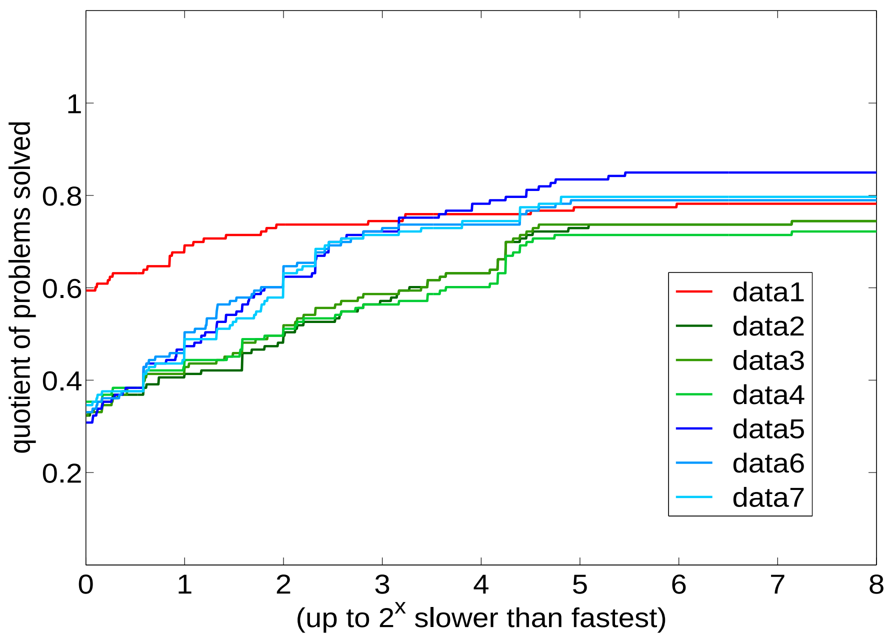


Figure 4.1: Comparison of different parameter settings for Algorithm 4.1

As explained before, these robustness informations can also be read out of Figure 4.1 by examining the values of $\rho_s(x)$ in the right-most position. However, Figure 4.1 also provides some further useful information. Examining $\rho_s(x)$ for different parameter variants, it becomes clear that using a starting parameter $t_0 = 10.0$ is not only more robust than using $t_0 = 1.00$, but also preferable in the context of small iteration numbers.

A large number of the test problems have strongly stationary solutions, where the additional constraints representing the complementarity condition is negligible for the degenerate components of the solution. Hence we can omit the associated derivatives of the additional constraints and solving $R(0)$ does not cause any differentiability problems. Therefore solving $R(0)$ in these cases is successful and mostly requires less iterations than solving a sequence of $R(t_k)$. This fact can be observed, by examining the values of $\rho_s(0)$, which correspond to the percentage of test problems that were solved by s within the smallest number of iterations (determined for each single test problem). Furthermore, it also denotes the probability that a given combination s solves a problem within the smallest number of iterations. Hence, since $\rho_s(0)$ is largest for data 1, where $\rho_s(0) \approx 0.6$, it seems to be the best choice concerning small iteration numbers. Comparing data 1 to data 5, we observe that for data 1 60% of the test problems were solved fastest, whereas for data 5 to solve 60% of the test problems Algorithm 4.1 requires at least up to 3.5 times more iterations than the best. However, consider the values of x corresponding to $\rho_s(x) = 0.75$, or in other words the maximum factor 2^x that is needed to solve 75% of the test problems, then we observe that for both parameter combinations data 1 and data 5 we obtain a maximum factor of approximately 10. If we consider yet larger percentages of test problems, that is values x with $\rho_s(x) > 0.75$, then the maximum factor for data 5 is even smaller than that one of data 1.

The next question we want to answer concerns the choice of the function $\theta(z)$ that we use for to formulate $R(t_k)$. As for the theory $\theta(z)$, only has to satisfy the Assumptions 2.1, there exist more than just the one choice we made above. Another possible function satisfying the Assumptions 2.1 for example corresponds to an interpolation polynomial, which additionally satisfies the conditions on the first and second order derivative. As a polynomial that satisfies these conditions, we can thus also use

$$\theta(z) = p(z) = \frac{1}{8}(-z^4 + 6z^2 + 3)\,.$$

We therefore undertake another numerical test to answer the question, whether either of these possibilities is preferable to the other and if the choice of $\theta(z)$ is a crucial factor for the performance of Algorithm 4.1. For this test, we chose according to our prior analysis, the parameter combination $t_0 = 10.0$ and $\sigma = 0.1$. The results for the choice $\theta(z) = p(z)$ can be found in the last column of Table A.1 in the Appendix.

In Table 4.5 we compare the failures of Algorithm 4.1 for both choices and Figure 4.2 displays their performance profiles.

type	$\theta(z) = s(z)$	$\theta(z) = p(z)$
1	9	14
2	10	9
3	1	1
sum	20	24
%	15	18

Table 4.5: Number, type and percentage of failures for $\theta(z) = s(z)$ and $\theta(z) = p(z)$

As can be seen from the tables and Figure 4.2 there is not such a significant difference in the performance for $\theta(z) = s(z)$ and for $\theta(z) = p(z)$ as for the different starting parameter t_0. Although, considering Table 4.5 and Figure 4.2, as the first choice for θ produces less failures and its performance profile $\rho_s(\eta)$ lies entirely above the one for $p(z)$, we will proceed with $\theta(z) = s(z)$.

Conclusions

The tests we made show that, although Algorithm 4.1 is kept very simple, most of the

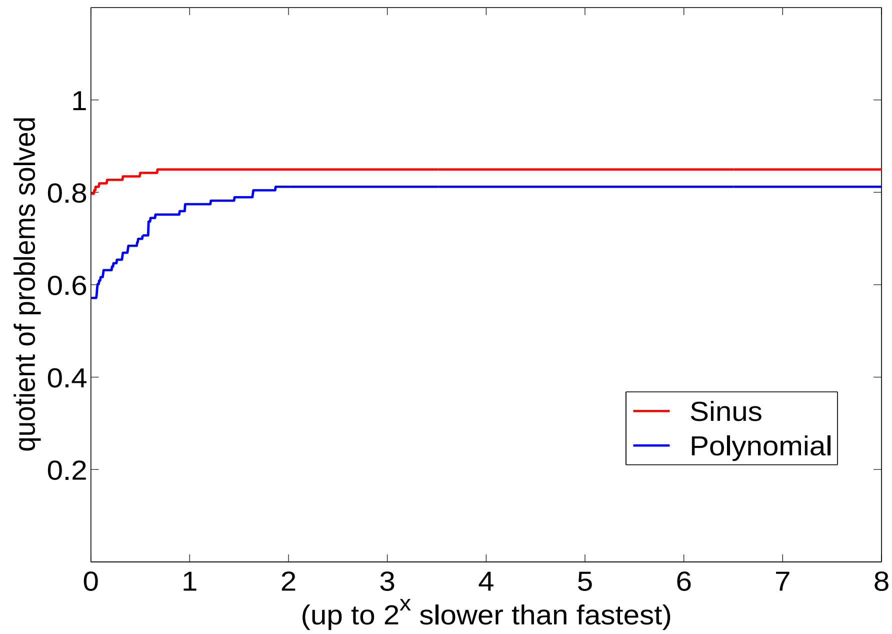


Figure 4.2: Comparison of Algorithm 4.1 for $\theta(z) = s(z)$ and $\theta(z) = p(z)$ with $t_0 = 10$ and $\sigma = 0.10$

MPECs could be solved. The best performance with $\theta(z) = s(z)$ concerning robustness and in parts concerning the iteration numbers was obtained for the variant using data 5. Moreover, for this variant $\theta(z) = s(z)$ outperformed $\theta(z) = p(z)$, which let us conjecture that this might also be true for the remaining data variants.

These observations lead us to the conclusion, that if our focus lies on a robust method to solve (1.19), then our choice should be solving a sequence $R(t_k)$ using the the parameter combination data 5 and $\theta(z) = s(z)$, rather than solving $R(0)$. As the reported failures are mainly of two different types, there might even be a chance to improve the simple outer algorithm by examining these failures and trying to derive some suitable modifications.

4.2.2 Failures

The first difficulty we have to deal with using Algorithm 4.1 to solve an MPEC in the form of (1.19) can be described as follows. Since we use an SQP method to solve $R(t_k)$, we solve a sequence of QPs, where we linearize the constraints of $R(t_k)$. However, the linearization

$$\Phi_j(x_1^k, x_2^k, t^k) + \nabla\Phi_j(x_1^k, x_2^k, t^k)^T d \leq 0 \tag{4.6}$$

of the constraint $\Phi_j(x_1^k, x_2^k, t^k) \leq 0$, in some cases might cause the steplength to converge to zero, hence the SQP algorithm stops, although we are not close enough to a solution yet.

The second difficulty concerns the case that, as we use the solution vector $x^*(t_k)$ as an initial point to solve $R(t_{k+1})$, which has a reduced feasible region compared to $R(t_k)$, our initial point $x^*(t_k)$ might not be feasible for $R(t_{k+1})$. Hence, the SQP solver then first has to solve a feasibility problem. If this could not be solved successfully, then the Outer Algorithm might get stuck in the infeasible point $x^*(t_k)$.

In the following we will explain both problems and how we can fix Algorithm 4.1 by some small modifications in more detail.

"Vertex Problem"

Consider a positive parameter $t_k > 0$ and a a pair (x^k_{1j}, x^k_{2j}) of the current iterate x^k with $x^k_{1j} < t_k$ and $x^k_{2j} = 0$ (the case that $x^k_{2j} < t_k$ and $x^k_{1j} = 0$ is similar). Hence it satisfies strict complementarity and is feasible with respect to the constraint $\Phi_j(x_1, x_2, t_k) \leq 0$. If we linearize

$$\Phi_j(x^k_1 + d_1, x^k_2 + d_2, t_k) \leq 0\,, \tag{4.7}$$

then we obtain (4.6), which is equivalent to

$$\alpha^k_j d_{1j} + (2 - \alpha^k_j)\, d_{2j} \leq -\Phi^k_j,$$

where α^k_j denotes the corresponding partial derivative of $\Phi^k_j = \Phi_j(x^k_1, x^k_2, t_k)$ (see Chapter 2). As $x^k_{1j} < t^k$ and $x^k_{2j} = 0$, Assumptions 2.1 on the function θ imply that

$$\theta'\left(\frac{x^k_{1j} - x^k_{2j}}{t_k}\right) < 1\,,$$

therefore

$$\alpha^k_j = 1 - \theta'\left(\frac{x^k_{1j} - x^k_{2j}}{t_k}\right) > 0\,.$$

It hence follows that a step d along the direction of e_{1j} is restricted by

$$d_{1j} \leq -\frac{\Phi^k_j}{\alpha^k_j}.$$

To evaluate $\Phi_j(x_1, x_2, t_k)$ along the direction of e_{1j}, we define

$$\tilde{\Phi}(x) := \Phi_j(x, 0, t) = x - t\theta\left(\frac{x}{t}\right)\,,$$

and approximate $\tilde{\Phi}(x)$ by its Taylor series expansion in $\hat{x} = t$ (see for example [Kön95]). Considering the Assumptions 2.1 we made on θ and additionally assuming that θ is at least three times continuously differentiable on $\mathcal{D} \subseteq \mathbb{R}$ with $[-1, 1] \subseteq \mathcal{D}$, the third-order Taylor approximation of $\tilde{\Phi}(x)$ in $\hat{x} = t$ is

$$T_3\tilde{\Phi}(x; \hat{x}) = \frac{1}{6}\tilde{\Phi}'''(\hat{x})(x - \hat{x})^3 = -\frac{1}{6t^2}\theta'''(1)(t - x)^3\,. \tag{4.8}$$

Hence, using this information to approximately determine the restriction of the steplength

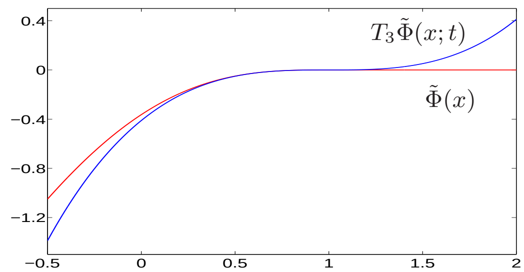


Figure 4.3: Functions and $T_3\tilde{\Phi}(x;t)$ for $t = 1.0$, θ as in (2.1) and $x \in [-0.5, 2.0]$.

of a step d we get

$$d_{1j} \leq -\frac{\Phi_j^k}{\alpha_j^k} \approx \frac{\frac{1}{6t^2}\theta'''(1)(t-x)^3}{3\frac{1}{6t^2}\theta'''(1)(t-x)^2} = \frac{1}{3}(t-x)\,.$$

Therefore, in the case that it would be beneficial to take a step d^k with $x_{1j}^{k+1} = x_{1j}^k + d_{1j}^k > t_k$ and $d_{2j}^k = 0$ the QP solver might produce a sequence (d^k) with

$$\lim_{k\to\infty} d_{1j}^k = 0$$

and $x_{1j}^k < t_k$ for all $k \in \mathbb{N}$, in other words the sequence (x_{1j}^k) will never "pass the boundary" t_k.

The explained behaviour can be observed for the following example:

Example 4.1.

$$\begin{array}{ll} \textit{minimize} & f(x) := (x_1-2)^2 + (x_2+2)^2 \\ \textit{subject to} & 0 \leq x_1 \quad \perp \quad x_2 \geq 0 \end{array}$$

We choose the starting vector $x^0 = (0,0)$ and $t_0 = 1.00$. In Table 4.6 we have listed the iteration number k and the corresponding values for d_1^k, x_1^k, the value Φ^k/α^k of the maximum steplength and the approximation of it. We obtained the values applying the MATLAB 7.5.0 built in QP solver `quadprog` [Ven02] iteratively to the quadratic approximations

$$\begin{array}{rc} \min_d & 2\,d_1^2 + 2\,d_2^2 + 2\,(x_1^k-2)\,d_1 + 2\,(x_2^k+2)\,d_2 \\ \text{subject to} & x_1^k + d_1 \geq 0 \\ & x_2^k + d_2 \geq 0 \\ & \Phi^k + \alpha^k d_1 + (2-\alpha^k)d_2 \leq 0 \end{array}$$

of $R(t_0)$ of our Example (4.1) with the starting vector $x^0 = (0,0)$ and $t_0 = 1.00$. The values of Table 4.6 and Figure 4.4, where we have plotted the contour of $\Phi(x_1, x_2, t) = 0$ and the produced iterates (x_1^k, x_2^k), again illustrates the behaviour we have just explained.

k	d_1^k	x_1^k	Φ^k/α^k	$1/3(t-x_1^k)$
0	0.363380228	0.000000000	0.363380228	0.333333333
1	0.219541466	0.363380228	0.219541466	0.212206591
2	0.141046105	0.582921693	0.141046105	0.139026102
3	0.092591232	0.723967798	0.092591232	0.092010734
4	0.061316726	0.816559031	0.061316726	0.061146990
5	0.040758081	0.877875757	0.040758081	0.040708081
6	0.027136830	0.918633839	0.027136831	0.027122054
7	0.018080817	0.945770669	0.018080817	0.018076444
8	0.012050800	0.963851486	0.012050800	0.012049505

Table 4.6: Results for Example 4.1

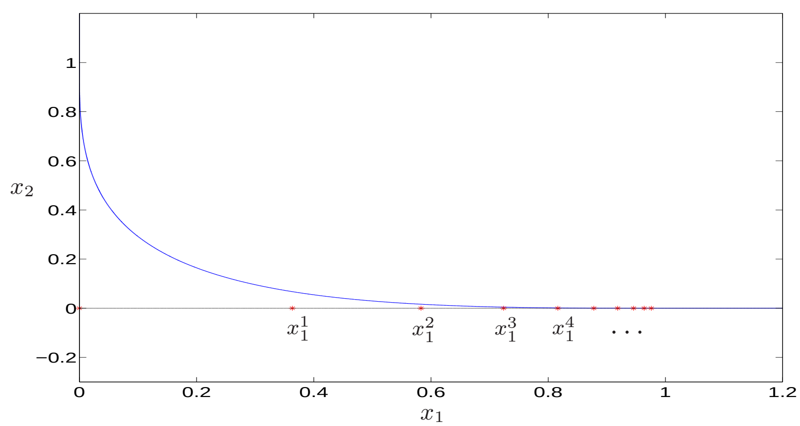


Figure 4.4: Illustration of the vertexproblem for Example 4.1

Modifications.
In Algorithm 4.1 we used as stopping criteria those ones of `filterSQP` for the solutions of $R(t_k)$ (hence for the inner iterations) and for the outer iteration loop we simply used a measure of the complementarity. Hence, since a small steplength induces `filterSQP` to stop in the case of the vertex problem that we just explained, both stopping criteria (the inner and the outer) are satisfied, although we are not sufficiently optimal yet.

One possibility to prevent the outer algorithm to stop in this case but to reduce t_k and solve $R(t_k)$ once again is to additionally check the norm of the KKT-residual and if it is not sufficiently small (according to the stopping tolerance of `filterSQP`), then the outer algorithm continues.

If we use `filterSQP` as a black box, then there is no other possibility than to react to the situation (as we cannot know in advance, that some iterates (x_{1j}^k, x_{2j}^k) will cause the vertex problem to appear). However, to wait for `filterSQP` to stop, because the steplength falls below the stopping tolerance, implies that the solver might carry out a lot of "useless" iterations (which has almost no further contribution than to reduce the steplength until the solver stops). Hence, in order to intervene in time, we have to decrease t_k in between the SQP iterations, thus directly after the solution of a QP: if for the current iterate x^k a pair (x^k_{1j}, x^k_{2j}) approaches either $(t_k, 0)$ or $(0, t_k)$, then we reduce t_k, such that $x^k_{ij} > t_k$ ($i = 1$ or 2, respectively). This however implies that we cannot use `filterSQP` as a black box any longer in order to omit the "useless" iterations.

However, as we do not want to interfere with the other constraints $\Phi_i(x_1, x_2, t_k) \leq 0$ with $i \in \{1, \ldots, p\} \backslash \{j\}$, we now use a parameter vector $t^k = (t^k_1, \ldots, t^k_p)$ instead of a scalar parameter, such that we have $\Phi_j(x_1, x_2, t^k_j) \leq 0$ for all $j \in \{1, \ldots, p\}$ and we can update t^k_j independently.

"Infeasible Initial Points"

Another problem, that we have to deal with, concerns the case that the solver does not find a feasible initial vector. If this problem arises in the first outer iteration, that is for problem $R(t_0)$, then this failure is not due to our relaxation but the choice of x_0 or the MPEC itself. However, if this does happen for some iteration $k+1$ later on, then it has to be due to the constraints $\Phi_j(x_1, x_2, t_{k+1}) \leq 0$, since $x^*(t_k)$ is feasible for $R(t_k)$ and therefore $h(x^*(t_k)) = 0$ and $g(x^*(t_k)) \geq 0$. Hence, the infeasiblity of $x^*(t_k)$ for $R(t_{k+1})$ refers to the situation that $\Phi_j(x^*_1(t_k), x^*_2(t_k), t_k) \leq 0$, but $\Phi_j(x^*_1(t_k), x^*_2(t_k), t_{k+1}) > 0$.

One example (that corresponds to `bard3` of MacMPEC), where we have observed this difficulty is

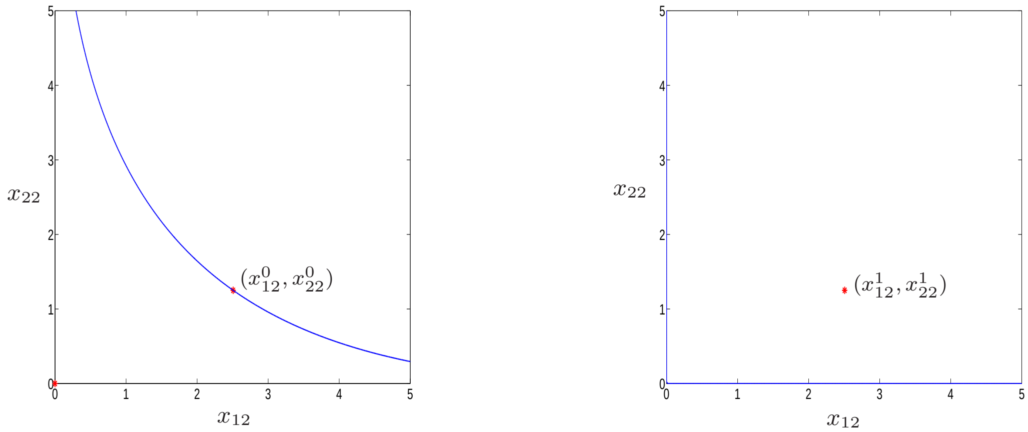
Example 4.2.

$$
\begin{array}{rr}
\textit{minimize} & f(x_0, x_1, x_2) := -x_{01}^2 - 3x_{02} - 4x_{03} + x_{04}^2 \\
\textit{subject to} & 2x_{03} + 2x_{21} - 3x_{22} = 0 \\
 & -5 - x_{21} + 4x_{22} = 0 \\
 & x_{01}^2 - 2x_{01} + x_{02}^2 - 2x_{03} + x_{04} + 3 - x_{11} = 0 \\
 & x_{02} + 3x_{03} - 4x_{04} - 4 - x_{12} = 0 \\
 & x_{01}^2 + 2x_{02} - 4 \leq 0 \\
 & 0 \leq x_{11} \quad \perp \quad x_{21} \geq 0 \\
 & 0 \leq x_{12} \quad \perp \quad x_{22} \geq 0
\end{array} .
$$

k	t_k	$(x^k_0, x^k_1, x^k_2) := x^*(t_k)$	$\lVert c(x^*(t_k)) \rVert$	$\text{compl}(x^*(t_k))$
0	10	(0.00, 2.00 ,1.88, 0.279, 3.53, 2.51, 0.00, 1.25)	0.00	1.25
1	0.1	(0.00, 2.00 ,1.88, 0.279, 3.53, 2.51, 0.00, 1.25)	0.00	1.25
2	1E-03	(0.00, 2.00 ,1.88, 0.279, 3.53, 2.51, 0.00, 1.25)	0.00	1.25
3	1E-05	(0.00, 2.00 ,1.88, 0.279, 3.53, 2.51, 0.00, 1.25)	0.00	1.25

Table 4.7: Results for Example 4.2, where $c(x)$ contains all constraints except for the complementarity constraints and $\text{compl}(x^k) = \sqrt{\sum_{j=1}^{2} \min(x^k_{1j}, x^k_{2j})^2}$

Table 4.7 and Figure 4.5 illustrates the infeasibility problem for Example 4.2. Starting with the initial point $x^0 = 0$, `filterSQP` successfully solves $R(10)$. However, trying



(a) Iteration $k = 0$, $t_0 = 10$

(b) Iteration $k = 1$, $t_1 = 0.1$

Figure 4.5: Illustration of the Infeasibility Problem for Example 4.2

to solve $R(0.1)$ the solver gets stuck in $x^*(10)$, as can be seen in Table 4.7. The pair $(x_{12}(10), x_{22}(10))$ is not feasible for $R(0.1)$ and the attempt to find a feasible point for $R(0.1)$ is unsuccessful, as can be seen on the right in Figure 4.5.

Modifications. As we will see later on, if we decrease t_k more gently, then we have a chance to find a feasible point for $R(t_{k+1})$. However, if we generally choose our update parameter σ larger, then we will loose the good performance concerning the iteration numbers for the 85% of problems that were solved without any difficulties, as can be noticed by Figure 4.1. Hence the idea is: first we try a more stringent parameter update and only in the case that $R(t_{k+1})$ could not be solved successfully (as no feasible starting point was found), we "re-update" t_k and use a gentler decrease for t_{k+1} (thus we enlarge t_{k+1}) and solve $R(t_{k+1})$ again. As $x^*(t_k)$ is feasible for $R(t_k)$, at least for $t_{k+1} = t_k$ we obtain a feasible initial vector. If $R(t_{k+1})$ was successfully solved for the gentler decrease, then for the next t_k, we try again the more stringent parameter update.

k	t_k	$(x_0^k, x_1^k, x_2^k) := x^*(t_k)$	$\|c(x^*(t_k))\|$	$\text{compl}(x^*(t_k))$
0	10.0	$(0.00, 2.00, 1.88, 0.28, 3.53, 2.51, 0.00, 1.25)$	0.00	1.25
1	5.00	$(0.00, 2.00, 1.88, 0.75, 4.00, 0.63, 0.00, 1.25)$	0.00	0.63
2	0.05	$(0.00, 2.00, 1.88, 0.91, 4.16, 0.00, 0.00, 1.25)$	0.00	0.00

Table 4.8: Results for Example 4.2, where $c(x)$ contains all constraints except for the complementarity constraints and $\text{compl}(x^k) = \sqrt{\sum_{j=1}^2 \min(x_{1j}^k x_{2j}^k)^2}$

Another idea in this context concerns the variable $x^*(t_k)$: since $x^*(t_k)$ is not feasible for $R(t^{k+1})$, it might be advantageous to check which constraints $\Phi_j(x_1^*(t_k),\, x_2^*(t_k),\, t_{k+1}) \leq 0$ are not satisfied by $x^*(t_k)$. Having detected the indices $j \in \{1, \ldots, p\}$, where an infeasi-

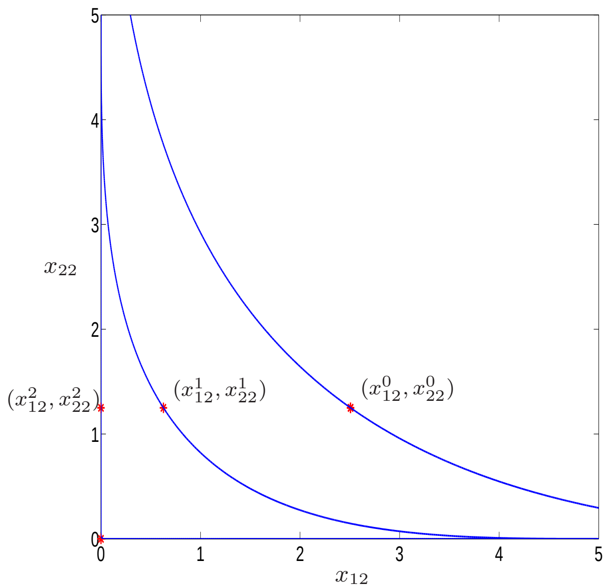


Figure 4.6: Illustration of the modified algorithm applied to Example 4.2

bility occurs, thus $\Phi_j(x_1^*(t_k),\ x_2^*(t_k),\ t_{k+1}) > 0$, we then project the corresponding pairs $(x_{1j}^*(t_k), x_{2j}^*(t_k))$ onto the positive axes. This intervention might cause the infeasibility of the new iterate concerning the remaining constraints, however, if this infeasibility is overcome easier than that one of the relaxed complementarity constraints, then the intervention benefits the behaviour of the outer algorithm.

In our modified algorithm, that will be presented as next, we combined the idea of the projection with the one of a gentler decrease in the case of an infeasible initial point $x^*(t_k)$ for problem $R(t^{k+1})$.

4.2.3 Modified Algorithm

In this section, we describe a modified outer algorithm, that includes the modifications we mentioned in the foregoing section. As the modification concerning the Vertex problem includes that we have to intervene in the SQP algorithm, we now incorporate the outer algorithm directly in `filterSQP`. We therefore also implemented the constraints $\Phi_j(x_1, x_2, t_k) \leq 0$ directly in `filterSQP`, in order to be able to change and compute these constraints independently of the others (as the rest of the constraints and their derivatives are computed automatically by `AMPL`). We then adapted `filterSQP` according to Algorithm 4.2, where we set the parameters as follows: $\varepsilon_C = 1.00E{-}08$, $\varepsilon_{SQP} = 1.00E{-}08$, $t_j^0 = 10.0$ for $j \in \{1, \ldots, p\}$, $\delta = 0.10$, $\delta_{\min} = 10^{-12}$, $\sigma_1 = 0.50$, $\sigma_2 = 0.10$, $\sigma_3 = 5.00$ and maxit= 15.

Algorithm 4.2: Modified Algorithm

---

1 *Choose* x^0, *the parameter vector* t^0 *and the tolerance and update parameter* ε_{SQP}, ε_C, δ, σ_1, σ_2 *and* σ_3, *a maximum iteration number* $maxit \geq 1$ *and the minimum parameter* $\delta_{\min}$.

2 $k \leftarrow 0, \quad t^{0,0} \leftarrow t^0$

repeat

  $l \leftarrow 0$

  repeat

3   Solve QP of $R(t^{k,l})$ for a step $d^{k,l}$ and update $x^{k,l}$.

4   **if** $(1-\delta)t_j^{k,l} < \max(x_{1j}^{k,l+1}, x_{2j}^{k,l+1}) < t_j^{k,l}$ *and* $x_{1j}^{k,l+1} x_{2j}^{k,l+1} < \varepsilon_C$ **then**

      $t_j^{k,l+1} \leftarrow \max(\sigma_1\, t_j^{k,l}, \delta_{\min})$

    $l \leftarrow l+1$

  until Stopping Criteria of `filterSQP` are satisfied.

5   Update solution x^k of $R(t^{k,l})$: $x^k \leftarrow x^{k,l}$

6   **if** *"Optimal solution found." for* $R(t^{k,l})$ **then**

    Decrease the parameters t_j^k:

    for $j = 1, \ldots, p$ **do**

      $t_j^{k+1,0} \leftarrow \max(\sigma_2\, t_j^{k,l}, \delta_{\min})$

  else

    Increase the parameters t_j^k:

    for $j = 1, \ldots, p$ **do**

      $t_j^{k+1,0} \leftarrow \sigma_3\, t_j^{k,l}$

7   **if** *"Nonlinear Constraints (locally) infeasible." for* $R(t^{k,l})$ **then**

    if $\Phi_j(x_1^k, x_2^k, t_j^{k+1,0}) > \varepsilon_C$ **then**

      Project infeasible pairs (x_{1j}^k, x_{2j}^k) onto axes:

      if $x_{1j}^k \leq x_{2j}^k$ **then**

        $x_{1j}^k \leftarrow 0$

      else

        $x_{2j}^k \leftarrow 0$

  $k \leftarrow k+1$

until $\|KKT_{res}\| \leq \varepsilon_{SQP}$, $\|c_{res}\| \leq \varepsilon_{SQP}$ *and* $compl(x^k) \leq \varepsilon_C$ *or* $k > maxit$

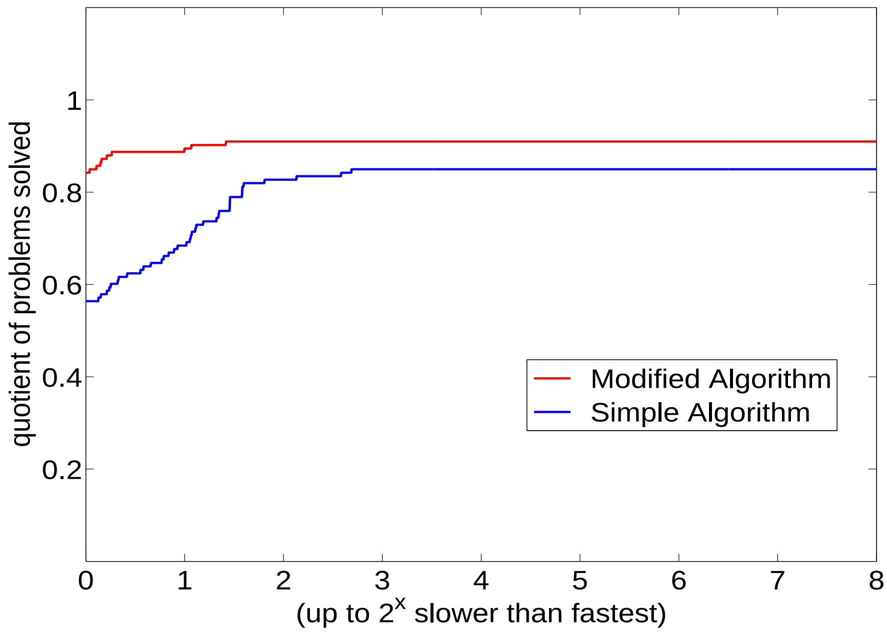


Figure 4.7: Comparison of Algorithm 4.2 and Algorithm 4.1

Our measure of complementarity is again

$$\operatorname{compl}(x) = \sqrt{\sum_{j=1}^{p} \min(x_{1j}, x_{2j})^2}\,.$$

Modified Algorithm compared to Simple Outer Algorithm

Before we compare our Modified Algorithm with the results we obtain for the relaxed and the exact bilinear approach, we first compare it with the Simple Outer Algorithm (Algorithm 4.1).

Figure 4.7 displays the performance profiles of Algorithm 4.1 and of Algorithm 4.2. Obviously the modifications we made not only improved the robustness of our method, but also the number of required iterations. We listed the number of discovered failures for the Modified Algorithm and again those ones of the Simple Outer Algorithm in Table 4.9. Considering Table 4.9, it becomes clear that the Modified Algorithm reduced the number of failures by 40% compared to the Simple Outer Algorithm. The improved robustness can also be deduced from Figure 4.7. Furthermore, according to Leyffer [Ley00], at least two failures of the first type are directly due an infeasible problem. Moreover, it can be seen in Figure 4.7, that the Modified Algorithm solves approximately 85% of the test problems as fastest and it needs at most approximately 2.7 times as many iterations as the Simple Outer Algorithm.

type	Modified Algorithm	Simple Algorithm
1	5	9
2	7	10
3	0	1
sum	12	20
%	9	15

Table 4.9: Number, type and percentage of failures for Algorithm 4.2 and Algorithm 4.1

4.3 Comparison

In this final section, we compare the numerical results we obtain for the Modified Algorithm (Algorithm 4.2) with the results for the exact and the relaxed bilinear solution approach for MPECs (see Section 1.4).

To obtain appropriate numerical results for these two approaches, comparable to those ones of Algorithm 4.2, we implemented the Outer Algorithm for the two bilinear approaches (Algorithm 4.3) directly in `filterSQP`.

Algorithm 4.3: Outer Algorithm for Bilinear Approach

Choose an initial vector $(x^0, \lambda^0, \mu^0, \nu_1^0, \nu_2^0, \xi^0)$*, an initial parameter* $t_0 \geq 0$*, an update parameter* $\sigma \in (0,1)$*, tolerance parameters* $\varepsilon_C > 0$ *and* $\varepsilon_{SQP} > 0$*, a maximum iteration number maxit* ≥ 1 *and a minimal parameter value* $\delta_{\min} > 0$

repeat

  Solve $NLP(t_k)$ *up to the given accuracy determined by* ε_{SQP}*.*

  Update the relaxation parameter t_k *:*
  $t_{k+1} \leftarrow \max(\sigma t_k, \delta_{\min})$

  Set $k \leftarrow k+1$

until $\text{compl}(x^*(t_k)) \leq \varepsilon_C$ or $k > \text{maxit}$

We used again

$$\text{compl}(x) = \sqrt{\sum_{j=1}^{p} \min(x_{1j}, x_{2j})^2}\,,$$

as a measure for the feasibility of an iterate x concerning the complementarity constraints. For the relaxed bilinear approach we chose the parameter values as for Algorithm 4.1 and data 5 (see also the Appendix): $t_0 = 10.0$, $\sigma = 0.10$, $\delta_{\min} = 1.00E-18$, maxit $= 20$, $\varepsilon_C = 1.00E-08$ and $\varepsilon_{SQP} = 1.00E-08$. The parameters $\delta_{\min}$ and maxit vary from those ones for Algorithm 4.2 as for the relaxed bilinear approach the parameter t_k might have to become smaller to obtain the same accuracy for the complementarity condition than for the new relaxation method (see Section 2.5)

For the exact bilinear approach we set $t_0 = 0.00$ and the maximum outer iteration number maxit $= 1$.

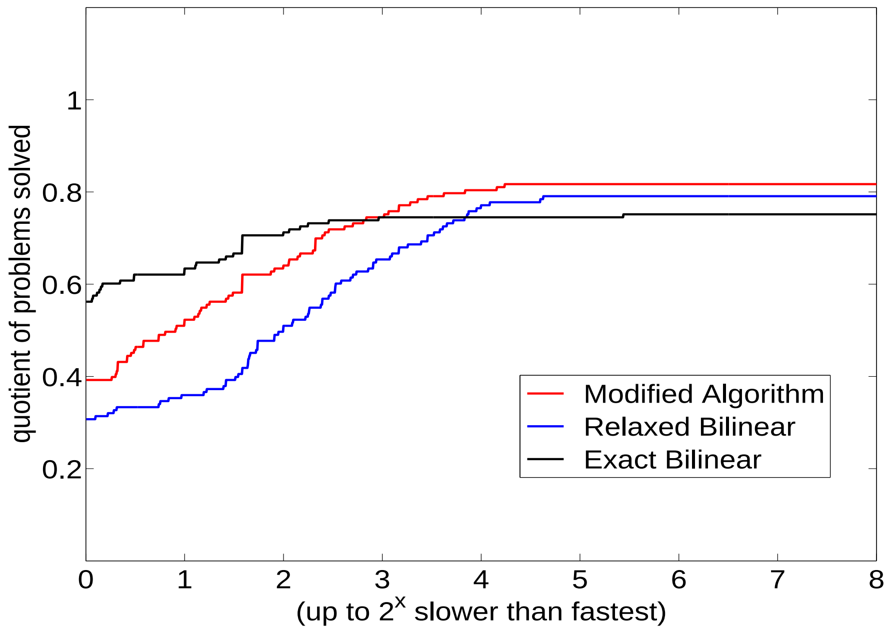


Figure 4.8: Comparison of the new relaxation method to the exact and the relaxed bilinear approach

For the comparison of this section, we enlarged the set of test problems by 20 large-scale MPECs of the MacMPEC test problem set. These are of size $n > 4000$ (here n denotes the number of variables). Hence our test problems set now contains 153 MPECS, which divide into 31 small ($n < 10$), 102 medium ($10 \leq n < 4000$) and 20 large-scale ($n > 4000$) problems. (We did not take all test problems of MacMPEC, because some of them seem to feature significant similarities.)

The detailed numerical results for all three approaches are listed in Table A.2, Table A.3 and Table A.4 in the Appendix. Moreover, Figure 4.8 summarizes the numerical results concerning the iteration numbers. It displays the performance of the new relaxation method (represented by Algorithm 4.2) compared to the exact and the relaxed bilinear approach.

In view of Figure 4.8 it becomes clear that the new relaxation method represents a good compromise between small iteration numbers and robustness of the method: using the exact bilinear solution approach most problems, where a satisfying solution is found, are solved within comparably few iterations, if not the smallest number of iterations. For 56% of the test problems the exact bilinear approach produces the smallest iteration number. Concerning the new relaxation method, only 39% of the test problems are solved within the smallest number of iterations. However, this value is worse for the relaxed bilinear approach, since for this approach $\rho(0) = 0.31$, which corresponds to 31% of the test problems. However, the advantage of the exact bilinear approach concerning small iteration numbers might mostly be due to the fact that most of the test problems have strongly stationary

type	new relaxation	relaxed bilinear	exact bilinear
1	6	3	7
2	20	22	26
3	1	6	5
sum	27	31	38
%	18	20	25

Table 4.10: Number, type and percentage of failures for the new relaxation method, the relaxed and the exact bilinear approach

solutions (confer Section 2.5 and the discussion further below).

For problems solved by both compared methods, the average factor concerning the required iterations of the new relaxation method compared to those ones required by the exact bilinear approach is approximately 2.7. The iteration numbers of the relaxed bilinear approach are in average approximately 1.8 times higher as those ones of the new relaxation method .

Concerning the robustness, from Figure 4.8 it becomes apparent, that the new relaxation method is the most robust one of all three approaches.

Hence, compared to the relaxed bilinear approach, the new relaxation method is more favourable, since it is more robust and in most cases it finds a solution in less iterations.

In Table 4.10 we summarized the failures of all three methods. In some cases the failures that occurred matches more than one type. In these cases we chose the seemingly more appropriate one.

The new relaxation fails to find an appropriate solution (that is the KKT- and the feasibility residual do not exceed ε_{SQP} and the condition on the complementarity measure is fulfilled) only for 27 test problems, which corresponds to 18% of all 153 test problems. The relaxed bilinear approach fails for 31 test problems that correspond to 20% of all test problems. Finally, the exact bilinear approach fails to find an appropriate solution for 38 MPECs, in other words it fails for 25% of the test problems.

Moreover, the set of MPECs for which all three methods fail consists of 23 test problems. Hence, the new relaxation method fails only on 4 problems, where one of the other two methods is successful. According to Leyffer [Ley00] at least four failures are due to infeasible problems. Other failures, especially of the second type, are due to infeasible QPs, that is mostly the case for problems where ifail = 4 or ifail = 5 (see the Appendix).

Concerning the complementarity condition, in view of Table A.2, Table A.3 and Table A.4 it becomes clear that for the new relaxation method the complementarity measures $\text{compl}_1(x)$ and $\text{compl}_2(x)$ are mostly at least as well satisfied as for the exact bilinear approach (only taking problems with appropriate solutions (see above) for both compared methods into account). It is theoretically clear, that for the relaxed bilinear approach the complementarity measures will be less well satisfied than for the other two methods, since for this approach exact complementarity can only be guaranteed to be satisfied for the limit $t_k \to 0$ (independently of the existence of degenerate components). This property is also distinguishable in Table A.3 in comparison to Table A.2 and Table A.4.

Comparing the final objective function values, it comes out that the relaxed bilinear approach finds a better objective function value than the new relaxation method for 8 test problems. The inverse direction holds only for one single MPEC. Comparing the objective function values obtained by the exact bilinear approach to those ones for the new relaxation method, we identify 6 test problems where the exact bilinear approach finds a better objective function value and 17 test problems where the new relaxation method finds a better solution value.

Finally, let us consider the theoretical statements we made in Section 2.5. We therefore examine the numerical results of three examples of the test problem set in more detail. The test problems we consider are

Example 4.3. `ex9.2.2`

$$\begin{array}{llcll}
\textit{min} & x_{01}^2 + (x_{02} - 10)^2 & & & \\
\textit{subject to} & 0 \le x_{01} & \le & 15 & : \lambda_1 \\
 & 0 \le x_{02} & & & : \lambda_2 \\
 & -x_{01} + x_{02} & \le & 0 & : \lambda_3 \\
 & x_{01} + x_{02} + x_{11} & = & 20 & : \mu_1 \\
 & -x_{02} + x_{12} & = & 0 & : \mu_2 \\
 & x_{02} + x_{13} & = & 20 & : \mu_3 \\
 & 2(x_{01} + 2x_{02} - 30) + x_{21} - x_{22} + x_{23} & = & 0 & : \mu_4 \\
 & 0 \le \; x_1 \;\perp\; x_2 & \ge & 0, & : \nu_1, \nu_2, \xi \,,
\end{array}$$

This problem has an M-stationary solution $x^* = (x_0^*, x_1^*, x_2^*) = (10, 10, 0, 10, 10, 0, 0, 0)$ with MPEC multipliers $(\lambda^*, \mu^*, \hat{\nu}_1, \hat{\nu}_2) = (0, 0, 13 \ 1/3, 0, 0, 0, 3 \ 1/3, 0, 0, 0, -3 \ 1/3, -3 \ 1/3, -3 \ 1/3)$.

Example 4.4. `ralph1`

$$\begin{array}{rll}
\textit{min} & f_1(x) = 2x_0 - x_2 & \\
 & f_2(x) = x_0 - x_2 & \\
\textit{subject to} & x_2 - x_0 - x_1 = 0 & : \mu \\
 & x_0 \le 0 & : \lambda \\
 & 0 \le \; x_1 \;\perp\; x_2 \; \ge 0, & : \nu_1, \nu_2, \xi
\end{array}$$

The M-stationary solution of this problem is $x^* = (x_0^*, x_1^*, x_2^*) = (0, 0, 0)$ with MPEC multipliers $(\lambda, \mu, \hat{\nu}_1, \hat{\nu}_2) = (2, 0, 0, -1)$ for objective function $f_1(x)$ and $(\lambda, \mu, \hat{\nu}_1, \hat{\nu}_2) = (1, 0, 0, -1)$ for objective function $f_2(x)$.

And finally `scholtes4` (which corresponds to Example 2.4):

$$\begin{array}{rll}
\text{min} & x_1 + x_2 - x_0 & \\
\text{subject to} & -4x_1 + x_0 \le 0 & : \lambda_1 \\
 & -4x_2 + x_0 \le 0 & : \lambda_2 \\
 & 0 \le \; x_1 \;\perp\; x_2 \; \ge 0, & : \nu_1, \nu_2, \xi \,.
\end{array}$$

According to Section 2.5, we assume that for the relaxed and the exact bilinear approach the generated multiplier sequences (ξ^k) of the complementarity conditions are unbounded.

	Exact Bilinear	Relaxed Bilinear	New Relaxation
KKT-residual	1.85 E+02	1.85 E+00	3.97 E -17
$\xi^*_{\max}$	1.78 E+09	9.39 E+08	9.95 E+00
inner(outer) Iterations	29(1)	91(20)	49(14)

Table 4.11: Values for Example 4.3

	Exact Bilinear	Relaxed Bilinear	New Relaxation
KKT-residual	2.36 E-01	0.00 E+00	1.57 E -14
$\xi^*_{\max}$	3.47 E+07	5.00 E+05	0.50 E+00
inner(outer) Iterations	33(1)	86(20)	10(10)

Table 4.12: Values for Example 4.4

Thus, in particular, the iterates ξ^k for the approximate solution of the problems will become very large. Moreover, in Section 2.5 we expected this unboundedness of the multipliers to cause numerical troubles. In Table 4.11 to Table 4.13 we summarized the significant values for these three examples. The values we obtained clearly support our assumptions.

Both methods, the relaxed and the exact bilinear approach, fail for all three problems. Furthermore, the corresponding numerical values of ξ^*, the multiplier corresponding to the complementarity conditions are very large, whereas they are of suitable sizes for the new relaxation method. Detailed numerical results we obtain for the relaxed bilinear approach and the new relaxation method for Example 4.3, Example 4.4 and Example 2.4 can be found in Table A.5 to Table A.10 in the Appendix.

Figure 4.9 and Figure 4.10 display the development of the significant complementarity multiplier ξ^k for the relaxed bilinear approach and the new relaxation method. It can be observed that for the relaxed bilinear approach the values of ξ^k diverge, whereas they stay constant for the new relaxation method. This might be one important reason, why the new relaxation method solves all three test problems without any difficulties.

Comparing the rate of convergence of the stationary points $x^*(t_k)$ for the relaxed bilinear approach to those ones we obtain for the relaxation method, it comes out that the numerical results support the theoretical statements we made in Section 2.5. From Figure 4.11, where we plotted the values $\|(x^k_{1j}, x^k_{2j}) - (x^*_{1j}, x^*_{2j})\|_2$ for the significant complementarity pairs (x^k_{1j}, x^k_{2j}) on a logarithmic scale, the rate of convergence we derived at the end of Section

	Exact Bilinear	Relaxed Bilinear	New Relaxation
KKT-residual	4.71 E-01	0.00 E+00	4.34 E -18
$\xi^*_{\max}$	3.47 E+08	5.00 E+05	1.00 E+00
inner(outer) Iterations	32(1)	85(20)	11(10)

Table 4.13: Values for Example 2.4

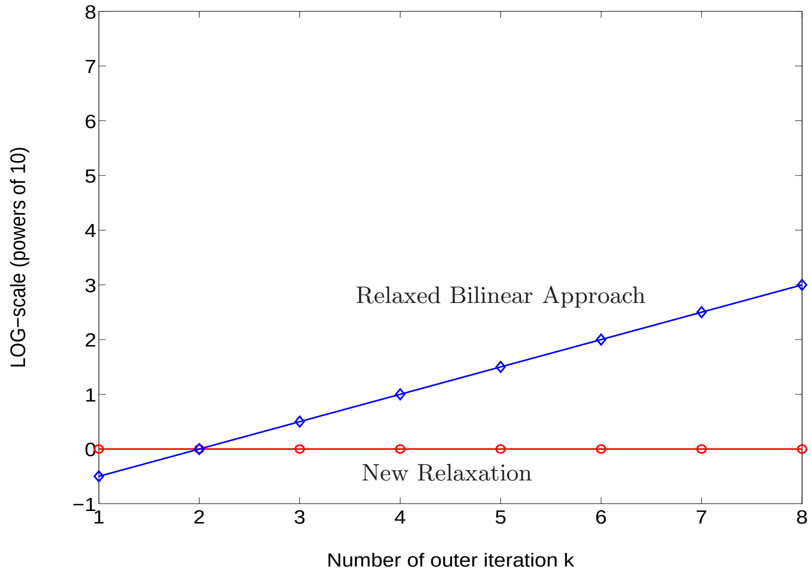


Figure 4.9: Comparison of complementarity multiplier ξ^k for $t^k = 0.1^{k-2}$ for Example 2.4 (similar to the results, for Example 4.4)

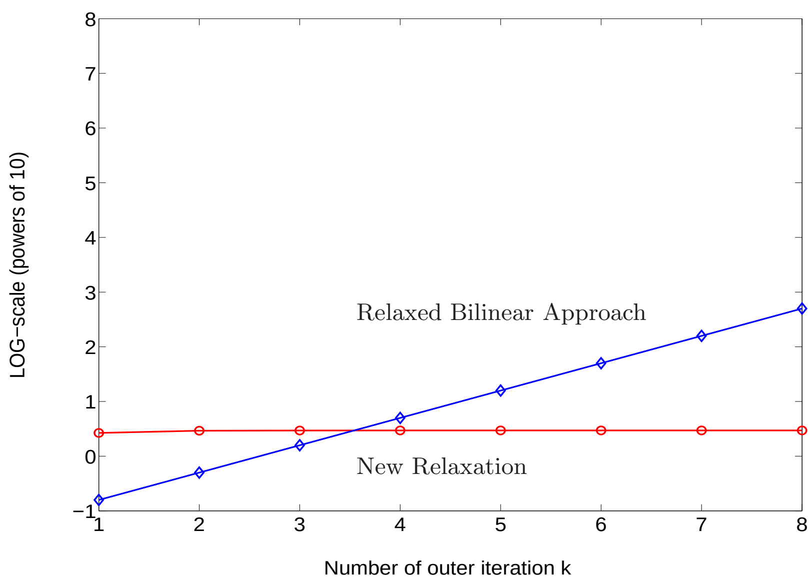


Figure 4.10: Comparison of complementarity multiplier ξ^k for $t^k = 0.1^{k-2}$, for Example 4.3

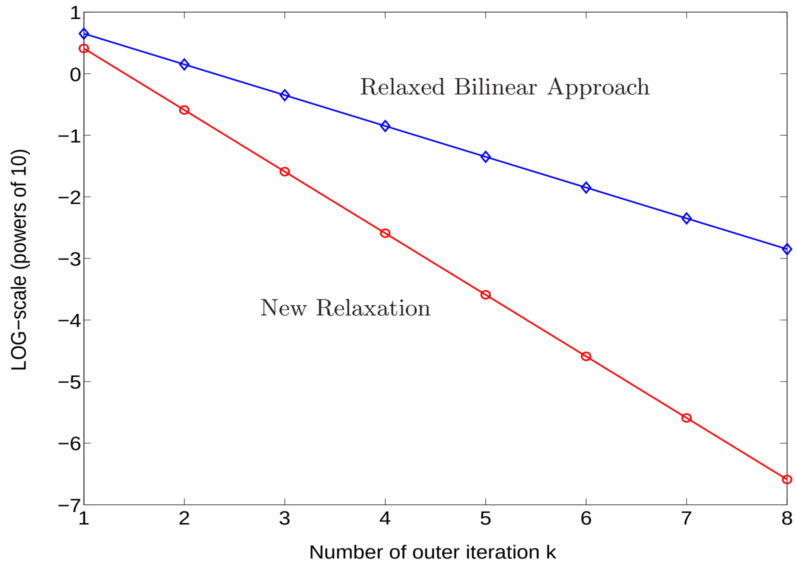


Figure 4.11: Comparison of Convergence $\|(x_1^k, x_2^k) - (x_1^*, x_2^*)\|_2$ for $t^k = 0.1^{k-2}$ and for Example 4.4

2.5 can directly be detected (since the resulting figures for all three test problems are almost identical we confine ourselves to present only the plot for Example 4.4).

Summary

In this doctoral thesis we introduced a new relaxation scheme for MPECs. We presented some basic properties of the relaxed nonlinear program we obtain applying the new relaxation scheme to the complementarity constraints. We also related the stationary points and solutions of the relaxed nonlinear program with those ones of the original MPEC. Our results resemble the corresponding result of [FLRS06], where the equality of the set of stationary points of the relaxed program and the set of strongly stationary points of the original MPEC is proved for the exact (relaxation-free) approach. Furthermore, our results form stronger results than comparable, existing ones for other regularization schemes.

Moreover, we presented some results concerning the convergence of a sequence of stationary points or solutions, respectively, of the relaxed nonlinear program to stationary points of the original MPEC. The first two convergence theorems form pendants to convergence results existing for regularization schemes for MPECs in the literature. However, having in mind the results of [Fle05], these convergence results (in particular those ones that resemble Theorem 2.3), are somewhat unsatisfactory, since they have strong assumptions and relatively weak conclusions. To our knowledge, there exist only a few comparable convergence results that do not make use of the relatively strong MPEC-LICQ. Hence, with the introduction of the MPEC-CRCQ and Theorem 2.4 we could present an improved convergence result, as we weakened the assumptions considerably and proved similar conclusions. This convergence result also shows that our relaxation method finds stationary points of MPECs that do not have strongly stationary solutions under suitable assumptions. However, as some examples and the results of the thesis of Flegel [Fle05] reveal, it might be a valuable task to get more insights in the convergence behaviour of a sequence of stationary points of the parametrized, relaxed nonlinear programs and to strengthen the existing convergence results even more.

In Chapter 3 we discussed the convergence behaviour of a general local SQP algorithm applied to the relaxed nonlinear program. We proved the local superlinear convergence of a sequence of iterates for a relaxed nonlinear program, where we assume that the relaxation parameter $t > 0$ is small enough. Open questions in this context concern for example the effect of globalization strategies, similar to the analysis in [Ani05a], and the case that the parameter $t > 0$ is not yet small enough, similar to the convergence results of Section 2.3 (specified for an SQP algorithm).

We then considered a modified two-sided relaxation scheme that combines our new relaxation scheme with the two-sided relaxation scheme of DeMiguel et al. [DFNS05] and we proved a local convergence result for an appropriate interior point algorithm similar to that stated in [DFNS05]. A subsequent task that deals with this modified two-sided relaxation scheme concerns a global convergence analysis of an appropriately adapted interior-point algorithm.

The last chapter of this thesis is devoted to the numerical experience we obtained for the new relaxation scheme. If we consider small iteration numbers as performance measure, then we have to admit that the exact bilinear approach of [FLRS06] outperformed the new relaxation method. However, it clearly came out, that the new relaxation scheme performs considerably better than the regularization scheme by Scholtes [Sch01]. Moreover, the new relaxation method seems to constitute a good compromise between small iteration numbers on the one hand and robustness on the other hand. Furthermore, since most of the MPECs of the test suite MacMPEC, thus of our test problem set, have strongly stationary solutions, one main advantage of the new relaxation method does not appear substantially in the numerical results we presented. It is therefore of interest, whether the relative numerical performance can be improved effectively, if we test the relaxation method on a test problem set that contains (more) MPECs that do not have strongly stationary solutions.

Finally, we would like to mention, that the theoretical properties and the convergence results of Chapter 2 might easily be further extended by applying the new relaxation scheme to so-called Equilibrium Problems with Equilibrium Constraints, or EPECs. These problems form an extension of MPECs and so far little is known about this challenging new field of nonlinear programming. Hence, some efficient solution methods for EPECs are still required.

Bibliography

[AAP04] S. Akella, M. Anitescu, and J. Peng. Mathematical Programs with Complementarity Constraints for Multiple Robot Coordination with Contact and Friction. *Proceedings of the IEEE International Conference in Robotics and Automation (ICRA 2004), New Orleans, LA*, pages 5224–5232, 2004.

[Ani05a] M. Anitescu. Global convergence of an elastic mode approach for a class of mathematical programs with complementarity constraints. *SIAM Journal on Optimization*, 16(1):120–145, 2005.

[Ani05b] M. Anitescu. On using the elastic mode in nonlinear programming approaches to mathematical programs with complementarity constraints. *SIAM Journal on Optimization*, 15(4):1203–1236, 2005.

[ATW07] M. Anitescu, P. Tseng, and S. J. Wright. Elastic-mode algorithms for mathematical programs with equilibrium constraints: global convergence and stationarity properties. *Mathematical Programming, Series A*, 110:337–371, 2007.

[AW05] J.S. Arora and Q. Wang. Review of formultations for structural and mechanical system optimization. *Structural Multidisciplinary Optimization*, 30:251–272, 2005.

[Bar98] J.F. Bard. *Practical Bilevel Optimization: Algorithms and Applications*. Kluwer Academic Publishers, Dordrecht, 1998.

[BR05] L.T. Biegler and A.U. Raghunathan. An interior point method for mathematical programs with complementarity constraints (mpccs). *SIAM Journal on Optimization*, 15(3):720–750, 2005.

[BRB07] A.T. Baumrucker, J.G. Renfro, and L.T. Biegler. MPEC Problem Formulations in Chemical Engineering Applications. preprint, June 2007.

[BS07] G. Bouza and G. Still. Mathematical programs with complementarity constraints: Convergence properties of a smoothing method. *Mathematics of Operations research*, 32(2):467–483, 2007.

[CGT00] A.R. Conn, N.I.M. Gould, and P.L. Toint. *Trust-Region Methods*. SIAM, 2000.

[CHLM06] Y. Chen, B.F. Hobbs, S. Leyffer, and T.S. Munson. Leader-follower equilibria for electric power and NO_X. *Computation Management and Science*, 3(4):307–330, 2006.

[CMS07] B. Colson, P. Marcotte, and G. Savard. An overview of bilevel optimization. *Annals of Operations Research*, 153:235–256, 2007.

[Dem02] S. Dempe. *Foundations of bilevel programming*. Nonconvex Optimization and Its Applications 61, Kluwer Academic Publishers, Dordrecht , 2002.

[DFNS05] V. DeMiguel, M. P. Friedlander, F. J. Nogales, and S. Scholtes. A two-sided relaxation scheme for mathematical programs with equilibrium constraints. *SIAM Journal on Optimization*, 16(2):587–609, 2005.

[DM02] E.D. Dolan and J.J. Moré. Benchmarking optimization software with performance profiles. *Mathematical Programming*, 91:201–213, 2002.

[FC68] A.V. Fiacco and G. P. Mc Cormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. J. Wiley and Sons, 1968.

[FFG99] M.C. Ferris, R. Fourer, and D. M. Gay. Expressing complementarity problems in an algebraic modeling language and communicating them to solvers. *SIAM Journal on Optimimization*, 9(4):991–1009, 1999.

[FGK90] R. Fourer, D.M. Gay, and B.W. Kernighan. A modeling language for mathematical programming. *Management Science*, 36:519–554, 1990.

[FGK03] R. Fourer, D.M. Gay, and B.W. Kernighan. *AMPL A Modeling Language for Mathematical Programming*. Thomson Brooks/Cole, second edition, 2003.

[FJQ99] F. Facchinei, H. Jiang, and L. Qi. A Smoothing Method for Mathematical Programs with Equilibrium Constraints. *Mathematical Programming*, 85(1):107–134, 1999.

[FK05a] M.L. Flegel and C. Kanzow. Abadie-type constraint qualification for mathematical programs with equilibrium constraints. *Journal of Optimization Theory and Applications*, 124(3):595–614, 2005.

[FK05b] M.L. Flegel and C. Kanzow. On M-stationary points for mathematical programs with equilibrium constraints. *Journal of Mathematical Analysis and Applications*, 310(1):286–302, 2005.

[FKO07] M.L. Flegel, C. Kanzow, and J.V. Outrata. Optimality conditions for disjunctive programs with application to mathematical programs with equilibrium constraints. *Set-Valued Analysis*, 15:139–162, 2007.

[FL98] R. Fletcher and S. Leyffer. User manual for filterSQP. *University of Dundee Numerical Analysis Report*, (NA\181), 1998.

[FL02a] R. Fletcher and S. Leyffer. Nonlinear programming without a penalty function. *Mathematical Programming, Series A*, (91):239–269, 2002.

[FL02b] R. Fletcher and S. Leyffer. Numerical experience with solving MPECs as NLPs. *Numerical Analysis Report NA/210, Department of Mathematics, University of Dundee*, 2002.

[FL04a] R. Fletcher and S. Leyffer. Solving mathematical programs with complementary constraints as nonlinear programs. *Optimization Methods and Software*, 19(1):15–40, 2004.

[FL04b] M. Fukushima and G.-H. Lin. Smoothing methods for mathematical programs with equilibrium constraints. *Proceedings of the International Conference on Informatics Research for Development of Knowledge Society Infrastructure*, pages 206–213, 2004.

[FL05] M. Fukushima and G.H. Lin. A modified relaxation scheme for mathematical programs with complementarity constraints. *Annals of Operations Research*, 133:63–84, 2005.

[Fle00] R. Fletcher. *Practical Methods of Optimization, Second Edition*. John Wiley & Sons, New York, 2000.

[Fle05] M.L. Flegel. *Constraint Qualifications and Stationarity Concepts for Mathematical Programs with Equilibrium Constraints*. PhD thesis, Julius-Maximilians-Universität Würzburg, 2005.

[FLP98] M. Fukushima, Z.-Q. Luo, and J.-S. Pang. A Globally Convergent Sequentional Quadratic Programming Algorithm for Mathmematical Programs with Linear Complementarity Constraints. *Computational Optimization and Applications*, 10:5–34, 1998.

[FLRS06] R. Fletcher, S. Leyffer, D. Ralph, and S. Scholtes. Local convergence of SQP methods for mathematical programs with equilibrium constraints. *SIAM Journal on Optimization*, 17(1):259–286, 2006.

[FLT07] R. Fletcher, S. Leyffer, and P.L. Toint. A Brief History of Filter Methods. *SIAG/Optimization Views-and-News*, 18(1):2–12, 2007.

[FP97] M.C. Ferris and J.S. Pang. Engineering and economic applications of complementarity problems. *SIAM Review*, 39(4):669–713, 1997.

[FP99] M. Fukushima and J.-S. Pang. Complementarity constraint qualifications and simplified B-stationary conditions for mathematical programs with equilibrium constraints. *Computational Optimization and Applications*, 13(1-3):111–136, 1999.

[FP03] F. Facchinei and J.S. Pang. *Finite-dimensional variational inequalities and complementarity problems*, volume 1. Springer Series in Operations Research. New York, Springer , 2003.

[FTL99] M.C. Ferris and F. Tin-Loi. On the solution of a minimum weight elastoplastic problem involving displacement and complementarity constraints. *Computer Methods in Applied Mechanics and Engineering*, 174:107–120, 1999.

[Gay97] D.M. Gay. Hooking your solver to ampl. *Technical Report*, (97-4-06), 1997.

[GK02] C. Geiger and C. Kanzow. *Theorie und Numerik restringierter Optimierungsaufgaben*. Springer, 2002.

[GMS02] P.E. Gill, W. Murray, and M.A. Saunders. Snopt: An sqp algorithm for large-scale constrained optimization. *SIAM Journal on Optimization*, 12(4):979–1006, 2002.

[GMS05] P.E. Gill, W. Murray, and M.A. Saunders. Snopt: An sqp algorithm for large-scale constrained optimization. *SIAM Review*, 47(1):99–131, 2005.

[HI06] M.J. Holler and G. Illing. *Einführung in die Spieltheorie*. Springer, 2006.

[HMP00] B.F. Hobbs, C.B. Metzler, and J.-S. Pang. Strategic gaming analysis for electric power systems: an mpec approach. *IEEE Transactions on Power Systems,*, 15(2):638 – 645, 2000.

[HP06] Y. Hollander and J.N. Prashker. The applicability of non-cooperative game theory in transport analysis. *Transportation*, 33(5):481–496, 2006.

[HR04] X.M. Hu and D. Ralph. Convergence of a penalty method for mathematical programming with complementarity constraints. *Journal of Optimization Theory and Applications*, 123(2):365–390, 2004.

[Jan84] R. Janin. Directional derivative of the marginal function in nonlinear programming. *Mathematical Programming Study*, 21:110–126, 1984.

[JKZ98] F. Jarre, M. Kočvara, and J. Zowe. Optimal truss design by interior-point methods. *SIAM Journal on Optimization*, 8(4):1084–1107, 1998.

[JLM06] J.-B. Jian, J.-L. Li, and X.-D. Mo. A Strongly and Superlinearly Convergent SQP Algorithm for Optimization Problems with Linear Complementarity Constraints. *Applied Mathematics and Optimization*, 54:17–46, 2006.

[JR99] H. Jiang and D. Ralph. QPECgen, a MATLAB generator for mathematical programs with quadratic objectives and affine variational inequality constraints. *Computational Optimization and Applications*, 13(1-3):25–59, 1999.

[JR00] H. Jiang and D. Ralph. Smooth SQP methods for mathematical programs with nonlinear complementarity constraints. *SIAM Journal on Optimization*, 10(3):779–808, 2000.

[JSRF07] J.J. Júdice, H.D. Sherali, I.M. Ribeiro, and A.M. Faustino. Complementarity Active Set Algorithm for Mathematical Programms with Equilibrium Constraints. *Journal of Optimization Theory and Applications*, 134:467–481, 2007.

[KM06] Y. Kanno and J.A.C. Martins. Arc-length method for frictional contact problems using mathematical programming with complementarity constraints. *Journal of Optimization Theory and Applications*, 131(1):89–113, 2006.

[Kön95] K. Königsberger. *Analysis 1*. Springer, third edition, 1995.

[Kos93] P. Kosmol. *Methoden zur numerischen Lösung nichtlinearer Gleichungen und Optimierungsaufgaben*. Teubner Studienbücher, 1993.

[KOZ98] M. Kočvara, J. Outrata, and J. Zowe. *Nonsmooth approach to optimization problems with equilibrium constraints. Theory, applications and numerical results.* Nonconvex Optimization and Its Applications. 28. Dordrecht: Kluwer Academic Publishers. , 1998.

[Ley00] S. Leyffer. MacMPEC: AMPL collection of MPECs. www.mcs.anl.gov/ leyffer/MacMPEC, 2000.

[Ley06] S. Leyffer. Complementarity constraints as nonlinear equations: theory and numerical experience. Dempe, Stephan (ed.) et al., Optimization with multivalued mappings. New York, NY: Springer. Springer Optimization and Its Applications 2, 169-208 (2006)., 2006.

[LF03a] G.-H. Lin and M. Fukushima. New Relaxation Method for Mathematical Programs with Complementarity Constraints. *Journal of Optimization Theory and Applications*, 118(1):81–116, 2003.

[LF03b] G.-H. Lin and M. Fukushima. Some Exact Penalty Resuls for Nonlinear Programs and Mathematical Programs with Equilibrium Constraints. *Journal of Optimization Theory and Applications*, 118(1):67–80, 2003.

[LH04] S. Lawphongpanich and D.W. Hearn. An MPEC approach to second-best toll pricing. *Mathematical Programming, Series B*, 101:33–55, 2004.

[LLCN06] S. Leyffer, G. López-Calva, and J. Nocedal. Interior methods for mathematical programs with complementarity constraints. *SIAM Journal on Optimization*, 17(1):52–77, 2006.

[LPR96] Z.-Q. Luo, J.-S. Pang, and D. Ralph. *Mathematical programs with equilibrium constraints.* Cambridge: Cambridge Univ. Press., 1996.

[LPRW96] Z.-Q. Luo, J.-S. Pang, D. Ralph, and S.Q. Wu. Exact penalization and stationarity conditions of mathematical programs with equilibrium constraints. *Mathematical Programming, Series A*, 75(1):19–76, 1996.

[LS04] X. Liu and J. Sun. Generalized stationary points and an interior-point method for mathematical programs with equilibrium constraints. *Mathematical Programming, Series B*, 101:231–261, 2004.

[LY07] G.-S. Liu and J.J. Ye. Merit Function Piecewise SQP Algorithm for Mathematical Programs with Equilibrium Constraints. *Journal of Optimization Theory and Applications*, 135:623–641, 2007.

[Pad99] M. Padberg. *Linear Optimization and Extensions*. Springer, second edition, 1999.

[RW04] D. Ralph and S.J. Wright. Some properties of regularization and penalization schemes for MPECs. *Optimization Methods and Software*, 19(5):527–556, 2004.

[Sch01] S. Scholtes. Convergence properties of a regularization scheme for mathematical programs with complementarity constraints. *SIAM Journal on Optimization*, 11(4):918–936, 2001.

[Sch04] W. Schlee. *Einführung in die Spieltheorie*. Vieweg, 2004.

[Spe98] P. Spellucci. An SQP method for general nonlinear programs using only equality constrained subproblems. *Mathematical Programming*, 82(3):413–448, 1998.

[SS99] S. Scholtes and M. Stöhr. Exact penalization of mathematical programs with equilibrium constraints. *SIAM Journal on Control and Optimization*, 37(2):617–652, 1999.

[SS00] H. Scheel and S. Scholtes. Mathematical programs with complementarity constraints: stationarity, optimality, and sensitivity. *Mathematical Operations Research*, 25(1):1–22, 2000.

[TLQ01] F. Tin-Loi and N.S. Que. Parameter Idenification of Quasi-brittle Materials as Mathematical Program with Equilibrium Constraints. *Computer Methods in Applied Mechanics and Engineering*, 190:5819–5836, 2001.

[Ven02] P. Venkataraman. *Applied optimization with MATLAB programming*. John Wiley, 2002.

[Ye99] J.J. Ye. Optimality conditions for optimization problems with complementarity constraints. *SIAM Journal on Optimization*, 9(2):374–387, 1999.

[Ye05] J.J. Ye. Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints. *Journal of Mathematical Analysis and Applications*, 307(1):350–369, 2005.

[YY96] H. Yamashita and H. Yabe. Superlinear and quadratic convergence of some primal-dual interior point methods for constrained optimization. *Mathematical Programming*, 75:377–397, 1996.

[ZLW04] J.Z. Zhang, G.S. Liu, and S.Y. Wang. A globally convergent approximately active search algorithm for solving mathematical programs with linear complementarity constraints. *Numerische Mathematik*, 98:539–558, 2004.

[ZLZ07] Z. Zhu, Z. Luo, and J. Zeng. A new smoothing technique for mathematical programs with equilibrium constraints. *Applied Mathematics and Mechanics (English Edition)*, 28(10):1407–1414, 2007.

A Numerical Results

A.1 Simple Outer Algorithm

Table A.1 : Results for the Simple Algorithm (Algorithm 4.1) - Iterationcounts

problem	data1	data2	data3	data4	data5	data6	data7	$\theta_p(z)$
bard1	9	22(1)	22(1)	22(1)	5(2)	5(2)	5(2)	5(2)
bard2	2	2(1)	2(1)	2(1)	3(1)	3(1)	3(1)	3(1)
bard3	4	23(1)	23(1)	23(1)	-1	-1	-1	-1
bar-truss	9	-2	-2	-2	63(1)	63(1)	63(1)	-1
bilevel1	4	4(1)	4(1)	4(1)	22(1)	22(1)	22(1)	24(2)
bilevel2	5	23(1)	23(1)	23(1)	20(2)	20(2)	20(2)	30(2)
bilevel3	6	26(2)	26(2)	26(2)	33(3)	32(2)	32(2)	103(3)
bilin	1	9(2)	9(2)	9(2)	9(3)	22(3)	21(2)	9(3)
dempe	2	22(1)	22(1)	22(1)	18(1)	18(1)	18(1)	18(1)
design-cent-1	5	6(1)	6(1)	6(1)	5(1)	5(1)	5(1)	5(1)
design-cent-4	-2	6(1)	6(1)	6(1)	-1	-1	-1	-1
desilva	5	5(1)	5(1)	5(1)	5(1)	5(1)	5(1)	5(1)
df1	2	25(1)	25(1)	25(1)	2(1)	2(1)	2(1)	2(1)
ex9.1.1	1	21(2)	21(2)	21(2)	24(1)	24(1)	24(1)	30(1)
ex9.1.2	1	3(1)	3(1)	3(1)	27(3)	2(2)	2(2)	35(3)
ex9.1.3	1	18(2)	18(2)	18(2)	17(3)	21(2)	21(2)	17(3)
ex9.1.4	-1	-1	-1	-1	-1	-1	-1	-1
ex9.1.5	1	23(2)	23(2)	23(2)	26(3)	7(2)	2(2)	36(3)
ex9.1.6	-1	-1	-1	-1	49(2)	49(2)	49(2)	-1
ex9.1.7	1	18(2)	18(2)	18(2)	19(3)	21(2)	21(2)	17(3)
ex9.1.8	1	19(1)	19(1)	19(1)	15(2)	4(2)	4(2)	29(2)
ex9.1.9	-1	37(2)	37(2)	37(2)	-1	-1	-1	-1
ex9.1.10	1	19(1)	19(1)	19(1)	15(2)	4(2)	4(2)	29(2)
ex9.2.1	9	20(1)	20(1)	20(1)	5(2)	5(2)	5(2)	5(2)
ex9.2.2	1	-2	-2	-2	44(9)	-2	-2	46(10)
ex9.2.3	2	3(1)	3(1)	3(1)	5(1)	5(1)	5(1)	4(1)
ex9.2.4	2	22(2)	22(2)	22(2)	4(2)	2(2)	2(2)	4(1)
ex9.2.5	1	1(1)	1(1)	1(1)	5(2)	5(2)	5(2)	5(2)
ex9.2.6	1	12(1)	12(1)	12(1)	22(2)	4(2)	4(2)	33(2)
ex9.2.7	9	20(1)	20(1)	20(1)	5(2)	5(2)	5(2)	5(2)
ex9.2.8	1	19(1)	19(1)	19(1)	22(2)	4(2)	4(2)	33(2)
ex9.2.9	1	1(1)	1(1)	1(1)	6(2)	6(2)	6(2)	6(2)

problem	(1)	(2)	(3)	(4)	(5)	(6)	(7)	$\theta_p(z)$
flp2	1	1(1)	1(1)	1(1)	3(3)	2(2)	2(2)	3(3)
flp4-1	2	283(1)	283(1)	283(1)	3(1)	3(1)	3(1)	3(1)
flp4-2	2	-2	-2	-2	3(1)	3(1)	3(1)	3(1)
flp4-3	2	-2	-2	-2	3(1)	3(1)	3(1)	3(1)
flp4-4	2	-2	-2	-2	3(1)	3(1)	3(1)	3(1)
gauvin	1	17(1)	17(1)	17(1)	5(2)	5(2)	5(2)	5(2)
gnash10	6	6(1)	6(1)	6(1)	6(1)	6(1)	6(1)	6(1)
gnash11	7	7(1)	7(1)	7(1)	7(1)	7(1)	7(1)	7(1)
gnash12	7	7(1)	7(1)	7(1)	7(1)	7(1)	7(1)	7(1)
gnash13	8	8(1)	8(1)	8(1)	23(1)	23(1)	23(1)	33(1)
gnash14	8	8(1)	8(1)	8(1)	26(1)	26(1)	26(1)	39(2)
gnash15	9	-1	-1	-1	17(3)	14(2)	14(2)	17(3)
gnash16	8	-1	-1	-1	15(2)	13(2)	13(2)	15(2)
gnash17	12	-1	-1	-1	13(2)	13(2)	13(2)	13(2)
gnash18	58	-1	-1	-1	22(4)	17(3)	-2	22(4)
gnash19	8	-1	-1	-1	25(3)	22(2)	22(2)	29(1)
hs044-i	2	59(2)	40(2)	40(2)	25(2)	60(2)	56(2)	38(2)
incid-set1-8	58	8(1)	8(1)	8(1)	8(1)	8(1)	8(1)	8(1)
incid-set1-16	120	13(1)	13(1)	13(1)	13(1)	13(1)	13(1)	13(1)
incid-set1c-8	47	5(1)	5(1)	5(1)	5(1)	5(1)	5(1)	5(1)
incid-set1c-16	78	22(1)	22(1)	22(1)	22(1)	22(1)	22(1)	22(1)
incid-set2-8	19	29(1)	29(1)	29(1)	29(1)	29(1)	29(1)	31(1)
incid-set2-16	25	28(1)	28(1)	28(1)	28(1)	28(1)	28(1)	28(1)
incid-set2c-8	37	29(1)	29(1)	29(1)	24(1)	24(1)	24(1)	25(1)
incid-set2c-16	29	24(1)	24(1)	24(1)	24(1)	24(1)	24(1)	24(1)
jr1	1	3(1)	3(1)	3(1)	1(1)	1(1)	1(1)	1(1)
jr2	-2	5(2)	5(2)	5(2)	6(3)	22(2)	22(2)	6(3)
kth1	1	2(1)	2(1)	2(1)	1(1)	1(1)	1(1)	1(1)
kth2	2	22(1)	22(1)	22(1)	1(1)	1(1)	1(1)	1(1)
kth3	-2	21(1)	21(1)	21(1)	22(2)	22(2)	22(2)	51(2)
liswet1-050	1	6(1)	6(1)	6(1)	1(1)	1(1)	1(1)	1(1)
liswet1-200	1	19(1)	19(1)	19(1)	5(1)	5(1)	5(1)	5(1)
liswet1-200	1	18(1)	18(1)	18(1)	4(1)	4(1)	4(1)	4(1)
nash1	1	1(1)	1(1)	1(1)	5(2)	21(2)	21(2)	6(2)
outrata31	8	54(3)	54(3)	53(3)	18(2)	16(2)	16(2)	17(2)
outrata32	9	71(3)	-2	-2	17(2)	-1	-1	18(2)
outrata33	7	49(2)	49(2)	-2	15(2)	-2	-2	15(2)
outrata34	7	-2	-2	-2	22(3)	-2	32(2)	23(3)
pack-comp1-8	68	5(4)	4(3)	3(2)	8(3)	7(2)	42(5)	8(3)
pack-comp1-16	-3	9(1)	9(1)	9(1)	8(1)	8(1)	8(1)	8(1)
pack-comp1c-8	92	5(4)	4(3)	3(2)	8(3)	7(2)	42(5)	8(3)
pack-comp1c-16	378	6(1)	6(1)	6(1)	6(1)	6(1)	6(1)	6(1)
pack-comp1p-8	-3	33(3)	32(2)	86(5)	48(4)	-1	112(3)	47(4)
pack-comp1p-16	-3	-2	-2	-2	-2	-2	-2	-2

problem	(1)	(2)	(3)	(4)	(5)	(6)	(7)	$\theta_p(z)$
pack-comp2-8	24	9(1)	9(1)	9(1)	9(1)	9(1)	9(1)	9(1)
pack-comp2-16	34	-2	16(1)	-2	-2	-2	16(1)	-2
pack-comp2c-8	7	6(1)	6(1)	6(1)	6(1)	6(1)	6(1)	6(1)
pack-comp2c-16	7	7(1)	7(1)	7(1)	7(1)	7(1)	7(1)	7(1)
pack-comp2p-8	-2	-2	-2	-2	-2	-2	-2	-2
pack-comp2p-16	-2	-2	-2	-2	-2	-2	-2	-2
pack-rig1-4	8	8(1)	8(1)	8(1)	8(1)	8(1)	8(1)	8(1)
pack-rig1-8	13	11(1)	11(1)	11(1)	11(1)	11(1)	11(1)	11(1)
pack-rig1-16	-2	-2	-1	-1	-1	-1	375(3)	-1
pack-rig1c-4	6	6(1)	6(1)	6(1)	6(1)	6(1)	6(1)	6(1)
pack-rig1c-8	11	6(1)	6(1)	6(1)	6(1)	6(1)	6(1)	6(1)
pack-rig1c-16	20	64(4)	56(3)	59(4)	107(5)	64(5)	56(3)	139(5)
pack-rig1p-4	4	2(1)	2(1)	2(1)	2(1)	2(1)	2(1)	2(1)
pack-rig1p-8	38	10(1)	10(1)	10(1)	10(1)	10(1)	10(1)	10(1)
pack-rig1p-16	-2	-2	-2	-2	-2	-1	-2	-1
pack-rig2-4	7	7(1)	7(1)	7(1)	7(1)	7(1)	7(1)	7(1)
pack-rig2-8	11	-1	44(3)	294(3)	-1	22(3)	44(3)	-1
pack-rig2-16	-3	-1	-1	-1	-1	-1	-1	-1
pack-rig2c-4	4	4(1)	4(1)	4(1)	4(1)	4(1)	4(1)	4(1)
pack-rig2c-8	32	83(5)	46(3)	49(3)	87(6)	14(3)	48(3)	102(6)
pack-rig2c-16	-3	-1	-1	-1	-1	-1	-1	-1
pack-rig2p-4	27	96(4)	25(3)	25(3)	33(5)	58(3)	-1	46(5)
pack-rig2p-8	-2	-2	66(3)	-2	-2	32(3)	63(3)	-1
pack-rig2p-16	-2	-2	-2	497(3)	872(6)	-1	-2	-1
pack-rig3-4	6	6(1)	6(1)	6(1)	6(1)	6(1)	6(1)	6(1)
pack-rig3-8	17	164(4)	45(3)	-2	202(5)	-2	-2	143(5)
pack-rig3-16	-2	-1	-1	-1	-1	-1	411(3)	-1
pack-rig3c-4	4	4(1)	4(1)	4(1)	4(1)	4(1)	4(1)	4(1)
pack-rig3c-8	7	62(4)	25(3)	49(3)	63(5)	9(3)	25(3)	61(5)
pack-rig3c-16	53	-2	-1	-1	-2	-1	175(3)	-2
portfl1	2	17(4)	4(3)	3(2)	18(5)	16(3)	4(3)	18(5)
portfl2	2	6(4)	5(3)	4(2)	6(5)	4(3)	4(3)	6(5)
portfl3	2	7(4)	21(3)	3(2)	8(5)	4(3)	21(3)	8(5)
portfl4	2	-2	10(3)	-2	-2	-2	10(3)	-2
portfl6	2	6(4)	10(3)	2(2)	7(5)	3(3)	10(3)	7(5)
qpec-100-1	-2	210(4)	52(2)	-2	254(5)	-2	-1	927(5)
qpec-100-2	68	-2	-2	200(2)	342(4)	436(3)	-1	372(5)
qpec-100-3	-2	172(3)	-2	-2	194(4)	309(3)	-2	242(4)
qpec-100-4	-2	-2	136(3)	90(2)	208(4)	227(3)	-2	648(4)
qpec-200-1	-2	274(4)	-2	-2	336(5)	-2	-2	625(5)
qpec-200-2	-3	- 3	-2	-2	-2	-2	-1	-3
qpec-200-3	-2	-2	-2	-2	488(5)	-2	-1	-2
qpec-200-4	135	126(3)	-2	-2	283(5)	-2	-1	-3
ralph1	-2	9(9)	5(5)	4(4)	10(10)	6(6)	4(4)	10(10)

problem	(1)	(2)	(3)	(4)	(5)	(6)	(7)	$\theta_p(z)$
ralph2	1	34(9)	24(5)	21(3)	39(10)	27(5)	28(4)	39(10)
ralphmod	25	-2	-1	-1	-1	-2	-2	-1
scholtes1	3	6(1)	6(1)	6(1)	2(1)	2(1)	2(1)	2(1)
scholtes2	2	3(1)	3(1)	3(1)	2(1)	2(1)	2(1)	2(1)
scholtes3	-2	20(2)	20(2)	20(2)	21(3)	43(3)	-2	33(3)
scholtes4	-2	2(1)	2(1)	2(1)	11(10)	7(5)	5(4)	11(10)
scholtes5	1	19(1)	19(1)	19(1)	1(1)	1(1)	1(1)	1(1)
sl1	1	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)
stackelberg1	4	4(1)	4(1)	4(1)	4(1)	4(1)	4(1)	4(1)
tap-09	12	-2	-2	-2	41(1)	41(1)	41(1)	43(2)
tap-15	-1	52(1)	52(1)	52(1)	56(2)	56(2)	56(2)	153(2)
water-net	108	-2	-2	214(2)	271(1)	271(1)	271(1)	170(1)
water-FL	-2	429(4)	412(3)	-2	-2	513(3)	406(3)	488(5)

A.2 Modified Outer Algorithm

The following part of the Appendix contains three tables concerning the numerical results we refer to in Section 4.2 and 4.3: first, the results for the modified outer algorithm, hence Algorithm 4.2, then our results concerning the relaxed bilinear reformulation (confer Section 1.4), thus Algorithm 4.3 and finally the results for the exact bilinear reformulation (confer Section 1.4), thus we apply Algorithm 4.3 only once to the reformulated problem $NLP(t)$ with $t = 0$.

Table A.2 :	Results for the Modified Algorithm
Table A.3 :	Results for the Relaxed Bilinear Reformulation
Table A.4 :	Results for the Exact Bilinear Reformulation

For our numerical tests we made the following choices for the required parameters:

Table A.2 :	$t_0 = 10.0$, $\sigma_1 = 0.50$, $\sigma_2 = 0.10$ and $\sigma_3 = 5.00$, $\varepsilon_C = 1.00E-08$, $\varepsilon_{SQP} = 1.00E-08$, $\delta_{\min} = 1.00E-12$ $\delta = 0.10$, maxit $= 15$
Table A.3 :	$t_0 = 10.0$, $\sigma = 0.10$, $\varepsilon_C = 1.00E-08$, $\varepsilon_{SQP} = 1.00E-08$, $\delta_{\min} = 1.00E-18$, maxit $= 20$
Table A.4 :	$t_0 = 0.00$, $\varepsilon_C = 1.00E-08$, $\varepsilon_{SQP} = 1.00E-08$

Finally, to shorten the notations, we use the following abbreviations:

n	: total number of variables
m	: total number of constraints (including complementarity constraints)
nc	: total number of complementarity constraints
ifail	: fail flag of `filterSQP` (see also [FL98]):
	ifail=0 *successful run*
	ifail=3 *nonlinear constraints locally infeasible*
	ifail=4 *h <= eps, but QP is infeasible*
	ifail=5 *termination with rho < eps*
it	: total number (sum) of inner (SQP) iterations
k	: total number of outer iterations (only for Tables A.2 and A.3)
fe	: total number (sum) of objective function evaluations
ce	: total number (sum) of constraint function evaluations
ge	: total number (sum) of gradient evaluations
he	: total number (sum) of hessian evaluations
uc	: total number (sum) of inner parameter updates (step 4 in Algorithm 4.2)
xc	: total number (sum) of projections of infeasible complementarity pairs
KKT	: norm of the KKT-residual (see also [FL98])
feas	: norm of the infeasibility residual (see also [FL98])
$\mathrm{compl}_1(x^*)$	: first complementarity measure : $\mathrm{compl}_1(x^*) := (\sum_{j=1}^{p} \min(x^*_{1j}, x^*_{2j})^2)^{1/2}$
$\mathrm{compl}_2(x^*)$	: second complementarity measure : $\mathrm{compl}_2(x^*) := (\sum_{j=1}^{p} (x^*_{1j}, x^*_{2j})^2)^{1/2}$
$\xi^*_{\max}$	: largest multiplier of the complementarity condition, thus of the conditions: $x_{1j}x_{2j} \leq 0$, $x_{1j}x_{2j} \leq t$ or $\Phi_j(x_1, x_2, t_j) \leq 0$
$f(x^*)$	: final value of the objective function (compare with values reported in [Ley00])
$(t_{\min}, t_{\max})$	: final values of the smallest/largest relaxation parameter t_j^k

Table A.2 : Results for the Modified Algorithm (Algorithm 4.2)

problem	n	m	nc	ifail	it(k)	fe	ce	ge	he	uc	xc
bard1	8	7	3	0	5 (2)	7	10	10	8	3	0
bard2	16	13	4	0	3 (1)	4	8	8	7	3	0
bard3	8	7	2	0	22 (4)	23	27	27	26	2	1
bar-truss	35	35	6	0	73 (2)	6	82	77	77	12	0
bilevel1	16	15	6	0	15 (1)	5	19	19	19	15	0
bilevel2	32	29	12	0	18 (2)	22	34	26	26	70	0
bilevel3	12	11	4	0	22 (2)	23	39	39	38	39	0
bilin	14	13	6	0	9 (3)	9	18	18	16	12	0
dempe	4	3	1	0	18 (1)	20	20	19	19	0	0
design-cent-1	15	15	3	0	5 (1)	6	7	7	6	0	0
design-cent-4	46	57	12	3	1673 (15)	2375	3262	1634	1622	2275	17
desilva	7	5	1	0	5 (1)	6	6	6	6	0	0
df1	3	4	1	0	2 (1)	3	3	3	3	0	0
ex9.1.1	13	12	5	0	4 (1)	5	8	8	7	2	0
ex9.1.2	10	9	4	0	13 (2)	14	25	25	25	30	0
ex9.1.3	23	21	6	0	10 (2)	8	16	16	16	11	0
ex9.1.4	10	9	4	3	39 (15)	34	57	57	45	0	5
ex9.1.5	13	12	5	0	14 (2)	15	23	23	23	16	0
ex9.1.6	14	13	6	0	14 (2)	15	29	29	28	42	0
ex9.1.7	17	15	6	0	10 (2)	8	16	16	16	11	0
ex9.1.8	12	11	4	0	5 (2)	7	10	10	9	6	0
ex9.1.9	12	11	5	0	18 (4)	20	32	28	26	13	1
ex9.1.10	12	11	4	0	5 (2)	7	10	10	9	6	0
ex9.2.1	10	9	4	0	5 (2)	7	10	10	8	3	0
ex9.2.2	10	19	4	0	49 (14)	62	114	114	110	93	0
ex9.2.3	16	15	6	0	4 (1)	4	7	7	7	8	0
ex9.2.4	8	7	2	0	4 (2)	5	7	7	7	2	0
ex9.2.5	8	7	3	0	6 (2)	8	12	12	10	6	0
ex9.2.6	12	10	4	0	8 (2)	9	12	12	12	4	0
ex9.2.7	10	9	4	0	5 (2)	7	10	10	8	3	0
ex9.2.8	6	5	2	0	8 (2)	9	12	12	12	2	0
ex9.2.9	9	8	3	0	16 (2)	14	22	16	16	2	0
flp2	6	4	2	0	1 (1)	2	3	3	2	0	0
flp4-1	110	90	30	0	3 (1)	4	5	5	4	0	0
flp4-2	170	170	60	0	3 (1)	4	5	5	4	0	0
flp4-3	210	240	70	0	3 (1)	4	5	5	4	0	0
flp4-4	300	350	100	0	3 (1)	4	5	5	4	0	0
gauvin	5	4	2	0	5 (2)	7	14	14	12	6	0
gnash10	17	16	8	0	6 (1)	5	10	10	10	24	0
gnash11	17	16	8	0	7 (1)	6	12	12	12	32	0
gnash12	17	16	8	0	7 (1)	6	12	12	12	32	0
gnash13	17	16	8	0	11 (1)	10	20	20	20	61	0
gnash14	17	16	8	0	12 (1)	10	21	21	21	55	0
gnash15	17	16	8	0	17 (3)	14	30	30	29	53	0
gnash16	17	16	8	0	15 (2)	10	23	23	23	36	0
gnash17	17	16	8	0	14 (2)	11	23	23	23	44	0
gnash18	17	16	8	0	22 (4)	23	47	47	44	116	0
gnash19	17	16	8	0	14 (3)	15	29	29	28	69	0
hs044-i	46	40	10	0	21 (2)	21	37	37	36	101	0

KKT	feas	$\text{compl}_1(x^*)$	$\text{compl}_2(x^*)$	$\xi^*_{\max}$	$f(x^*)$	$(t_{\min}, t_{\max})$
0.0000	0.4441E-15	0.0000	0.0000	1.3333	17.000	(0.500 , 0.500)
0.22469E-13	0.1776E-14	0.0000	0.0000	0.0000	-6598.0	(1.25 , 10.0)
0.22204E-15	0.000	0.0000	0.0000	0.22656	-12.679	(0.250 , 0.250)
0.22650E-12	0.7532E-12	0.0000	0.0000	417.86	10167.	(6.25 , 50.0)
0.0000	0.4619E-13	0.0000	0.0000	0.50000	-10.000	(1.25 , 2.50)
0.54113E-13	0.7216E-14	0.0000	0.0000	0.0000	-6600.0	(0.977E-03 , 1.00)
0.22204E-15	0.8882E-15	0.0000	0.0000	0.66016	-12.679	(0.610E-04 , 0.625E-01)
0.0000	0.1332E-14	0.0000	0.0000	38.000	-14.600	(0.625E-02 , 0.500E-01)
0.49145E-08	0.2457E-08	0.0000	0.0000	0.0000	31.250	(10.0 , 10.0)
0.95416E-14	0.8604E-15	0.0000	0.0000	0.0000	-1.8606	(10.0 , 10.0)
0.0000	6.500	1.7500	35.889	1.0000	-3341.0	(0.100E-11 , 3.91)
0.44787E-13	0.4441E-15	0.0000	0.0000	0.0000	-1.0000	(10.0 , 10.0)
0.0000	0.000	0.0000	0.0000	0.0000	0.0000	(10.0 , 10.0)
0.0000	0.1776E-14	0.0000	0.0000	0.0000	-13.000	(2.50 , 10.0)
0.0000	0.000	0.0000	0.0000	0.0000	-6.2500	(0.977E-03 , 0.250)
0.0000	0.1221E-14	0.0000	0.0000	7.7500	-23.000	(0.625E-01 , 1.00)
0.0000	0.6667	0.33333	10.667	1.0000	-63.000	(3.91 , 3.91)
0.0000	0.3331E-15	0.0000	0.0000	0.0000	-1.0000	(0.781E-02 , 1.00)
0.20015E-15	0.7550E-14	0.0000	0.0000	0.10000	-49.000	(0.244E-03 , 0.500)
0.0000	0.1221E-14	0.0000	0.0000	6.5000	-23.000	(0.625E-01 , 1.00)
0.0000	0.4441E-15	0.0000	0.0000	0.25000	-3.2500	(0.250 , 1.00)
0.55511E-16	0.1332E-14	0.0000	0.0000	0.22222	3.1111	(0.391E-02 , 0.500)
0.0000	0.4441E-15	0.0000	0.0000	0.25000	-3.2500	(0.250 , 1.00)
0.0000	0.1332E-14	0.0000	0.0000	1.3333	17.000	(0.500 , 1.00)
0.39721E-14	0.000	0.0000	0.0000	9.9474	100.00	(0.100E-11 , 0.100E-11)
0.0000	0.1066E-13	0.0000	0.0000	0.62500	5.0000	(2.50 , 10.0)
0.0000	0.000	0.0000	0.0000	0.50000	0.50000	(0.500 , 0.500)
0.0000	0.2220E-15	0.0000	0.0000	3.0000	9.0000	(0.250 , 0.250)
0.0000	0.000	0.0000	0.0000	0.25000	-1.0000	(0.250 , 1.00)
0.0000	0.1332E-14	0.0000	0.0000	1.3333	17.000	(0.500 , 1.00)
0.0000	0.000	0.0000	0.0000	0.25000	1.5000	(0.250 , 1.00)
0.0000	0.8882E-15	0.0000	0.0000	0.0000	2.0000	(0.500 , 1.00)
0.0000	0.3553E-14	0.0000	0.0000	0.0000	0.0000	(10.0 , 10.0)
0.0000	0.1090E-11	0.0000	0.0000	0.0000	0.0000	(10.0 , 10.0)
0.0000	0.2726E-11	0.0000	0.0000	0.0000	0.0000	(10.0 , 10.0)
0.0000	0.5246E-11	0.0000	0.0000	0.0000	0.0000	(10.0 , 10.0)
0.0000	0.1327E-10	0.0000	0.0000	0.0000	0.0000	(10.0 , 10.0)
0.17764E-14	0.000	0.0000	0.0000	0.50000	20.000	(0.312E-01 , 0.500)
0.59403E-11	0.7816E-12	0.0000	0.0000	4.8683	-230.82	(1.25 , 1.25)
0.10631E-14	0.3553E-14	0.0000	0.0000	3.5475	-129.91	(0.625 , 0.625)
0.25197E-08	0.5529E-09	0.0000	0.0000	1.7999	-36.933	(0.625 , 0.625)
0.45754E-14	0.3109E-13	0.0000	0.0000	0.74948	-7.0618	(0.391E-01 , 0.312)
0.39307E-11	0.1288E-12	0.0000	0.0000	0.10735	-0.17905	(0.391E-01 , 10.0)
0.20192E-12	0.7994E-14	0.0000	0.0000	5.7633	-354.70	(0.195E-03 , 0.250E-01)
0.21175E-10	0.7603E-12	0.0000	0.0000	4.7030	-241.44	(0.156E-01 , 0.125)
0.25106E-12	0.6306E-13	0.0000	0.0000	2.8920	-90.749	(0.781E-02 , 0.625E-01)
0.22123E-11	0.1190E-12	0.0000	0.0000	1.4679	-25.698	(0.381E-07 , 0.250E-02)
0.26295E-10	0.2495E-11	0.0000	0.0000	0.70029	-6.1167	(0.488E-04 , 0.250E-01)
0.20351E-14	0.1444E-13	0.0000	0.0000	17.542	17.090	(0.122E-03 , 1.00)

problem	n	m	nc	ifail	it(k)	fe	ce	ge	he	uc	xc
incid-set1-8	247	249	49	0	8 (1)	9	10	10	9	0	0
incid-set1-16	999	1005	225	0	26 (1)	40	41	17	16	0	0
incid-set1-32	4039	4053	961	0	6005 (7)	9476	9483	3473	3466	0	0
incid-set1c-8	247	256	49	0	5 (1)	6	7	7	6	0	0
incid-set1c-16	999	1020	225	0	22 (1)	32	33	17	16	0	0
incid-set1c-32	4039	4084	961	0	40 (1)	57	58	27	26	0	0
incid-set2-8	247	249	49	0	29 (1)	46	46	19	19	0	0
incid-set2-16	999	1005	225	0	31 (1)	48	48	19	19	0	0
incid-set2-32	4039	4053	961	0	58 (1)	79	79	38	39	0	0
incid-set2c-8	247	256	49	0	24 (1)	37	37	15	15	0	0
incid-set2c-16	999	1020	225	0	22 (1)	30	30	16	16	0	0
incid-set2c-32	4039	4084	961	0	45 (1)	68	68	27	27	0	0
jr1	3	2	1	0	1 (1)	2	2	2	2	0	0
jr2	3	2	1	0	6 (3)	9	11	11	10	1	0
kth1	2	2	1	0	1 (1)	2	2	2	2	0	0
kth2	2	2	1	0	1 (1)	2	2	2	2	0	0
kth3	2	2	1	0	8 (2)	10	12	12	12	2	0
liswet1-050	202	153	50	0	1 (1)	2	3	3	2	0	0
liswet1-100	402	303	100	0	2 (1)	3	6	6	5	146	0
liswet1-200	802	603	200	0	3 (1)	4	8	8	7	527	0
nash1	7	5	1	0	5 (2)	7	9	9	7	0	0
outrata31	17	16	4	0	18 (2)	19	27	25	25	12	0
outrata32	17	16	4	0	17 (2)	17	24	24	25	12	0
outrata33	17	16	4	0	15 (2)	14	22	22	22	10	0
outrata34	17	16	4	0	19 (2)	18	29	29	30	22	0
pack-comp1-8	237	251	49	0	8 (3)	10	10	10	11	0	0
pack-comp1-16	981	1025	225	0	8 (1)	9	9	9	9	0	0
pack-comp1-32	4005	4157	961	5	293 (15)	216	319	263	287	9067	0
pack-comp1c-8	237	258	49	0	8 (3)	10	10	10	11	0	0
pack-comp1c-16	981	1040	225	0	6 (1)	7	7	7	7	0	0
pack-comp1c-32	4005	4188	961	0	5 (1)	6	6	6	6	0	0
pack-comp1p-8	237	236	49	0	48 (4)	86	88	40	39	10	0
pack-comp1p-16	981	980	225	0	60 (15)	100	113	73	72	2488	0
pack-comp1p-32	4005	4004	961	0	597 (15)	748	1183	664	707	25949	3
pack-comp2-8	237	251	49	0	9 (1)	12	12	9	9	0	0
pack-comp2-16	981	1025	225	0	40 (7)	29	33	33	33	730	0
pack-comp2-32	4005	4157	961	4	87 (15)	48	57	47	57	0	0
pack-comp2c-8	237	258	49	0	6 (1)	7	7	7	7	0	0
pack-comp2c-16	981	1040	225	0	7 (1)	8	8	8	8	0	0
pack-comp2c-32	4005	4188	961	4	39 (15)	28	47	40	31	0	0
pack-comp2p-8	237	236	49	0	88 (15)	142	156	93	95	531	0
pack-comp2p-16	981	980	225	0	130 (15)	181	214	120	127	2514	0
pack-comp2p-32	4005	4004	961	5	344 (15)	315	594	276	336	73	0
pack-rig1-4	57	62	9	0	8 (1)	9	10	10	9	0	0
pack-rig1-8	237	251	49	0	11 (1)	13	14	12	11	0	0
pack-rig1-16	981	1025	225	0	246 (6)	134	368	248	270	4553	0
pack-rig1-32	4005	4157	961	4	78 (15)	78	98	58	50	0	0
pack-rig1c-4	57	65	9	0	6 (1)	7	8	8	7	0	0
pack-rig1c-8	237	258	49	0	6 (1)	7	8	8	7	0	0
pack-rig1c-16	981	1040	225	0	39 (5)	44	64	62	58	2018	0
pack-rig1c-32	4005	4188	961	4	57 (15)	30	50	45	35	0	0
pack-rig1p-4	57	56	9	0	2 (1)	3	4	4	3	0	0

KKT	feas	$\text{compl}_1(x^*)$	$\text{compl}_2(x^*)$	$\xi^*_{\max}$	$f(x^*)$	$(t_{\min}, t_{\max})$
0.0000	0.8762E-12	0.0000	0.0000	0.0000	0.38164E-16	(10.0 , 10.0)
0.14209E-12	0.2070E-08	0.0000	0.0000	0.0000	0.12403E-15	(10.0 , 10.0)
0.71235E-08	0.3511E-13	0.0000	0.0000	0.0000	0.22908E-06	(0.156E+06 , 0.156E+06)
0.0000	0.1712E-08	0.0000	0.0000	0.0000	0.38164E-16	(10.0 , 10.0)
0.0000	0.2318E-13	0.0000	0.0000	0.0000	0.12013E-15	(10.0 , 10.0)
0.25099E-09	0.6894E-09	0.0000	0.0000	0.0000	0.98370E-05	(10.0 , 10.0)
0.44476E-09	0.6490E-11	0.0000	0.0000	0.0000	0.45179E-02	(10.0 , 10.0)
0.25939E-09	0.1131E-08	0.0000	0.0000	0.0000	0.29978E-02	(10.0 , 10.0)
0.47172E-08	0.7085E-12	0.0000	0.0000	0.0000	0.17695E-02	(10.0 , 10.0)
0.27921E-08	0.6982E-10	0.0000	0.0000	0.0000	0.54713E-02	(10.0 , 10.0)
0.30468E-11	0.1465E-10	0.0000	0.0000	0.0000	0.35996E-02	(10.0 , 10.0)
0.30020E-08	0.2064E-09	0.0000	0.0000	0.0000	0.24357E-02	(10.0 , 10.0)
0.0000	0.000	0.0000	0.0000	0.0000	0.50000	(10.0 , 10.0)
0.0000	0.000	0.0000	0.0000	0.50000	0.50000	(0.500E-01 , 0.500E-01)
0.0000	0.000	0.0000	0.0000	0.0000	0.0000	(10.0 , 10.0)
0.0000	0.000	0.0000	0.0000	0.0000	0.0000	(10.0 , 10.0)
0.0000	0.000	0.0000	0.0000	0.50000	0.50000	(0.250 , 0.250)
0.72768E-14	0.5537E-13	0.0000	0.0000	0.0000	0.13994E-01	(10.0 , 10.0)
0.95093E-14	0.3409E-12	0.0000	0.0000	0.0000	0.13734E-01	(2.50 , 10.0)
0.18169E-12	0.2359E-11	0.0000	0.0000	0.0000	0.17009E-01	(1.25 , 10.0)
0.0000	0.3553E-14	0.0000	0.0000	0.0000	0.0000	(1.00 , 1.00)
0.43351E-08	0.5893E-08	0.0000	0.0000	0.12190	3.2077	(0.156E-01 , 1.00)
0.25676E-13	0.1146E-12	0.0000	0.0000	0.11744	3.4494	(0.156E-01 , 1.00)
0.26040E-10	0.1767E-10	0.0000	0.0000	0.43120	4.6043	(0.312E-01 , 1.00)
0.45776E-15	0.2665E-14	0.0000	0.0000	0.92654	6.5927	(0.195E-02 , 1.00)
0.0000	0.6983E-15	0.0000	0.0000	0.0000	0.60000	(0.100 , 0.100)
0.35402E-08	0.9800E-14	0.0000	0.0000	0.0000	0.61695	(10.0 , 10.0)
0.21220-313	0.6981E-09	0.0000	0.0000	0.0000	0.65298	(0.977E-07 , 195.313)
0.0000	0.6137E-15	0.0000	0.0000	0.0000	0.60000	(0.100 , 0.100)
0.58199E-13	0.5823E-14	0.0000	0.0000	0.0000	0.62304	(10.0 , 10.0)
0.87098E-08	0.5540E-09	0.0000	0.0000	0.0000	0.66144	(10.0 , 10.0)
0.74630E-09	0.6327E-15	0.0000	0.0000	0.0000	0.60000	(0.500E-02 , 0.100E-01)
0.37216E-06	0.2073E-14	0.0000	0.0000	0.0000	0.61695	(0.100E-11 , 0.100E-11)
0.23859E-06	0.6162E-14	0.0000	0.0000	0.0000	0.65298	(0.100E-11 , 0.250E-09)
0.22501E-08	0.2190E-09	0.0000	0.0000	0.0000	0.67312	(10.0 , 10.0)
0.94769E-08	0.3422E-14	0.0000	0.0000	0.0000	0.72714	(0.625E-06 , 0.100E-04)
0.69532-309	0.000	2.6647	6.9181	0.0000	0.71341	(0.122E+10 , 0.122E+10)
0.16093E-09	0.1630E-10	0.0000	0.0000	0.0000	0.67346	(10.0 , 10.0)
0.72487E-08	0.7865E-14	0.0000	0.0000	0.0000	0.72747	(10.0 , 10.0)
0.44466-322	0.000	0.0000	0.0000	0.0000	0.78294	(0.122E+10 , 0.122E+10)
0.47749E-06	0.8936E-15	0.0000	0.0000	0.0000	0.67312	(0.100E-11 , 0.100E-11)
0.21587E-06	0.2534E-14	0.0000	0.0000	0.0000	0.72714	(0.100E-11 , 0.100E-11)
0.21220-313	0.000	0.0000	0.0000	0.0000	0.78260	(0.488E-06 , 0.488E+06)
0.58477E-10	0.8676E-12	0.0000	0.0000	0.0000	0.71889	(10.0 , 10.0)
0.63142E-08	0.3154E-14	0.0000	0.0000	0.0000	0.78793	(10.0 , 10.0)
0.24611E-08	0.9680E-15	0.0000	0.0000	0.0000	0.82601	(0.100E-11 , 0.100E-03)
0.69532-309	0.000	0.13456E-02	0.31290E-05	0.0000	0.85089	(0.122E+10 , 0.122E+10)
0.12857E-12	0.2776E-15	0.0000	0.0000	0.0000	0.72101	(10.0 , 10.0)
0.86160E-09	0.5196E-13	0.0000	0.0000	0.0000	0.78830	(10.0 , 10.0)
0.25153E-08	0.2302E-14	0.0000	0.0000	0.0000	0.82650	(0.305E-07 , 0.100E-02)
0.44466-322	0.000	0.0000	0.0000	0.0000	0.85164	(0.122E+10 , 0.122E+10)
0.0000	0.8327E-16	0.0000	0.0000	0.0000	0.60000	(10.0 , 10.0)

problem	n	m	nc	ifail	it(k)	fe	ce	ge	he	uc	xc
pack-rig1p-8	237	236	49	0	10 (1)	11	12	12	11	0	0
pack-rig1p-16	981	980	225	0	231 (15)	213	350	288	279	11081	0
pack-rig1p-32	4005	4004	961	0	628 (15)	427	960	562	607	52207	29
pack-rig2-4	57	62	9	0	7 (1)	8	8	8	8	0	0
pack-rig2-8	237	251	49	0	107 (6)	49	171	117	129	297	0
pack-rig2-16	981	1025	225	3	35 (15)	15	40	35	35	0	0
pack-rig2-32	4005	4157	961	4	46 (15)	15	31	21	27	0	0
pack-rig2c-4	57	65	9	0	4 (1)	5	5	5	5	0	0
pack-rig2c-8	237	258	49	0	43 (6)	39	63	58	53	205	0
pack-rig2c-16	981	1040	225	3	34 (15)	15	36	34	34	0	0
pack-rig2c-32	4005	4188	961	5	46 (15)	15	31	24	27	0	0
pack-rig2p-4	57	56	9	0	28 (5)	25	39	37	40	60	0
pack-rig2p-8	237	236	49	0	150 (6)	58	199	146	151	39	1
pack-rig2p-16	981	980	225	3	1610 (15)	219	3068	1238	1658	14507	1
pack-rig2p-32	4005	4004	961	5	53 (15)	39	44	38	39	0	0
pack-rig3-4	57	62	9	0	6 (1)	7	7	7	7	0	0
pack-rig3-8	237	251	49	0	117 (6)	35	194	110	134	51	1
pack-rig3-16	981	1025	225	0	913 (15)	167	1644	718	971	7381	2
pack-rig3-32	4005	4157	961	4	220 (15)	214	263	145	144	0	0
pack-rig3c-4	57	65	9	0	4 (1)	5	5	5	5	0	0
pack-rig3c-8	237	258	49	0	50 (6)	45	73	63	59	209	3
pack-rig3c-16	981	1040	225	0	75 (6)	70	120	106	106	3564	0
pack-rig3c-32	4005	4188	961	5	52 (15)	25	43	39	29	0	0
portfl1	111	49	12	0	18 (5)	23	43	43	38	57	0
portfl2	111	49	12	0	6 (5)	11	19	19	14	22	0
portfl3	111	49	12	0	8 (5)	15	25	23	18	37	0
portfl4	111	49	12	0	41 (10)	51	100	100	90	368	0
portfl6	111	49	12	0	7 (5)	12	22	22	17	37	0
qpec-100-1	205	202	100	0	58 (4)	66	116	106	104	2830	0
qpec-100-2	210	202	100	0	53 (3)	54	103	103	101	2067	0
qpec-100-3	210	204	100	0	64 (6)	67	128	126	122	3641	0
qpec-100-4	220	204	100	0	132 (4)	98	191	167	165	1858	0
qpec-200-1	410	404	200	0	160 (5)	129	227	199	198	5279	0
qpec-200-2	420	404	200	0	243 (11)	248	432	347	338	21895	0
qpec-200-3	420	408	200	0	214 (5)	171	310	285	284	10842	0
qpec-200-4	440	408	200	0	9 (2)	11	12	10	9	0	0
ralph1	3	5	1	0	10 (10)	20	20	20	20	0	0
ralph2	2	2	1	0	40 (10)	50	50	50	50	0	0
ralphmod	204	200	100	5	251 (15)	165	409	373	372	8569	6
scholtes1	4	3	1	0	2 (1)	3	3	3	3	0	0
scholtes2	4	3	1	0	2 (1)	3	3	3	3	0	0
scholtes3	2	2	1	0	44 (2)	46	48	30	30	2	0
scholtes4	3	5	1	0	11 (10)	21	21	21	21	0	0
scholtes5	7	6	2	0	1 (1)	2	3	3	2	0	0
sl1	11	8	3	0	1 (1)	2	3	3	2	0	0
stackelberg1	3	2	1	0	4 (1)	5	10	10	9	4	0
tap-09	162	142	36	0	19 (1)	16	34	34	34	234	0
tap-15	1560	2100	675	0	22 (1)	20	39	35	35	691	0
tollmpec	6051	6100	1824	0	1398 (15)	1237	2806	1579	1622	211172	465
tollmpec1	6051	6100	1824	3	923 (15)	159	1442	976	1022	36001	126
water-net	180	214	64	0	107 (1)	87	130	56	70	0	0
water-FL	1737	2424	784	0	663 (5)	477	804	410	519	962	0

KKT	feas	$\text{compl}_1(x^*)$	$\text{compl}_2(x^*)$	$\xi^*_{\max}$	$f(x^*)$	$(t_{\min}, t_{\max})$
0.71639E-08	0.2793E-15	0.0000	0.0000	0.0000	0.78793	(10.0 , 10.0)
0.60913E-04	0.8270E-15	0.0000	0.0000	0.0000	0.82601	(0.100E-11 , 0.100E-11)
0.30906	0.4148E-14	0.0000	0.0000	0.45969E-02	0.85089	(0.100E-11 , 0.125E-07)
0.45874E-13	0.4276E-15	0.0000	0.0000	0.0000	0.69453	(10.0 , 10.0)
0.96271E-08	0.1738E-14	0.0000	0.0000	0.0000	0.78040	(0.954E-08 , 0.500E-02)
0.16036E-13	0.1408E-03	0.81137E-02	0.15367E-03	0.0000	0.99917	(0.610E+11 , 0.610E+11)
0.69533-309	0.000	0.11503E-01	0.14561E-03	0.0000	0.99978	(0.610E+11 , 0.610E+11)
0.32004E-09	0.1429E-11	0.0000	0.0000	0.0000	0.71153	(10.0 , 10.0)
0.27492E-10	0.3067E-12	0.0000	0.0000	0.0000	0.79931	(0.156E-03 , 0.500E-02)
0.13381E-13	0.1408E-03	0.81137E-02	0.15367E-03	0.0000	0.99917	(0.610E+11 , 0.610E+11)
0.69532-309	0.4866E-03	0.11483E-01	0.15530E-03	0.0000	0.99978	(0.610E+11 , 0.610E+11)
0.0000	0.3892E-16	0.0000	0.0000	0.0000	0.60000	(0.391E-05 , 0.100E-02)
0.86553E-08	0.4265E-15	0.0000	0.0000	0.0000	0.78040	(0.250E-02 , 0.500E-02)
2041.7	0.5643	0.39317E-01	0.28899E-02	0.20892E+06	13.240	(0.100E-11 , 0.977E+04)
0.69533-309	0.000	0.95509E-02	0.11963E-03	0.0000	1.1359	(0.977E+04 , 0.977E+04)
0.87962E-11	0.1510E-12	0.0000	0.0000	0.0000	0.66619	(10.0 , 10.0)
0.58069E-13	0.2680E-15	0.0000	0.0000	0.0000	0.73520	(0.122E-05 , 0.500E-02)
0.46030E-05	0.1835E-13	0.0000	0.0000	0.0000	0.80043	(0.100E-11 , 0.250E-09)
0.69533-309	0.000	0.29268	0.83663E-01	0.0000	0.87516	(0.244E+08 , 0.244E+08)
0.29579E-09	0.1402E-11	0.0000	0.0000	0.0000	0.68277	(10.0 , 10.0)
0.42143E-08	0.5007E-10	0.0000	0.0000	0.0000	0.75347	(0.156E-03 , 0.500E-02)
0.65407E-13	0.7314E-14	0.0000	0.0000	0.0000	0.81860	(0.149E-11 , 0.100E-03)
0.69533-309	0.000	0.14090E-01	0.32381E-03	0.0000	0.90891	(0.122E+10 , 0.122E+10)
0.55004E-16	0.3886E-15	0.0000	0.0000	0.73342E-02	0.15024E-04	(0.305E-07 , 0.500E-03)
0.42065E-16	0.2290E-15	0.0000	0.0000	0.10817E-01	0.14573E-04	(0.125E-03 , 0.500E-03)
0.24126E-16	0.2671E-15	0.0000	0.0000	0.41389E-02	0.62650E-05	(0.313E-04 , 0.500E-03)
0.91730E-08	0.1292E-15	0.0000	0.0000	0.40921E-02	0.21773E-05	(0.100E-11 , 0.100E-11)
0.22045E-16	0.1440E-15	0.0000	0.0000	0.48488E-02	0.23613E-05	(0.313E-04 , 0.500E-03)
0.15395E-14	0.1046E-13	0.0000	0.0000	0.47791	0.99003E-01	(0.100E-11 , 0.100E-01)
0.23780E-14	0.3131E-13	0.0000	0.0000	0.49641	-6.2605	(0.100E-11 , 0.100)
0.39161E-10	0.4063E-11	0.0000	0.0000	0.78961	-5.4506	(0.100E-11 , 0.313E-05)
0.56373E-14	0.2085E-13	0.0000	0.0000	1.6882	-4.0511	(0.100E-11 , 0.100E-01)
0.84431E-14	0.2975E-11	0.0000	0.0000	4.0984	-1.9348	(0.100E-11 , 0.100E-02)
0.17736E-10	0.7233E-11	0.0000	0.0000	1.7455	-23.833	(0.100E-11 , 0.100E-11)
0.89033E-14	0.1118E-09	0.0000	0.0000	2.7084	-1.9478	(0.100E-11 , 0.100E-02)
0.43448E-07	0.1062E-12	5.2276	7.7628	0.51040E-01	-17.906	(1.00 , 1.00)
0.15701E-15	0.000	0.18169E-08	0.33011E-17	0.50000	-0.18169E-08	(0.100E-07 , 0.100E-07)
0.76896E-08	0.000	0.0000	0.0000	0.29713E-08	0.25704E-17	(0.100E-07 , 0.100E-07)
1.5742	0.000	0.61272E-03	0.61075E-04	0.0000	-114.55	(0.312E-08 , 195.313)
0.0000	0.000	0.0000	0.0000	0.0000	2.0000	(10.0 , 10.0)
0.0000	0.000	0.0000	0.0000	0.0000	15.000	(10.0 , 10.0)
0.0000	0.000	0.0000	0.0000	0.50000	0.50000	(0.250 , 0.250)
0.0000	0.000	0.18169E-08	0.33011E-17	1.0000	-0.36338E-08	(0.100E-07 , 0.100E-07)
0.0000	0.000	0.0000	0.0000	0.0000	1.0000	(10.0 , 10.0)
0.43368E-18	0.2006E-14	0.0000	0.0000	0.0000	0.10000E-03	(10.0 , 10.0)
0.17764E-14	0.000	0.0000	0.0000	0.0000	-3266.7	(0.625 , 0.625)
0.19008E-08	0.3375E-12	0.0000	0.0000	0.0000	109.13	(0.610E-03 , 10.0)
0.25939E-08	0.1063E-11	0.0000	0.0000	0.16725E-01	184.50	(0.610E-03 , 10.0)
48.947	0.2530E-08	112.14	8154.2	20.474	-1187.3	(0.100E-11 , 195.313)
1.7484	53.69	9.7673	120.84	47599.	789.78	(0.100E-11 , 3.91)
0.12257E-08	0.7105E-14	0.0000	0.0000	0.0000	1023.4	(10.0 , 10.0)
0.99045E-08	0.3009E-12	0.0000	0.0000	27.802	3410.2	(0.100E-11 , 0.100E-02)

Table A.3 : Results for the Relaxed Bilinear Reformulation

problem	n	m	nc	ifail	it(k)	fe	ce	ge	he
bard1	8	7	3	0	29 (10)	39	48	48	39
bard2	16	13	4	0	2 (1)	3	4	4	3
bard3	8	7	2	0	57 (10)	137	142	62	60
bar-truss	35	35	6	0	30 (11)	35	49	49	41
bilevel1	16	15	6	0	23 (9)	32	41	41	32
bilevel2	32	29	12	0	24 (10)	34	44	44	34
bilevel3	12	11	4	0	88 (11)	173	184	96	88
bilin	14	13	6	0	34 (11)	45	56	56	45
dempe	4	3	1	0	18 (1)	20	20	19	19
design-cent-1	15	15	3	0	5 (1)	6	7	7	6
design-cent-4	46	57	12	0	7980 (20)	12197	12369	4374	4363
desilva	7	5	1	0	5 (1)	6	6	6	6
df1	3	4	1	0	2 (1)	3	3	3	3
ex9.1.1	13	12	5	0	9 (9)	18	27	27	18
ex9.1.2	10	9	4	0	11 (11)	22	23	23	22
ex9.1.3	23	21	6	0	31 (11)	42	52	52	42
ex9.1.4	10	9	4	0	30 (11)	40	51	51	41
ex9.1.5	13	12	5	0	6 (3)	9	10	10	9
ex9.1.6	14	13	6	0	32 (10)	41	51	51	42
ex9.1.7	17	15	6	0	32 (11)	43	53	53	43
ex9.1.8	12	11	4	0	11 (10)	21	22	22	21
ex9.1.9	12	11	5	0	23 (11)	34	42	42	34
ex9.1.10	12	11	4	0	11 (10)	21	22	22	21
ex9.2.1	10	9	4	0	30 (10)	40	50	50	40
ex9.2.2	10	19	4	0	91 (20)	110	129	127	109
ex9.2.3	16	15	6	0	21 (9)	30	37	37	30
ex9.2.4	8	7	2	0	27 (10)	37	46	46	37
ex9.2.5	8	7	3	0	39 (11)	50	60	60	50
ex9.2.6	12	10	4	0	24 (11)	35	44	44	35
ex9.2.7	10	9	4	0	30 (10)	40	50	50	40
ex9.2.8	6	5	2	0	12 (11)	23	24	24	23
ex9.2.9	9	8	3	0	10 (10)	20	29	29	20
flp2	6	4	2	0	1 (1)	2	3	3	2
flp4-1	110	90	30	0	1 (1)	2	3	3	2
flp4-2	170	170	60	0	1 (1)	2	3	3	2
flp4-3	210	240	70	0	1 (1)	2	3	3	2
flp4-4	300	350	100	0	1 (1)	2	3	3	2
gauvin	5	4	2	0	23 (10)	33	38	38	33
gnash10	17	16	8	0	20 (9)	27	31	31	29
gnash11	17	16	8	0	22 (9)	29	35	35	31
gnash12	17	16	8	0	20 (9)	27	32	32	29
gnash13	17	16	8	0	21 (9)	28	33	33	30
gnash14	17	16	8	0	25 (9)	31	42	42	34
gnash15	17	16	8	0	38 (11)	47	56	55	48
gnash16	17	16	8	0	34 (10)	41	50	49	43
gnash17	17	16	8	0	33 (10)	42	51	49	42
gnash18	17	16	8	0	48 (12)	58	70	69	59
gnash19	17	16	8	0	40 (11)	50	58	58	51
hs044-i	46	40	10	0	42 (11)	51	62	62	53

KKT	feas	$\text{compl}_1(x^*)$	$\text{compl}_2(x^*)$	$\xi^*_{\max}$	$f(x^*)$	t_k
0.60469E-18	0.2220E-15	0.33077E-08	0.14142E-07	0.76190	17.000	0.100E-07
0.14211E-13	0.2220E-14	0.0000	0.0000	0.0000	-6598.0	10.0
0.29733E-15	0.8882E-15	0.80000E-08	0.10000E-07	0.36250	-12.679	0.100E-07
0.44715E-08	0.1421E-12	0.24903E-11	0.17321E-08	1.4516	10167.	0.100E-08
0.12636E-15	0.2665E-13	0.60093E-08	0.14142E-06	0.15000	-10.000	0.100E-06
0.0000	0.1518E-13	0.62048E-08	0.20000E-07	0.0000	-6600.0	0.100E-07
0.59508E-16	0.9325E-15	0.14636E-08	0.17321E-08	1.0927	-12.679	0.100E-08
0.14656E-13	0.8882E-15	0.39520E-08	0.20000E-08	40.000	-14.600	0.100E-08
0.46521E-08	0.2326E-08	0.0000	0.0000	0.0000	31.250	10.0
0.95416E-14	0.8604E-15	0.0000	0.0000	0.0000	-1.8606	10.0
0.20174E-07	0.1646E-13	0.54264E-17	0.22361E-17	3.6427	-3.0792	0.100E-17
0.44787E-13	0.4441E-15	0.0000	0.0000	0.0000	-1.0000	10.0
0.0000	0.000	0.0000	0.0000	0.0000	0.0000	10.0
0.0000	0.1110E-15	0.71429E-08	0.10000E-06	0.0000	-13.000	0.100E-06
0.0000	0.2068E-24	0.10000E-08	0.10000E-08	0.0000	-6.2500	0.100E-08
0.25122E-14	0.1221E-14	0.30092E-08	0.17321E-08	3.2000	-29.200	0.100E-08
0.11102E-15	0.3664E-14	0.30009E-08	0.14142E-08	2.4375	-37.000	0.100E-08
0.0000	0.000	0.0000	0.0000	0.0000	-1.0000	0.100
0.28101E-08	0.1844E-13	0.35721E-08	0.20000E-07	0.66667E-01	-49.000	0.100E-07
0.17764E-14	0.1110E-14	0.30092E-08	0.17321E-08	2.6667	-26.000	0.100E-08
0.0000	0.4441E-15	0.25000E-08	0.10000E-07	0.12500	-3.2500	0.100E-07
0.55592E-16	0.4441E-15	0.10138E-08	0.14142E-08	0.44444	3.1111	0.100E-08
0.0000	0.4441E-15	0.25000E-08	0.10000E-07	0.12500	-3.2500	0.100E-07
0.89050E-15	0.4441E-15	0.33077E-08	0.14142E-07	0.76190	17.000	0.100E-07
1.8596	0.7108E-14	0.11481E-08	0.40790E-17	0.93944E+09	100.000	0.100E-17
0.53051E-08	0.1156E-13	0.83333E-08	0.14142E-06	0.83333E-01	5.0000	0.100E-06
0.81002E-12	0.2613E-15	0.10000E-07	0.10000E-07	1.00000	0.50000	0.100E-07
0.41875E-14	0.3109E-14	0.11180E-08	0.14142E-08	6.0000	9.0000	0.100E-08
0.64031E-08	0.1635E-15	0.14142E-08	0.14142E-08	0.50000	-1.0000	0.100E-08
0.89050E-15	0.4441E-15	0.33077E-08	0.14142E-07	0.76190	17.000	0.100E-07
0.41370E-16	0.2068E-24	0.10000E-08	0.10000E-08	0.50000	1.5000	0.100E-08
0.0000	0.9116E-16	0.16667E-08	0.10000E-07	0.0000	2.0000	0.100E-07
0.0000	0.3553E-14	0.0000	0.0000	0.0000	0.0000	10.0
0.0000	0.1097E-11	0.0000	0.0000	0.0000	0.0000	10.0
0.0000	0.2725E-11	0.0000	0.0000	0.0000	0.0000	10.0
0.0000	0.5238E-11	0.0000	0.0000	0.0000	0.0000	10.0
0.72155E-11	0.1327E-10	0.0000	0.0000	0.0000	0.46551E-23	10.0
0.56879E-08	0.7594E-16	0.25000E-08	0.10000E-07	0.25000	20.000	0.100E-07
0.33100E-14	0.3412E-13	0.22560E-08	0.20000E-06	0.14211	-230.82	0.100E-06
0.53745E-08	0.1903E-13	0.20610E-08	0.20000E-06	0.91802E-01	-129.91	0.100E-06
0.21779E-08	0.1243E-13	0.18269E-08	0.20000E-06	0.39663E-01	-36.933	0.100E-06
0.35341E-14	0.3845E-13	0.16923E-08	0.20000E-06	0.14908E-01	-7.0618	0.100E-06
0.12358E-09	0.4441E-14	0.16052E-08	0.20000E-06	0.19886E-02	-0.17905	0.100E-06
0.20290E-14	0.000	0.21060E-08	0.20000E-08	7.6472	-354.70	0.100E-08
0.35145E-14	0.1638E-13	0.71328E-08	0.20000E-07	1.9472	-241.44	0.100E-07
0.10246E-14	0.9770E-14	0.96833E-08	0.20000E-07	1.6742	-90.749	0.100E-07
0.12666E-10	0.5329E-14	0.13732E-08	0.22361E-09	12.712	-25.698	0.100E-09
0.15059E-10	0.1250E-13	0.20533E-08	0.22361E-08	2.8027	-6.1167	0.100E-08
0.42518E-08	0.4004E-14	0.50349E-08	0.20000E-08	5.6908	15.618	0.100E-08

problem	n	m	nc	ifail	it(k)	fe	ce	ge	he
incid-set1-8	247	249	49	0	8 (1)	9	10	10	9
incid-set1-16	999	1005	225	0	26 (1)	40	41	17	16
incid-set1-32	4039	4053	961	0	6005 (7)	9476	9483	3473	3466
incid-set1c-8	247	256	49	0	5 (1)	6	7	7	6
incid-set1c-16	999	1020	225	0	22 (1)	32	33	17	16
incid-set1c-32	4039	4084	961	0	40 (1)	57	58	27	26
incid-set2-8	247	249	49	0	31 (1)	47	47	19	19
incid-set2-16	999	1005	225	0	31 (1)	48	48	19	19
incid-set2-32	4039	4053	961	0	47 (1)	63	63	29	30
incid-set2c-8	247	256	49	0	20 (1)	28	28	13	13
incid-set2c-16	999	1020	225	0	22 (1)	30	30	16	16
incid-set2c-32	4039	4084	961	0	45 (1)	68	68	27	27
jr1	3	2	1	0	1 (1)	2	2	2	2
jr2	3	2	1	0	28 (11)	39	46	46	39
kth1	2	2	1	0	1 (1)	2	2	2	2
kth2	2	2	1	0	1 (1)	2	2	2	2
kth3	2	2	1	0	19 (10)	29	29	29	29
liswet1-050	202	153	50	0	1 (1)	2	3	3	2
liswet1-100	402	303	100	0	1 (1)	2	3	3	2
liswet1-200	802	603	200	0	1 (1)	2	3	3	2
nash1	7	5	1	0	26 (11)	37	47	47	37
outrata31	17	16	4	0	25 (10)	35	39	39	35
outrata32	17	16	4	0	24 (10)	34	39	39	34
outrata33	17	16	4	0	27 (10)	37	41	41	37
outrata34	17	16	4	0	30 (11)	41	46	46	41
pack-comp1-8	237	251	49	0	30 (13)	43	43	43	43
pack-comp1-16	981	1025	225	0	8 (1)	9	9	9	9
pack-comp1-32	4005	4157	961	5	89 (20)	58	67	46	55
pack-comp1c-8	237	258	49	0	30 (13)	43	43	43	43
pack-comp1c-16	981	1040	225	0	6 (1)	7	7	7	7
pack-comp1c-32	4005	4188	961	0	5 (1)	6	6	6	6
pack-comp1p-8	237	236	49	0	94 (17)	177	178	92	91
pack-comp1p-16	981	980	225	0	65 (20)	127	127	66	66
pack-comp1p-32	4005	4004	961	0	591 (20)	857	1022	396	447
pack-comp2-8	237	251	49	0	9 (1)	12	12	9	9
pack-comp2-16	981	1025	225	0	42 (9)	33	33	33	33
pack-comp2-32	4005	4157	961	4	262 (20)	124	244	166	199
pack-comp2c-8	237	258	49	0	6 (1)	7	7	7	7
pack-comp2c-16	981	1040	225	0	7 (1)	8	8	8	8
pack-comp2c-32	4005	4188	961	5	49 (20)	33	59	50	38
pack-comp2p-8	237	236	49	0	94 (20)	159	161	92	91
pack-comp2p-16	981	980	225	0	327 (20)	412	476	242	253
pack-comp2p-32	4005	4004	961	4	616 (20)	805	998	410	456
pack-rig1-4	57	62	9	0	8 (1)	9	10	10	9
pack-rig1-8	237	251	49	0	11 (1)	13	14	12	11
pack-rig1-16	981	1025	225	0	287 (15)	330	347	252	251
pack-rig1-32	4005	4157	961	4	314 (20)	249	328	250	240
pack-rig1c-4	57	65	9	0	6 (1)	7	8	8	7
pack-rig1c-8	237	258	49	0	6 (1)	7	8	8	7
pack-rig1c-16	981	1040	225	0	104 (14)	119	133	121	109
pack-rig1c-32	4005	4188	961	5	67 (20)	35	62	55	42
pack-rig1p-4	57	56	9	0	2 (1)	3	4	4	3

KKT	feas	$\text{compl}_1(x^*)$	$\text{compl}_2(x^*)$	$\xi^*_{\max}$	$f(x^*)$	t_k
0.0000	0.8762E-12	0.0000	0.0000	0.0000	0.38164E-16	10.0
0.14209E-12	0.2070E-08	0.0000	0.0000	0.0000	0.12403E-15	10.0
0.71235E-08	0.3511E-13	0.0000	0.0000	0.0000	0.22908E-06	0.100E-04
0.0000	0.1712E-08	0.0000	0.0000	0.0000	0.38164E-16	10.0
0.0000	0.2318E-13	0.0000	0.0000	0.0000	0.12013E-15	10.0
0.25099E-09	0.6894E-09	0.0000	0.0000	0.0000	0.98370E-05	10.0
0.16722E-09	0.1058E-10	0.0000	0.0000	0.0000	0.45179E-02	10.0
0.25939E-09	0.1131E-08	0.0000	0.0000	0.0000	0.29978E-02	10.0
0.44404E-08	0.3747E-13	0.0000	0.0000	0.0000	0.17497E-02	10.0
0.18098E-10	0.2327E-10	0.0000	0.0000	0.0000	0.54713E-02	10.0
0.30468E-11	0.1465E-10	0.0000	0.0000	0.0000	0.35996E-02	10.0
0.30020E-08	0.2064E-09	0.0000	0.0000	0.0000	0.24357E-02	10.0
0.0000	0.000	0.0000	0.0000	0.0000	0.50000	10.0
0.24236E-15	0.1121E-15	0.20000E-08	0.10000E-08	2.0000	0.50000	0.100E-08
0.0000	0.000	0.0000	0.0000	0.0000	0.0000	10.0
0.0000	0.000	0.0000	0.0000	0.0000	0.0000	10.0
0.45126E-14	0.000	0.10000E-07	0.10000E-07	1.00000	0.50000	0.100E-07
0.72768E-14	0.5537E-13	0.0000	0.0000	0.0000	0.13994E-01	10.0
0.10898E-13	0.2688E-12	0.0000	0.0000	0.0000	0.13734E-01	10.0
0.24796E-12	0.1827E-11	0.0000	0.0000	0.0000	0.17009E-01	10.0
0.0000	0.3735E-14	0.20000E-08	0.10000E-08	0.0000	0.0000	0.100E-08
0.23843E-15	0.9663E-15	0.67243E-08	0.10000E-07	0.16393	3.2077	0.100E-07
0.14649E-15	0.2686E-14	0.71430E-08	0.10000E-07	0.16778	3.4494	0.100E-07
0.16959E-14	0.1661E-13	0.82801E-08	0.10000E-07	0.71407	4.6043	0.100E-07
0.29013E-15	0.8249E-16	0.11183E-08	0.10000E-08	2.0723	6.5927	0.100E-08
0.0000	0.8353E-15	0.27704E-08	0.14142E-10	0.0000	0.60000	0.100E-10
0.35402E-08	0.9674E-14	0.0000	0.0000	0.0000	0.61695	10.0
0.69533-309	0.000	3.0881	14.142	1.0000	0.64681	0.100E-17
0.0000	0.9439E-15	0.27704E-08	0.14142E-10	0.0000	0.60000	0.100E-10
0.58199E-13	0.5823E-14	0.0000	0.0000	0.0000	0.62304	10.0
0.87098E-08	0.5540E-09	0.0000	0.0000	0.0000	0.66144	10.0
0.26132E-09	0.5375E-11	0.56847E-09	0.22190E-11	0.0000	0.60000	0.100E-14
0.58561E-07	0.1888E-14	0.0000	0.0000	0.0000	0.61695	0.100E-17
0.19205E-06	0.6150E-14	0.0000	0.0000	0.0000	0.65298	0.100E-17
0.22501E-08	0.2190E-09	0.0000	0.0000	0.0000	0.67312	10.0
0.91120E-08	0.8583E-10	0.0000	0.0000	0.0000	0.72714	0.100E-06
0.21220-313	0.000	0.0000	0.0000	0.0000	0.78260	0.100E-17
0.16093E-09	0.1630E-10	0.0000	0.0000	0.0000	0.67346	10.0
0.72487E-08	0.7865E-14	0.0000	0.0000	0.0000	0.72747	10.0
0.21220-313	0.000	0.0000	0.0000	0.0000	0.78294	0.100E-17
0.24934E-06	0.8435E-15	0.0000	0.0000	0.0000	0.67312	0.100E-17
0.20820E-06	0.2518E-14	0.0000	0.0000	0.0000	0.72714	0.100E-17
0.21220-313	0.000	0.0000	0.0000	0.0000	0.78260	0.100E-17
0.58477E-10	0.8676E-12	0.0000	0.0000	0.0000	0.71889	10.0
0.63142E-08	0.3154E-14	0.0000	0.0000	0.0000	0.78793	10.0
0.41090E-08	0.1350E-10	0.99280E-09	0.38781E-11	0.0000	0.82601	0.100E-12
0.69533-309	0.000	0.67457E-04	0.57766E-07	1.0000	0.85089	0.100E-17
0.12857E-12	0.2776E-15	0.0000	0.0000	0.0000	0.72101	10.0
0.86160E-09	0.5196E-13	0.0000	0.0000	0.0000	0.78830	10.0
0.23414E-08	0.5760E-13	0.11177E-08	0.43661E-11	0.0000	0.82650	0.100E-11
0.21220-313	0.000	0.0000	0.0000	0.0000	0.85164	0.100E-17
0.0000	0.8327E-16	0.0000	0.0000	0.0000	0.60000	10.0

problem	n	m	nc	ifail	it(k)	fe	ce	ge	he
pack-rig1p-8	237	236	49	0	10 (1)	11	12	12	11
pack-rig1p-16	981	980	225	0	132 (4)	169	173	91	90
pack-rig1p-32	4005	4004	961	5	102 (20)	114	141	78	65
pack-rig2-4	57	62	9	0	7 (1)	8	8	8	8
pack-rig2-8	237	251	49	0	32 (12)	44	44	44	44
pack-rig2-16	981	1025	225	3	49 (20)	20	57	50	49
pack-rig2-32	4005	4157	961	5	56 (20)	20	38	26	34
pack-rig2c-4	57	65	9	0	4 (1)	5	5	5	5
pack-rig2c-8	237	258	49	0	67 (12)	89	100	88	77
pack-rig2c-16	981	1040	225	3	50 (20)	20	52	50	50
pack-rig2c-32	4005	4188	961	4	56 (20)	20	39	29	35
pack-rig2p-4	57	56	9	0	34 (13)	47	47	47	47
pack-rig2p-8	237	236	49	0	86 (12)	121	126	95	91
pack-rig2p-16	981	980	225	0	50 (15)	69	70	56	55
pack-rig2p-32	4005	4004	961	4	61 (20)	43	53	41	49
pack-rig3-4	57	62	9	0	6 (1)	7	7	7	7
pack-rig3-8	237	251	49	0	32 (12)	44	45	45	44
pack-rig3-16	981	1025	225	0	277 (20)	232	294	224	227
pack-rig3-32	4005	4157	961	4	321 (20)	202	404	187	243
pack-rig3c-4	57	65	9	0	4 (1)	5	5	5	5
pack-rig3c-8	237	258	49	0	59 (12)	96	107	81	70
pack-rig3c-16	981	1040	225	5	100 (20)	134	154	116	97
pack-rig3c-32	4005	4188	961	4	62 (20)	30	56	49	37
portfl1	111	49	12	0	86 (17)	104	121	119	102
portfl2	111	49	12	0	116 (20)	140	167	144	125
portfl3	111	49	12	0	116 (20)	142	162	150	130
portfl4	111	49	12	0	97 (20)	101	121	121	101
portfl6	111	49	12	0	60 (15)	75	90	90	75
qpec-100-1	205	202	100	0	78 (12)	66	102	100	89
qpec-100-2	210	202	100	0	63 (12)	76	88	86	74
qpec-100-3	210	204	100	0	171 (17)	120	213	189	178
qpec-100-4	220	204	100	0	173 (18)	174	228	188	175
qpec-200-1	410	404	200	0	159 (14)	107	197	170	160
qpec-200-2	420	404	200	0	214 (12)	193	271	208	200
qpec-200-3	420	408	200	5	299 (20)	231	361	282	270
qpec-200-4	440	408	200	0	64 (3)	77	79	63	61
ralph1	3	5	1	0	86 (20)	106	106	106	106
ralph2	2	2	1	0	58 (20)	78	78	75	75
ralphmod	204	200	100	3	418 (20)	329	472	363	353
scholtes1	4	3	1	0	2 (1)	3	3	3	3
scholtes2	4	3	1	0	2 (1)	3	3	3	3
scholtes3	2	2	1	0	23 (10)	33	33	33	33
scholtes4	3	5	1	0	85 (20)	105	105	105	105
scholtes5	7	6	2	0	1 (1)	2	3	3	2
sl1	11	8	3	0	1 (1)	2	3	3	2
stackelberg1	3	2	1	0	4 (1)	5	6	6	5
tap-09	162	142	36	0	34 (10)	49	59	44	35
tap-15	1560	2100	675	0	37 (10)	48	59	47	38
tollmpec	6051	6100	1824	5	405 (20)	767	832	243	229
tollmpec1	6051	6100	1824	5	160 (20)	152	176	138	124
water-net	180	214	64	0	333 (13)	486	583	221	246
water-FL	1737	2424	784	0	1602 (13)	1307	1740	830	1081

KKT	feas	$compl_1(x^*)$	$compl_2(x^*)$	ξ^*_{max}	$f(x^*)$	t_k
0.71639E-08	0.2793E-15	0.0000	0.0000	0.0000	0.78793	10.0
0.70897E-08	0.2509E-14	0.0000	0.0000	0.0000	0.82601	0.100E-01
0.69533-309	0.000	0.85402	10.000	0.0000	0.78080	0.100E-17
0.45874E-13	0.4276E-15	0.0000	0.0000	0.0000	0.69453	10.0
0.84685E-11	0.1348E-13	0.64008E-08	0.10001E-09	0.0000	0.78040	0.100E-09
0.19962E-08	0.1408E-03	0.75373E-10	0.34731E-12	0.0000	0.99917	0.100E-17
0.69533-309	0.000	0.11503E-01	0.14561E-03	1.0000	0.99978	0.100E-17
0.32004E-09	0.1429E-11	0.0000	0.0000	0.0000	0.71153	10.0
0.75775E-11	0.9688E-13	0.64000E-08	0.10000E-09	0.0000	0.79931	0.100E-09
0.13384E-13	0.1408E-03	0.16000E-09	0.74999E-12	0.0000	0.99917	0.100E-17
0.69532-309	0.000	0.11483E-01	0.15530E-03	1.0000	0.99978	0.100E-17
0.0000	0.6362E-15	0.58451E-08	0.10001E-10	0.0000	0.60000	0.100E-10
0.36865E-08	0.1023E-13	0.55658E-08	0.10001E-09	0.0000	0.78040	0.100E-09
0.12797E-08	0.1070E-10	0.75289E-09	0.33912E-11	0.0000	1.0851	0.100E-12
0.69533-309	0.000	0.95509E-02	0.11963E-03	1.0000	1.1359	0.100E-17
0.87962E-11	0.1510E-12	0.0000	0.0000	0.0000	0.66619	10.0
0.29100E-08	0.2345E-15	0.58182E-08	0.10000E-09	0.0000	0.73520	0.100E-09
0.13134E-06	0.1300E-10	0.50003E-06	0.35355E-11	0.0000	0.80043	0.100E-17
0.69532-309	0.000	0.15813E-01	0.50880E-03	1.0000	0.69868	0.100E-17
0.29579E-09	0.1402E-11	0.0000	0.0000	0.0000	0.68277	10.0
0.16317E-08	0.2317E-10	0.58182E-08	0.10000E-09	0.0000	0.75347	0.100E-09
0.69532-309	0.9758E-08	0.50000E-06	0.27000E-11	0.0000	0.81860	0.100E-17
0.69532-309	0.000	0.14090E-01	0.32381E-03	1.0000	0.90891	0.100E-17
0.64728E-08	0.4606E-11	0.0000	0.0000	0.89651	0.15024E-04	0.100E-14
0.13874E-07	0.1643E-11	0.0000	0.0000	0.68203	0.14573E-04	0.100E-17
0.87597E-08	0.2307E-11	0.0000	0.0000	0.21907	0.62650E-05	0.100E-17
0.38752E-08	0.3450E-11	0.0000	0.0000	0.35009	0.21773E-05	0.100E-17
0.36551E-08	0.1086E-11	0.10421E-09	0.26458E-12	0.14880	0.23613E-05	0.100E-12
0.78854E-09	0.1767E-13	0.47666E-08	0.74833E-09	10.092	0.99003E-01	0.100E-09
0.10487E-08	0.3641E-13	0.12964E-08	0.60828E-09	1.8542	-6.5907	0.100E-09
0.62590E-08	0.1000E-11	0.0000	0.0000	10.170	-5.4817	0.100E-14
0.46966E-08	0.6832E-10	0.0000	0.0000	2.8843	-4.0648	0.100E-15
0.61026E-09	0.6863E-10	0.56319E-09	0.95917E-11	158.43	-1.9348	0.100E-11
0.71892E-08	0.1548E-12	0.17598E-08	0.92736E-09	1.4198	-24.036	0.100E-09
0.69532-309	0.5514E-08	0.33767E-08	0.14956E-08	0.0000	-1.9534	0.100E-17
0.60098E-07	0.5003E-10	4.6775	5.0137	0.41814	-16.877	0.100
0.0000	0.1000E-11	0.10000E-05	0.10000E-11	0.50000E+06	-0.10000E-05	0.100E-17
0.0000	0.2583E-12	0.50818E-06	0.25825E-12	2.0000	-0.51650E-12	0.100E-17
1.0362	0.2491E-07	0.0000	0.0000	0.14763E-03	-683.03	0.100E-17
0.0000	0.000	0.0000	0.0000	0.0000	2.0000	10.0
0.0000	0.000	0.0000	0.0000	0.0000	15.000	10.0
0.54523E-16	0.000	0.10000E-07	0.10000E-07	1.0000	0.50000	0.100E-07
0.0000	0.1000E-11	0.10000E-05	0.10000E-11	0.10000E+07	-0.20000E-05	0.100E-17
0.0000	0.000	0.0000	0.0000	0.0000	1.0000	10.0
0.43368E-18	0.2006E-14	0.0000	0.0000	0.0000	0.10000E-03	10.0
0.17764E-14	0.000	0.0000	0.0000	0.0000	-3266.7	10.0
0.66637E-10	0.7512E-13	0.39108E-08	0.32267E-07	0.0000	109.13	0.100E-07
0.41223E-10	0.4103E-12	0.82334E-08	0.52202E-07	0.0000	184.29	0.100E-07
0.69533-309	0.000	0.45430E-02	0.61571E-01	0.0000	-208.42	0.100E-17
0.69533-309	0.000	0.10218E-02	0.24550E-03	0.0000	979.38	0.100E-17
0.79077E-08	0.1157E-13	0.11252E-08	0.10000E-10	483.43	1039.4	0.100E-10
0.83131E-09	0.1431E-11	0.36212E-08	0.28284E-10	6776.6	3368.1	0.100E-10

Table A.4 : Results for Exact Bilinear Reformulation

problem	n	m	nc	ifail	it	fe	ce	ge	he
bard1	8	7	3	0	4	5	6	6	5
bard2	16	13	4	0	2	3	4	4	3
bard3	8	7	2	0	4	5	6	6	5
bar-truss	35	35	6	0	9	4	10	10	10
bilevel1	16	15	6	0	4	5	6	6	5
bilevel2	32	29	12	0	6	7	8	8	7
bilevel3	12	11	4	0	6	7	8	8	7
bilin	14	13	6	0	2	3	4	4	3
dempe	4	3	1	0	2	4	4	3	3
design-cent-1	15	15	3	0	5	6	7	7	6
design-cent-4	46	57	12	0	37	46	47	8	7
desilva	7	5	1	0	5	6	6	6	6
df1	3	4	1	0	2	3	3	3	3
ex9.1.1	13	12	5	0	3	4	5	5	4
ex9.1.2	10	9	4	0	2	3	4	4	3
ex9.1.3	23	21	6	0	2	3	4	4	3
ex9.1.4	10	9	4	0	5	4	6	6	6
ex9.1.5	13	12	5	0	2	3	4	4	3
ex9.1.6	14	13	6	0	5	5	6	6	6
ex9.1.7	17	15	6	0	2	3	4	4	3
ex9.1.8	12	11	4	0	1	2	3	3	2
ex9.1.9	12	11	5	0	2	3	4	4	3
ex9.1.10	12	11	4	0	1	2	3	3	2
ex9.2.1	10	9	4	0	4	5	6	6	5
ex9.2.2	10	19	4	0	29	31	32	30	29
ex9.2.3	16	15	6	0	4	5	6	6	5
ex9.2.4	8	7	2	0	3	4	5	5	4
ex9.2.5	8	7	3	0	6	7	8	8	7
ex9.2.6	12	10	4	0	2	3	4	4	3
ex9.2.7	10	9	4	0	4	5	6	6	5
ex9.2.8	6	5	2	0	3	4	5	5	4
ex9.2.9	9	8	3	0	3	4	5	5	4
flp2	6	4	2	0	4	5	6	6	5
flp4-1	110	90	30	0	3	4	5	5	4
flp4-2	170	170	60	0	3	4	5	5	4
flp4-3	210	240	70	0	3	4	5	5	4
flp4-4	300	350	100	0	3	4	5	5	4
gauvin	5	4	2	0	3	4	5	5	4
gnash10	17	16	8	0	6	5	7	7	7
gnash11	17	16	8	0	7	6	8	8	8
gnash12	17	16	8	0	7	6	8	8	8
gnash13	17	16	8	0	8	7	9	9	9
gnash14	17	16	8	0	8	6	9	9	9
gnash15	17	16	8	0	12	10	13	13	13
gnash16	17	16	8	0	8	6	8	7	8
gnash17	17	16	8	0	10	9	11	11	11
gnash18	17	16	8	0	10	9	11	11	11
gnash19	17	16	8	0	6	6	7	7	7
hs044-i	46	40	10	0	4	5	6	6	5

KKT	feas.	$\text{compl}_1(x)$	$\text{compl}_2(x)$	ξ^*_{max}	$f(x^*)$
0.0000	0.1110E-14	0.0000	0.0000	0.76190	17.000
0.14211E-13	0.2220E-14	0.0000	0.0000	0.0000	-6598.0
0.0000	0.000	0.0000	0.0000	0.36250	-12.679
0.19934E-12	0.4263E-12	0.0000	0.0000	1.4516	10167.
0.0000	0.8882E-14	0.0000	0.0000	0.15000	-10.0000
0.0000	0.3027E-08	0.60543E-08	0.30272E-08	0.0000	-6600.0
0.44409E-15	0.1776E-14	0.0000	0.0000	1.0927	-12.679
0.88818E-15	0.5551E-15	0.0000	0.0000	22.000	-5.6000
0.0000	0.000	0.0000	0.0000	15.273	49.000
0.95416E-14	0.8604E-15	0.0000	0.0000	0.0000	-1.8606
4.2426	0.1422E-15	0.55879E-08	0.31225E-16	6.0000	-0.33528E-07
0.44787E-13	0.4441E-15	0.0000	0.0000	0.0000	-1.0000
0.0000	0.000	0.0000	0.0000	0.0000	0.0000
0.50877E-15	0.1110E-15	0.0000	0.0000	0.0000	-13.000
0.0000	0.000	0.0000	0.0000	6.5000	-3.0000
0.0000	0.1110E-15	0.0000	0.0000	34.000	-6.0000
0.0000	0.1799E-13	0.0000	0.0000	2.4375	-37.000
0.0000	0.000	0.0000	0.0000	15.000	4.0000
0.0000	0.1776E-14	0.0000	0.0000	1.5556	-21.000
0.0000	0.1110E-15	0.0000	0.0000	34.000	-6.0000
0.0000	0.000	0.0000	0.0000	0.12500	-3.2500
0.55511E-16	0.2220E-14	0.0000	0.0000	2.7778	9.7778
0.0000	0.000	0.0000	0.0000	0.12500	-3.2500
0.44409E-15	0.4441E-14	0.0000	0.0000	0.76190	17.000
18.511	0.3843E-16	0.24130E-08	0.38428E-16	0.17717E+10	100.00
0.0000	0.7105E-14	0.0000	0.0000	0.83333E-01	5.0000
0.0000	0.000	0.0000	0.0000	1.0000	0.50000
0.0000	0.8882E-15	0.0000	0.0000	6.0000	9.0000
0.15701E-15	0.4441E-15	0.0000	0.0000	0.50000	-1.0000
0.44409E-15	0.4441E-14	0.0000	0.0000	0.76190	17.000
0.0000	0.000	0.0000	0.0000	0.50000	1.5000
0.0000	0.000	0.0000	0.0000	0.16667	2.0000
0.0000	0.000	0.0000	0.0000	0.0000	0.0000
0.0000	0.1092E-11	0.0000	0.0000	0.0000	0.0000
0.0000	0.2727E-11	0.0000	0.0000	0.0000	0.0000
0.0000	0.5243E-11	0.0000	0.0000	0.0000	0.0000
0.0000	0.1329E-10	0.0000	0.0000	0.0000	0.0000
0.0000	0.000	0.0000	0.0000	0.25000	20.000
0.39504E-10	0.7816E-12	0.0000	0.0000	0.14211	-230.82
0.16251E-14	0.3553E-14	0.0000	0.0000	0.91802E-01	-129.91
0.26278E-08	0.5529E-09	0.0000	0.0000	0.39663E-01	-36.933
0.14873E-14	0.6217E-14	0.0000	0.0000	0.14908E-01	-7.0618
0.34057E-13	0.1243E-13	0.0000	0.0000	0.19886E-02	-0.17905
0.11809E-11	0.1688E-13	0.0000	0.0000	7.6472	-354.70
0.97925E-10	0.1776E-13	0.0000	0.0000	1.9472	-241.44
0.30670E-09	0.2798E-12	0.0000	0.0000	1.6742	-90.749
0.12888E-14	0.6661E-14	0.0000	0.0000	12.712	-25.698
0.84045E-08	0.6217E-14	0.0000	0.0000	2.8027	-6.1167
0.25121E-14	0.1043E-13	0.0000	0.0000	7.8190	17.090

problem	n	m	nc	ifail	it	fe	ce	ge	he
incid-set1-8	247	249	49	0	38	49	50	23	24
incid-set1-16	999	1005	225	0	56	62	63	29	31
incid-set1-32	4039	4053	961	6	1000	1582	1583	611	616
incid-set1c-8	247	256	49	0	39	44	45	22	24
incid-set1c-16	999	1020	225	0	62	73	74	38	42
incid-set1c-32	4039	4084	961	0	24	26	27	20	20
incid-set2-8	247	249	49	0	29	36	36	21	21
incid-set2-16	999	1005	225	0	25	28	28	17	17
incid-set2-32	4039	4053	961	0	66	68	68	44	46
incid-set2c-8	247	256	49	0	22	23	23	16	18
incid-set2c-16	999	1020	225	0	28	33	33	20	21
incid-set2c-32	4039	4084	961	0	50	63	63	33	34
jr1	3	2	1	0	1	2	2	2	2
jr2	3	2	1	0	6	7	7	7	7
kth1	2	2	1	0	1	2	2	2	2
kth2	2	2	1	0	3	4	4	4	4
kth3	2	2	1	0	4	5	5	5	5
liswet1-050	202	153	50	0	1	2	3	3	2
liswet1-100	402	303	100	0	1	2	3	3	2
liswet1-200	802	603	200	0	1	2	3	3	2
nash1	7	5	1	0	7	8	9	9	8
outrata31	17	16	4	0	8	9	10	10	9
outrata32	17	16	4	0	9	10	11	11	10
outrata33	17	16	4	0	7	8	9	9	8
outrata34	17	16	4	0	7	8	9	9	8
pack-comp1-8	237	251	49	0	36	55	55	25	26
pack-comp1-16	981	1025	225	0	44	35	51	32	41
pack-comp1-32	4005	4157	961	5	256	2	414	131	214
pack-comp1c-8	237	258	49	0	17	15	16	14	17
pack-comp1c-16	981	1040	225	0	25	19	25	22	25
pack-comp1c-32	4005	4188	961	0	217	19	361	136	199
pack-comp1p-8	237	236	49	0	54	61	61	32	34
pack-comp1p-16	981	980	225	0	69	86	88	47	51
pack-comp1p-32	4005	4004	961	0	146	242	244	94	98
pack-comp2-8	237	251	49	0	9	11	11	10	10
pack-comp2-16	981	1025	225	0	18	19	19	15	16
pack-comp2-32	4005	4157	961	5	318	6	527	164	269
pack-comp2c-8	237	258	49	0	6	7	7	7	7
pack-comp2c-16	981	1040	225	0	7	8	8	8	8
pack-comp2c-32	4005	4188	961	5	197	2	286	103	162
pack-comp2p-8	237	236	49	0	39	42	42	26	26
pack-comp2p-16	981	980	225	0	56	69	70	36	38
pack-comp2p-32	4005	4004	961	0	132	187	189	87	91
pack-rig1-4	57	62	9	0	8	9	10	10	9
pack-rig1-8	237	251	49	0	28	26	28	23	25
pack-rig1-16	981	1025	225	0	131	32	245	88	126
pack-rig1-32	4005	4157	961	0	80	69	80	50	55
pack-rig1c-4	57	65	9	0	6	7	8	8	7
pack-rig1c-8	237	258	49	0	13	12	13	13	14
pack-rig1c-16	981	1040	225	0	104	17	189	76	105
pack-rig1c-32	4005	4188	961	0	57	20	55	49	52
pack-rig1p-4	57	56	9	0	4	4	5	5	5

KKT	feas.	$\text{compl}_1(x)$	$\text{compl}_2(x)$	$\xi^*_{\max}$	$f(x^*)$
0.0000	0.9182E-11	0.0000	0.0000	0.0000	0.38164E-16
0.55511E-16	0.3448E-08	0.0000	0.0000	0.0000	0.12186E-15
0.30254E-06	0.1035E-07	0.0000	0.0000	0.0000	0.27683E-06
0.0000	0.5086E-13	0.0000	0.0000	0.0000	0.41633E-16
0.0000	0.5973E-14	0.0000	0.0000	0.0000	0.12143E-15
0.58022E-09	0.3797E-08	0.0000	0.0000	0.0000	0.11712E-04
0.17429E-08	0.5961E-09	0.0000	0.0000	0.0000	0.45179E-02
0.51579E-08	0.2890E-13	0.0000	0.0000	0.0000	0.32168E-02
0.83992E-08	0.2982E-10	0.0000	0.0000	0.0000	0.17692E-02
0.75621E-10	0.5349E-12	0.0000	0.0000	0.0000	0.56301E-02
0.40468E-09	0.1629E-13	0.0000	0.0000	0.0000	0.36158E-02
0.26102E-08	0.2051E-12	0.0000	0.0000	0.0000	0.24404E-02
0.0000	0.000	0.0000	0.0000	0.0000	0.50000
0.0000	0.000	0.0000	0.0000	2.0000	0.50000
0.0000	0.000	0.0000	0.0000	0.0000	0.0000
0.0000	0.000	0.0000	0.0000	0.0000	0.0000
0.0000	0.000	0.0000	0.0000	1.0000	0.50000
0.72768E-14	0.5537E-13	0.0000	0.0000	0.0000	0.13994E-01
0.10898E-13	0.2688E-12	0.0000	0.0000	0.0000	0.13734E-01
0.24796E-12	0.1827E-11	0.0000	0.0000	0.0000	0.17009E-01
0.0000	0.5763E-11	0.11515E-10	0.57573E-11	0.0000	0.0000
0.11057E-13	0.6217E-14	0.0000	0.0000	0.16393	3.2077
0.32368E-15	0.5107E-14	0.0000	0.0000	0.16778	3.4494
0.19766E-09	0.6163E-10	0.0000	0.0000	0.71407	4.6043
0.30593E-14	0.2331E-14	0.0000	0.0000	2.0723	6.5927
0.0000	0.6072E-15	0.0000	0.0000	0.0000	0.60000
0.12233E-09	0.5327E-14	0.0000	0.0000	0.0000	0.61695
0.18169E-01	0.9242E-04	0.45461E-02	0.13184E-04	0.81081E+07	0.62379
0.0000	0.3492E-08	0.42759E-04	0.25534E-08	0.0000	0.60000
0.69618E-13	0.5694E-14	0.0000	0.0000	0.0000	0.62304
0.51402E-12	0.2214E-13	0.0000	0.0000	0.0000	0.66144
0.15245E-08	0.5331E-15	0.0000	0.0000	0.0000	0.60000
0.25577E-06	0.1895E-14	0.0000	0.0000	0.0000	0.61695
0.41047E-06	0.9625E-14	0.0000	0.0000	0.0000	0.65298
0.30674E-09	0.1324E-14	0.0000	0.0000	0.0000	0.67312
0.68060E-08	0.3119E-14	0.0000	0.0000	0.0000	0.72714
88.942	0.8098E-03	0.29820E-03	0.19994E-06	665.53	0.88173
0.12719E-09	0.1395E-10	0.0000	0.0000	0.0000	0.67346
0.14338E-08	0.3819E-14	0.0000	0.0000	0.0000	0.72747
0.84989E-01	0.4003E-03	0.13687E-01	0.22093E-03	0.59486E+06	0.78991
0.43108E-06	0.8728E-15	0.0000	0.0000	0.0000	0.67312
0.19803E-06	0.2561E-14	0.0000	0.0000	0.0000	0.72714
0.31988E-06	0.1726E-13	0.0000	0.0000	0.0000	0.78260
0.58477E-10	0.8676E-12	0.0000	0.0000	0.0000	0.71889
0.58445E-08	0.4655E-13	0.0000	0.0000	0.0000	0.78793
0.14043E-06	0.1112E-10	0.0000	0.0000	0.0000	0.82601
0.29048E-07	0.7896E-14	0.0000	0.0000	12.561	0.85089
0.12857E-12	0.2776E-15	0.0000	0.0000	0.0000	0.72101
0.76018E-09	0.3313E-13	0.0000	0.0000	0.0000	0.78830
0.50253E-08	0.2052E-12	0.0000	0.0000	0.0000	0.82650
0.13264E-07	0.1167E-11	0.20843E-11	0.20354E-14	0.0000	0.85164
0.0000	0.5378E-16	0.0000	0.0000	0.0000	0.60000

problem	n	m	nc	ifail	it	fe	ce	ge	he
pack-rig1p-8	237	236	49	0	43	31	43	34	42
pack-rig1p-16	981	980	225	0	374	115	619	230	342
pack-rig1p-32	4005	4004	961	0	146	141	186	89	103
pack-rig2-4	57	62	9	0	7	8	8	8	8
pack-rig2-8	237	251	49	0	11	12	12	12	12
pack-rig2-16	981	1025	225	3	10	1	11	9	10
pack-rig2-32	4005	4157	961	3	84	1	112	54	70
pack-rig2c-4	57	65	9	0	4	5	5	5	5
pack-rig2c-8	237	258	49	0	7	8	8	8	8
pack-rig2c-16	981	1040	225	3	10	1	11	9	10
pack-rig2c-32	4005	4188	961	3	105	1	156	67	89
pack-rig2p-4	57	56	9	0	12	11	12	12	13
pack-rig2p-8	237	236	49	0	15	16	16	16	16
pack-rig2p-16	981	980	225	0	54	44	72	36	44
pack-rig2p-32	4005	4004	961	5	38	2	2	2	3
pack-rig3-4	57	62	9	0	6	7	7	7	7
pack-rig3-8	237	251	49	0	14	15	15	11	12
pack-rig3-16	981	1025	225	0	30	29	30	24	27
pack-rig3-32	4005	4157	961	5	189	2	252	90	144
pack-rig3c-4	57	65	9	0	4	5	5	5	5
pack-rig3c-8	237	258	49	0	7	8	8	8	8
pack-rig3c-16	981	1040	225	0	18	19	21	17	19
pack-rig3c-32	4005	4188	961	5	131	16	162	84	105
portfl1	111	49	12	0	10	11	12	12	11
portfl2	111	49	12	0	14	15	16	16	15
portfl3	111	49	12	0	15	16	17	17	16
portfl4	111	49	12	0	4	5	6	6	5
portfl6	111	49	12	0	12	13	14	14	13
qpec-100-1	205	202	100	0	14	15	16	16	15
qpec-100-2	210	202	100	0	12	13	14	14	13
qpec-100-3	210	204	100	0	13	14	15	15	14
qpec-100-4	220	204	100	0	7	8	9	9	8
qpec-200-1	410	404	200	0	13	12	14	14	14
qpec-200-2	420	404	200	0	17	18	19	19	18
qpec-200-3	420	408	200	0	12	13	14	14	13
qpec-200-4	440	408	200	0	7	8	9	9	8
ralph1	3	5	1	0	33	35	35	31	32
ralph2	2	2	1	0	15	16	16	16	16
ralphmod	204	200	100	0	31	32	33	32	32
scholtes1	4	3	1	0	4	5	5	5	5
scholtes2	4	3	1	0	2	3	3	3	3
scholtes3	2	2	1	0	4	5	5	5	5
scholtes4	3	5	1	0	32	34	34	31	32
scholtes5	7	6	2	0	3	4	5	5	4
sl1	11	8	3	0	1	2	3	3	2
stackelberg1	3	2	1	0	4	5	6	6	5
tap-09	162	142	36	0	20	7	21	21	21
tap-15	1560	2100	675	0	23	8	24	24	24
tollmpec	6051	6100	1824	0	169	640	656	95	95
tollmpec1	6051	6100	1824	0	32	15	32	32	33
water-net	180	214	64	0	118	94	115	59	80
water-FL	1737	2424	784	0	278	249	325	143	180

KKT	feas.	$\text{compl}_1(x)$	$\text{compl}_2(x)$	$\xi^*_{\max}$	$f(x^*)$
0.18321E-07	0.6167E-15	0.0000	0.0000	0.0000	0.78793
0.64317E-07	0.3533E-12	0.0000	0.0000	0.0000	0.82601
0.41954E-06	0.5746E-10	0.0000	0.0000	12.561	0.85089
0.45945E-13	0.4224E-15	0.0000	0.0000	0.0000	0.69453
0.60569E-08	0.1435E-14	0.0000	0.0000	0.0000	0.78040
0.36220E-12	0.1408E-03	0.31301E-06	0.14676E-08	0.0000	1.0000
0.74789E-11	0.4866E-03	0.19841E-05	0.15104E-08	0.0000	1.0000
0.32004E-09	0.1429E-11	0.0000	0.0000	0.0000	0.71153
0.41307E-13	0.1828E-12	0.11676E-10	0.18243E-12	0.0000	0.79931
0.36220E-12	0.1408E-03	0.31301E-06	0.14676E-08	0.0000	1.0000
0.90942E-11	0.4866E-03	0.0000	0.0000	0.0000	1.0000
0.77036E-13	0.2349E-10	0.13729E-07	0.23490E-10	0.0000	0.60000
0.17555E-07	0.2515E-15	0.0000	0.0000	0.0000	0.78040
0.16779E-08	0.9257E-15	0.0000	0.0000	0.0000	1.0851
0.13834-321	0.000	0.32608E-01	0.81603E-03	5092.9	1.2492
0.87949E-11	0.1510E-12	0.0000	0.0000	0.0000	0.66619
0.82557E-08	0.1321E-13	0.0000	0.0000	0.0000	0.73520
0.14087E-07	0.3692E-12	0.17864E-10	0.90714E-13	0.0000	0.80043
0.50116E-02	0.2717E-04	0.11462E-02	0.31410E-05	4.6454	0.77463
0.29579E-09	0.1402E-11	0.0000	0.0000	0.0000	0.68277
0.14093E-09	0.3732E-11	0.50551E-10	0.86885E-12	0.0000	0.75347
0.18292E-10	0.3613E-12	0.17514E-10	0.88938E-13	0.0000	0.81860
1.2068	0.1149E-03	0.0000	0.0000	0.33264E+06	0.91028
0.34544E-07	0.3990E-15	0.0000	0.0000	0.89651	0.15024E-04
0.31968E-07	0.2186E-15	0.0000	0.0000	0.68202	0.14573E-04
0.42816E-07	0.3331E-15	0.0000	0.0000	0.21907	0.62650E-05
0.66253E-16	0.2194E-15	0.0000	0.0000	0.35009	0.21773E-05
0.40796E-07	0.2619E-15	0.0000	0.0000	0.14880	0.23613E-05
0.95194E-15	0.1489E-13	0.0000	0.0000	139.66	0.33588
0.40120E-14	0.2034E-13	0.0000	0.0000	45.983	-2.9409
0.49563E-14	0.2315E-13	0.0000	0.0000	4.0685	-5.4450
0.27973E-14	0.2347E-13	0.0000	0.0000	7.3103	-1.4361
0.27560E-08	0.8426E-13	0.0000	0.0000	158.43	-1.9348
0.86144E-14	0.1078E-12	0.0000	0.0000	4.8761	-22.262
0.55042E-14	0.6466E-13	0.0000	0.0000	35.452	-1.9534
0.81391E-14	0.9048E-13	0.0000	0.0000	10.169	-6.0254
0.23570	0.9224E-16	0.96043E-08	0.92242E-16	0.34707E+08	-0.96043E-08
0.26342E-08	0.9313E-09	0.30518E-04	0.93132E-09	1.9999	-0.18626E-08
0.72621E-07	0.3985E-10	0.0000	0.0000	0.11953E-03	-683.03
0.0000	0.000	0.0000	0.0000	0.0000	2.0000
0.0000	0.000	0.0000	0.0000	0.0000	15.000
0.0000	0.000	0.0000	0.0000	1.0000	0.50000
0.47140	0.3690E-17	0.19209E-08	0.36897E-17	0.34707E+09	-0.38417E-08
0.0000	0.000	0.0000	0.0000	0.0000	1.0000
0.43368E-18	0.2006E-14	0.0000	0.0000	0.0000	0.10000E-03
0.17764E-14	0.000	0.0000	0.0000	0.0000	-3266.7
0.27920E-11	0.4798E-10	0.0000	0.0000	0.16349E-02	109.14
0.13500E-09	0.6467E-08	0.0000	0.0000	0.25730E-02	184.50
0.26411E-07	0.6493E-11	0.0000	0.0000	5.0120	-208.26
0.48721E-08	0.9712E-11	0.0000	0.0000	164.51	979.39
0.15592E-07	0.1243E-12	0.0000	0.0000	3030.6	997.58
0.75096E-08	0.1013E-12	0.0000	0.0000	16830.	3324.7

A.3 Results for M-Stationary Examples

A.3.1 Results for the New Relaxation Method

Table A.5 : Results for Example 4.3

k	t_1^k	(x_{11}^k, x_{21}^k)	$\|(x_{11}^k, x_{21}^k)\|_2$	$1/2\sqrt{2}\,\theta(0)\,t_1^k$	$x_{11}^k x_{21}^k$	$\frac{1}{4}(t_1^k)^2$	ξ_1^k
1	10	(0.9839,2.952)	3.1112	2.5695	2.9040	25.0	2.658
2	1.0	(0.9839,2.952)E-01	3.1112E-01	2.5695E-01	2.9040E-02	2.5E-01	2.919
3	1.0E-01	(0.9839,2.952)E-02	3.1112E-02	2.5695E-02	2.9040E-04	2.5E-03	2.946
4	1.0E-02	(0.9839,2.952)E-03	3.1112E-03	2.5695E-03	2.9040E-06	2.5E-05	2.948
5	1.0E-03	(0.9839,2.952)E-04	3.1112E-04	2.5695E-04	2.9040E-08	2.5E-07	2.948
6	1.0E-04	(0.9839,2.952)E-05	3.1112E-05	2.5695E-05	2.9040E-10	2.5E-09	2.948
7	1.0E-05	(0.9839,2.952)E-06	3.1112E-06	2.5695E-06	2.9040E-12	2.5E-11	2.948
8	1.0E-06	(0.9839,2.952)E-07	3.1112E-07	2.5695E-07	2.9040E-14	2.5E-13	2.948
9	1.0E-07	(0.9839,2.952)E-08	3.1112E-08	2.5695E-08	2.9040E-16	2.5E-15	2.952
10	1.0E-08	(0.9839,2.952)E-09	3.1112E-09	2.5695E-09	2.9040E-18	2.5E-17	3.214

Table A.6 : Results for Example 4.4

k	t_k	(x_1^k, x_2^k)	$\|(x_1^k, x_2^k)\|_2$	$1/2\sqrt{2}\,\theta(0)\,t_k$	$x_1^k x_2^k$	$\frac{1}{4}t_k^2$	ξ^k
1	10	(1.817,1.817)	2.5695	2.5695	3.3011	25.0	0.5
2	1.0	(1.817,1.817)E-01	2.5695E-01	2.5695E-01	3.3011E-02	2.5E-01	0.5
3	1.0E-01	(1.817,1.817)E-02	2.5695E-02	2.5695E-02	3.3011E-04	2.5E-03	0.5
4	1.0E-02	(1.817,1.817)E-03	2.5695E-03	2.5695E-03	3.3011E-06	2.5E-05	0.5
5	1.0E-03	(1.817,1.817)E-04	2.5695E-04	2.5695E-04	3.3011E-08	2.5E-07	0.5
6	1.0E-04	(1.817,1.817)E-05	2.5695E-05	2.5695E-05	3.3011E-10	2.5E-09	0.5
7	1.0E-05	(1.817,1.817)E-06	2.5695E-06	2.5695E-06	3.3011E-12	2.5E-11	0.5
8	1.0E-06	(1.817,1.817)E-07	2.5695E-07	2.5695E-07	3.3011E-14	2.5E-13	0.5
9	1.0E-07	(1.817,1.817)E-08	2.5695E-08	2.5695E-08	3.3011E-16	2.5E-15	0.5
10	1.0E-08	(1.817,1.817)E-09	2.5695E-09	2.5695E-09	3.3011E-18	2.5E-17	0.5

Table A.7 : Results for Example 2.4

k	t_k	(x_1^k, x_2^k)	$\|(x_1^k, x_2^k)\|_2$	$1/2\sqrt{2}\,\theta(0)\,t_k$	$x_1^k x_2^k$	$\frac{1}{4}t_k^2$	ξ^k
1	10	(1.817,1.817)	2.5695	2.5695	3.3011	25.0	1.0
2	1.0	(1.817,1.817)E-01	2.5695E-01	2.5695E-01	3.3011E-02	2.5E-01	1.0
3	1.0E-01	(1.817,1.817)E-02	2.5695E-02	2.5695E-02	3.3011E-04	2.5E-03	1.0
4	1.0E-02	(1.817,1.817)E-03	2.5695E-03	2.5695E-03	3.3011E-06	2.5E-05	1.0
5	1.0E-03	(1.817,1.817)E-04	2.5695E-04	2.5695E-04	3.3011E-08	2.5E-07	1.0
6	1.0E-04	(1.817,1.817)E-05	2.5695E-05	2.5695E-05	3.3011E-10	2.5E-09	1.0
7	1.0E-05	(1.817,1.817)E-06	2.5695E-06	2.5695E-06	3.3011E-12	2.5E-11	1.0
8	1.0E-06	(1.817,1.817)E-07	2.5695E-07	2.5695E-07	3.3011E-14	2.5E-13	1.0
9	1.0E-07	(1.817,1.817)E-08	2.5695E-08	2.5695E-08	3.3011E-16	2.5E-15	1.0
10	1.0E-08	(1.817,1.817)E-09	2.5695E-09	2.5695E-09	3.3011E-18	2.5E-17	1.0

A.3.2 Results for the Relaxed Bilinear Approach

Table A.8 : Results for Example 4.3

k	t^k	(x_1^k, x_2^k)	$\|(x_1^k, x_2^k)\|_2$	$\sqrt{2t^k}$	$x_1^k x_2^k$	ξ^k
1	10	(1.984,5.041)	5.4174	4.4721	1.0E+01	7.239E-01
2	1.0	(5.938E-01,1.684)	1.7858	1.4142	1.0	8.959
3	1.0E-01	(1.842E-01,5.428E-01) E-01	5.7320E-01	4.4721E-01	1.0E-01	2.870E+01
4	1.0E-02	(5.790E-02,1.727E-01) E-01	1.8215E-01	1.4142E-01	1.0E-02	9.112E+01
5	1.0E-03	(1.827E-02,5.472E-02)E-02	5.7693E-02	4.4721E-02	1.0E-03	2.885E+02
6	1.0E-04	(5.775E-03,1.732E-02)E-02	1.8253E-02	1.4142E-02	1.0E-04	9.127E+02
7	1.0E-05	(1.826E-03,5.477E-03)E-03	5.7731E-03	4.4721E-03	1.0E-05	2.887E+03
8	1.0E-06	(5.774E-04,1.732E-03)E-03	1.8257E-03	1.4142E-03	1.0E-06	9.128E+03
9	1.0E-07	(1.826E-04,5.477E-04)E-04	5.7734E-04	4.4721E-04	1.0E-07	2.886E+04
10	1.0E-08	(5.77E-05,1.732E-04)E-04	1.8256E-04	1.4142E-04	1.0E-08	9.129E+04

Table A.9 : Results for Example 4.4

k	t^k	(x_1^k, x_2^k)	$\|(x_1^k, x_2^k)\|_2$	$\sqrt{2t^k}$	$x_1^k x_2^k$	ξ^k
1	10	(3.162,3.162)	4.4721	4.4721	1.0E+01	1.581E-01
2	1.0	(1.000,1.000)	1.4142	1.4142	1.0	5.000E-01
3	1.0E-01	(3.162,3.162) E-01	4.4721E-01	4.4721E-01	1.0E-01	1.581
4	1.0E-02	(1.000,1.000) E-01	1.4142E-01	1.4142E-01	1.0E-02	5.000
5	1.0E-03	(3.162,3.162)E-02	4.4721E-02	4.4721E-02	1.0E-03	1.581E+01
6	1.0E-04	(1.000,1.000)E-02	1.4142E-02	1.4142E-02	1.0E-04	5.000E+01
7	1.0E-05	(3.162,3.162)E-03	4.4721E-03	4.4721E-03	1.0E-05	1.581E+02
8	1.0E-06	(1.000,1.000)E-03	1.4142E-03	1.4142E-03	1.0E-06	5.000E+02
9	1.0E-07	(3.162,3.162)E-04	4.4721E-04	4.4721E-04	1.0E-07	1.581E+03
10	1.0E-08	(1.000,1.000)E-04	1.4142E-04	1.4142E-04	1.0E-08	5.000E+03

Table A.10 : Results for Example 2.4

k	t^k	(x_1^k, x_2^k)	$\|(x_1^k, x_2^k)\|_2$	$\sqrt{2t^k}$	$x_1^k x_2^k$	ξ^k
1	10	(3.162,3.162)	4.4721	4.4721	1.0E+01	3.162E-01
2	1.0	(1.000,1.000)	1.4142	1.4142	1.0	1.000
3	1.0E-01	(3.162,3.162) E-01	4.4721E-01	4.4721E-01	1.0E-01	3.162
4	1.0E-02	(1.000,1.000) E-01	1.4142E-01	1.4142E-01	1.0E-02	1.000E+01
5	1.0E-03	(3.162,3.162)E-02	4.4721E-02	4.4721E-02	1.0E-03	3.162E+01
6	1.0E-04	(1.000,1.000)E-02	1.4142E-02	1.4142E-02	1.0E-04	1.000E+02
7	1.0E-05	(3.162,3.162)E-03	4.4721E-03	4.4721E-03	1.0E-05	3.162E+02
8	1.0E-06	(1.000,1.000)E-03	1.4142E-03	1.4142E-03	1.0E-06	1.000E+03
9	1.0E-07	(3.162,3.162)E-04	4.4721E-04	4.4721E-04	1.0E-07	3.162E+03
10	1.0E-08	(1.000,1.000)E-04	1.4142E-04	1.4142E-04	1.0E-08	1.000E+04