



universität
wien

DISSERTATION

Titel der Dissertation

A feasible second order bundle algorithm for nonsmooth, nonconvex optimization problems with inequality constraints and its application to certificates of infeasibility

Verfasser

Hannes Fendl

angestrebter akademischer Grad

Doktor der Naturwissenschaften (Dr. rer. nat.)

Wien, im Mai 2011

Studienkennzahl lt. Studienblatt: A 091 405

Dissertationsgebiet lt. Studienblatt: Mathematik

Betreuer: Ao. Prof. Dipl.-Ing. Dr. Hermann Schichl

„Mutig in die neuen Zeiten.“

Paula von Preradović

Abstract

This thesis extends the SQP-approach of the well-known bundle-Newton method for nonsmooth unconstrained minimization to a feasible second order bundle algorithm for nonsmooth, nonconvex optimization problems with inequality constraints: Instead of using a penalty function or a filter or an improvement function to deal with the presence of constraints, the search direction is determined by solving a convex quadratically constrained quadratic program to obtain good iteration points. Moreover, we investigate certain versions of the search direction problem, we justify the applicability of this approach numerically by using different solvers for the computation of the search direction and we show global convergence of the method under certain assumptions.

Furthermore, we present an important application of nonsmooth optimization to constraint satisfaction problems: We introduce a certificate of infeasibility for finding exclusion boxes by solving a linearly constrained nonsmooth optimization problem. Additionally, the constructed certificate can be used to enlarge an exclusion box by solving a nonlinearly constrained nonsmooth optimization problem.

Finally, the good performance of the second order bundle algorithm is demonstrated by comparison with test results of other solvers on examples of the Hock-Schittkowski collection, on custom examples that arise in the context of finding exclusion boxes for constraint satisfaction problems, and on higher dimensional piecewise quadratic examples.

Abstract (German)

Diese Arbeit erweitert den SQP-Zugang des Bundle-Newton-Verfahrens für nichtglatte, unrestringierte Optimierungsprobleme zu einem zulässigen Bundle-Algorithmus zweiter Ordnung für nichtglatte, nichtkonvexe Optimierungsprobleme mit Ungleichungsnebenbedingungen. An Stelle der Verwendung einer Straffunktion oder eines Filters oder einer Improvement-Funktion zur Behandlung der Nebenbedingungen, wird die Suchrichtung durch Lösen eines konvexen quadratischen Optimierungsproblems mit quadratischen Nebenbedingungen bestimmt, um gute Iterationspunkte zu erhalten. Außerdem untersuchen wir einige Varianten des Suchrichtungsproblems, wir geben eine numerische Rechtfertigung für die Anwendbarkeit des vorgestellten Zugangs, indem wir die Effektivität von verschiedener Lösungssoftware für die Berechnung der Suchrichtung vergleichen, und wir weisen die globale Konvergenz der Methode unter bestimmten Voraussetzungen nach.

Weiters stellen wir eine wichtige Anwendung der nichtglatte Optimierung für Zulässigkeitsprobleme vor: Dazu führen wir ein Unzulässigkeitszertifikat ein, welches das Auffinden von Ausschlussboxen durch Lösen eines nichtglatte Optimierungsproblems mit linearen Nebenbedingungen ermöglicht. Zusätzlich kann dieses Zertifikat verwendet werden, um eine Ausschlussbox durch Lösen eines nichtglatte Optimierungsproblems mit nichtlinearen Nebenbedingungen zu vergrößern.

Schließlich besprechen wir noch die im Vergleich zu anderer Lösungssoftware guten Testergebnisse von unserem Bundle-Algorithmus zweiter Ordnung für einige Hock-Schittkowski-Beispiele, für Beispiele die im Zusammenhang mit der Auffindung von Ausschlussboxen in Zulässigkeitsproblemen auftreten und für höher dimensionale stückweise quadratische Beispiele.

Acknowledgements

Ich bedanke mich bei meinen Eltern für ihre Unterstützung. Ich danke Prof. Hermann Schichl und Prof. Arnold Neumaier für die Stellung des Themas und die Hilfestellungen zur Anfertigung dieser Arbeit. Ich danke meinen beiden ehemaligen Lehrern im Gymnasium Mag. Fritz Fassler und Mag. Andreas Sulzgruber für ihren hervorragenden Unterricht.

Contents

Abstract	iii
Abstract (German)	iv
Acknowledgements	v
Contents	vii
1 Introduction	1
1.1 Outline	1
1.2 Organization	4
1.3 Notation	4
2 Basics of optimization theory	7
2.1 Introduction	7
2.2 Smooth optimization	9
2.2.1 Smooth Optimality conditions	9
2.2.2 SQP	11
2.3 Nonsmooth optimization	15
2.3.1 Nonsmooth Optimality conditions	15
2.3.2 Problems of smooth solvers on nonsmooth optimization problems	18
2.3.3 Bundle algorithms	18
2.3.4 Other algorithms for nonsmooth optimization	21
3 A feasible second order bundle algorithm for nonsmooth, nonconvex optimization problems with inequality constraints	23
3.1 Derivation of the method	23
3.2 The dual search direction problem	31
3.3 Presentation of the algorithm	35
3.4 Presentation of the line search	42
3.5 Convergence of the line search	51
3.6 Global convergence	52
4 Implementation issues	109
4.1 Introduction	109
4.2 Second-order cone programming	110
4.3 The reduced problem	112
4.3.1 Reduction of problem size	112
4.3.2 SOCP-reformulation	116
4.3.3 Strict feasibility of the SOCP	118
4.3.4 Connection between the Lagrange multipliers	123

4.4	Third party software	126
4.4.1	Comparison of the QCQP-solvers	126
4.4.2	socp	129
4.4.3	IPOPT	130
4.4.4	MOSEK	132
5	Certificates of infeasibility via nonsmooth optimization	135
5.1	Interval arithmetic	135
5.2	Certificate of infeasibility	136
5.3	Properties of the certificate for quadratic F	144
5.4	Exclusion boxes for constraint satisfaction problems	149
5.5	Implementation issues for finding exclusion boxes	153
5.5.1	Basics	153
5.5.2	Linearly constrained case	154
5.5.3	Nonlinearly constrained case	158
5.5.4	Starting point	160
6	Numerical results	167
6.1	Introduction	167
6.2	Hock-Schittkowski Test-set	172
6.2.1	Reduced algorithm vs. Full algorithm	174
6.2.2	Reduced algorithm vs. MPBNGC	175
6.2.3	Reduced algorithm vs. SolvOpt	175
6.3	Exclusion boxes	176
6.3.1	Notation	176
6.3.2	Overview of the results	178
6.3.3	Linearly constrained case (fixed box)	181
6.3.3.1	BNLC vs. MPBNGC	181
6.3.3.2	BNLC vs. SolvOpt	183
6.3.4	Linearly constrained case (variable box)	184
6.3.4.1	BNLC vs. MPBNGC	184
6.3.4.2	BNLC vs. SolvOpt	186
6.3.5	Nonlinearly constrained case	187
6.3.5.1	Reduced algorithm vs. MPBNGC	187
6.3.5.2	Reduced algorithm vs. SolvOpt	189
6.4	Higher dimensional piecewise quadratic examples	191
6.4.1	Easy examples with N/2 constraint components	195
6.4.1.1	Reduced algorithm vs. MPBNGC	196
6.4.1.2	Reduced algorithm vs. SolvOpt	196
6.4.2	Easy examples with N constraint components	196
6.4.2.1	Reduced algorithm vs. MPBNGC	197
6.4.2.2	Reduced algorithm vs. SolvOpt	197
6.4.3	Difficult examples with N/2 constraint components	198
6.4.3.1	Reduced algorithm vs. MPBNGC	198
6.4.3.2	Reduced algorithm vs. SolvOpt	199
6.4.4	Difficult examples with N constraint components	199
6.4.4.1	Reduced algorithm vs. MPBNGC	200
6.4.4.2	Reduced algorithm vs. SolvOpt	200
7	Perspectives	203

A Result tables	205
Bibliography	236
Curriculum Vitae	237

Chapter 1

Introduction

1.1 Outline

An important area of modern research is global optimization as it occurs very frequently in applications (extensive surveys on global optimization can be found in NEUMAIER [83], FLOUDAS [23, 24], HANSEN [34], and KEARFOTT [48]). A method for solving global optimization problems efficiently is by using a branch and bound algorithm (as, e.g., BARON by SAHINIDIS [89, 90], the COCONUT environment by SCHICHL [91, 92, 93], or LINGO by SCHRAGE [97]), which divides the feasible set into smaller regions and then tries to exclude regions that cannot contain a global optimizer. Therefore, it is important to have tools which allow to identify such regions. In this thesis we will present a method which is able to find such regions for a CSP (constraint satisfaction problem), i.e. for a global optimization problem with a constant objective function.

Certificate of infeasibility. For this purpose we consider the CSP

$$\begin{aligned} F(x) &\in \mathbf{F} \\ x &\in \mathbf{x} \end{aligned} \tag{1.1}$$

and we assume that a solver, which is able to solve a CSP, takes the box $\mathbf{u} := [\underline{u}, \bar{u}] \subseteq \mathbf{x}$ into consideration during the solution process. We will present a certificate of infeasibility f (that is based upon an idea of Arnold Neumaier), which is a nondifferentiable and nonconvex function in general, with the following property: If there exists a vector y with

$$f(y, \underline{u}, \bar{u}) < 0, \tag{1.2}$$

then the CSP (1.1) has no feasible point in $\mathbf{u}$ and consequently this box can be excluded for the rest of the solution process. Therefore, a box $\mathbf{u}$ for which (1.2) holds is called an **exclusion box**.

Easy examples immediately show that there exist CSPs which have boxes that satisfy (1.2), so it is worth to pursue this approach further.

Exclusion boxes. The obvious way for finding an exclusion box for the CSP (1.1) is to minimize f

$$\min_y f(y, \underline{u}, \bar{u}) \quad (1.3)$$

and stop the minimization if a negative function value occurs. Since modern solvers offer many other possibilities for treating a box, we do not want to spend too much time for this minimization problem. Therefore, the idea is to let a nonsmooth solver only perform a few steps for solving (1.3).

To find at least an exclusion box $\mathbf{v} := [\underline{v}, \bar{v}] \subseteq \mathbf{u}$ with $\underline{v} + r \leq \bar{v}$, where $r \in (0, \bar{u} - \underline{u})$ is fixed, we can try to solve the linearly constrained problem

$$\begin{aligned} \min_{y, \underline{v}, \bar{v}} f(y, \underline{v}, \bar{v}) \\ \text{s.t. } [\underline{v} + r, \bar{v}] \subseteq \mathbf{u} . \end{aligned} \quad (1.4)$$

Another important aspect in this context is to enlarge an exclusion box $\mathbf{v}$ by solving

$$\begin{aligned} \max_{y, \underline{v}, \bar{v}} \mu(\underline{v}, \bar{v}) \\ \text{s.t. } f(y, \underline{v}, \bar{v}) \leq \delta \\ [\underline{v}, \bar{v}] \subseteq \mathbf{u} , \end{aligned} \quad (1.5)$$

where $\delta < 0$ is given and μ measures the magnitude of the box $\mathbf{v}$ (e.g., $\mu(\underline{v}, \bar{v}) := |\bar{v} - \underline{v}|_1$). Since only feasible points of (1.5) are useful for enlarging an exclusion box and we only want to perform a few steps of a nonsmooth solver as before, we expect benefits from a nonsmooth solver that only creates feasible iterates because then the current best point can always be used for our purpose.

Linearly constrained nonsmooth optimization. For solving the optimization problems (1.3), (1.4) and (1.5), we need a nonsmooth solver. There exist different types of nonsmooth solvers like the R-algorithm by SHOR [100], gradient sampling algorithms (e.g., by BURKE et al. [15]) or bundle methods. We will concentrate on the latter ones as they proved to be quite efficient.

One of the few publicly available bundle methods is the bundle-Newton method for nonsmooth, non-convex unconstrained minimization by LUKŠAN & VLČEK [64]. We sum up its key features:

- It is the only method which I know of that uses second order information of the objective function, which results in faster convergence (in particular it was shown in LUKŠAN & VLČEK [64, p. 385, Section 4] that the bundle-Newton method converges superlinearly for strongly convex, twice times continuously differentiable functions).
- The search direction is computed by solving a convex quadratic program (QP) (based on an SQP-approach in some sense).
- It uses a line search concept for deciding whether a serious step or a null step is performed.
- Its implementation PNEW, which is described in LUKŠAN & VLČEK [62], is written in FORTRAN.

Therefore, we can use the bundle-Newton method for solving (1.3) and (1.4) (as the linear constraints can just be inserted into the QP without any additional difficulties).

For dealing with (1.5) we need to solve a constrained optimization problem. In general, every nonsmooth solver for unconstrained optimization can treat constrained problems via penalty functions. Nevertheless, choosing the penalty parameter well is a highly nontrivial task. Furthermore, we proclaimed above (when discussing exclusion boxes) that we only want the nonsmooth solver to perform a few steps and consequently we would like to achieve a feasible descent within these steps.

Nonlinearly constrained nonsmooth optimization. Therefore, we extend the bundle-Newton method to the constrained case in a very special way:

- We use second order information of the constraint (without loss of generality, we can always assume in nonsmooth optimization that we have a single constraint since the maximum of finitely many locally Lipschitz continuous function is locally Lipschitz continuous — and this is the important class of nonsmooth functions we deal with).
- We use the SQP-approach of the bundle-Newton method for computing the search direction for the constrained case and combine it with the idea of quadratic constraint approximation, as it is used, e.g., in the sequential quadratically constrained quadratic programming method by SOLODOV [101] (this method is not a bundle method), in the hope to obtain good feasible iterates, where we only accept strictly feasible points as serious steps. Therefore, we have to solve a strictly feasible convex QCQP (quadratically constrained quadratic program) for computing the search direction.
- Using such a QCQP for computing the search direction yields a line search condition for accepting infeasible points as trial points (which is different to that in, e.g., MIFFLIN [76]).
- One of the most important properties of the convex QP (that is used to determine the search direction) with respect to a bundle method is its strong duality (e.g., for a meaningful termination criterion, for global convergence,...) which is also true in the case of strictly feasible convex QCQPs.

Since there exist only a few solvers specialized in solving QCQPs (all written in MATLAB or C, none in FORTRAN), we realized the following implementations:

- A pure MATLAB code (for easy understanding, modifying and testing new ideas concerning the algorithm).
- A MATLAB code in which the main parts of the algorithm are split into several subroutines, where every subroutine can either be called as pure MATLAB code or via a C `mex`-file (this is useful for partially speeding up the algorithm, but still keeping it simple enough for modifying and testing many examples of the modified version).
- A pure C code (for performance).

In the unconstrained case, all three versions produce the same results as the original FORTRAN bundle-Newton method by LUKŠAN & VLČEK [64]. The ability of treating nonsmooth constraints was successfully tested on many examples of SCHITTKOWSKI [95, 96] with inequality constraints, where the inequality constraints were formulated as a single nonsmooth inequality constraint.

Furthermore, the algorithm was integrated into an experimental version of GloptLab by DOMES [18], which is a configurable MATLAB framework for computing the global solution of a quadratic CSP, as a custom method (including a version using analytic derivative information of the certificate of infeasibility), where reasonable results were achieved, too.

1.2 Organization

The thesis is organized as follows: In Chapter 2 we recall the basics of an SQP-method which is a common technique in smooth optimization and we summarize the most important facts about nonsmooth optimization theory. In Chapter 3 we give the theoretical foundation of the second order bundle algorithm and afterwards we present the algorithm and the line search in detail. Furthermore, we show the convergence of the line search and the global convergence of the algorithm (cf. FENDL & SCHICHL [20]). In Chapter 4 we discuss some aspects that arise when using a convex QCQP for the computation of the search direction problem like the reduction of its dimension and the existence of a strictly feasible starting point for its SOCP-reformulation. Furthermore, we justify the approach of determining the search direction by solving a QCQP numerically by comparing the results of some well-known solvers for our search direction problem (cf. FENDL & SCHICHL [21, p. 11 ff, Section 3]). In Chapter 5 we first recall the basic facts of interval analysis which are necessary for introducing the certificate of infeasibility which is done afterwards. Then we discuss some important properties of the certificate and we explain in detail how the certificate is used for obtaining exclusion boxes in a CSP by applying a nonsmooth solver. Furthermore, the optimization problems to which we apply the nonsmooth solver as well as how we obtain a starting point for it are discussed (cf. FENDL et al. [19]). In Chapter 6 we give numerical results of the second order bundle algorithm for some examples of the Hock-Schittkowski collection by SCHITTKOWSKI [95, 96], for custom examples that arise in the context of finding exclusion boxes for a quadratic CSP in GloptLab by DOMES [18] as well as for higher dimensional piecewise quadratic examples, and finally we compare these results to those of MPBNGC by MÄKELÄ [70] and SolvOpt by KAPPEL & KUNTSEVICH [41] to emphasize the good performance of the algorithm on constrained problems (cf. FENDL & SCHICHL [21, p. 22 ff, Section 4]). Finally, we discuss some possible, future development plans, which might be worth considering, in Chapter 7.

1.3 Notation

Throughout the thesis we use the following conventions:

- We denote the non-negative real numbers by

$$\mathbb{R}_{\geq 0} := \{x \in \mathbb{R} : x \geq 0\}$$

(and analogously for $\leq$ as well as $>$).

- For $x \in \mathbb{R}^n$ we denote the Euclidean norm of x by

$$|x| := \left(\sum_{i=1}^n x_i^2 \right)^{\frac{1}{2}},$$

which will be the most frequently used norm in the following. In a few situations we will write $|x|_2$ instead of $|x|$ to emphasize the Euclidean norm. Furthermore, we denote the 1-norm of x by

$$|x|_1 := \sum_{i=1}^n |x_i|,$$

and we denote the maximum-norm of x by

$$|x|_\infty := \max_{i=1,\dots,n} |x_i|.$$

- Let $x \in \mathbb{R}^n$ and $1 \leq i \leq j \leq n$. Then we define the (MATLAB-like) colon operator

$$x_{i:j} := \begin{pmatrix} x_i \\ \vdots \\ x_j \end{pmatrix}.$$

- We denote the vector space of the real symmetric $n \times n$ -matrices by $\mathbb{R}_{\text{sym}}^{n \times n}$.
- For $A \in \mathbb{R}_{\text{sym}}^{n \times n}$ we denote the spectral norm of A by

$$|A| := \sqrt{\max_{1 \leq i \leq n} |\lambda_i|}, \tag{1.6}$$

where λ_i denotes the i th eigenvalue of A (Note that for real symmetric matrices all eigenvalues are real).

- We denote the smallest resp. the largest eigenvalue of a positive definite matrix $A \in \mathbb{R}_{\text{sym}}^{n \times n}$ by $\lambda_{\min}(A)$ resp. $\lambda_{\max}(A)$.

Chapter 2

Basics of optimization theory

In the following chapter we summarize the basics of an SQP-method, since we will approximate a nonsmooth problem by a sequence of smooth problems to derive our algorithm in Section 3.1 and hence we will need some facts about smooth optimization, and we present the most important facts about nonsmooth optimization theory.

2.1 Introduction

We recall the Cauchy-Schwarz inequality.

Proposition 2.1. *Let $x, y \in \mathbb{R}^n$. Then*

$$|x^T y| \leq |x| \cdot |y| . \quad (2.1)$$

Proof. HEUSER [37, p. 96, 12.3 Cauchy-Schwarzsche Ungleichung]. □

We recall the following property of the spectral norm.

Proposition 2.2. *Let $A \in \mathbb{R}^{n \times n}$ be positive definite. Then all eigenvalues of A are positive and*

$$|A| = \sqrt{\lambda_{\max}(A)} . \quad (2.2)$$

Proof. The proof of the positivity of the eigenvalues can be found in GOLUB & VAN LOAN [27, p. 394, Follow up of Theorem 8.1.2]. Combining this with (1.6) directly yields (2.2). □

We will use the following tools from real analysis.

Proposition 2.3. *Let $X \subseteq \mathbb{R}$ be bounded from below. Then*

$$\sup(-X) = -\inf X , \quad (2.3)$$

where

$$-X := \{-x : x \in X\} .$$

Proof. HEUSER [37, p. 76, Exercise 4(a)]. □

Proposition 2.4. *Let $\{x_n\}$ and $\{y_n\}$ be sequences in $\mathbb{R}^n$. If $x_n \rightarrow x$ is convergent and y_n is bounded, then*

$$(x_n - x)y_n \rightarrow 0 . \quad (2.4)$$

Proof. Clear. □

Theorem 2.5 (Bolzano-Weierstraß). *Each bounded sequence in $\mathbb{R}^n$ has a convergent subsequence.*

Proof. HEUSER [38, p. 22, Proof of 109.15 Satz]. □

We recall the following basics from convex analysis.

Definition 2.6. We define the convex hull of $S \subseteq \mathbb{R}^n$ by

$$\text{ch}(S) := \left\{ \sum_{i=1}^k \lambda_i x_i : x_i \in S, \lambda_i \geq 0, \sum_{i=1}^k \lambda_i = 1, k = 1, 2, \dots \right\} ,$$

i.e. the convex hull consists of all convex combinations of elements in S .

Definition 2.7. A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is called convex, if for all $x, y \in \mathbb{R}^n$ and $t \in [0, 1]$

$$f((1-t)x + ty) \leq (1-t)f(x) + tf(y) \quad (2.5)$$

holds, i.e. f is convex, if the function values of the connection line are below the connection line of the function values.

Proposition 2.8. *A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex, if and only if*

$$f\left(\sum_{i=1}^k \lambda_i x_i\right) \leq \sum_{i=1}^k \lambda_i f(x_i) \quad \text{for all } x_i \in M, \lambda_i \geq 0 \text{ for } i = 1, \dots, k \geq 2, \sum_{i=1}^k \lambda_i = 1 , \quad (2.6)$$

i.e. f is convex, if and only if the image of each convex combination is not greater than the convex combination of the image.

Proof. This is a standard result from convex analysis which can be looked up in, e.g., BOYD & VANDENBERGHE [13, p. 77, Section 3.1.8]. □

Example 2.9. Let $\gamma > 0$ and $\omega \geq 1$.

1. The functions $h_1 : \mathbb{R} \rightarrow \mathbb{R}$ and $h_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$

$$h_1(x) := \gamma|x|^\omega$$

$$h_2\left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}\right) := \max(x_1, y_1)$$

are convex (by checking (2.5)).

2. Let $x_i, y_i \in \mathbb{R}$ and $t_i \geq 0$ for $i = 1, \dots, k$ with $\sum_{i=1}^k t_i = 1$, then (2.6) and the convexity of h_1 and h_2 imply

$$\gamma \left(\sum_{i=1}^k t_i |x_i| \right)^\omega \leq \sum_{i=1}^k t_i (\gamma |x_i|^\omega) \quad (2.7)$$

$$\max \left(\sum_{i=1}^k t_i x_i, \sum_{i=1}^k t_i y_i \right) \leq \sum_{i=1}^k t_i \max(x_i, y_i) . \quad (2.8)$$

2.2 Smooth optimization

2.2.1 Smooth Optimality conditions

Theorem 2.10. *Let $f, F_i : \mathbb{R}^n \rightarrow \mathbb{R}$ (with $i = 1, \dots, m$) be C^1 and $\hat{x} \in \mathbb{R}^n$ be a solution of the smooth optimization problem*

$$\begin{aligned} \min f(x) \\ \text{s.t. } F_i(x) \leq 0 \quad \text{for all } i = 1, \dots, m . \end{aligned} \quad (2.9)$$

Then there exists $\hat{z} \geq 0$ with

$$\begin{aligned} \kappa \nabla f(\hat{x})^T + \sum_{i=1}^m \nabla F_i(\hat{x})^T \hat{z}_i &= 0 \\ \hat{z}_i F_i(\hat{x}) &= 0 \quad \text{for all } i = 1, \dots, m \\ \kappa &= 1 \text{ or } (\kappa = 0, \hat{z} \neq 0) . \end{aligned} \quad (2.10)$$

If all occurring functions are convex, then the following additional statements hold:

- *If there exists a strictly feasible x (i.e. $F(x) < 0$), then we can always assume $\kappa = 1$.*
- *The conditions (2.10) are sufficient for a feasible $\hat{x}$ being a minimizer of (2.9).*

Proof. Combine, e.g., SCHICHL & NEUMAIER [94, p. 19, 4.1 Theorem] and BOYD & VANDENBERGHE [13, p. 243, 5.5.3 KKT optimality conditions]. □

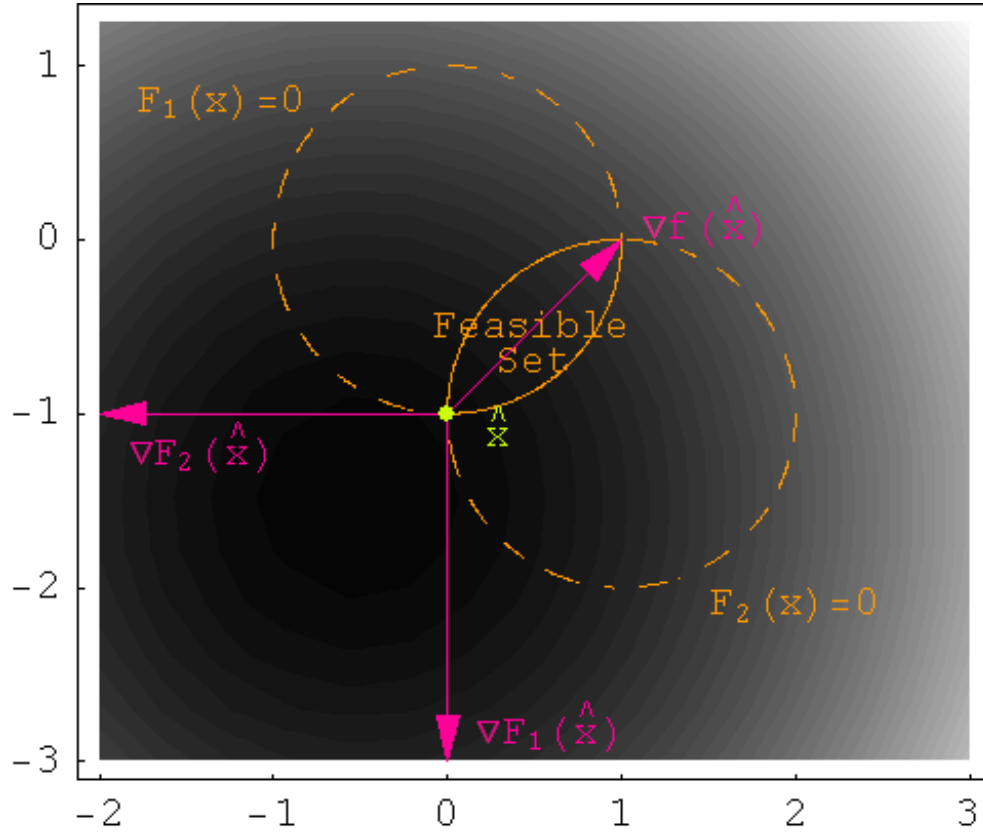
Example 2.11. We define the convex functions $f, \hat{F}_1, \hat{F}_2, \hat{F}_3 : \mathbb{R}^2 \rightarrow \mathbb{R}$ by

$$\begin{aligned} f(x) &:= \left(x_1 + \frac{1}{2}\right)^2 + \left(x_2 + \frac{3}{2}\right)^2 \\ \hat{F}_1(x) &:= x_1^2 + x_2^2 - 1 \\ \hat{F}_2(x) &:= (x_1 - 1)^2 + (x_2 + 1)^2 - 1 \\ \hat{F}_3(x) &:= (x_1 - 1)^2 - x_2 - 1 . \end{aligned}$$

1. Convex constraints: We consider optimization problem (2.9) with

$$F_i(x) := \hat{F}_i(x) \quad \text{for } i = 1, 2 ,$$

which yields the following graphical illustration (including the contour lines of f)



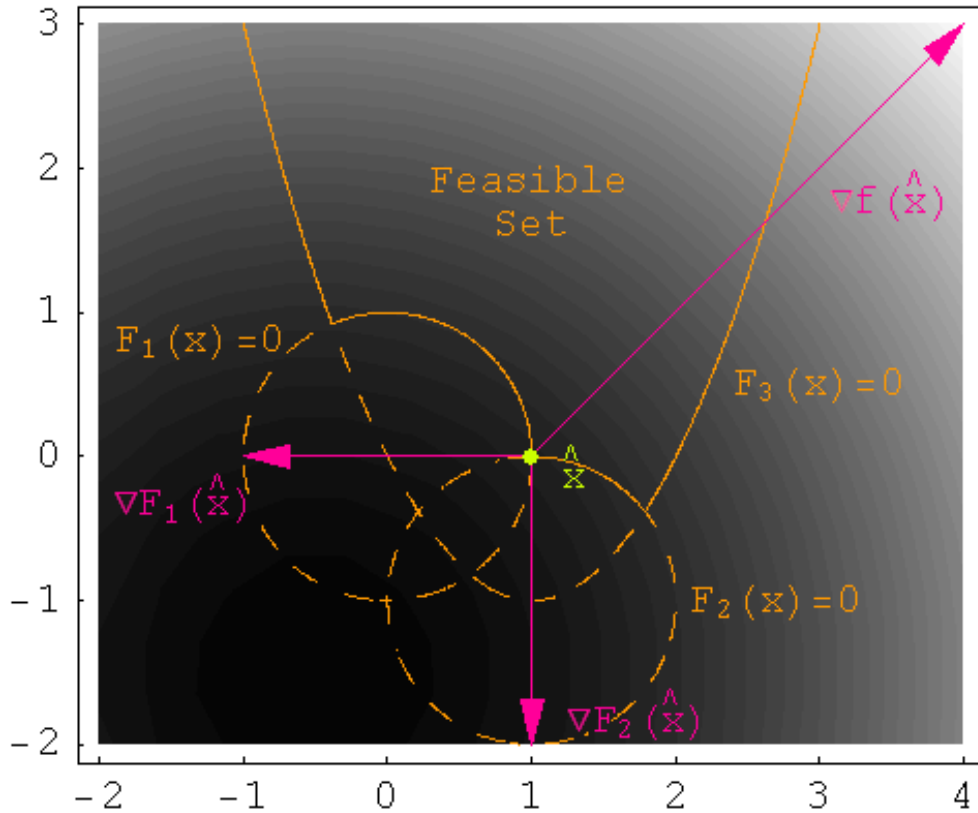
then the (global) minimizer $\hat{x} = (0, -1)$ satisfies the optimality conditions (2.10) with $\hat{z} = (\frac{1}{2}, \frac{1}{2})$.

2. Nonconvex constraints: We consider optimization problem (2.9) with

$$F_i(x) := -\hat{F}_i(x) \quad \text{for } i = 1, 2$$

$$F_3(x) := \hat{F}_3(x),$$

which yields the following plot (including the contour lines of f)



then the local minimizer $\hat{x} = (1, 0)$ satisfies the optimality conditions (2.10) with $\hat{z} = (\frac{3}{2}, \frac{3}{2})$.

2.2.2 SQP

Motivation 2.12. One possibility to find a solution of the optimization problem (2.9) is using an SQP-method (sequential quadratic programming). An SQP-method minimizes the quadratic approximation of the Lagrangian subject to linearizations of the constraints and then it uses the obtained minimizer as the new iteration point (or it performs a line search between the current iteration point and the obtained minimizer to determine the new iteration point). Since quadratic information is necessary for this approach, we demand $f, F_i : \mathbb{R}^n \rightarrow \mathbb{R}$ (with $i = 1, \dots, m$) to be C^2 in this subsection. Furthermore, we use the following compact notation for the constraints

$$F(x) := \begin{pmatrix} F_1(x) \\ \vdots \\ F_m(x) \end{pmatrix} .$$

Definition 2.13.

1. We define the Lagrangian $L : \mathbb{R}^n \times \mathbb{R}_{\geq 0}^m \rightarrow \mathbb{R}$ of optimization problem (2.9) by

$$L(x, \lambda) := f(x) + \sum_{i=1}^m F_i(x) \lambda_i . \quad (2.11)$$

2. We define the dual problem of (2.9) by

$$\begin{aligned} \max_{\lambda} \quad & L(x, \lambda) \\ \text{s.t.} \quad & \nabla_x L(x, \lambda) = 0 . \end{aligned} \quad (2.12)$$

Proposition 2.14.

1. The Lagrangian L is an underestimation of the objective function f on the feasible set for all $\lambda \geq 0$

$$L(x, \lambda) \leq f(x) . \quad (2.13)$$

2. The optimization problem

$$\begin{aligned} \min \quad & L(x, \lambda) \\ \text{s.t.} \quad & F_i(x) \leq 0 \quad \text{for all } i = 1, \dots, m \\ & \lambda \geq 0 \end{aligned} \quad (2.14)$$

is a relaxation for (2.9), i.e. (2.14) has a larger feasible set than (2.9) and the objective function of (2.14) is an underestimation of the objective function of (2.9) on the feasible set of (2.9).

Proof.

1. (2.13) holds due to

$$L(x, \lambda) \stackrel{(2.11)}{=} f(x) + \sum_{i=1}^m \underbrace{F_i(x)}_{\leq 0} \underbrace{\lambda_i}_{\geq 0} \leq f(x) .$$

2. (2.14) has a larger feasible set than (2.9) (because of the additional variable λ) and L is an underestimation of f on the feasible set due to (2.13). $\square$

Proposition 2.15. *We have*

$$\nabla_{(x,\lambda)} L(x, \lambda) = (\nabla f(x) + \lambda^T \nabla F(x), F(x)^T) \quad (2.15)$$

$$\nabla_{(x,\lambda)}^2 L(x, \lambda) = \begin{pmatrix} \nabla_{xx}^2 L(x, \lambda) & \nabla F(x)^T \\ \nabla F(x) & 0 \end{pmatrix}, \quad (2.16)$$

where

$$\nabla_{xx}^2 L(x, \lambda) = \nabla^2 f(x) + \sum_{i=1}^m \nabla^2 F_i(x) \lambda_i. \quad (2.17)$$

Proof. By direct calculation. $\square$

Assumption 2.16. Let the following two properties be satisfied (cf. NOCEDAL & WRIGHT [84, p. 531, Assumption 18.1]):

1. “Constraint Qualification”: The matrix $\nabla F(x) \in \mathbb{R}^{m \times n}$ (gradient of the constraints) has full rank.
2. The Hessian of the Lagrangian with respect to the x -components $\nabla_{xx}^2 L(x, \lambda) \in \mathbb{R}^{n \times n}$ from (2.17) is positive definite on the tangent space of the constraints

$$d^T \nabla_{xx}^2 L(x, \lambda) d > 0 \text{ for all } d \in \mathbb{R}^n \text{ with } d \neq 0 \text{ and } \nabla F(x) d = 0 \quad (2.18)$$

(cf. the sufficient second order optimality conditions in, e.g., NOCEDAL & WRIGHT [84, p. 333, Theorem 12.6]).

Proposition 2.17. *Let $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}_{\geq 0}^m$ be fixed and $d_x \in \mathbb{R}^n$, $d_\lambda \in \mathbb{R}_{\geq 0}^m$ and introduce the notation*

$\doteq$:= “Is equal up to error terms of first order”

$\ddot{=}$:= “Is equal up to error terms of second order” .

1. *Let Assumption 2.16 be satisfied. Then the quadratic approximation of L at the point (x, λ) is given by*

$$L(x + d_x, \lambda + d_\lambda) \ddot{=} L(x, \lambda) + \nabla f(x) d_x + F(x)^T d_\lambda + \frac{1}{2} d_x^T \nabla_{xx}^2 L(x, \lambda) d_x. \quad (2.19)$$

2. *The linear approximation of F_i (for $i = 1, \dots, m$) at the point $x \in \mathbb{R}^n$ is given by*

$$F_i(x + d_x) \doteq F_i(x) + \nabla F_i(x) d_x. \quad (2.20)$$

Proof.

1. Because of

$$\begin{aligned} \nabla L(x, \lambda) \begin{pmatrix} d_x \\ d_\lambda \end{pmatrix} &\stackrel{(2.15)}{=} \nabla f(x) d_x + \lambda^T (\nabla F(x) d_x) + F(x)^T d_\lambda \\ &\stackrel{(2.18)}{=} \nabla f(x) d_x + F(x)^T d_\lambda \end{aligned} \quad (2.21)$$

and

$$\begin{aligned} \begin{pmatrix} d_x \\ d_\lambda \end{pmatrix}^T \nabla^2 L(x, \lambda) \begin{pmatrix} d_x \\ d_\lambda \end{pmatrix} &\stackrel{(2.16)}{=} \begin{pmatrix} d_x \\ d_\lambda \end{pmatrix}^T \begin{pmatrix} \nabla_{xx}^2 L(x, \lambda) d_x + \nabla F(x)^T d_\lambda \\ \nabla F(x) d_x \end{pmatrix} \\ &\stackrel{(2.18)}{=} d_x^T \nabla_{xx}^2 L(x, \lambda) d_x, \end{aligned} \quad (2.22)$$

Taylor's theorem implies (2.19)

$$\begin{aligned} L(x + d_x, \lambda + d_\lambda) &\doteq L(x, \lambda) + \nabla L(x, \lambda) \begin{pmatrix} d_x \\ d_\lambda \end{pmatrix} + \frac{1}{2} \begin{pmatrix} d_x \\ d_\lambda \end{pmatrix}^T \nabla^2 L(x, \lambda) \begin{pmatrix} d_x \\ d_\lambda \end{pmatrix} \\ &\stackrel{(2.21), (2.22)}{\doteq} L(x, \lambda) + \nabla f(x) d_x + F(x)^T d_\lambda + \frac{1}{2} d_x^T \nabla_{xx}^2 L(x, \lambda) d_x. \end{aligned}$$

2. (2.20) holds due to Taylor's theorem. □

Proposition 2.18. *Let Assumption 2.16 be satisfied. Then the QP*

$$\begin{aligned} &f(x) + \min_{d_x} \nabla f(x) d_x + \frac{1}{2} d_x^T \nabla_{xx}^2 L(x, \lambda) d_x \\ &\text{s.t. } F_i(x) + \nabla F_i(x) d_x \leq 0 \quad \text{for all } i = 1, \dots, m \end{aligned} \quad (2.23)$$

is a (local) approximation of the optimization problem (2.14) (i.e. for a relaxation of optimization problem (2.9)), where its solution d_x is called the **SQP-step** for the optimization problem (2.14) at x .

Proof. For $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}_{\geq 0}^m$, $d_\lambda \in \mathbb{R}^m$ (all fixed) with

$$\lambda + d_\lambda \geq 0 \quad (2.24)$$

and (the variable) $d_x \in \mathbb{R}^n$, optimization problem (2.14) reads

$$\begin{aligned} &\min_{d_x} L(x + d_x, \lambda + d_\lambda) \\ &\text{s.t. } F_i(x + d_x) \leq 0 \quad \text{for all } i = 1, \dots, m \\ &\quad \lambda + d_\lambda \geq 0 \\ &\stackrel{(2.24)}{=} \min_{d_x} L(x + d_x, \lambda + d_\lambda) \\ &\text{s.t. } F_i(x + d_x) \leq 0 \quad \text{for all } i = 1, \dots, m \\ &\stackrel{(2.19)}{\doteq} \min_{d_x} L(x, \lambda) + \nabla f(x) d_x + F(x)^T d_\lambda + \frac{1}{2} d_x^T \nabla_{xx}^2 L(x, \lambda) d_x \\ &\text{s.t. } F_i(x + d_x) \leq 0 \quad \text{for all } i = 1, \dots, m \\ &\stackrel{(2.11)}{=} \sum_{i=1}^m F_i(x) \lambda_i + F(x)^T d_\lambda + f(x) + \min_{d_x} \nabla f(x) d_x + \frac{1}{2} d_x^T \nabla_{xx}^2 L(x, \lambda) d_x \\ &\quad \text{s.t. } F_i(x + d_x) \leq 0 \quad \text{for all } i = 1, \dots, m \\ &\stackrel{(2.20)}{\doteq} \sum_{i=1}^m F_i(x) \lambda_i + F(x)^T d_\lambda + f(x) + \min_{d_x} \nabla f(x) d_x + \frac{1}{2} d_x^T \nabla_{xx}^2 L(x, \lambda) d_x \\ &\quad \text{s.t. } F_i(x) + \nabla F_i(x) d_x \leq 0 \quad \text{for all } i = 1, \dots, m. \end{aligned} \quad \square$$

Remark 2.19.

1. According to Motivation 2.12, we determine the search direction $d_x \in \mathbb{R}^n$ by solving the QP (2.23).
2. A difficulty of an infeasible SQP-method (e.g., SNOPT by GILL et al. [26]) — i.e. infeasible iteration points x_k may occur — is that the linear constraints of the QP (2.23) can be infeasible: Consider, e.g., the optimization problem (2.9) with the constraint $F : \mathbb{R} \rightarrow \mathbb{R}^2$

$$F(x) := \begin{pmatrix} x-1 \\ -x^2+4 \end{pmatrix} \leq 0 \quad (2.25)$$

(cf. NOCEDAL & WRIGHT [84, p. 535, 18.3 Algorithmic development]), where the feasible set is colored orange (cf. Figure 2.1). If we linearize the constraint (2.25) at $x_k := 1$ (Note: This x_k is infeasible for (2.25)), we obtain for the QP (2.23) the linear constraints

$$\begin{aligned} 0 &\geq \begin{pmatrix} F_1(x_k) + \nabla F_1(x_k)(x-x_k) \\ F_2(x_k) + \nabla F_2(x_k)(x-x_k) \end{pmatrix} \\ &= \begin{pmatrix} x-1 \\ 3-2(x-1) \end{pmatrix} \\ &\iff x \leq 1 \quad \wedge \quad x \geq \frac{5}{2}, \end{aligned}$$

which are inconsistent (and therefore we have no orange area in Figure 2.2).

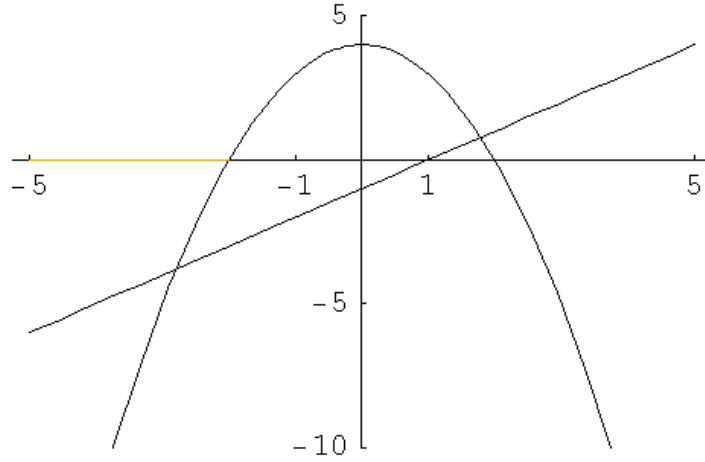


FIGURE 2.1: Feasible set of the optimization problem

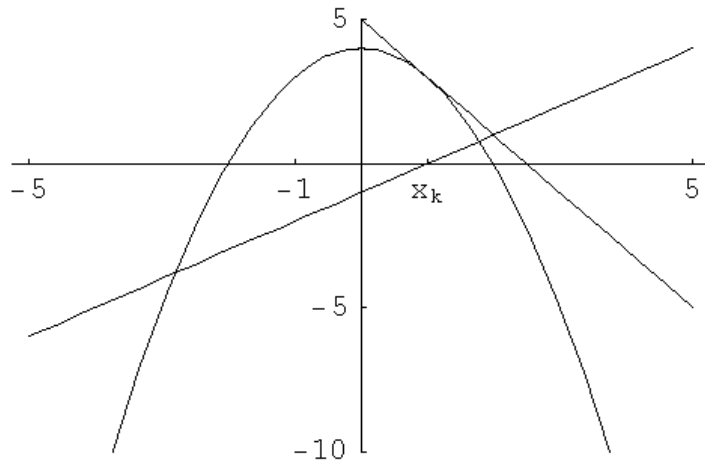


FIGURE 2.2: Linearizations of the constraints at x_k

3. Note that the difficulty described above does not arise for a feasible SQP-method (e.g., FSQP by LAWRENCE & TITS [55]) — i.e. only feasible iteration points x_k are accepted — as then $d_x = 0$ is always feasible for the QP (2.23). Nevertheless, in this case it can be difficult to obtain feasible points that make good progress towards a solution (cf. Remark 3.11).

2.3 Nonsmooth optimization

2.3.1 Nonsmooth Optimality conditions

Motivation 2.20. We gather information on the optimality conditions of the nonsmooth optimization problem

$$\begin{aligned} \min f(x) \\ \text{s.t. } F_i(x) \leq 0 \quad \text{for all } i = 1, \dots, m, \end{aligned} \quad (2.26)$$

where $f, F_i : \mathbb{R}^n \rightarrow \mathbb{R}$ are locally Lipschitz continuous. For this purpose, we closely follow the exposition in BORWEIN & LEWIS [12].

Definition 2.21. Let $U \subseteq \mathbb{R}^n$ be open and $f : \mathbb{R}^n \rightarrow \mathbb{R}$.

1. f is called Lipschitz continuous on U , if there exists $K > 0$ with

$$|f(y) - f(x)| \leq K|y - x|$$

for all $x, y \in U$.

2. f is called locally Lipschitz continuous around $x \in \mathbb{R}^n$, if f is Lipschitz continuous on a neighborhood U of x .

Definition 2.22. Let $U \subseteq \mathbb{R}^n$ be open and $f : \mathbb{R}^n \rightarrow \mathbb{R}$.

1. We define the **Clarke directional derivative** in $x \in U$ in direction $d \in \mathbb{R}^n$ by

$$f^0(x, d) := \limsup_{h \rightarrow 0, t \downarrow 0} \frac{f(x + h + td) - f(x + h)}{t}.$$

2. We define the **subdifferential** $\partial f(x) \subseteq \mathbb{R}^n$ of f in $x \in U$ by

$$\partial f(x) := \text{ch}\{g \in \mathbb{R}^n : g^T d \leq f^0(x, d) \text{ for all } d \in \mathbb{R}^n\}. \quad (2.27)$$

The elements of $\partial f(x)$ are called **subgradients**.

3. We define the set $\partial^2 f(x) \subseteq \mathbb{R}_{\text{sym}}^{n \times n}$ of the substitutes for the Hessian of f at x by

$$\partial^2 f(x) := \begin{cases} \{G\} & \text{if the Hessian } G \text{ of } f \text{ at } x \text{ exists} \\ \mathbb{R}_{\text{sym}}^{n \times n} & \text{else.} \end{cases} \quad (2.28)$$

Remark 2.23.

- In contrast to the definition of the subdifferential $\partial f(x)$ in (2.27), the definition of $\partial^2 f(x)$ in (2.28) is a purely formal one.
- In practice, if the Hessian of f at x does not exist, then we choose the matrix $G(x) \in \partial^2 f(x)$ similar to the Hessian at a nearby point at which it exists. If, e.g., f is a max-function

$$f(x) = \max_{i=1,\dots,k} f_i(x)$$

with twice times continuously differentiable functions $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$, then $G(x) \in \partial^2 f(x)$ is given by, e.g.,

$$G(x) = \nabla^2 f_i(x) ,$$

where i is any index for which $f_i(x) = f(x)$ holds.

Notation 2.24. For a set S we denote the power set of S by $\mathcal{P}(S) := \{X : X \subseteq S\}$.

We summarize the most important properties of the Clarke directional derivative and the subdifferential.

Proposition 2.25.

1. The Clarke directional derivative has the alternative representation

$$f^0(x, d) = \max_{g \in \partial f(x)} g^T d .$$

2. Locally Lipschitz continuous functions are differentiable almost everywhere, i.e. everywhere except on a set $U \subseteq \mathbb{R}^n$ with measure zero, and the subdifferential $\partial f(x)$ has the alternative representation

$$\partial f(x) = \text{ch} \left\{ \lim_{k \rightarrow \infty} \nabla f(x_k) : x_k \xrightarrow{k \rightarrow \infty} x, x_k \notin U \right\} . \quad (2.29)$$

3. For the subdifferential $\partial f(x)$ the following statements are valid:

$$\partial f(x) \text{ is non-empty, convex and compact.} \quad (2.30)$$

If $f \in C^1(\mathbb{R}^n)$, then $\partial f(x) = \{\nabla f(x)\}$.

$$\text{If } f \text{ is convex, then } \partial f(x) = \{g \in \mathbb{R}^n : f(y) - f(x) \geq g^T(y - x) \text{ for all } y \in \mathbb{R}^n\} . \quad (2.31)$$

4. The subdifferential (interpreted as a set-valued map) $\partial f : \mathbb{R}^n \rightarrow \mathcal{P}(\mathbb{R}^n)$ is

(a) locally bounded, i.e.:

$$\text{If } U \subseteq \mathbb{R}^n \text{ is bounded, then } \{g \in \partial f(y) : y \in U\} \text{ is bounded} . \quad (2.32)$$

(b) upper semicontinuous, i.e.: Let $\{x_i\}$ be a sequence in $\mathbb{R}^n$ and $g_i \in \partial f(x_i)$ for all i .

$$\begin{aligned} &\text{If } \bar{x} \in \mathbb{R}^n \text{ is an accumulation point of } \{x_i\} \text{ and} \\ &\bar{g} \in \mathbb{R}^n \text{ is an accumulation point of } \{g_i\}, \text{ then } \bar{g} \in \partial f(\bar{x}) . \end{aligned} \quad (2.33)$$

Proof.

1. BORWEIN & LEWIS [12, p. 145, Corollary 6.1.2].
2. (2.29) holds due to BORWEIN & LEWIS [12, p. 154, Theorem 6.2.5].
3. The first property holds due to BORWEIN & LEWIS [12, p. 145, Corollary 6.1.2]. The second property follows directly from (2.29) and the assumption of the continuous differentiability of f . The third property (2.31) follows from BORWEIN & LEWIS [12, p. 152, Theorem 6.2.2 and p. 44, inequality (3.1.4)].
4. The local boundedness result (2.32) can be found in BORWEIN & LEWIS [12, p. 216, beginning of Section 8.2 with proof on p. 220, Exercise 5]. The result (2.33) on upper semicontinuity can be found in BORWEIN & LEWIS [12, p. 156, Exercise 12(b)(ii)]. $\square$

Theorem 2.26 (First order nonsmooth optimality conditions). *Let $\hat{x}$ be a local minimizer of (2.26) and $f, F_i : \mathbb{R}^n \rightarrow \mathbb{R}$ (with $i = 1, \dots, m$) be Lipschitz continuous in a neighborhood of $\hat{x}$ and*

$$J(x) := \{j \in \{1, \dots, m\} : F_j(x) = 0\} ,$$

then there exist $\lambda_0 \geq 0, \lambda_j \geq 0$ for $j \in J(\hat{x})$ with $\lambda_k > 0$ for at least one $k \in \{0\} \cup J(\hat{x})$, so that

$$0 \in \lambda_0 \partial f(\hat{x}) + \sum_{j \in J(\hat{x})} \lambda_j \partial F_j(\hat{x}) .$$

Furthermore, if there exists a direction $d \in \mathbb{R}^n$ that satisfies the (nonsmooth) constraint qualification

$$F_j^\circ(\hat{x}, d) < 0 \quad \text{for all } j \in J(\hat{x}) , \tag{2.34}$$

then we can always set $\lambda_0 = 1$.

Proof. BORWEIN & LEWIS [12, p. 146, Proof of Theorem 6.1.8]. $\square$

Remark 2.27. Since

$$F_i(x) \leq 0 \quad \text{for all } i = 1, \dots, m \iff F(x) \leq 0 ,$$

where

$$F(x) := \max_{i=1, \dots, m} c_i F_i(x)$$

with constants $c_i > 0$, and since F is still locally Lipschitz continuous (cf., e.g., MIFFLIN [75, p. 969, Theorem 6 (a)]), we can always assume $m = 1$ in (2.26). Since we do not take scaling problems of the constraints into account here, we choose $c_i = 1$ for all $i = 1, \dots, m$ and therefore we always consider the nonsmooth optimization problem with a single nonsmooth constraint

$$\begin{aligned} \min & f(x) \\ \text{s.t. } & F(x) \leq 0 , \end{aligned} \tag{2.35}$$

where $f, F : \mathbb{R}^n \rightarrow \mathbb{R}$ are locally Lipschitz continuous, instead of (2.26).

Corollary 2.28. *Let the constraint qualification (2.34) be satisfied for (2.35), then the optimality condition for (2.35) reads as follows: There exists $\lambda \geq 0$ with*

$$\begin{aligned} 0 &\in \partial f(\hat{x}) + \lambda \partial F(\hat{x}) \\ \lambda F(\hat{x}) &= 0 \\ F(\hat{x}) &\leq 0. \end{aligned} \tag{2.36}$$

Proof. Inserting into Theorem 2.26 with $m = 1$. □

Remark 2.29. The algorithms in MIFFLIN [74, 75, 76] (for solving nonlinearly constrained nonsmooth optimization problems) use a fixed point theorem about certain upper semicontinuous point to set mappings by MERRILL [73] as optimality condition which is different to an approach with the optimality conditions in Theorem 2.26 or Corollary 2.28.

2.3.2 Problems of smooth solvers on nonsmooth optimization problems

Already the attempt to solve an unconstrained nonsmooth optimization problem by a smooth solver (e.g., by a line search algorithm or by a trust region method) by just replacing the gradient by a subgradient, fails in general (cf., e.g., ZOWE [111, p. 461-462]):

- If $g \in \partial f(x)$, then the search direction $-g$ does not need to be a direction of descent (contrary to the behavior of the gradient of a differentiable function).
- It can happen that $\{x_k\}$ converges towards a minimizer $\hat{x}$, although the sequence of gradients $\{\nabla f(x_k)\}$ does not converge towards 0 and therefore we cannot identify $\hat{x}$ as a minimizer.
- It can happen that $\{x_k\}$ converges towards a point $\hat{x}$, but $\hat{x}$ is not stationary for f .

The reason for these problems can be characterized as follows: If f is not differentiable at x , the gradient ∇f is discontinuous at x and therefore $\nabla f(x)$ does not give any information about the behavior of ∇f in a neighborhood of x . One approach for dealing with this difficulty is to use information of f at points near x as the so called bundle algorithms do (cf. Subsection 2.3.3).

Not surprisingly, like in smooth optimization, the presence of constraints adds additional complexity, since constructing a descent sequence whose limit satisfies the constraints is (both theoretically and numerically) much more difficult than achieving this aim without the requirement of satisfying any restrictions.

2.3.3 Bundle algorithms

Bundle algorithms are iterative methods for solving nonsmooth optimization problems. They only need to compute one element g of the subdifferential $\partial f(x)$ per iteration, which in practice is easily computable by algorithmic differentiation (cf., e.g., GRIEWANK & CORLISS [30]). For computing the search direction, they collect information about the function (e.g., subgradients) from previous

iterations. This collected information is referred to as “the bundle”. Bundle algorithms work according to the following scheme which reads at iteration k , where we describe the unconstrained case for easiness of exposition (cf., e.g., SCHRAMM [98, p. 13]):

- Compute a subgradient $g_k \in \partial f(x_k)$.
- Compute a search direction d_k by using all information about the function which has been accumulated already.
- By using a line search or a trust region approach, determine the next iteration point x_{k+1} and enlarge/improve the information about the function:
 - If we can find $t > 0$ such that f achieves sufficient descent at the trial point $y_{k+1} := x_k + td_k$, we perform a “serious step”:
 - * Set $x_{k+1} := y_{k+1}$ (i.e. we take the trial point as new iteration point).
 - * Add a new subgradient $g_{k+1} \in \partial f(x_{k+1})$ to the bundle.
 - If we cannot find such a $t > 0$, we perform a “null step”:
 - * Set $x_{k+1} := x_k$ (i.e. we stay at the current iteration point).
 - * Determine $t > 0$ such that adding a new subgradient $g_{k+1} \in \partial f(x_{k+1})$ to the bundle yields a “meaningful extension of the bundle”.

Remark 2.30. As in smooth optimization, convex nonsmooth optimization is much easier than nonconvex nonsmooth optimization as well in theory as in practice because of the following reasons:

- Convex functions only have global minimizers.
- The cutting plane approximation of a convex function always yields an underestimation (cf. (2.31)) which in particular simplifies convergence analysis.

A good introduction to nonsmooth optimization which treats the convex, unconstrained case in great detail is BONNANS et al. [11, p. 106 ff]. Moreover, very detailed standard references for nonsmooth nonconvex optimization are KIWIŁ [49] and MÄKELÄ & NEITTAANMÄKI [71], which both in particular discuss constrained problems extensively.

Now we give a brief overview over a few bundle algorithms. We start this overview with the following bundle algorithms that support nonconvex constraints:

- The multiobjective proximal bundle method for nonconvex nonsmooth optimization (MPBNGC) by MÄKELÄ [70] is a first order method that uses the improvement function

$$h_{x_k}(x) := \max (f(x) - f(x_k), F(x)) . \quad (2.37)$$

for the handling of the constraints. Further details about the proximal bundle method can be found in MÄKELÄ & NEITTAANMÄKI [71].

- The algorithms in MIFFLIN [74, 75, 76] support a nonconvex objective function as well as nonconvex constraints (cf. Remark 2.29).

- NOA by KIWIEL & STACHURSKI [52] is a nonsmooth optimization algorithm that handles non-convex constraints by using a penalty function or an improvement function, while in the special case of convex constraints it offers an alternative treatment by the constraint linearization technique by KIWIEL [50].
- The limited memory bundle algorithm for inequality constrained nondifferentiable optimization by KARMITSA et al. [47] combines LMBM by HAARALA [31] with the feasible directions interior point technique by HERSKOVITS [35], HERSKOVITS & SANTOS [36] for dealing with the constraints. The search direction is determined by solving a linear system.

In addition a few bundle algorithms can only handle convex constraints:

- The bundle trust algorithm by SCHRAMM [98], SCHRAMM & ZOWE [99], which also supports a nonconvex objective function, handles the constraints by using the constraint linearization technique by KIWIEL [50].
- The bundle filter algorithm by FLETCHER & LEYFFER [22] is only applicable to convex optimization problems and it computes the search direction by solving a linear program.
- The bundle-filter method for nonsmooth convex constrained optimization by KARAS et al. [42] is based on the improvement function from (2.37).
- The infeasible bundle method for nonsmooth convex constrained optimization by SAGASTIZÁBAL & SOLODOV [88] is also based on the improvement function from (2.37), but it uses neither a penalty function nor a filter.

Moreover, there are some bundle algorithms that support at most linear constraints:

- The variable metric bundle method PVAR by LUKŠAN & VLČEK [65], VLČEK & LUKŠAN [107] can solve nonsmooth linearly constrained problems with a nonconvex objective function.
- The implementation PBUN of the proximal bundle method by LUKŠAN & VLČEK [62, 66], VLČEK [106] optimizes a nonconvex objective function, where the feasible set is given by linear constraints.
- The proximal bundle method by KIWIEL [51], which is based on a restricted step concept, can handle a nonconvex objective function and linear constraints.
- The focus of the limited memory bundle method LMBM by HAARALA [31], HAARALA et al. [32, 33] is the treatment of large-scale nonsmooth nonconvex unconstrained optimization problems. This is done by combining ideas from the variable metric bundle method LUKŠAN & VLČEK [65], VLČEK & LUKŠAN [107] and limited memory variable metric methods by, e.g., BYRD et al. [16]. Its bound constraint version is presented in KARMITSA & MÄKELÄ [45, 46].

All algorithms mentioned above only use first order information of the objective function and the constraints as input. Nevertheless, there are some very interesting bundle methods, since they are Newton-like methods (at least in some sense) and which only support the handling of linear constraints yet as far as I know (except for putting the objective function and the constraints into a penalty

function with a fixed penalty parameter and then applying the unconstrained algorithm to the penalty function):

- The quasi-Newton bundle-type method for nondifferentiable convex optimization by MIFFLIN et al. [77] generalizes the idea of Quasi-Newton methods to nonsmooth optimization and it converges superlinearly for strongly convex functions (and some additional technical assumptions).
- The bundle-Newton method for nonsmooth unconstrained minimization by LUKŠAN & VLČEK [64] supports a nonconvex objective function, it is based on an SQP-approach, and it is the only method for solving nonsmooth optimization problems that I know which uses Hessian information. Furthermore, its rate of convergence is superlinear for strongly convex, twice times continuously differentiable functions. Moreover, a description of the implementation PNEW of the bundle-Newton method can be found in LUKŠAN & VLČEK [62]).

Remark 2.31. As already described on page 3, we will extend the SQP-approach of the bundle-Newton method to a second order bundle algorithm for nonsmooth, nonconvex inequality constraints by using quadratic information in a very special form (for theoretical details s. Chapter 3 and for details concerning the implementation s. Section 4.3).

2.3.4 Other algorithms for nonsmooth optimization

There exist several other methods for solving nonsmooth optimization problems that are not based on the bundle approach or that are no bundle algorithms in the sense as described in Subsection 2.3.3 (recall from there that bundle algorithms use a concept of serious and null steps, and they compute exactly one subgradient at each iteration). A few representatives of these methods that support at most linear constraints are:

- The algorithm PMIN by LUKŠAN & VLČEK [63], which is based on LUKŠAN [61], solves linearly constrained minimax optimization problems, i.e. the objective function must be maximum of twice times continuously differentiable functions.
- The robust gradient sampling algorithm for nonsmooth nonconvex optimization by BURKE et al. [15] approximates the whole subdifferential at each iteration (cf. BURKE et al. [14]) and does not make null steps. Moreover, its authors recommend to use MOSEK by ANDERSEN et al. [3] for solving the QP that must be solved for the computation of the search direction in this algorithm.
- The MATLAB-code HANSO by OVERTON [85] combines ideas from BFGS algorithms (cf. LEWIS & OVERTON [56]) and from the gradient sampling algorithm by BURKE et al. [15] for solving nonsmooth unconstrained optimization problems.
- The derivative-free bundle method (DFBM) by BAGIROV [4], where “derivate-free” means that no derivate information is used explicitly, can solve linearly constrained nonsmooth problems. The subgradients are approximated by finite differences in this algorithm (cf. BAGIROV [5]). DFBM is an essential part of the programming library for global and non-smooth optimization GANSO by BAGIROV et al. [6].

- The discrete gradient method DGM for nonsmooth nonconvex unconstrained optimization by BAGIROV et al. [8] is a bundle-like method that does not compute subgradients, but approximates them by discrete gradients.
- The quasisecant method QSM for minimizing nonsmooth nonconvex functions by BAGIROV & GANJEHLOU [7] combines ideas both from bundle methods and from the gradient sampling method by BURKE et al. [15].

Furthermore, we want to mention the following solver for nonsmooth convex optimization problems:

- The oracle based optimization engine OBOE by VIAL & SAWHNEY [105] is based on the analytic center cutting plane method by NESTEROV & VIAL [81], which is an interior point framework.

Finally, we list a few algorithms that can also handle nonconvex constraints:

- The robust sequential quadratic programming algorithm extends the gradient sampling algorithm by CURTIS & OVERTON [17] for nonconvex, nonsmooth constrained optimization.
- SolvOpt by KAPPEL & KUNTSEVICH [41] is an implementation of the R-algorithm by SHOR [100]. It handles the constraints by automatically adapting the penalty parameter.
- `ralg` by KROSHKO [53] is another implementation of the R-algorithm by SHOR [100] that is only available in (the interpreted programming language) Python. The constraints are handled by a filter technique.

Remark 2.32. KARMITSA et al. [44] gives a brief, excellent description of the main ideas (including very good readable pseudo code) of many of the unconstrained methods resp. the unconstrained versions of the methods which we mentioned in Subsection 2.3.3 and in this subsection (for further information visit the online decision tree for nonsmooth optimization software by KARMITSA [43]).

Chapter 3

A feasible second order bundle algorithm for nonsmooth, nonconvex optimization problems with inequality constraints

In this chapter we discuss the theoretical basics of our second order bundle algorithm and we give a detailed presentation of the algorithm and the line search. Furthermore, we prove the convergence of the line search and we show the global convergence of the algorithm.

3.1 Derivation of the method

Assumption 3.1. We make the following assumptions in this chapter:

- The functions $f, F : \mathbb{R}^n \rightarrow \mathbb{R}$ are locally Lipschitz continuous.
- $g_j \in \partial f(y_j)$, $\hat{g}_j \in \partial F(y_j)$.
- $G_j \in \partial^2 f(y_j)$, $\hat{G}_j \in \partial^2 F(y_j)$ (cf. (2.28)).

Motivation 3.2. We want to find a local minimizer for the nonsmooth optimization problem (2.35)

$$\begin{aligned} \min_{x \in \mathbb{R}^n} f(x) \\ \text{s.t. } F(x) \leq 0 . \end{aligned}$$

Therefore, to attain this goal, we want to find a point that satisfies the first order optimality conditions (2.36). To achieve this aim, we will propose an extension to the bundle-Newton method for nonsmooth unconstrained minimization by LUKŠAN & VLČEK [64]: If we are in the optimization problem (2.35) at the iteration point $x_k \in \mathbb{R}^n$ (with iteration index k), we want to compute the next trial point (i.e. the search direction) by approximating both the objective function f and the constraint F at x_k by a piecewise quadratic function and then perform a single SQP-step, as defined in Proposition 2.18, to the resulting optimization problem.

Definition 3.3. Let $J_k \subseteq \{1, \dots, k\}$.

1. We define a quadratic approximation of f resp. F in $y_j \in \mathbb{R}^n$ with damping parameter ρ_j resp. $\hat{\rho}_j \in [0, 1]$ for $j \in J_k$ by

$$f_j^\sharp(x) := f(y_j) + g_j^T(x - y_j) + \frac{1}{2}\rho_j(x - y_j)^T G_j(x - y_j) \quad (3.1)$$

$$F_j^\sharp(x) := F(y_j) + \hat{g}_j^T(x - y_j) + \frac{1}{2}\hat{\rho}_j(x - y_j)^T \hat{G}_j(x - y_j) \quad (3.2)$$

and the corresponding gradients by

$$g_j^\sharp(x) := \nabla f_j^\sharp(x)^T \stackrel{(3.1)}{=} g_j + \rho_j G_j(x - y_j) \quad (3.3)$$

$$\hat{g}_j^\sharp(x) := \nabla F_j^\sharp(x)^T \stackrel{(3.2)}{=} \hat{g}_j + \hat{\rho}_j \hat{G}_j(x - y_j) . \quad (3.4)$$

2. We define the piecewise quadratic approximation of f resp. F in $x_k \in \mathbb{R}^n$ by

$$f_k^\square(x) := \max_{j \in J_k} f_j^\sharp(x) \quad (3.5)$$

$$F_k^\square(x) := \max_{j \in J_k} F_j^\sharp(x) . \quad (3.6)$$

Remark 3.4. According to Motivation 3.2 we approximate the objective function f at x_k by $f_k^\square$ from (3.5) and the constraint F at x_k by $F_k^\square$ from (3.6) in the optimization problem (2.35) and then we perform a single SQP-step to the resulting optimization problem

$$\begin{aligned} \min_{x \in \mathbb{R}^n} f_k^\square(x) \\ \text{s.t. } F_k^\square(x) \leq 0 . \end{aligned} \quad (3.7)$$

Proposition 3.5. Define

$$\hat{f}(x, u) := u \quad (3.8)$$

$$\hat{F}_j(x, u) := f_j^\sharp(x) - u \quad (3.9)$$

$$\check{F}_j(x, u) := F_j^\sharp(x) , \quad (3.10)$$

then

$$\nabla \hat{f}(x, u) \begin{pmatrix} d \\ s \end{pmatrix} = s \quad (3.11)$$

$$\hat{F}_j(x, u) + \nabla \hat{F}_j(x, u) \begin{pmatrix} d \\ s \end{pmatrix} = f_j^\sharp(x) - u + \nabla f_j^\sharp(x) d - s \quad (3.12)$$

$$\check{F}_j(x, u) + \nabla \check{F}_j(x, u) \begin{pmatrix} d \\ s \end{pmatrix} = F_j^\sharp(x) + \nabla F_j^\sharp(x) d \quad (3.13)$$

$$\begin{pmatrix} d \\ s \end{pmatrix}^T \nabla_{(x,u),(x,u)}^2 L(x, u, \lambda^{k-1}, \mu^{k-1}) \begin{pmatrix} d \\ s \end{pmatrix} = d^T W^k d , \quad (3.14)$$

where

$$\begin{aligned} L(x, u, \lambda, \mu) &:= \hat{f}(x, u) + \sum_j \lambda_j \hat{F}_j(x, u) + \sum_j \mu_j \check{F}_j(x, u) \\ W^k &:= \sum_{j \in J_{k-1}} \lambda_j^{k-1} \rho_j G_j + \sum_{j \in J_{k-1}} \mu_j^{k-1} \hat{\rho}_j \hat{G}_j \end{aligned} \quad (3.15)$$

and λ_j^{k-1} resp. μ_j^{k-1} denote the Lagrange multipliers with respect to f resp. F at iteration $k-1$ for $j \in J_{k-1}$.

Proof.

- We have

$$\nabla \hat{f}(x, u) \stackrel{(3.8)}{=} (0, 1) \quad (3.16)$$

$$\implies \nabla^2 \hat{f}(x, u) \stackrel{(3.16)}{=} 0. \quad (3.17)$$

Therefore we obtain (3.11)

$$\nabla \hat{f}(x, u) \left(\begin{smallmatrix} d \\ s \end{smallmatrix} \right) \stackrel{(3.16)}{=} s.$$

- We have

$$\nabla \hat{F}_j(x, u) \stackrel{(3.9)}{=} (\nabla_x f_j^\#(x), -1) \quad (3.18)$$

$$\implies \nabla^2 \hat{F}_j(x, u) \stackrel{(3.18), (3.1)}{\downarrow} \begin{pmatrix} \rho_j G_j & 0 \\ 0 & 0 \end{pmatrix}. \quad (3.19)$$

- We have

$$\nabla \check{F}_j(x, u) \stackrel{(3.10)}{=} (\nabla_x F_j^\#(x), 0) \quad (3.20)$$

$$\implies \nabla^2 \check{F}_j(x, u) \stackrel{(3.20), (3.2)}{\downarrow} \begin{pmatrix} \hat{\rho}_j \hat{G}_j & 0 \\ 0 & 0 \end{pmatrix}. \quad (3.21)$$

- (3.12) holds because of

$$\begin{aligned} \hat{F}_j(x, u) + \nabla \hat{F}_j(x, u) \left(\begin{smallmatrix} d \\ s \end{smallmatrix} \right) &\stackrel{(3.9), (3.18)}{\downarrow} f_j^\#(x) - u + (\nabla_x f_j^\#(x), -1) \left(\begin{smallmatrix} d \\ s \end{smallmatrix} \right) \\ &= f_j^\#(x) - u + \nabla f_j^\#(x) d - s. \end{aligned}$$

- (3.13) holds because of

$$\begin{aligned} \check{F}_j(x, u) + \nabla \check{F}_j(x, u) \left(\begin{smallmatrix} d \\ s \end{smallmatrix} \right) &\stackrel{(3.10), (3.20)}{\downarrow} F_j^\#(x) + (\nabla_x F_j^\#(x), 0) \left(\begin{smallmatrix} d \\ s \end{smallmatrix} \right) \\ &= F_j^\#(x) + \nabla F_j^\#(x) d. \end{aligned}$$

- (3.14) holds because we have at step $k-1$

$$\begin{aligned} \nabla_{(x,u),(x,u)}^2 L(x, u, \lambda^{k-1}, \mu^{k-1}) &\stackrel{(2.17)}{=} \nabla^2 \hat{f}(x, u) + \sum_{j \in J_{k-1}} \nabla^2 \hat{F}_j(x, u) \lambda_j + \sum_{j \in J_{k-1}} \nabla^2 \check{F}_j(x, u) \mu_j \\ &\stackrel{(3.17), (3.19), (3.21)}{\downarrow} \sum_{j \in J_{k-1}} \begin{pmatrix} \rho_j G_j & 0 \\ 0 & 0 \end{pmatrix} \lambda_j + \sum_{j \in J_{k-1}} \begin{pmatrix} \hat{\rho}_j \hat{G}_j & 0 \\ 0 & 0 \end{pmatrix} \mu_j \\ &\stackrel{(3.15)}{=} \begin{pmatrix} W^k & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned} \quad \square$$

Proposition 3.6. The SQP-step $(d, \hat{v}) \in \mathbb{R}^{n+1}$ for (3.7) is given by the solution of the QP

$$\begin{aligned} & f(x_k) + \min_{d, \hat{v}} \hat{v} + \frac{1}{2} d^T W^k d \\ & \text{s.t.} \quad - (f(x_k) - f_j^k) + d^T g_j^k \leq \hat{v} \quad \text{for all } j \in J_k \\ & \quad F(x_k) - (F(x_k) - F_j^k) + d^T \hat{g}_j^k \leq 0 \quad \text{for all } j \in J_k, \end{aligned} \quad (3.22)$$

where

$$f_j^k := f_j^\#(x_k) \quad (3.23)$$

$$g_j^k := g_j^\#(x_k) \stackrel{(3.3)}{=} g_j + \rho_j G_j(x_k - y_j) \quad (3.24)$$

$$F_j^k := F_j^\#(x_k) \quad (3.25)$$

$$\hat{g}_j^k := \hat{g}_j^\#(x_k) \stackrel{(3.4)}{=} \hat{g}_j + \hat{\rho}_j \hat{G}_j(x_k - y_j). \quad (3.26)$$

Proof. Since

$$\begin{aligned} f_k^\square(x) = u \quad \wedge \quad F_k^\square(x) \leq 0 & \stackrel{(3.5), (3.6)}{\iff} \max_{j \in J_k} f_j^\#(x) = u \quad \wedge \quad \max_{j \in J_k} F_j^\#(x) \leq 0 \\ & \stackrel{(3.9), (3.10)}{\iff} (\hat{F}_j(x, u) \leq 0 \quad \wedge \quad \check{F}_j(x, u) \leq 0) \quad \text{for all } j \in J_k, \end{aligned} \quad (3.27)$$

we can rewrite (3.7) as a smooth optimization problem

$$\begin{aligned} & \min_x f_k^\square(x) \\ & \text{s.t.} \quad F_k^\square(x) \leq 0 \\ & = \min_{x, u} u \\ & \text{s.t.} \quad f_k^\square(x) = u \\ & \quad F_k^\square(x) \leq 0 \\ & \stackrel{(3.27)}{=} \min_{x, u} \hat{f}(x, u) \\ & \text{s.t.} \quad \hat{F}_j(x, u) \leq 0 \quad \text{for all } j \in J_k \\ & \quad \check{F}_j(x, u) \leq 0 \quad \text{for all } j \in J_k. \end{aligned}$$

If we are at the iteration point $(x_k, u_k) \in \mathbb{R}^n \times \mathbb{R}$ in this optimization problem, then, according to (2.23) the SQP-step for this problem is given by the solution of

$$\begin{aligned} & \hat{f}(x_k, u_k) + \min_{d, \hat{v}} \nabla \hat{f}(x_k, u_k) \left(\frac{d}{\hat{v}} \right) + \frac{1}{2} \left(\frac{d}{\hat{v}} \right)^T \nabla_{(x, u), (x, u)}^2 L(x_k, u_k, \lambda^{k-1}, \mu^{k-1}) \left(\frac{d}{\hat{v}} \right) \\ & \text{s.t.} \quad \hat{F}_j(x_k, u_k) + \nabla \hat{F}_j(x_k, u_k) \left(\frac{d}{\hat{v}} \right) \leq 0 \quad \text{for all } j \in J_k \\ & \quad \check{F}_j(x_k, u_k) + \nabla \check{F}_j(x_k, u_k) \left(\frac{d}{\hat{v}} \right) \leq 0 \quad \text{for all } j \in J_k \\ & \stackrel{(3.8), (3.11), (3.14), (3.12), (3.13)}{\stackrel{\downarrow}{=}} u_k + \min_{d, \hat{v}} \hat{v} + \frac{1}{2} d^T W^k d \\ & \text{s.t.} \quad - (u_k - f_j^\#(x_k)) + \nabla f_j^\#(x_k) d \leq \hat{v} \quad \text{for all } j \in J_k \\ & \quad F(x_k) - (F(x_k) - F_j^\#(x_k)) + \nabla F_j^\#(x_k) d \leq 0 \quad \text{for all } j \in J_k. \end{aligned}$$

Since we are at the iterate x_k , we choose $u_k := f(x_k)$, which is the only meaningful choice for u_k . Now, using (3.23), (3.24), (3.25) and (3.26) yields the desired result. $\square$

Since $f_j^\#$ resp. $F_j^\#$ are only global underestimators for convex f resp. F and $\rho_j = \hat{\rho}_j = 0$ and since $f_k^\square$ resp. $F_k^\square$ approximate f resp. F only well for trial points close to x_k , we decrease the activity of non local information (e.g., non local subgradients) by the following definition.

Definition 3.7. We define the localized approximation errors of f resp. F by

$$\alpha_j^k := \max(|f(x_k) - f_j^k|, \gamma_1(s_j^k)^{\omega_1}) \quad (3.28)$$

$$A_j^k := \max(|F(x_k) - F_j^k|, \gamma_2(\hat{s}_j^k)^{\omega_2}) , \quad (3.29)$$

where

$$s_j^k := |y_j - x_j| + \sum_{i=j}^{k-1} |x_{i+1} - x_i| \quad (3.30)$$

$$\hat{s}_j^k := |y_j - x_j| + \sum_{i=j}^{k-1} |x_{i+1} - x_i| \quad (3.31)$$

denotes a locality measure for $j = 1, \dots, k$ with fixed parameters $\gamma_i > 0$ and $\omega_i \geq 1$ for $i = 1, 2$.

Remark 3.8.

- The locality measure s_j^k describes the distance of the path $\overrightarrow{y_j x_j \dots x_k}$.
- The only intention of the distinction between s_j^k and $\hat{s}_j^k$ is to make the investigation and implementation of future development plans (as, e.g., distinguishing between a bundle for the objective function and a bundle for the constraint) easier. Therefore, $\hat{s}_j^k$ is replaced by s_j^k in FENDL & SCHICHL [20, 21].

Proposition 3.9. The locality measure s_j^k has the following properties

$$s_j^k + |x_{k+1} - x_k| = s_j^{k+1} \quad \text{for all } j = 1, \dots, k \quad (3.32)$$

$$s_j^k \geq |y_j - x_k| \quad \text{for all } j = 1, \dots, k \quad (3.33)$$

$$s_k^k = |y_k - x_k| .$$

Proof. Straightforward calculations. $\square$

Like the bundle-Newton method by LUKŠAN & VLČEK [64], our algorithm uses a convex search direction problem and therefore we modify (3.22) in the following sense.

Proposition 3.10. If we generalize (3.22) by

- using the localized approximation errors (3.28) and (3.29)
- and replacing W^k by a positive definite modification $\bar{W}_p^k$ (e.g., the Gill-Murray factorization by GILL & MURRAY [25]),

then the generalized version of (3.22) reads

$$\begin{aligned} & f(x_k) + \min_{d, \hat{v}} \hat{v} + \frac{1}{2} d^T \bar{W}_p^k d \\ & \text{s.t.} \quad -\alpha_j^k + d^T g_j^k \leq \hat{v} \quad \text{for all } j \in J_k \\ & \quad \quad F(x_k) - A_j^k + d^T \hat{g}_j^k \leq 0 \quad \text{for all } j \in J_k . \end{aligned} \quad (3.34)$$

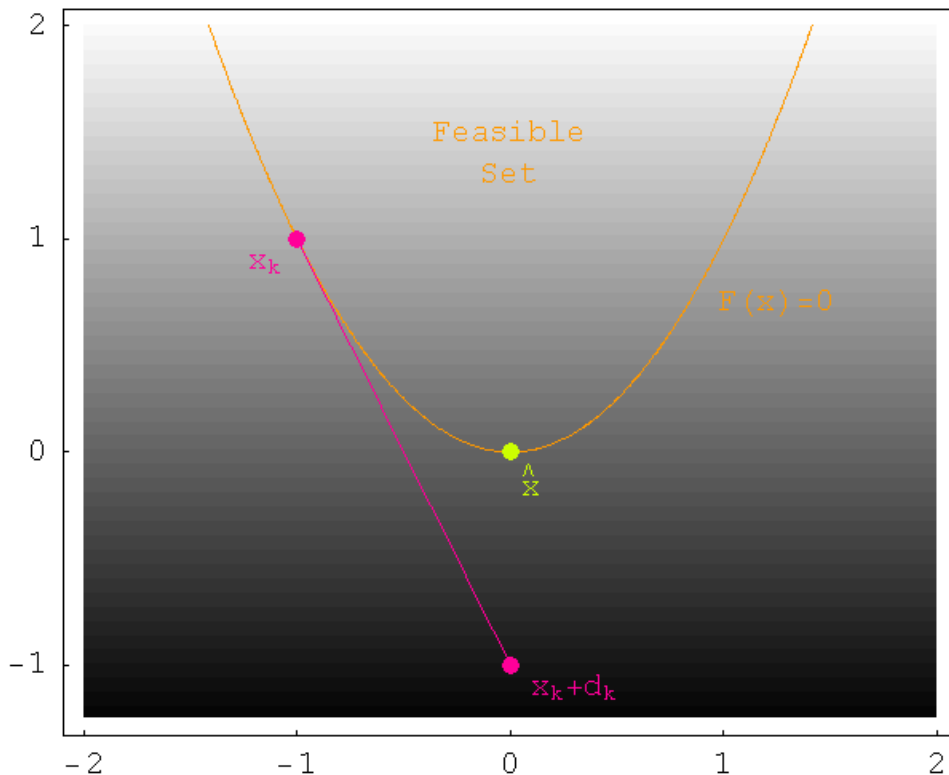
Proof. Replace $f(x_k) - f_j^k$ by α_j^k , $F(x_k) - F_j^k$ by A_j^k and W^k by $\bar{W}_p^k$ in (3.22). $\square$

Remark 3.11. The standard SQP approach for smooth optimization problems suffers from the Maratos effect MARATOS [72], which, in general, prevents infeasible SQP-methods from getting a descent in the merit function and feasible SQP-methods from finding (good) feasible points (cf. TITS [103, p. 1003] and Example 3.12). Some well known techniques for avoiding the Maratos effect are replacing the merit function by an augmented Lagrangian, using second order corrections, using a watchdog technique (which is a non-monotone line search) (cf., e.g., NOCEDAL & WRIGHT [84, p. 440, 15.5 The Maratos effect]), or a quadratic approximation of the constraints (cf. SOLODOV [101]). We will choose the quadratic constraint approximation approach to avoid the Maratos effect, which makes the search direction problem slightly more difficult to solve than a QP, but, as we will see, still guarantees strong duality which is necessary for proving convergence of our bundle method.

Example 3.12. Consider the optimization problem (2.35) with $f, F : \mathbb{R}^2 \rightarrow \mathbb{R}$

$$\begin{aligned} f(x) &:= x_2 \\ F(x) &:= x_1^2 - x_2 . \end{aligned}$$

Then this problem has the (global) minimizer $\hat{x} = 0$. Furthermore, it is smooth and consequently its SQP-direction, which is obtained by solving the QP (2.23), at the iteration $k = 0$ at the iteration point $x_k := (-1, 1 + 10^{-8})$ and $\lambda_k := 1$, which implies that x_k is close to the boundary, is given by $d_k = (1, -2)$, which we illustrate in the following figure (including the contour lines of f):



Since we have for $t \in [0, 1]$

$$F(x_k + td_k) \leq 0 \iff t \leq 10^{-4} ,$$

a feasible SQP-method can only make a tiny step towards the solution $\hat{x}$ on the standard SQP-direction in this example at the iteration $k = 0$, and similar observations can be made for any other point x_k with $k \neq 0$ that is close to the boundary (Note that the objective function f has no impact on the Hessian of the Lagrangian in the QP (2.23) in this example).

Remark 3.11 leads to the following idea.

Idea 3.13. Let $\bar{G}_j^k \in \mathbb{R}_{\text{sym}}^{n \times n}$ be positive definite (e.g., positive definite modifications of $\hat{G}_j \in \partial^2 F(y_j)$). Then we can try to determine the search direction by solving the convex QCQP

$$\begin{aligned} f(x_k) + \min_{d, \hat{v}} \hat{v} + \frac{1}{2} d^T \bar{W}_p^k d \\ \text{s.t. } -\alpha_j^k + d^T g_j^k \leq \hat{v} & \quad \text{for all } j \in J_k \\ F(x_k) - A_j^k + d^T \hat{g}_j^k + \frac{1}{2} d^T \bar{G}_j^k d \leq 0 & \quad \text{for all } j \in J_k \end{aligned} \quad (3.35)$$

instead of the QP (3.34), i.e. instead of just demanding that the first order approximations of the constraint are feasible, we demand that the first order approximations of the constraint must be the more feasible, the more we move away from x_k .

Remark 3.14. If $F(x_k) \leq 0$, ((3.34) as well as) (3.35) is always feasible and therefore we do not have to deal with infeasible search direction problems as they occur in infeasible SQP-methods (cf. Remark 2.19). Nevertheless, we have to demand $F(x_k) < 0$, since otherwise it can happen that $d_k = 0$ is the only feasible point and therefore the solution of (3.35), but x_k is not stationary for (2.35) as Example 3.15 shows. This is similar to difficulties arising in smooth problems at saddle points of the constraints.

Example 3.15. We consider the optimization problem (2.35) with

$$\begin{aligned} f(x) &:= x_2 \\ F(x) &:= \max \left(\min(F_1(x), F_2(x)), F_3(x) \right) \\ &= \begin{cases} \min(F_1(x), F_2(x)) & \min(F_1(x), F_2(x)) \geq F_3(x) \\ F_3(x) & \min(F_1(x), F_2(x)) < F_3(x) \end{cases} \\ &= \begin{cases} F_1(x) & \min(F_1(x), F_2(x)) \geq F_3(x) \wedge F_1(x) \leq F_2(x) \\ F_2(x) & \min(F_1(x), F_2(x)) \geq F_3(x) \wedge F_1(x) > F_2(x) \\ F_3(x) & \min(F_1(x), F_2(x)) < F_3(x) , \end{cases} \end{aligned} \quad (3.36)$$

where

$$\begin{aligned} F_1(x) &:= x_1^2 + x_2^2 \\ F_2(x) &:= -x_1 + x_2^2 \\ F_3(x) &:= x_1 - 2 , \end{aligned} \quad (3.37)$$

and we assume that we are at the iteration point $x_k := 0$.

- Since

$$\hat{F}(x) := \max(F_2(x), F_3(x)) \quad (3.38)$$

is convex, and since an easy examination yields that

$$F(x) \leq 0 \stackrel{(3.36), (3.38)}{\iff} \hat{F}(x) \leq 0, \quad (3.39)$$

the feasible set of our optimization problem (2.35) is convex. Therefore, the linearity of f implies that our optimization problem has the unique minimizer $\hat{x} = (2, -\sqrt{2})$. These facts are summarized in Figure 3.1 (including the contour lines of f).

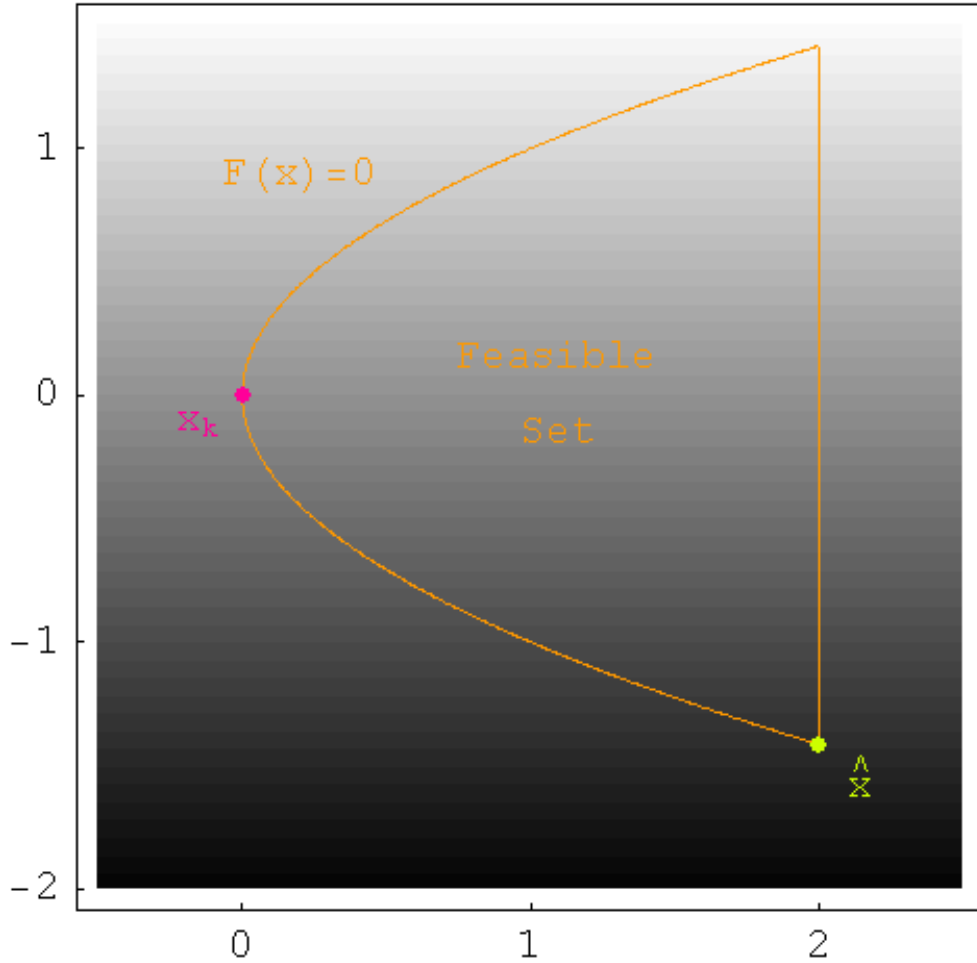


FIGURE 3.1: Illustration of the optimization problem

- Due to (3.36), the quadratic approximation of F with respect to x_k in the QCQP (3.35) reads

$$\begin{aligned} F_1(x_k + d) &\leq 0 \\ \stackrel{(3.37)}{\iff} d_1^2 + d_2^2 &\leq 0 \\ \iff d &= 0, \end{aligned}$$

i.e. $d = 0$ is the only feasible point for the QCQP (3.35) and therefore its solution, although $x_k = 0$ is not a stationary point for our optimization problem (for this consider f), resp. much less a minimizer (since $\hat{x}$ is the unique minimizer of our optimization problem).

As it can be seen, e.g., from considering the restriction of F to $x_2 = 0$ in Figure 3.2, the reason for the occurrence of $d = 0$ at x_k is the nonconvexity of F (which is a result of the presence of the min-function in F), although the feasible set, which is colored orange like in Figure 3.1, is convex.

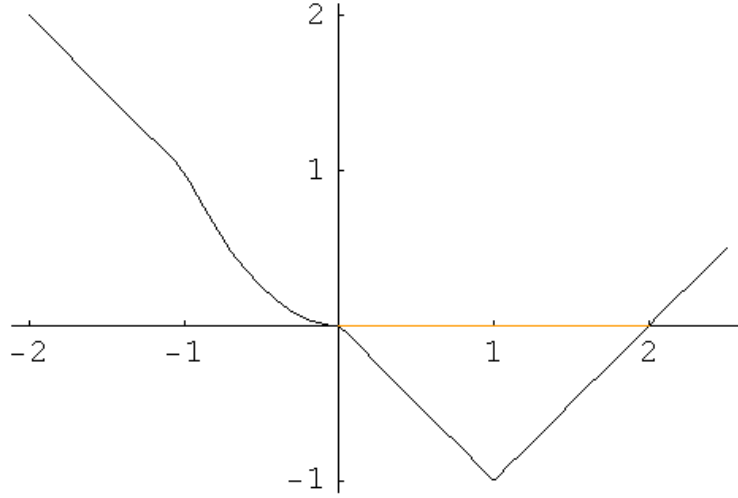


FIGURE 3.2: Sectional drawing of $F(x_1, 0)$

- Notice that if we substitute F from (3.36) by $\hat{F}$ from (3.38) in the constraint of our optimization problem, which yields the same feasible set due to (3.39), the difficulty which we described above does not occur.

3.2 The dual search direction problem

In this section we investigate the dual problem of the QCQP (3.35) which plays an important role for proving the global convergence of the method (cf. Section 3.6).

Theorem 3.16. *We consider the optimization problem*

$$\begin{aligned} & \min_{x \in \mathbb{R}^n} f_0(x) \\ & \text{s.t. } Ax \leq b \\ & \quad f_i(x) \leq 0 \quad \text{for all } i = 1, \dots, q \end{aligned} \tag{3.40}$$

with $f_0 : \mathbb{R}^n \rightarrow \mathbb{R}$ convex, $A \in \mathbb{R}^{p \times n}$, $b \in \mathbb{R}^p$, and $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ convex as well as not affine for all $i = 1, \dots, q$. If the Slater constraint qualification is satisfied, i.e. there exists a point x in the relative interior of $\mathcal{D}$, where

$$\mathcal{D} := \bigcap_{j=0}^q \text{dom} f_j$$

and $\text{dom} f_j$ denotes the domain of f_j , with

$$\begin{aligned} & Ax \leq b \\ & f_i(x) < 0 \quad \text{for all } i = 1, \dots, q, \end{aligned}$$

then strong duality holds (i.e. the optimal function values of the primal problem (3.40) and the corresponding dual problem

$$\begin{aligned} \max_{\lambda, \mu} \quad & f_0(x) + \lambda^T(Ax - b) + \sum_{i=1}^q \mu_i f_i(x) \\ \text{s.t.} \quad & \nabla f_0(x) + \lambda^T A + \sum_{i=1}^q \mu_i \nabla f_i(x) = 0 \\ & \lambda \geq 0, \mu \geq 0 \end{aligned}$$

(cf. (2.12)) coincide), and the (common) optimal function value is attained and finite.

Proof. BOYD & VANDENBERGHE [13, p. 234, Section 5.3.2]. □

Proposition 3.17. *If $(\hat{v}_k, d_k)$ solves the QCQP (3.35), then there exists $(\lambda^k, \mu^k) \in \mathbb{R}_{\geq 0}^{2|J_k|}$ with*

$$\begin{aligned} (\bar{W}_p^k + \sum_{j \in J_k} \bar{G}_j^k \mu_j^k) d_k + \sum_{j \in J_k} \lambda_j^k g_j^k + \sum_{j \in J_k} \mu_j^k \hat{g}_j^k &= 0 \\ \sum_{j \in J_k} \lambda_j^k &= 1 \\ -\alpha_j^k + d_k^T g_j^k - \hat{v}_k &= 0 \quad \wedge \quad \lambda_j^k \geq 0 \\ -\alpha_j^k + d_k^T g_j^k - \hat{v}_k &< 0 \quad \wedge \quad \lambda_j^k = 0 \\ F(x_k) - A_j^k + d_k^T \hat{g}_j^k + \frac{1}{2} d_k^T \bar{G}_j^k d_k &= 0 \quad \wedge \quad \mu_j^k \geq 0 \\ F(x_k) - A_j^k + d_k^T \hat{g}_j^k + \frac{1}{2} d_k^T \bar{G}_j^k d_k &< 0 \quad \wedge \quad \mu_j^k = 0. \end{aligned} \tag{3.41}$$

Proof. Insert (3.35) into (2.10). □

Proposition 3.18. *The dual problem of the QCQP (3.35) is given by*

$$\begin{aligned} f(x_k) - \min_{\lambda, \mu} \frac{1}{2} \left| H_k(\mu) \left(\sum_{j \in J_k} \lambda_j g_j^k + \mu_j \hat{g}_j^k \right) \right|^2 + \sum_{j \in J_k} \lambda_j \alpha_j^k + \mu_j A_j^k - \left(\sum_{j \in J_k} \mu_j \right) F(x_k) \\ \text{s.t.} \quad \lambda_j \geq 0, \mu_j \geq 0 \quad \text{for all } j \in J_k, \quad \sum_{j \in J_k} \lambda_j = 1, \end{aligned} \tag{3.42}$$

where

$$H_k(\mu) := (\bar{W}_p^k + \sum_{j \in J_k} \mu_j \bar{G}_j^k)^{-\frac{1}{2}}. \tag{3.43}$$

If $F(x_k) < 0$, then the duality gap is zero, and, furthermore, if we denote the minimizer of the dual problem ((3.42)) by (λ^k, μ^k) , then the minimizer $(d_k, \hat{v}_k)$ of the primal QCQP (3.35) satisfies

$$\begin{aligned} d_k &= -(\bar{W}_p^k + \sum_{j \in J_k} \mu_j^k \bar{G}_j^k)^{-1} \left(\sum_{j \in J_k} \lambda_j^k g_j^k + \sum_{j \in J_k} \mu_j^k \hat{g}_j^k \right) \\ \hat{v}_k &= \left(\sum_{j \in J_k} \lambda_j^k g_j^k \right)^T d_k - \sum_{j \in J_k} \lambda_j^k \alpha_j^k \\ &= -d_k^T \bar{W}_p^k d_k - \frac{1}{2} d_k^T \left(\sum_{j \in J_k} \mu_j^k \bar{G}_j^k \right) d_k - \sum_{j \in J_k} \lambda_j^k \alpha_j^k - \sum_{j \in J_k} \mu_j^k A_j^k - \left(\sum_{j \in J_k} \mu_j^k \right) (-F(x_k)) \\ &\leq 0. \end{aligned}$$

Proof. For $u : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ and $U : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^{2m}$, where $m := |J_k|$, defined by

$$\begin{aligned} u(d, \hat{v}) &:= \hat{v} + \frac{1}{2} d^T \bar{W}_p^k d \\ U(d, \hat{v}) &:= \left(-\alpha_1^k + d^T g_1^k - \hat{v}, \dots, -\alpha_m^k + d^T g_m^k - \hat{v}, F(x_k) - A_1^k + d^T \hat{g}_1^k + \frac{1}{2} d^T \bar{G}_1^k d, \dots, F(x_k) - A_m^k + d^T \hat{g}_m^k + \frac{1}{2} d^T \bar{G}_m^k d \right)^T \end{aligned} \quad (3.44)$$

we obtain

$$\begin{aligned} \nabla u(d, \hat{v}) &\stackrel{(3.44)}{=} (d^T \bar{W}_p^k, 1) \in \mathbb{R}^{1 \times (n+1)} \\ \nabla U(d, \hat{v}) &\stackrel{(3.44)}{=} \begin{pmatrix} g_1^k & \dots & g_m^k & \hat{g}_1^k + \bar{G}_1^k d & \dots & \hat{g}_m^k + \bar{G}_m^k d \\ -1 & \dots & -1 & 0 & \dots & 0 \end{pmatrix}^T \in \mathbb{R}^{(m+m) \times (n+1)}. \end{aligned} \quad (3.45)$$

Therefore, the equality constraint (containing the gradient of the Lagrangian) of the dual problem of the QCQP (3.35) reads due to (2.12) and (2.11)

$$0 = \nabla u(d, \hat{v}) + \left(\frac{\lambda}{\mu} \right)^T \nabla U(d, \hat{v}) \stackrel{(3.45)}{=} \left(d^T (\bar{W}_p^k + \sum_{j=1}^m \mu_j \bar{G}_j^k) + \sum_{j=1}^m \lambda_j g_j^k + \sum_{j=1}^m \mu_j \hat{g}_j^k, 1 - \sum_{j=1}^m \lambda_j \right) \quad (3.46)$$

and the first n components of (3.46) yield due to the symmetry of $\bar{W}_p^k$ and $\bar{G}_j^k$

$$d = -(\bar{W}_p^k + \sum_{j=1}^m \mu_j \bar{G}_j^k)^{-1} \left(\sum_{j=1}^m \lambda_j g_j^k + \sum_{j=1}^m \mu_j \hat{g}_j^k \right), \quad (3.47)$$

which implies

$$\begin{aligned} d^T (\bar{W}_p^k + \sum_{j=1}^m \mu_j \bar{G}_j^k) d &\stackrel{(3.46)}{=} - \left(\sum_{j=1}^m \lambda_j g_j^k + \sum_{j=1}^m \mu_j \hat{g}_j^k \right)^T d \\ &\stackrel{(3.47)}{=} |(\bar{W}_p^k + \sum_{j=1}^m \mu_j \bar{G}_j^k)^{-\frac{1}{2}} \left(\sum_{j=1}^m \lambda_j g_j^k + \sum_{j=1}^m \mu_j \hat{g}_j^k \right)|^2 \\ &\stackrel{(3.43)}{=} |H_k(\mu) \left(\sum_{j=1}^m \lambda_j g_j^k + \sum_{j=1}^m \mu_j \hat{g}_j^k \right)|^2. \end{aligned} \quad (3.48)$$

Consequently, we obtain for the objective function of the dual problem of the QCQP (3.35) (cf. (2.12) and (2.11))

$$\begin{aligned} u(d, \hat{v}) + \left(\frac{\lambda}{\mu} \right)^T U(d, \hat{v}) &\stackrel{(3.44)}{=} \hat{v} + \frac{1}{2} d^T \bar{W}_p^k d + \sum_{j=1}^m \lambda_j (-\alpha_j^k + d^T g_j^k - \hat{v}) + \sum_{j=1}^m \mu_j (F(x_k) - A_j^k + d^T \hat{g}_j^k + \frac{1}{2} d^T \bar{G}_j^k d) \\ &\stackrel{(3.46)}{=} \frac{1}{2} d^T \bar{W}_p^k d + \sum_{j=1}^m \lambda_j (-\alpha_j^k + d^T g_j^k) + \sum_{j=1}^m \mu_j (F(x_k) - A_j^k + d^T \hat{g}_j^k + \frac{1}{2} d^T \bar{G}_j^k d) \\ &= -\frac{1}{2} d^T (\bar{W}_p^k + \sum_{j=1}^m \mu_j \bar{G}_j^k) d + d^T \left((\bar{W}_p^k + \sum_{j=1}^m \mu_j \bar{G}_j^k) d + \sum_{j=1}^m \lambda_j g_j^k + \sum_{j=1}^m \mu_j \hat{g}_j^k \right) \\ &\quad - \sum_{j=1}^m \lambda_j \alpha_j^k - \sum_{j=1}^m \mu_j (A_j^k - F(x_k)) \\ &\stackrel{(3.46)}{=} -\frac{1}{2} d^T (\bar{W}_p^k + \sum_{j=1}^m \mu_j \bar{G}_j^k) d - \sum_{j=1}^m \lambda_j \alpha_j^k - \sum_{j=1}^m \mu_j (A_j^k - F(x_k)) \end{aligned} \quad (3.49)$$

$$\stackrel{(3.48)}{=} -\frac{1}{2} |H_k(\mu) \left(\sum_{j=1}^m \lambda_j g_j^k + \sum_{j=1}^m \mu_j \hat{g}_j^k \right)|^2 - \sum_{j=1}^m \lambda_j \alpha_j^k - \sum_{j=1}^m \mu_j A_j^k - \left(\sum_{j=1}^k \mu_j \right) (-F(x_k)). \quad (3.50)$$

As a result, the dual problem reads

$$\begin{aligned}
 & f(x_k) + \max_{\lambda, \mu} u(d, \hat{v}) + \left(\frac{\lambda}{\mu}\right)^T U(d, \hat{v}) \\
 & \text{s.t. } \nabla u(d, \hat{v}) + \left(\frac{\lambda}{\mu}\right)^T \nabla U(d, \hat{v}) = 0, \lambda \geq 0, \mu \geq 0 \\
 & \stackrel{(2.3), (3.50), (3.46), (3.47)}{=} f(x_k) - \min_{\lambda, \mu} \frac{1}{2} |H_k(\mu) \left(\sum_{j=1}^m \lambda_j g_j^k + \sum_{j=1}^m \mu_j \hat{g}_j^k \right)|^2 + \sum_{j=1}^m \lambda_j \alpha_j^k + \sum_{j=1}^m \mu_j A_j^k + \left(\sum_{j=1}^m \mu_j \right) (-F(x_k)) \\
 & \text{s.t. } \sum_{j=1}^m \lambda_j = 1, \lambda \geq 0, \mu \geq 0.
 \end{aligned}$$

Now let $(d_k, \hat{v}_k)$ be the primal solution and (λ^k, μ^k) the dual solution. Since the QCQP (3.35) is convex and (because of the assumption $F(x_k) < 0$) strictly feasible, the optimal primal and dual objective function values coincide according to Theorem 3.16

$$\begin{aligned}
 & f(x_k) - \left(\frac{1}{2} |H_k(\mu^k) \left(\sum_{j=1}^m \lambda_j^k g_j^k + \sum_{j=1}^m \mu_j^k \hat{g}_j^k \right)|^2 + \sum_{j=1}^m \lambda_j^k \alpha_j^k + \sum_{j=1}^m \mu_j^k A_j^k + \left(\sum_{j=1}^m \mu_j^k \right) (-F(x_k)) \right) \\
 & \stackrel{(3.49)}{=} f(x_k) - \frac{1}{2} d_k^T (\bar{W}_p^k + \sum_{j=1}^m \mu_j^k \bar{G}_j^k) d_k - \sum_{j=1}^m \lambda_j^k \alpha_j^k - \sum_{j=1}^m \mu_j^k (A_j^k - F(x_k)) \\
 & \stackrel{\text{Theorem 3.16}}{=} f(x_k) + \hat{v}_k + \frac{1}{2} d_k^T \bar{W}_p^k d_k,
 \end{aligned}$$

which is (by isolating $\hat{v}_k$ in the two last equations of the above chain of equations) equivalent to

$$\hat{v}_k = -d_k^T \bar{W}_p^k d_k - \frac{1}{2} d_k^T \left(\sum_{j=1}^m \mu_j^k \bar{G}_j^k \right) d_k - \sum_{j=1}^m \lambda_j^k \alpha_j^k - \sum_{j=1}^m \mu_j^k (A_j^k - F(x_k)). \quad (3.51)$$

From (3.51) we obtain the following two representations of $\hat{v}_k$

$$\hat{v}_k \stackrel{(3.51)}{=} -d_k^T \bar{W}_p^k d_k - \frac{1}{2} d_k^T \left(\sum_{j=1}^m \mu_j^k \bar{G}_j^k \right) d_k - \sum_{j=1}^m \lambda_j^k \alpha_j^k - \sum_{j=1}^m \mu_j^k A_j^k - \left(\sum_{j=1}^m \mu_j^k \right) (-F(x_k)) \stackrel{F(x_k) \leq 0}{\leq} 0$$

and

$$\begin{aligned}
 & \hat{v}_k \stackrel{(3.51)}{=} -d_k^T \bar{W}_p^k d_k - \frac{1}{2} d_k^T \left(\sum_{j=1}^m \mu_j^k \bar{G}_j^k \right) d_k - \sum_{j=1}^m \lambda_j^k \alpha_j^k - \sum_{j=1}^m \mu_j^k (A_j^k - F(x_k)) \\
 & = -\left(d_k^T (\bar{W}_p^k + \sum_{j=1}^m \mu_j^k \bar{G}_j^k) \right) d_k - \sum_{j=1}^m \lambda_j^k \alpha_j^k + \sum_{j=1}^m \mu_j^k (F(x_k) - A_j^k + \frac{1}{2} d_k^T \bar{G}_j^k d_k) \\
 & \stackrel{(3.46)}{=} \left(\sum_{j=1}^m \lambda_j^k g_j^{kT} + \sum_{j=1}^m \mu_j^k \hat{g}_j^{kT} \right) d_k - \sum_{j=1}^m \lambda_j^k \alpha_j^k + \sum_{j=1}^m \mu_j^k (F(x_k) - A_j^k + \frac{1}{2} d_k^T \bar{G}_j^k d_k) \\
 & = \left(\sum_{j=1}^m \lambda_j^k g_j^k \right)^T d_k - \sum_{j=1}^m \lambda_j^k \alpha_j^k + \sum_{j=1}^m \mu_j^k (F(x_k) - A_j^k + \hat{g}_j^{kT} d_k + \frac{1}{2} d_k^T \bar{G}_j^k d_k) \\
 & \stackrel{(3.41)}{=} \left(\sum_{j=1}^m \lambda_j^k g_j^k \right)^T d_k - \sum_{j=1}^m \lambda_j^k \alpha_j^k. \quad \square
 \end{aligned}$$

Remark 3.19. Analogous to Idea 3.13, we could have added positive definite terms to the first order approximations of the objective function in the QCQP (3.35) which would result in additional convex combinations of the corresponding matrices in the dual problem (3.42) (cf. FENDL & SCHICHL [20]). Nevertheless, we do not make use of this as our intention in Idea 3.13 is to obtain feasible iteration points.

3.3 Presentation of the algorithm

The method described in Algorithm 3.20 works according to the following scheme: After choosing a strictly feasible starting point $x_1 \in \mathbb{R}^n$ and setting up a few positive definite matrices, we compute the localized approximation errors. Then we solve a convex QCQP to determine the search direction, where the quadratic constraints of the QCQP serve to obtain preferably feasible points that yield a good descent. After computing the aggregated data and the predicted descent as well as testing the termination criterion, we perform a line search (s. Algorithm 3.25) on the ray given by the search direction. This yields a trial point y_{k+1} that has the following property: Either y_{k+1} is strictly feasible and the objective function achieves sufficient descent (serious step) or y_{k+1} is strictly feasible and the model of the objective function changes sufficiently (null step with respect to the objective function) or y_{k+1} is not strictly feasible and the model of the constraint changes sufficiently (null step with respect to the constraint). Afterwards we update the iteration point x_{k+1} and the information stored in the bundle. Now, we repeat this procedure until the termination criterion is satisfied.

Algorithm 3.20.

0. *Initialization:* Choose the following parameters, which will not be changed during the algorithm:

TABLE 3.1: Initial parameters

General	Default	Description
$x_1 \in \mathbb{R}^n$		Strictly feasible initial point
$y_1 = x_1$		Initial trial point
$\varepsilon \geq 0$		Final optimality tolerance
$M \geq 2$	$M = n + 3$	Maximal bundle dimension
$t_0 \in (0, 1)$	$t_0 = 0.001$	Initial lower bound for step size of serious step in line search
$\hat{t}_0 \in (0, 1)$	$\hat{t}_0 = 0.001$	Scaling parameter for t_0
$m_L \in (0, \frac{1}{2})$	$m_L = 0.01$	Descent parameter for serious step in line search
$m_R \in (m_L, 1)$	$m_R = 0.5$	Parameter for change of model of objective function for short serious and null steps in line search
$m_F \in (0, 1)$	$m_F = 0.01$	Parameter for change of model of constraint for short serious and null steps in line search
$\zeta \in (0, \frac{1}{2})$	$\zeta = 0.01$	Coefficient for interpolation in line search
$\vartheta \geq 1$	$\vartheta = 1$	Exponent for interpolation in line search
$C_S > 0$	$C_S = 10^{50}$	Upper bound of the distance between x_k and y_k
$C_G > 0$	$C_G = 10^{50}$	Upper bound of the norm of the damped matrices $\{\rho_j G_j\}$ ($ \rho_j G_j \leq C_G$)
$\hat{C}_G > 0$	$\hat{C}_G = C_G$	Upper bound of the norm of the damped matrices $\{\hat{\rho}_j \hat{G}_j\}$ ($ \hat{\rho}_j \hat{G}_j \leq \hat{C}_G$)
$\bar{\bar{C}}_G > 0$	$\bar{\bar{C}}_G = C_G$	Upper bound of the norm of the matrices $\{\bar{\bar{G}}_j^k\}$ and $\{\bar{\bar{G}}^k\}$ ($\max(\bar{\bar{G}}_j^k , \bar{\bar{G}}^k) \leq \bar{\bar{C}}_G$)
$i_\rho \geq 0$	$i_\rho = 3$	Selection parameter for ρ_{k+1} (cf. Remark 3.21)

TABLE 3.1: Initial parameters (continued)

General	Default	Description
$i_l \geq 0$		Line search selection parameter (cf. Remark 3.21)
$i_m \geq 0$		Matrix selection parameter (cf. Remark 3.21)
$i_r \geq 0$		Bundle reset parameter (cf. Remark 3.21)
$\gamma_1 > 0$	$\gamma_1 = 1$	Coefficient for locality measure for objective function
$\gamma_2 > 0$	$\gamma_2 = 1$	Coefficient for locality measure for constraint
$\omega_1 \geq 1$	$\omega_1 = 2$	Exponent for locality measure for objective function
$\omega_2 \geq 1$	$\omega_2 = 2$	Exponent for locality measure for constraint

Set the initial values of the data which gets changed during the algorithm:

$$\begin{aligned} i_n &= 0 \quad (\# \text{ subsequent null and short steps}) \\ i_s &= 0 \quad (\# \text{ subsequent serious steps}) \\ J_1 &= \{1\} \quad (\text{set of bundle indices}) . \end{aligned}$$

Compute the following information at the initial trial point

$$f_p^1 = f_1^1 = f(y_1) \tag{3.52}$$

$$g_p^1 = g_1^1 = g(y_1) \in \partial f(y_1) \tag{3.53}$$

$$G_p^1 = G_1 = G(y_1) \in \partial^2 f(y_1)$$

$$F_p^1 = F_1^1 = F(y_1) < 0 \quad (y_1 \text{ is strictly feasible according to assumption}) \tag{3.54}$$

$$\hat{g}_p^1 = \hat{g}_1^1 = \hat{g}(y_1) \in \partial F(y_1) \tag{3.55}$$

$$\hat{G}_p^1 = \hat{G}_1 = \hat{G}(y_1) \in \partial^2 F(y_1) \tag{3.56}$$

and set

$$\hat{s}_p^1 = \hat{s}_1^1 = s_p^1 = s_1^1 = 0 \quad (\text{locality measure}) \tag{3.57}$$

$$\hat{\rho}_1 = \rho_1 = 1 \quad (\text{damping parameter})$$

$$\bar{\kappa}^1 = 1 \quad (\text{Lagrange multiplier for optimality condition})$$

$$k = 1 \quad (\text{iterator}) .$$

1. *Determination of the matrices for the QCQP:*

$$\text{if } (\text{step } k-1 \text{ and } k-2 \text{ were serious steps}) \wedge (\lambda_{k-1}^{k-1} = 1 \vee \underbrace{i_s > i_r}_{\text{bundle reset}})$$

$$W = G_k + \bar{\kappa}^k \hat{G}_k \tag{3.58}$$

else

$$W = G_p^k + \bar{\kappa}^k \hat{G}_p^k \tag{3.59}$$

end

if $i_n \leq i_m + i_l$ (i.e. $\#$ of subsequent null and short steps $\leq$ the fixed number $i_m + i_l$)

$\bar{W}_p^k$ = “positive definite modification of W ”

else (i.e. more than $i_m + i_l$ subsequent null and short steps were executed)

$$\bar{W}_p^k = \bar{W}_p^{k-1} \quad (3.60)$$

end

if $i_n < i_m + i_l$ (i.e. # of subsequent null and short steps $<$ the fixed number $i_m + i_l$)

$$\begin{aligned} \bar{\hat{G}}^k &= \text{"positive definite modification of } \hat{G}_p^k\text{"} \\ \bar{\hat{G}}_j^k &= \text{"positive definite modification of } \hat{G}_j^k\text{"} \quad \text{for all } j \in J_k \end{aligned} \quad (3.61)$$

else if $i_n = i_m + i_l$

$$\begin{aligned} \bar{\hat{G}}^k &= \text{"positive definite modification of } \hat{G}_p^k\text{"} \\ \bar{\hat{G}}_j^k &= \bar{G}^k \quad \text{for all } j \in J_k \end{aligned} \quad (3.62)$$

else (i.e. at least $i_m + i_l$ subsequent null and short steps were executed)

$$\begin{aligned} \bar{\hat{G}}^k &= \bar{\hat{G}}^{k-1} \\ \bar{\hat{G}}_j^k &= \bar{\hat{G}}^{k-1} \quad \text{for all } j \in J_k \end{aligned} \quad (3.63)$$

end

2. *Computation of the localized approximation errors:*

$$\alpha_j^k := \max(|f(x_k) - f_j^k|, \gamma_1(s_j^k)^{\omega_1}) \quad (3.64)$$

$$\alpha_p^k := \max(|f(x_k) - f_p^k|, \gamma_1(s_p^k)^{\omega_1}) \quad (3.65)$$

$$A_j^k := \max(|F(x_k) - F_j^k|, \gamma_2(\hat{s}_j^k)^{\omega_2}) \quad (3.66)$$

$$A_p^k := \max(|F(x_k) - F_p^k|, \gamma_2(\hat{s}_p^k)^{\omega_2}) . \quad (3.67)$$

3. *Determination of the search direction:* Compute the solution $(d_k, \hat{v}_k) \in \mathbb{R}^{n+1}$ of the (convex) QCQP

$$\begin{aligned} \min_{d, \hat{v}} \quad & \hat{v} + \frac{1}{2} d^T \bar{W}_p^k d, \\ \text{s.t.} \quad & -\alpha_j^k + d^T g_j^k \leq \hat{v} \quad \text{for } j \in J_k \\ & -\alpha_p^k + d^T g_p^k \leq \hat{v} \quad \text{if } i_s \leq i_r \\ & F(x_k) - A_j^k + d^T \hat{g}_j^k + \frac{1}{2} d^T \bar{\hat{G}}_j^k d \leq 0 \quad \text{for } j \in J_k \\ & F(x_k) - A_p^k + d^T \hat{g}_p^k + \frac{1}{2} d^T \bar{\hat{G}}^k d \leq 0 \quad \text{if } i_s \leq i_r \end{aligned} \quad (3.68)$$

and its corresponding Lagrange multiplier $(\lambda^k, \lambda_p^k, \mu^k, \mu_p^k) \in \mathbb{R}_{\geq 0}^{2(|J_k|+1)}$, i.e.

$$d_k = -H_k^2 \left(\sum_{j \in J_k} \lambda_j^k g_j^k + \lambda_p^k g_p^k + \sum_{j \in J_k} \mu_j^k \hat{g}_j^k + \mu_p^k \hat{g}_p^k \right) \quad (3.69)$$

$$\begin{aligned} \hat{v}_k = & -d_k^T \bar{W}_p^k d_k - \frac{1}{2} d_k^T \left(\sum_{j \in J_k} \mu_j^k \bar{\hat{G}}_j^k + \mu_p^k \bar{\hat{G}}^k \right) d_k - \sum_{j \in J_k} \lambda_j^k \alpha_j^k - \lambda_p^k \alpha_p^k \\ & - \sum_{j \in J_k} \mu_j^k A_j^k - \mu_p^k A_p^k - \left(\sum_{j \in J_k} \mu_j^k + \mu_p^k \right) (-F(x_k)) \end{aligned} \quad (3.70)$$

(due to Proposition 3.18), where

$$H_k := (\bar{W}_p^k + \sum_{j \in J_k} \mu_j^k \bar{G}_j^k + \mu_p^k \bar{G}^k)^{-\frac{1}{2}}. \quad (3.71)$$

Set

$$\bar{\kappa}^{k+1} := \sum_{j \in J_k} \mu_j^k + \mu_p^k \quad (3.72)$$

$$\kappa_j^k := \begin{cases} \frac{\mu_j^k}{\bar{\kappa}^{k+1}} & \text{for } \bar{\kappa}^{k+1} > 0 \\ 0 & \text{for } \bar{\kappa}^{k+1} = 0 \end{cases} \quad (3.73)$$

$$\kappa_p^k := \begin{cases} \frac{\mu_p^k}{\bar{\kappa}^{k+1}} & \text{for } \bar{\kappa}^{k+1} > 0 \\ 0 & \text{for } \bar{\kappa}^{k+1} = 0. \end{cases} \quad (3.74)$$

if $i_s > i_r$ (i.e. # subsequent serious steps > the fixed parameter i_r)

$i_s = 0$ (bundle reset)

end

4. *Aggregation:* We set for the aggregation of information of the objective function

$$\tilde{f}_p^k = \sum_{j \in J_k} \lambda_j^k f_j^k + \lambda_p^k f_p^k \quad (3.75)$$

$$\tilde{g}_p^k = \sum_{j \in J_k} \lambda_j^k g_j^k + \lambda_p^k g_p^k \quad (3.76)$$

$$G_p^{k+1} = \sum_{j \in J_k} \lambda_j^k \rho_j G_j^k + \lambda_p^k G_p^k \quad (3.77)$$

$$\tilde{s}_p^k = \sum_{j \in J_k} \lambda_j^k s_j^k + \lambda_p^k s_p^k \quad (3.78)$$

$$\tilde{\alpha}_p^k = \max(|f(x_k) - \tilde{f}_p^k|, \gamma_1(\tilde{s}_p^k)^{\omega_1}) \quad (3.79)$$

and for the aggregation of information of the constraint

$$\tilde{F}_p^k = \sum_{j \in J_k} \kappa_j^k F_j^k + \kappa_p^k F_p^k \quad (3.80)$$

$$\tilde{g}_p^k = \sum_{j \in J_k} \kappa_j^k \hat{g}_j^k + \kappa_p^k \hat{g}_p^k \quad (3.81)$$

$$\hat{G}_p^{k+1} = \sum_{j \in J_k} \kappa_j^k \hat{\rho}_j \hat{G}_j^k + \kappa_p^k \hat{G}_p^k \quad (3.82)$$

$$\tilde{s}_p^k = \sum_{j \in J_k} \kappa_j^k \hat{s}_j^k + \kappa_p^k \hat{s}_p^k \quad (3.83)$$

$$\tilde{A}_p^k = \max(|F(x_k) - \tilde{F}_p^k|, \gamma_2(\tilde{s}_p^k)^{\omega_2}) \quad (3.84)$$

and we set

$$v_k = -d_k^T \bar{W}_p^k d_k - \frac{1}{2} d_k^T \left(\sum_{j \in J_k} \mu_j^k \bar{G}_j^k + \mu_p^k \bar{G}^k \right) d_k - \tilde{\alpha}_p^k - \bar{\kappa}^{k+1} \tilde{A}_p^k - \bar{\kappa}^{k+1} (-F(x_k)) \quad (3.85)$$

$$w_k = \frac{1}{2} |H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k)|^2 + \tilde{\alpha}_p^k + \bar{\kappa}^{k+1} \tilde{A}_p^k + \bar{\kappa}^{k+1} (-F(x_k)). \quad (3.86)$$

5. *Termination criterion:*

if $w_k \leq \varepsilon$
 stop
 end

6. *Line search:* We compute step sizes $0 \leq t_L^k \leq t_R^k \leq 1$ and $t_0^k \in (0, t_0]$ by using the line search described in Algorithm 3.25

$$[t_L^k, t_R^k, t_0^k] = \text{Linesearch}(v_k, \dots)$$

and we set

$$x_{k+1} = x_k + t_L^k d_k \quad (\text{is created strictly feasible by the line search}) \quad (3.87)$$

$$y_{k+1} = x_k + t_R^k d_k \quad (3.88)$$

$$f_{k+1} = f(y_{k+1}), \quad g_{k+1} = g(y_{k+1}) \in \partial f(y_{k+1}), \quad G_{k+1} = G(y_{k+1}) \in \partial^2 f(y_{k+1}) \quad (3.89)$$

$$F_{k+1} = F(y_{k+1}), \quad \hat{g}_{k+1} = \hat{g}(y_{k+1}) \in \partial F(y_{k+1}), \quad \hat{G}_{k+1} = \hat{G}(y_{k+1}) \in \partial^2 F(y_{k+1}) .$$

7. *Update:*

if $i_n \leq i_\rho$ (i.e. # of the subsequent null and short steps $\leq i_\rho$)

$$\rho_{k+1} = \min(1, \frac{C_G}{|G_{k+1}|}) \quad (3.90)$$

else ($\rightsquigarrow$ linear term for measuring approximation error)

$$\rho_{k+1} = 0$$

end

We set

$$\hat{\rho}_{k+1} = \min(1, \frac{\hat{C}_G}{|\hat{G}_{k+1}|}) . \quad (3.91)$$

if $t_L^k \geq t_0^k$ (serious step)

$$i_n = 0$$

$$i_s = i_s + 1$$

else (no serious step, i.e. null or short step)

$$i_n = i_n + 1 \quad (3.92)$$

end

Compute the updates of the locality measure for the objective function

$$s_j^{k+1} = s_j^k + |x_{k+1} - x_k| \quad \text{for } j \in J_k \quad (3.93)$$

$$s_{k+1}^{k+1} = |x_{k+1} - y_{k+1}| \quad (3.94)$$

$$s_p^{k+1} = \tilde{s}_p^k + |x_{k+1} - x_k| \quad (3.95)$$

and for the constraint

$$\hat{s}_j^{k+1} = \hat{s}_j^k + |x_{k+1} - x_k| \quad \text{for } j \in J_k \quad (3.96)$$

$$\hat{s}_{k+1}^{k+1} = |x_{k+1} - y_{k+1}| \quad (3.97)$$

$$\hat{s}_p^{k+1} = \tilde{\hat{s}}_p^k + |x_{k+1} - x_k| . \quad (3.98)$$

Compute the updates for the objective function approximation

$$\begin{aligned} f_j^{k+1} &= f_j^k + g_j^{kT}(x_{k+1} - x_k) + \frac{1}{2}\rho_j(x_{k+1} - x_k)^T G_j(x_{k+1} - x_k) \quad \text{for } j \in J_k \\ f_{k+1}^{k+1} &= f_{k+1}^k + g_{k+1}^{kT}(x_{k+1} - y_{k+1}) + \frac{1}{2}\rho_{k+1}(x_{k+1} - y_{k+1})^T G_{k+1}(x_{k+1} - y_{k+1}) \end{aligned} \quad (3.99)$$

$$f_p^{k+1} = \tilde{f}_p^k + \tilde{g}_p^{kT}(x_{k+1} - x_k) + \frac{1}{2}(x_{k+1} - x_k)^T G_p^{k+1}(x_{k+1} - x_k) \quad (3.100)$$

and for the constraint

$$\begin{aligned} F_j^{k+1} &= F_j^k + \hat{g}_j^{kT}(x_{k+1} - x_k) + \frac{1}{2}\hat{\rho}_j(x_{k+1} - x_k)^T \hat{G}_j(x_{k+1} - x_k) \quad \text{for } j \in J_k \\ F_{k+1}^{k+1} &= F_{k+1}^k + \hat{g}_{k+1}^{kT}(x_{k+1} - y_{k+1}) + \frac{1}{2}\hat{\rho}_{k+1}(x_{k+1} - y_{k+1})^T \hat{G}_{k+1}(x_{k+1} - y_{k+1}) \end{aligned} \quad (3.101)$$

$$F_p^{k+1} = \tilde{F}_p^k + \tilde{g}_p^{kT}(x_{k+1} - x_k) + \frac{1}{2}(x_{k+1} - x_k)^T \hat{G}_p^{k+1}(x_{k+1} - x_k) . \quad (3.102)$$

Compute the updates for the subgradient of the objective function approximation

$$g_j^{k+1} = g_j^k + \rho_j G_j(x_{k+1} - x_k) \quad \text{for } j \in J_k \quad (3.103)$$

$$g_{k+1}^{k+1} = g_{k+1}^k + \rho_{k+1} G_{k+1}(x_{k+1} - y_{k+1}) \quad (3.104)$$

$$g_p^{k+1} = \tilde{g}_p^k + G_p^{k+1}(x_{k+1} - x_k) \quad (3.105)$$

and for the constraint

$$\hat{g}_j^{k+1} = \hat{g}_j^k + \hat{\rho}_j \hat{G}_j(x_{k+1} - x_k) \quad \text{for } j \in J_k \quad (3.106)$$

$$\hat{g}_{k+1}^{k+1} = \hat{g}_{k+1}^k + \hat{\rho}_{k+1} \hat{G}_{k+1}(x_{k+1} - y_{k+1}) \quad (3.107)$$

$$\hat{g}_p^{k+1} = \tilde{g}_p^k + \hat{G}_p^{k+1}(x_{k+1} - x_k) . \quad (3.108)$$

Choose $J_{k+1} \subseteq \{k - M + 2, \dots, k + 1\} \cap \{1, 2, \dots\}$ with $k + 1 \in J_{k+1}$.

$k = k + 1$

Go to 1

Remark 3.21.

- Like in the original unconstrained bundle-Newton method by LUKŠAN & VLČEK [64], the parameters i_m and i_r as well as the additional parameter i_l are only needed for proving convergence. Since in practice we usually terminate an algorithm in particular, if a maximal number of iterations `Nit_max` is exceeded, we always choose $i_m = i_n = i_l = \text{Nit_max} + 1$ in our implementation of Algorithm 3.20.
- The case distinction for the choice of W according to (3.58) resp. (3.59) is only necessary for showing the superlinear convergence of the original unconstrained bundle-Newton method for strongly convex, twice times continuously differentiable functions (cf. LUKŠAN & VLČEK [64, p. 385, Section 4]).
- As the choice $i_\rho = 3$ (cf. the initialization of Algorithm 3.20) for the case distinction $i_n \leq i_\rho$ for ρ_{k+1} from (3.90) is due to empirical observations in the original unconstrained bundle-Newton method (cf. LUKŠAN & VLČEK [64, p. 378]), the fact that we make no case distinction for $\hat{\rho}_{k+1}$ from (3.91) was also found out numerically.

- A numerically meaningful choice of the matrices $\tilde{G}_j^k$ and $\tilde{G}^k$ that occur in (3.61) is discussed in Section 4.3.

Proposition 3.22. *We have for all $k \geq 0$*

$$|H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1}\tilde{g}_p^k)|^2 = d_k^T (\bar{W}_p^k + \sum_{j \in J_k} \mu_j^k \tilde{G}_j^k + \mu_p^k \tilde{G}^k) d_k \quad (3.109)$$

$$w_k = -\frac{1}{2} d_k^T \bar{W}_p^k d_k - v_k. \quad (3.110)$$

Proof. Because of

$$\begin{aligned} H_k &\stackrel{(3.71)}{=} (\bar{W}_p^k + \sum_{j \in J_k} \mu_j^k \tilde{G}_j^k + \mu_p^k \tilde{G}^k)^{-\frac{1}{2}} \\ &\stackrel{\wedge(-2)}{\Longleftrightarrow} H_k^{-2} = \bar{W}_p^k + \sum_{j \in J_k} \mu_j^k \tilde{G}_j^k + \mu_p^k \tilde{G}^k \end{aligned} \quad (3.111)$$

and

$$\begin{aligned} d_k &\stackrel{(3.69)}{=} -H_k^2 \left(\sum_{j \in J_k} \lambda_j^k g_j^k + \lambda_p^k g_p^k + \sum_{j \in J_k} \mu_j^k \hat{g}_j^k + \mu_p^k \hat{g}_p^k \right) \\ &\stackrel{(3.72), (3.73), (3.74)}{\stackrel{\downarrow}{=}} -H_k^2 \left(\sum_{j \in J_k} \lambda_j^k g_j^k + \lambda_p^k g_p^k + \bar{\kappa}^{k+1} \left(\sum_{j \in J_k} \kappa_j^k \hat{g}_j^k + \kappa_p^k \hat{g}_p^k \right) \right) \\ &\stackrel{(3.76), (3.81)}{\stackrel{\downarrow}{=}} -H_k^2 (\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k) \end{aligned} \quad (3.112)$$

we obtain (3.109)

$$\begin{aligned} d_k^T (\bar{W}_p^k + \sum_{j \in J_k} \mu_j^k \tilde{G}_j^k + \mu_p^k \tilde{G}^k) d_k &\stackrel{(3.111)}{=} d_k H_k^{-2} d_k \\ &\stackrel{(3.112)}{=} ((\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k)^T H_k^2) H_k^{-2} (H_k^2 (\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k)) \\ &= (\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k)^T H_k^2 (\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k) \\ &= |H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k)|^2. \end{aligned}$$

Furthermore, (3.110) holds due to

$$\begin{aligned} w_k &\stackrel{(3.86)}{=} \frac{1}{2} |H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k)|^2 + \tilde{\alpha}_p^k + \bar{\kappa}^{k+1} \tilde{A}_p^k + \bar{\kappa}^{k+1} (-F(x_k)) \\ &\stackrel{(3.109)}{=} \frac{1}{2} d_k^T (\bar{W}_p^k + \sum_{j \in J_k} \mu_j^k \tilde{G}_j^k + \mu_p^k \tilde{G}^k) d_k + \tilde{\alpha}_p^k + \bar{\kappa}^{k+1} \tilde{A}_p^k + \bar{\kappa}^{k+1} (-F(x_k)) \\ &= d_k^T \bar{W}_p^k d_k - \frac{1}{2} d_k^T \bar{W}_p^k d_k + \frac{1}{2} d_k^T \left(\sum_{j \in J_k} \mu_j^k \tilde{G}_j^k + \mu_p^k \tilde{G}^k \right) d_k + \tilde{\alpha}_p^k + \bar{\kappa}^{k+1} \tilde{A}_p^k + \bar{\kappa}^{k+1} (-F(x_k)) \\ &\stackrel{(3.85)}{=} -\frac{1}{2} d_k^T \bar{W}_p^k d_k - v_k. \end{aligned} \quad \square$$

Remark 3.23. If we consider a nonsmooth unconstrained optimization problem (i.e. we drop the constraint $F(x) \leq 0$ in optimization problem (2.35)), then our formula for v_k from (3.85) reduces to the formula for v_k in the unconstrained bundle-Newton method (cf. LUKŠAN & VLČEK [64, p. 377, formula (13)])

$$\begin{aligned} v_k &\stackrel{(3.85), (3.109)}{\stackrel{\downarrow}{=}} -|H_k \tilde{g}_p^k|^2 - \tilde{\alpha}_p^k. \end{aligned}$$

3.4 Presentation of the line search

Motivation 3.24. We extend the line search of the bundle-Newton method for nonsmooth unconstrained minimization to the constrained case in the line search described in Algorithm 3.25. For obtaining a clear arrangement of the line search, we compute data concerning the objective function in `ComputeObjectiveData` and data concerning the constraint in `ComputeConstraintData`. Before formulating the line search in detail, we give a brief overview of its functionality:

- Starting with the step size $t = 1$, we check if the point $x_k + td_k$ is strictly feasible.
 - If so and if additionally the objective function decreases sufficiently in this point and t is not too small, then we take $x_k + td_k$ as new iteration point in Algorithm 3.20 (serious step).
 - Otherwise, if the point $x_k + td_k$ is strictly feasible and the model of the objective function changes sufficiently, we take $x_k + td_k$ as new trial point (short/null step with respect to the objective function).
- If $x_k + td_k$ is not strictly feasible, but the model of the constraint changes sufficiently (in particular here the quadratic approximation of the constraint comes into play), we take $x_k + td_k$ as new trial point (short/null step with respect to the constraint).
- After choosing a new step size $t \in [0, 1]$ by interpolation, we iterate this procedure.

Algorithm 3.25.

function $[t_L, t_R, t_0] = \text{Linesearch}(v_k, \dots)$

0. *Initialization:* Choose $\zeta \in (0, \frac{1}{2})$ as well as $\vartheta \geq 1$ and set $t_L = 0$ as well as $t = t_U = 1$

1. *Modification of either t_L or t_U :*

```

    if  $F(x_k + td_k) < 0$  (strictly feasible)
        if  $f(x_k + td_k) \leq f(x_k) + m_L v_k \cdot t$ 
             $t_L = t$ 
        else if  $f(x_k + td_k) > f(x_k) + m_L v_k \cdot t$ 
             $t_U = t$ 
        end
    else if  $F(x_k + td_k) \geq 0$  (not strictly feasible)
         $t_U = t$ 
         $t_0 = \hat{t}_0 t_U$ 
    end
    if  $t_L \geq t_0$ 
         $t_R = t_L$ 
        return (serious step)
    end

```

(3.113)

2. *Decision of return:*

```

if  $i_n < i_l$ 
  if  $F(x_k + td_k) < 0$  (strictly feasible)
     $[g, G, \dots] = \text{ComputeObjectiveData}(t, \dots)$ 
    if  $Z = \text{true}$ 
       $t_R = t$ 
      return (short/null step: change of model of the objective function)
    end
  else if  $F(x_k + td_k) \geq 0$  (not strictly feasible)
     $[\hat{g}, \hat{G}, \dots] = \text{ComputeConstraintData}(t, \dots)$ 
    if  $\hat{Z} = \text{true}$ 
       $t_R = t$ 
      return (short/null step: change of model of the constraint)
    end
  end
else if  $i_n \geq i_l$ 
   $[g, G, \dots] = \text{ComputeObjectiveData}(t, \dots)$ 
  if  $F(x_k + td_k) < 0$  and  $Z = \text{true}$ 
     $t_R = t$ 
    return (Short-Serious-/Null-Step: change of model of the objective function)
  end
end

```

3. *Interpolation:* Choose

$$t \in [t_L + \zeta(t_U - t_L)^\vartheta, t_U - \zeta(t_U - t_L)^\vartheta] . \quad (3.114)$$

4. *Loop:* Go to 1

function $[g, G, \dots] = \text{ComputeObjectiveData}(t, \dots)$

$$g = g(x_k + td_k) \in \partial f(x_k + td_k)$$

$$G = G(x_k + td_k) \in \partial^2 f(x_k + td_k)$$

$$\rho = \begin{cases} \min(1, \frac{C_G}{|G|}) & \text{for } i_n \leq 3 \\ 0 & \text{else} \end{cases}$$

$$f = f(x_k + td_k) + (t_L - t)g^T d_k + \frac{1}{2}\rho(t_L - t)^2 d_k^T G d_k \quad (3.115)$$

$$\beta = \max(|f(x_k + t_L d_k) - f|, \gamma_1 |t_L - t|^{\omega_1} |d_k|^{\omega_1}) \quad (3.116)$$

$$Z = -\beta + d_k^T (g + \rho(t_L - t)Gd_k) \geq m_R v_k \text{ and } (t - t_L)|d_k| \leq C_S \quad (3.117)$$

function $[\hat{g}, \hat{G}, \dots] = \text{ComputeConstraintData}(t, \dots)$

$$\hat{g} = \hat{g}(x_k + td_k) \in \partial F(x_k + td_k)$$

$$\hat{G} = \hat{G}(x_k + td_k) \in \partial^2 F(x_k + td_k)$$

$$\hat{\rho} = \min(1, \frac{\hat{C}_G}{|\hat{G}|})$$

$$F = F(x_k + td_k) + (t_L - t)\hat{g}^T d_k + \frac{1}{2}\rho(t_L - t)^2 d_k^T \hat{G} d_k \quad (3.118)$$

$$\hat{\beta} = \max(|F(x_k + t_L d_k) - F|, \gamma_2 |t_L - t|^{\omega_2} |d_k|^{\omega_2}) \quad (3.119)$$

$$\bar{\bar{G}} = \text{"positive definite modification of } \hat{G}\text{"} \quad (3.120)$$

$$\hat{Z} = F(x_k + t_L d_k) - \hat{\beta} + d_k^T (\hat{g} + \hat{\rho}(t_L - t)\hat{G}d_k) \geq m_F \cdot (-\frac{1}{2}d_k^T \bar{\bar{G}}d_k) \text{ and } (t - t_L)|d_k| \leq C_S \quad (3.121)$$

Remark 3.26.

1. The parameter i_l is only necessary for proving global convergence of Algorithm 3.20 (to be more precise, it is only needed to show (3.392)). If we choose $i_l = 0$, then only a change of the model of the objective function yields a short or null step. In fact we have i_l steps in Algorithm 3.20 in which we can use any meaningful criterion for terminating the line search (even for the unconstrained case as it is partially done in the implementation of the original unconstrained bundle-Newton method anyway).
2. (3.113) is due to the following observation:
 - (a) Consider the line search (Algorithm 3.25) without (3.113) (i.e. t_0 is fixed, e.g., $t_0 := 0.5 \in (0, 1)$, where this large, but legal value for t_0 is only chosen to obtain a better graphical illustration in Figure 3.3). It can happen (in particular) at the beginning of Algorithm 3.20 that the search direction d_k is bad as we have no knowledge on the behavior of f and F yet. Consequently, the following situation can occur:
 - The model of the objective function f does not change (e.g., if f is linear on $x_k + td_k$ with $t \in [0, 1]$).
 - There are no step sizes $t > t_0$ which yield feasible $x_k + td_k$ that are colored orange in Figure 3.3 (this is in particular possible, if we are near the boundary of the feasible set).

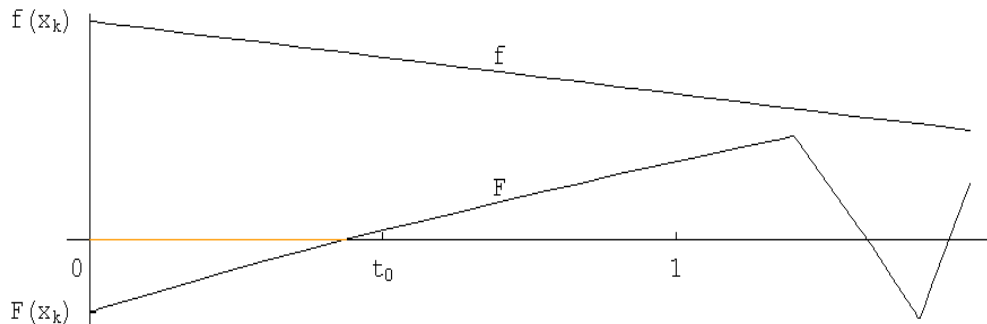


FIGURE 3.3: Line search with fixed t_0

In this situation the line search will not terminate for fixed t_0 (in particular in the case $i_n < i_l$ the model of F does not need to even satisfy (3.121) for infeasible $x_k + td_k$).

- (b) Therefore, we need to decrease t_0 to have at least one feasible step in the line search for which a descent of f is enough for terminating the line search (similar to the unconstrained case). As the convergence analysis will show, this must not be done too often (cf. (3.318) and Remark 3.56). Because we use the quadratic terms in the constraint approximation to obtain as much feasibility as possible on the search path $t \mapsto x_k + td_k$ with $t \in [0, 1]$ (cf. Idea 3.13), we expect that this should be true. Indeed, in practice t_0 turns out to be only modified at the beginning of Algorithm 3.20 (cf. Chapter 6).
- (c) In particular, if $F(x_k + td_k) < 0$ for all $t \in [0, 1]$ (e.g., if F is constant and negative on $\mathbb{R}^n$ which in fact yields an unconstrained optimization problem), the case (3.113) will never occur and therefore t_0 will not get changed (this is the reason why t_0 is constant in the bundle-Newton method for nonsmooth unconstrained minimization).

3. The step sizes which the line search returns correspond to the points

$$x_{k+1} = x_k + t_L^k d_k, \quad y_{k+1} = x_k + td_k = x_k + t_R^k d_k. \quad (3.122)$$

4. Only strictly feasible iteration points are accepted in the line search

$$F(x_k + t_L^k d_k) < 0. \quad (3.123)$$

Nevertheless, trial points may be infeasible (if $i_n < i_l$).

5. Similar to the line search in the bundle-Newton method for nonsmooth unconstrained minimization by LUKŠAN & VLČEK [64], we want to choose a new point in the interval $[t_L + \zeta(t_U - t_L)^\vartheta, t_U - \zeta(t_U - t_L)^\vartheta]$ by interpolation (cf. (3.114)). For this purpose, we set up a polynomial p passing through

$$(t_L, f(x_k + t_L d_k)) \text{ and } (t_U, f(x_k + t_U d_k))$$

as well as a polynomial q passing through

$$(t_L, F(x_k + t_L d_k)) \text{ and } (t_U, F(x_k + t_U d_k)).$$

Now we minimize p subject to the constraint $q(t) \leq 0$ on $[t_L + \zeta(t_U - t_L)^\vartheta, t_U - \zeta(t_U - t_L)^\vartheta]$

$$\begin{aligned} \min p(t) \\ \text{s.t. } q(t) \leq 0, t \in [t_L + \zeta(t_U - t_L)^\vartheta, t_U - \zeta(t_U - t_L)^\vartheta] \end{aligned} \quad (3.124)$$

and we use a solution $\hat{t}$ of (3.124) as the new point. The degree of the polynomial should be chosen in a way that solving (3.124) is easy (e.g., if we choose p and q as quadratic polynomials, then solving (3.124) consists of solving a one-dimensional linear equation, a one-dimensional quadratic equation and a few case distinctions).

We only need the estimate (3.131) of the following proposition in context with the line search (Algorithm 3.25) (in the proof of Proposition 3.29). The other properties are needed for showing global convergence of Algorithm 3.20 in Section 3.6. Nevertheless, as the proofs of these properties are very

similar, we already formulate them here.

Proposition 3.27. *Let*

$$\hat{\alpha}_p^k := \sum_{j \in J_k} \lambda_j^k \alpha_j^k + \lambda_p^k \alpha_p^k \quad (3.125)$$

$$\hat{A}_p^k := \sum_{j \in J_k} \kappa_j^k A_j^k + \kappa_p^k A_p^k \quad (3.126)$$

$$\hat{w}_k := \frac{1}{2} |H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k)|^2 + \hat{\alpha}_p^k + \bar{\kappa}^{k+1} \hat{A}_p^k + \bar{\kappa}^{k+1} (-F(x_k)) \quad (3.127)$$

(Note: $\hat{w}_k$ is the optimal function value of the dual problem (3.42)). Then we have at iteration k of Algorithm 3.20

$$\tilde{\alpha}_p^k \leq \hat{\alpha}_p^k \quad (3.128)$$

$$\tilde{A}_p^k \leq \hat{A}_p^k \quad \text{if } \bar{\kappa}^{k+1} > 0 \quad (3.129)$$

$$0 \leq w_k \leq \hat{w}_k \quad (3.130)$$

$$\hat{v}_k \leq v_k \leq 0. \quad (3.131)$$

Proof.

- Since $\lambda_j^k \geq 0$ for $j \in J_k$ and $\lambda_p^k \geq 0$ holds for the solution of the dual problem (3.42) of the QCQP (3.68), we have

$$1 = \sum_{j \in J_k} \lambda_j^k + \lambda_p^k \xrightarrow{\cdot f(x_k)} f(x_k) = \sum_{j \in J_k} \lambda_j^k f(x_k) + \lambda_p^k f(x_k) \quad (3.132)$$

and hence (3.128) follows from

$$\begin{aligned} \tilde{\alpha}_p^k &\stackrel{(3.79)}{=} \max(|\tilde{f}_p^k - f(x_k)|, \gamma_1(\tilde{s}_p^k)^{\omega_1}) \\ &\stackrel{(3.75), (3.78)}{\downarrow} \max(|\sum_{j \in J_k} \lambda_j^k f_j^k + \lambda_p^k f_p^k - f(x_k)|, \gamma_1(\sum_{j \in J_k} \lambda_j^k s_j^k + \lambda_p^k s_p^k)^{\omega_1}) \\ &\stackrel{(3.132)}{=} \max(|\sum_{j \in J_k} \lambda_j^k f_j^k + \lambda_p^k f_p^k - \sum_{j \in J_k} \lambda_j^k f(x_k) - \lambda_p^k f(x_k)|, \gamma_1(\sum_{j \in J_k} \lambda_j^k s_j^k + \lambda_p^k s_p^k)^{\omega_1}) \\ &\stackrel{(2.7)}{\leq} \max\left(\sum_{j \in J_k} \lambda_j^k |f_j^k - f(x_k)| + \lambda_p^k |f_p^k - f(x_k)|, \sum_{j \in J_k} \lambda_j^k \gamma_1(s_j^k)^{\omega_1} + \lambda_p^k \gamma_1(s_p^k)^{\omega_1}\right) \\ &\stackrel{(2.8)}{\leq} \sum_{j \in J_k} \lambda_j^k \max(|f_j^k - f(x_k)|, \gamma_1(s_j^k)^{\omega_1}) + \lambda_p^k \max(|f_p^k - f(x_k)|, \gamma_1(s_p^k)^{\omega_1}) \\ &\stackrel{(3.64), (3.65)}{\downarrow} \sum_{j \in J_k} \lambda_j^k \alpha_j^k + \lambda_p^k \alpha_p^k \\ &\stackrel{(3.125)}{=} \hat{\alpha}_p^k. \end{aligned}$$

- If $\bar{\kappa}^{k+1} > 0$, we have

$$1 \stackrel{(3.73)}{=} \sum_{j \in J_k} \kappa_j^k + \kappa_p^k \xrightarrow{\cdot F(x_k)} F(x_k) = \sum_{j \in J_k} \kappa_j^k F(x_k) + \kappa_p^k F(x_k) \quad (3.133)$$

and hence (3.129) follows from

$$\begin{aligned}
\tilde{A}_p^k &\stackrel{(3.84)}{=} \max(|\tilde{F}_p^k - F(x_k)|, \gamma_2(\tilde{\hat{s}}_p^k)^{\omega_2}) \\
&\stackrel{(3.80), (3.83)}{\downarrow} \max(|\sum_{j \in J_k} \kappa_j^k F_j^k + \kappa_p^k F_p^k - F(x_k)|, \gamma_2(\sum_{j \in J_k} \kappa_j^k \hat{s}_j^k + \kappa_p^k \hat{s}_p^k)^{\omega_2}) \\
&\stackrel{(3.133)}{=} \max(|\sum_{j \in J_k} \kappa_j^k F_j^k + \kappa_p^k F_p^k - \sum_{j \in J_k} \kappa_j^k F(x_k) - \kappa_p^k F(x_k)|, \gamma_2(\sum_{j \in J_k} \kappa_j^k \hat{s}_j^k + \kappa_p^k \hat{s}_p^k)^{\omega_2}) \\
&\stackrel{(2.7)}{\leq} \max\left(\sum_{j \in J_k} \kappa_j^k |F(x_k) - F_j^k| + \kappa_p^k |F(x_k) - F_p^k|, \sum_{j \in J_k} \kappa_j^k \gamma_2(\hat{s}_j^k)^{\omega_2} + \kappa_p^k \gamma_2(\hat{s}_p^k)^{\omega_2}\right) \\
&\stackrel{(2.8)}{\leq} \sum_{j \in J_k} \kappa_j^k \max(|F_j^k - F(x_k)|, \gamma_2(\hat{s}_j^k)^{\omega_2}) + \kappa_p^k \max(|F_p^k - F(x_k)|, \gamma_2(\hat{s}_p^k)^{\omega_2}) \\
&\stackrel{(3.66), (3.67)}{\downarrow} \sum_{j \in J_k} \kappa_j^k A_j^k + \kappa_p^k A_p^k \\
&\stackrel{(3.126)}{=} \hat{A}_p^k.
\end{aligned}$$

- For $\bar{\kappa}^{k+1} \geq 0$ we have

$$\bar{\kappa}^{k+1} \tilde{A}_p^k \leq \bar{\kappa}^{k+1} \hat{A}_p^k \quad (3.134)$$

due to (3.129). Adding up (3.128) and (3.134) yields

$$\begin{aligned}
&\stackrel{(3.79), (3.72), (3.84)}{\downarrow} \\
0 &\leq \tilde{\alpha}_p^k + \bar{\kappa}^{k+1} \tilde{A}_p^k \leq \hat{\alpha}_p^k + \bar{\kappa}^{k+1} \hat{A}_p^k. \quad (3.135)
\end{aligned}$$

- We obtain (3.130) due to

$$\begin{aligned}
w_k &\stackrel{(3.86)}{=} \frac{1}{2} |H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k)|^2 + \tilde{\alpha}_p^k + \bar{\kappa}^{k+1} \tilde{A}_p^k + \bar{\kappa}^{k+1} (-F(x_k)) \\
&\stackrel{(3.135)}{\leq} \frac{1}{2} |H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k)|^2 + \hat{\alpha}_p^k + \bar{\kappa}^{k+1} \hat{A}_p^k + \bar{\kappa}^{k+1} (-F(x_k)) \\
&\stackrel{(3.127)}{=} \hat{w}_k
\end{aligned}$$

and

$$w_k \stackrel{(3.86)}{=} \frac{1}{2} |H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k)|^2 + \tilde{\alpha}_p^k + \bar{\kappa}^{k+1} \tilde{A}_p^k + \underbrace{\bar{\kappa}^{k+1}}_{\stackrel{(3.72)}{\geq 0}} \underbrace{(-F(x_k))}_{<0} \stackrel{(3.135)}{\geq} 0.$$

- Since we have

$$\begin{aligned}
0 &\stackrel{(3.135)}{\leq} \tilde{\alpha}_p^k + \bar{\kappa}^{k+1} \tilde{A}_p^k \\
&\stackrel{(3.135)}{\leq} \hat{\alpha}_p^k + \bar{\kappa}^{k+1} \hat{A}_p^k \\
&\stackrel{(3.125), (3.126)}{\downarrow} \sum_{j \in J_k} \lambda_j^k \alpha_j^k + \lambda_p^k \alpha_p^k + \bar{\kappa}^{k+1} \left(\sum_{j \in J_k} \kappa_j^k A_j^k + \kappa_p^k A_p^k \right) \\
&\stackrel{(3.73), (3.74)}{\downarrow} \sum_{j \in J_k} \lambda_j^k \alpha_j^k + \lambda_p^k \alpha_p^k + \sum_{j \in J_k} \mu_j^k A_j^k + \mu_p^k A_p^k \\
\Rightarrow 0 &\geq -\tilde{\alpha}_p^k - \bar{\kappa}^{k+1} \tilde{A}_p^k \geq -\sum_{j \in J_k} \lambda_j^k \alpha_j^k - \lambda_p^k \alpha_p^k - \sum_{j \in J_k} \mu_j^k A_j^k - \mu_p^k A_p^k, \quad (3.136)
\end{aligned}$$

we obtain (3.131) due to

$$\begin{aligned}
 v_k &\stackrel{(3.85)}{=} -d_k^T \bar{W}_p^k d_k - \frac{1}{2} d_k^T \left(\sum_{j \in J_k} \mu_j^k \bar{G}_j^k + \mu_p^k \bar{G}^k \right) d_k - \tilde{\alpha}_p^k - \bar{\kappa}^{k+1} \tilde{A}_p^k - \bar{\kappa}^{k+1} (-F(x_k)) \\
 &\stackrel{(3.136)}{\geq} -d_k^T \bar{W}_p^k d_k - \frac{1}{2} d_k^T \left(\sum_{j \in J_k} \mu_j^k \bar{G}_j^k + \mu_p^k \bar{G}^k \right) d_k - \sum_{j \in J_k} \lambda_j^k \alpha_j^k - \lambda_p^k \alpha_p^k \\
 &\quad - \sum_{j \in J_k} \mu_j^k A_j^k - \mu_p^k A_p^k - \bar{\kappa}^{k+1} (-F(x_k)) \\
 &\stackrel{(3.72)}{=} -d_k^T \bar{W}_p^k d_k - \frac{1}{2} d_k^T \left(\sum_{j \in J_k} \mu_j^k \bar{G}_j^k + \mu_p^k \bar{G}^k \right) d_k - \sum_{j \in J_k} \lambda_j^k \alpha_j^k - \lambda_p^k \alpha_p^k \\
 &\quad - \sum_{j \in J_k} \mu_j^k A_j^k - \mu_p^k A_p^k - \left(\sum_{j \in J_k} \mu_j^k + \mu_p^k \right) (-F(x_k)) \\
 &\stackrel{(3.70)}{=} \hat{v}_k
 \end{aligned}$$

and

$$\begin{aligned}
 &\bar{W}_p^k \text{ positive definite, (3.130)} \\
 v_k &\stackrel{(3.110)}{=} -\frac{1}{2} d_k^T \bar{W}_p^k d_k - w_k \stackrel{\downarrow}{\leq} 0 . \quad \square
 \end{aligned}$$

Proposition 3.28. *If the line search is entered at iteration k of Algorithm 3.20, then*

$$v_k < 0 . \quad (3.137)$$

Furthermore, if there occurs a step size t with $F(x_k + td_k) \geq 0$ in the line search, then

$$-\frac{1}{2} d_k^T \bar{G}_{x_k + td_k} d_k < 0 . \quad (3.138)$$

Proof.

- If the line search is entered at iteration k (cf. step 6 of Algorithm 3.20), then no termination occurred at step 5 of Algorithm 3.20 at iteration k , i.e.

$$0 \leq \varepsilon < w_k , \quad (3.139)$$

and therefore we obtain (3.137)

$$\begin{aligned}
 &\stackrel{(3.139), (3.110)}{0} \stackrel{\downarrow}{<} -\frac{1}{2} d_k^T \bar{W}_p^k d_k - v_k \stackrel{\downarrow}{\leq} -v_k . \\
 &\bar{W}_p^k \text{ positive definite}
 \end{aligned}$$

- (By contradiction) Suppose (3.138) does not hold, i.e. $d_k = 0$ due to (3.120). Then, since all iteration points x_k are strictly feasible due to (3.123), we obtain the following contradiction

$$\begin{aligned}
 &\stackrel{\text{assumption}}{0} \stackrel{\downarrow}{\leq} F(x_k + td_k) = F(x_k) \stackrel{x_k \text{ strictly feasible}}{<} 0 . \quad \square
 \end{aligned}$$

Proposition 3.29.

1. *The condition*

$$-\beta + d_k^T (g + \rho(t_L - t) G d_k) \geq m_R v_k$$

from (3.117) of the line search (Algorithm 3.25) corresponds to

$$-\alpha_{k+1}^{k+1} + d_k^T g_{k+1}^{k+1} \geq m_R v_k \quad (3.140)$$

resp. in explicit detail

$$-\alpha_{x_k+td_k}^{x_k+t_L d_k} + d_k^T g_{x_k+td_k}^{x_k+t_L d_k} \geq m_R v_k . \quad (3.141)$$

2. If the line search (Algorithm 3.25) terminates with condition (3.117), then the old search direction d_k and the old predicted descent v_k (of iteration k) are sufficiently infeasible for the new QCQP (3.68) (at iteration $k+1$) in Algorithm 3.20 (i.e. the old search direction d_k cannot occur as search direction at iteration $k+1$ and therefore we obtain a different search direction at iteration $k+1$ — and this is the formalized version of the demand for a “meaningful extension of the bundle” as mentioned in Subsection 2.3.3).

3. The condition

$$F(x_k + t_L d_k) - \hat{\beta} + d_k^T (\hat{g} + \hat{\rho}(t_L - t) \hat{G} d_k) \geq m_F \cdot (-\frac{1}{2} d_k^T \bar{\bar{G}} d_k)$$

from (3.121) of the line search (Algorithm 3.25) corresponds to

$$F(x_k + t_L d_k) - A_{x_k+td_k}^{x_k+t_L d_k} + d_k^T g_{x_k+td_k}^{x_k+t_L d_k} \geq m_F \cdot (-\frac{1}{2} d_k^T \bar{\bar{G}}_{x_k+td_k} d_k) . \quad (3.142)$$

4. If the line search (Algorithm 3.25) terminates with condition (3.121), then the old search direction d_k (of iteration k) is sufficiently infeasible for the new QCQP (3.68) (at iteration $k+1$) in Algorithm 3.20 (i.e. using a QCQP also yields a “meaningful extension of the bundle” in the constrained case).

5. The condition

$$(t - t_L) |d_k| \leq C_S$$

in (3.121) resp. (3.117) of the line search (Algorithm 3.25) corresponds to

$$|y_{k+1} - x_{k+1}| \leq C_S . \quad (3.143)$$

Proof.

1. Because of

$$\begin{aligned} & f \stackrel{(3.115)}{=} f(x_k + t d_k) + (t_L - t) g^T d_k + \frac{1}{2} \rho(t_L - t)^2 d_k^T G d_k \\ & = f(x_k + t d_k) + g^T ((x_k + t_L d_k) - (x_k + t d_k)) \\ & \quad + \frac{1}{2} \rho((x_k + t_L d_k) - (x_k + t d_k))^T G ((x_k + t_L d_k) - (x_k + t d_k)) \\ & \stackrel{(3.122)}{=} f(y_{k+1}) + g^T (x_{k+1} - y_{k+1}) + \frac{1}{2} \rho(x_{k+1} - y_{k+1})^T G (x_{k+1} - y_{k+1}) \\ & \quad \left\{ \begin{array}{l} \stackrel{(3.99)}{=} f_{k+1}^{k+1} \\ \stackrel{(3.1)}{=} f_{y_{k+1}}^\#(x_{k+1}) \stackrel{(3.122)}{=} f_{x_k+td_k}^\#(x_k + t_L d_k) \end{array} \right. \end{aligned} \quad (3.144)$$

and

$$\begin{aligned}
\beta &\stackrel{(3.116)}{=} \max(|f - f(x + t_L d)|, \gamma_1 |t_L - t|^\omega |d_k|^{\omega_1}) \\
&= \max(|f - f(x + t_L d)|, \gamma_1 |x_k + t_L d_k - (x_k + t d_k)|^{\omega_1}) \\
&\stackrel{(3.122)}{=} \max(|f - f(x_{k+1})|, \gamma_1 |x_{k+1} - y_{k+1}|^{\omega_1}) \\
&\stackrel{(3.94)}{=} \max(|f - f(x_{k+1})|, \gamma_1 (s_{k+1}^{k+1})^{\omega_1}) \\
&\stackrel{(3.144)}{=} \max(|f_{k+1}^{k+1} - f(x_{k+1})|, \gamma_1 (s_{k+1}^{k+1})^{\omega_1}) \\
&\stackrel{(3.64)}{=} \alpha_{k+1}^{k+1}, \tag{3.145}
\end{aligned}$$

we obtain (3.140) as well as (3.141)

$$\begin{aligned}
m_R v_k &\stackrel{(3.117)}{\leq} -\beta + d_k^T (g + \rho(t_L - t) G d_k) \\
&\stackrel{(3.145)}{=} -\alpha_{k+1}^{k+1} + d_k^T (g + \rho G(t_L d_k - t d_k)) \\
&= -\alpha_{k+1}^{k+1} + d_k^T (g + \rho G(x_k + t_L d_k - (x_k + t d_k))) \\
&\stackrel{(3.122)}{=} -\alpha_{k+1}^{k+1} + d_k^T (g + \rho G(x_{k+1} - y_{k+1})) \\
&\left\{ \begin{array}{l} \stackrel{(3.104)}{=} -\alpha_{k+1}^{k+1} + d_k^T g_{k+1}^{k+1} \\ \stackrel{(3.3), (3.122)}{\downarrow} -\alpha_{x_k + t_L d_k}^{x_k + t_L d_k} + d_k^T g_{x_k + t_L d_k}^\#(x_k + t_L d_k) \end{array} \right.
\end{aligned}$$

2. Due to the initialization of Algorithm 3.20 and (3.137), we have

$$0 < m_R < 1 \quad \xrightarrow{v_k \leq 0} \quad m_R v_k > v_k. \tag{3.146}$$

Since the line search (Algorithm 3.25) terminates with condition (3.117) due to assumption, we obtain

$$\begin{aligned}
-\alpha_{x_k + t_L d_k}^{x_k + t_L d_k} + d_k^T g_{x_k + t_L d_k}^{x_k + t_L d_k} &\stackrel{(3.141)}{\geq} m_R v_k \\
&\stackrel{(3.146)}{>} v_k \\
&\stackrel{(3.131)}{\geq} \hat{v}_k,
\end{aligned}$$

i.e. d_k is sufficiently infeasible for the new QCQP (3.68) (with respect to the approximation of the objective function) at iteration $k + 1$.

3. Because of

$$\begin{aligned}
F &\stackrel{(3.118)}{=} F(x_k + t d_k) + (t_L - t) \hat{g}^T d_k + \frac{1}{2} \hat{\rho} (t_L - t)^2 d_k^T \hat{G} d_k \\
&= F(x_k + t d_k) + \hat{g}^T ((x_k + t_L d_k) - (x_k + t d_k)) \\
&\quad + \frac{1}{2} \hat{\rho} ((x_k + t_L d_k) - (x_k + t d_k))^T \hat{G} ((x_k + t_L d_k) - (x_k + t d_k)) \\
&\stackrel{(3.122)}{=} F(y_{k+1}) + \hat{g}^T (x_{k+1} - y_{k+1}) + \frac{1}{2} \hat{\rho} (x_{k+1} - y_{k+1})^T \hat{G} (x_{k+1} - y_{k+1}) \\
&\left\{ \begin{array}{l} \stackrel{(3.101)}{=} F_{k+1}^{k+1} \\ \stackrel{(3.2)}{=} F_{y_{k+1}}^\#(x_{k+1}) \stackrel{(3.122)}{=} F_{x_k + t_L d_k}^\#(x_k + t_L d_k) \end{array} \right. \tag{3.147}
\end{aligned}$$

and

$$\begin{aligned}
\hat{\beta} &\stackrel{(3.119)}{=} \max(|F - F(x + t_L d)|, \gamma_2 |t_L - t|^\omega |d_k|^{\omega_2}) \\
&= \max(|F - F(x + t_L d)|, \gamma_2 |x_k + t_L d_k - (x_k + t d_k)|^{\omega_2}) \\
&\stackrel{(3.122)}{=} \max(|F - F(x_{k+1})|, \gamma_2 |x_{k+1} - y_{k+1}|^{\omega_2}) \\
&\stackrel{(3.97)}{=} \max(|F - F(x_{k+1})|, \gamma_2 (\hat{s}_{k+1}^{k+1})^{\omega_2}) \\
&\stackrel{(3.147)}{=} \max(|F_{k+1}^{k+1} - F(x_{k+1})|, \gamma_2 (\hat{s}_{k+1}^{k+1})^{\omega_2}) \\
&\stackrel{(3.66)}{=} A_{k+1}^{k+1}, \tag{3.148}
\end{aligned}$$

we obtain (3.142)

$$\begin{aligned}
&m_F \cdot \left(-\frac{1}{2} d_k \bar{\tilde{G}}_{x_k + t d_k} d_k\right) \\
&\stackrel{(3.121)}{\leq} F(x_k + t_L d_k) - \hat{\beta} + d_k^T (\hat{g} + \hat{\rho}(t_L - t) \hat{G} d_k) \\
&\stackrel{(3.148)}{=} F(x_k + t_L d_k) - A_{k+1}^{k+1} + d_k^T (\hat{g} + \hat{\rho} \hat{G}(t_L d_k - t d_k)) \\
&= F(x_k + t_L d_k) - A_{k+1}^{k+1} + d_k^T (\hat{g} + \hat{\rho} \hat{G}(x_k + t_L d_k - (x_k + t d_k))) \\
&\stackrel{(3.122)}{=} F(x_{k+1}) - A_{k+1}^{k+1} + d_k^T (\hat{g} + \hat{\rho} \hat{G}(x_{k+1} - y_{k+1})) \\
&\left\{ \begin{array}{l} \stackrel{(3.107)}{=} F(x_{k+1}) - A_{k+1}^{k+1} + d_k^T g_{k+1}^{k+1} \\ \stackrel{(3.4), (3.122)}{\downarrow} F(x_k + t_L d_k) - A_{x_k + t d_k}^{x_k + t_L d_k} + d_k^T \hat{g}_{x_k + t d_k}^\#(x_k + t_L d_k) \end{array} \right.
\end{aligned}$$

4. Due to the initialization of Algorithm 3.20 and (3.138), we have

$$0 < m_F < 1 \quad \cdot \left(-\frac{1}{2} d_k \bar{\tilde{G}}_{x_k + t d_k} d_k\right) < 0 \quad \xrightarrow{\quad} \quad m_F \cdot \left(-\frac{1}{2} d_k \bar{\tilde{G}}_{x_k + t d_k} d_k\right) > -\frac{1}{2} d_k \bar{\tilde{G}}_{x_k + t d_k} d_k. \tag{3.149}$$

Since the line search (Algorithm 3.25) terminates with condition (3.121) due to assumption, we obtain

$$\begin{aligned}
F(x_k + t_L d_k) - A_{x_k + t d_k}^{x_k + t_L d_k} + d_k^T g_{x_k + t d_k}^{x_k + t_L d_k} &\stackrel{(3.142)}{\geq} m_F \cdot \left(-\frac{1}{2} d_k \bar{\tilde{G}}_{x_k + t d_k} d_k\right) \\
&\stackrel{(3.149)}{>} -\frac{1}{2} d_k \bar{\tilde{G}}_{x_k + t d_k} d_k,
\end{aligned}$$

i.e. d_k is sufficiently infeasible for the new QCQP (3.68) (with respect to the approximation of the constraint) at iteration $k + 1$.

5. (3.143) holds due to

$$|y_{k+1} - x_{k+1}| \stackrel{(3.122)}{=} (t - t_L^k) |d_k| \stackrel{(3.117), (3.121)}{\downarrow} \leq C_S. \quad \square$$

3.5 Convergence of the line search

For proving the convergence of the line search (Algorithm 3.25) we have to identify a large subclass of locally Lipschitz continuous functions, which is the class of weakly upper semismooth functions (that

contains, e.g., functions that are the pointwise maximum of finitely many continuously differentiable functions due to MIFFLIN [75, p. 963, Theorem 2]).

Definition 3.30. A locally Lipschitz continuous function $f : \mathbb{R}^N \rightarrow \mathbb{R}$ is called weakly upper semismooth, if

$$\limsup_{i \rightarrow \infty} \bar{g}_i^T d \geq \liminf_{i \rightarrow \infty} \frac{f(x_k + t_i d) - f(x_k)}{t_i} \quad (3.150)$$

holds for all $x \in \mathbb{R}^N$, $d \in \mathbb{R}^N$, $\{\bar{g}_i\}_i \subset \mathbb{R}^N$ with $\bar{g}_i \in \partial f(x + t_i d)$ and $\{t_i\}_i \subset \mathbb{R}_+$ with $t_i \searrow 0$.

Proposition 3.31. Let $f : \mathbb{R}^N \rightarrow \mathbb{R}$ be weakly upper semismooth, then the line search (Algorithm 3.25) terminates after finitely many steps with

$$t_L^k = t_L, \quad t_R^k = t, \quad t_0^k > 0.$$

Proof. If $F(x_k + t d_k) < 0$ for all $t \in [0, 1]$, then this is exactly the same situation as in the line search of the unconstrained bundle-Newton method which terminates after finitely many iterations due to LUKŠAN & VLČEK [64, p. 379, Proof of Lemma 2.3].

Otherwise, since F is continuous and $F(x_k) < 0$, there exists a largest $\tilde{t} > 0$ with

$$\begin{aligned} F(x_k + d_k \tilde{t}) &= 0 \\ F(x_k + d_k s) &< 0 \quad \text{for all } s < \tilde{t}. \end{aligned}$$

Therefore, after sufficiently many iterations in the line search (Algorithm 3.25) (Note that the interval $[t_L, t_U]$ is shrinking at each iteration of the line search due to (3.114)), there only occur t_L, t_0, t_U with

$$\begin{aligned} 0 &\leq t_L < t_U < \tilde{t} \\ 0 &< t_0 < t_U < \tilde{t} \end{aligned}$$

(i.e. from now on all $x_k + t d_k$ with $t \in \{t_L, t_U\}$ are feasible) and consequently t_0 (where $x_k + t_0 d_k$ is also feasible,) does not change anymore (cf. (3.113)). Hence, here we also have exactly the same situation as in the line search of the unconstrained bundle-Newton method which terminates after finitely many iterations due to LUKŠAN & VLČEK [64, p. 379, Proof of Lemma 2.3]. $\square$

Remark 3.32. The proof of Proposition 3.31 only relies on f satisfying (3.150), the continuity of F and the strict feasibility of x_k . In particular, F does not need to be weakly upper semismooth.

3.6 Global convergence

Assumption 3.33. For investigating the global convergence of Algorithm 3.20 we will follow closely the proof of global convergence of the bundle-Newton method for nonsmooth unconstrained minimization in LUKŠAN & VLČEK [64, p. 380-385, Section 3] with modifications which concern the constrained case and the use of determining the search direction by solving a QCQP, where we will work out everything in great detail so that it is obvious to see which passages of the proof are similar to the

unconstrained case resp. which passages require a careful examination. Therefore, we set

$$\begin{aligned} \varepsilon &= 0 \\ (\lambda_j^k, \mu_j^k) &= 0 \quad \text{for all } j \notin J_k. \end{aligned} \tag{3.151}$$

A main difference to the proof of convergence of the unconstrained bundle-Newton method is that here H_k from (3.71) depends on the Lagrange multipliers (μ^k, μ_p^k) of the QCQP (3.68), which implies that so do the search direction d_k from (3.69) (and consequently the new iteration point x_{k+1} from (3.87) as well as the new trial point y_{k+1} from (3.88)) and the termination criterion w_k from (3.86) in particular. Furthermore, this dependence does not allow us to achieve the equality $H_{k+1} = H_k$ in the proof of Theorem 3.55 in contrast to LUKŠAN & VLČEK [64, top of page 385, Proof of Theorem 3.8], which extends the complexity of the already quite involved proof of the unconstrained bundle-Newton method.

Hence we give a brief overview of the main steps of the proof:

1. Proposition 3.34 serves as an extension of LUKŠAN & VLČEK [64, Lemma 3.1] in the constrained case for expressing the p -tilde data (as, e.g., $\tilde{g}_p^k, \tilde{\hat{g}}_p^k, \dots$) as convex combinations in which no p -data (as, e.g., $g_p^k, \hat{g}_p^k, \dots$) occurs.
2. In Proposition 3.35 we recall LUKŠAN & VLČEK [64, Lemma 3.2], which offers a sufficient condition to identify a vector as an element of the subdifferential.
3. In Theorem 3.36 we show that if Algorithm 3.20 stops at iteration k , then the current iteration point x_k is stationary for the optimization problem (2.35) (cf. LUKŠAN & VLČEK [64, Lemma 3.3]).
4. From then on we assume that the algorithm does not terminate (cf. (3.189)).
5. After summarizing some properties of positive definite matrices, we deduce bounds for $\{(\bar{W}_p^k)^{-1}\}$ and $\{\bar{W}_p^k + \bar{\kappa}^{k+1} \bar{G}^k\}$ in Corollary 3.43, which will be essential in the following.
6. In Proposition 3.44 we show that if some boundedness assumptions are satisfied and the limit inferior of the sequence $\{\max(w_k, |x_k - \bar{x}|\}\}$ is zero, where $\bar{x}$ denotes any accumulation point of the sequence of iteration points $\{x_k\}$, then $\bar{x}$ is stationary for the optimization problem (2.35), where the proof relies on Carathéodory's theorem (cf., e.g., NEUMAIER [82]) as well as on the local boundedness (2.32) and the upper semicontinuity (2.33) of the subdifferentials ∂f and ∂F (cf. LUKŠAN & VLČEK [64, Lemma 3.6]).
7. The statement $t_L^k v_k \rightarrow 0$ in Proposition 3.45 is taken without any modification from LUKŠAN & VLČEK [64, Lemma 3.5(ii)], since that proof only depends on the negativity of v_k , which holds due to (3.131).
8. Although the properties of the shifted sequences $\{x_{k+i}\}$, $\{w_{k+i}\}$ and $\{t_L^{k+i}\}$ in Proposition 3.46 look very similar to those in LUKŠAN & VLČEK [64, Lemma 3.7], we have to take care of the dependence of (μ^k, μ_p^k) , which we noticed before, in the proof.

9. The estimation of a certain quadratic function on the interval $[0, 1]$ in Proposition 3.47 is directly taken from LUKŠAN & VLČEK [64, p. 381, Lemma 3.4].
10. After recalling the differentiability of matrix valued functions to give a formula for the derivative of the matrix square root in Proposition 3.51 and after formulating the mean value theorem for vector valued functions on a convex set in Theorem 3.53, we combine these two results to obtain a Lipschitz estimate for the inverse matrix square root in Proposition 3.54, which serves as replacement for the property $H_{k+1} = H_k$ of the proof of the unconstrained bundle-Newton method as mentioned above.
11. Finally, we prove that under some additional boundedness assumptions the limit inferior of the sequence $\{\max(w_k, |x_k - \bar{x}|)\}$ is always zero and therefore Proposition 3.44 yields Theorem 3.55 (cf. LUKŠAN & VLČEK [64, Theorem 3.8]) which states that each accumulation point $\bar{x}$ of the sequence of iteration points $\{x_k\}$ is stationary for the optimization problem (2.35).

Proposition 3.34. *If Algorithm 3.20 has not stopped before iteration k with $k \geq 1$, then the following statements hold:*

- There exist $\hat{\lambda}_j^k \in \mathbb{R}$ for $j = 1, \dots, k$ with

$$1 = \sum_{j=1}^k \hat{\lambda}_j^k, \quad \hat{\lambda}_j^k \geq 0 \quad (3.152)$$

$$G_p^{k+1} = \sum_{j=1}^k \hat{\lambda}_j^k \rho_j G_j \quad (3.153)$$

$$\tilde{g}_p^k = \sum_{j=1}^k \hat{\lambda}_j^k g_j^k \quad (3.154)$$

$$\tilde{s}_p^k = \sum_{j=1}^k \hat{\lambda}_j^k s_j^k. \quad (3.155)$$

- If $\bar{\kappa}^{k+1} > 0$, then there exist $\hat{\kappa}_j^k \in \mathbb{R}$ for $j = 1, \dots, k$ with

$$1 = \sum_{j=1}^k \hat{\kappa}_j^k, \quad \hat{\kappa}_j^k \geq 0 \quad (3.156)$$

$$\hat{G}_p^{k+1} = \sum_{j=1}^k \hat{\kappa}_j^k \hat{\rho}_j \hat{G}_j \quad (3.157)$$

$$\tilde{\hat{g}}_p^k = \sum_{j=1}^k \hat{\kappa}_j^k \hat{g}_j^k \quad (3.158)$$

$$\tilde{\hat{s}}_p^k = \sum_{j=1}^k \hat{\kappa}_j^k \hat{s}_j^k. \quad (3.159)$$

- If $\bar{\kappa}^{k+1} = 0$, then (3.157), (3.158) and (3.159) hold with

$$\hat{\kappa}_j^k := 0 \quad \text{for all } j = 1, \dots, k. \quad (3.160)$$

Proof. (by induction)

1. *Claim.:* For iteration $k = 1$ of Algorithm 3.20 we have

$$\begin{aligned}
 \lambda_1^k &= 1 \\
 \lambda_p^k &= 0 \\
 \bar{\kappa}^{k+1} &= \mu_1^k \\
 \kappa_1^k &= \begin{cases} 1 & \text{for } \bar{\kappa}^{k+1} > 0 \\ 0 & \text{for } \bar{\kappa}^{k+1} = 0 \end{cases} \\
 \kappa_p^k &= 0 .
 \end{aligned} \tag{3.161}$$

Proof of claim.: Since

$$\begin{aligned}
 g_p^1 &\stackrel{(3.53)}{=} g_1^1 \\
 \alpha_p^1 &\stackrel{(3.65)}{=} \max(|f_p^1 - f(x_1)|, \gamma_1(s_p^1)^{\omega_1}) \\
 &\stackrel{(3.52), (3.57)}{\leq} \max(|f_1^1 - f(x_1)|, \gamma_1(s_1^1)^{\omega_1}) \\
 &\stackrel{(3.64)}{=} \alpha_1^1
 \end{aligned}$$

and

$$\begin{aligned}
 \hat{g}_p^1 &\stackrel{(3.55)}{=} \hat{g}_1^1 \\
 A_p^1 &\stackrel{(3.67)}{=} \max(|F_p^1 - F(x_1)|, \gamma_2(\hat{s}_p^1)^{\omega_2}) \\
 &\stackrel{(3.54), (3.57)}{\leq} \max(|F_1^1 - F(x_1)|, \gamma_2(\hat{s}_1^1)^{\omega_2}) \\
 &\stackrel{(3.66)}{=} A_1^1 \\
 \bar{G}^1 &\stackrel{(3.61), (3.62), (3.56)}{\leq} \bar{G}_1^1 ,
 \end{aligned}$$

the QCQP (3.68) at iteration $k = 1$ reads

$$\begin{aligned}
 \min_{d, \hat{v}} \quad & \hat{v} + \frac{1}{2} d^T \bar{W}_p^k d , \\
 \text{s.t.} \quad & -\alpha_1^1 + d^T g_1^1 \leq \hat{v} && \text{for } j \in J_k \\
 & -\alpha_1^1 + d^T g_1^1 \leq \hat{v} && \text{if } i_s \leq i_r \\
 & F(x_1) - A_1^1 + d^T \hat{g}_1^1 + \frac{1}{2} d^T \bar{G}_1^1 d \leq 0 && \text{for } j \in J_k \\
 & F(x_1) - A_1^1 + d^T \hat{g}_1^1 + \frac{1}{2} d^T \bar{G}_1^1 d \leq 0 && \text{if } i_s \leq i_r .
 \end{aligned}$$

Since the aggregated (p-)constraint coincides with the corresponding bundle constraint, we can drop the aggregated (p-)constraint and consequently the dual problem (3.42) has only two variables λ_1^1 and μ_1^1 , where

$$\lambda_1^1 = 1$$

must hold, so that the equality constraint of the dual problem (3.42) is satisfied. Now, if we set

$$\begin{aligned}
 \lambda_p^1 &= 0 \\
 \mu_p^1 &= 0 ,
 \end{aligned}$$

then the dual solution does not change. Furthermore, we have

$$\bar{\kappa}^2 \stackrel{(3.72)}{=} \mu_1^1 + \mu_p^1 = \mu_1^1$$

and therefore we obtain

$$\begin{aligned} \kappa_j^1 &\stackrel{(3.73)}{=} \begin{cases} 1 & \text{for } \bar{\kappa}^2 > 0 \\ 0 & \text{for } \bar{\kappa}^2 = 0 \end{cases} \\ \kappa_p^1 &\stackrel{(3.74)}{=} 0 . \end{aligned} \quad \blacksquare$$

2. *Claim.:* The base case is satisfied for $k = 1$ with

$$\hat{\lambda}_1^k := 1 \tag{3.162}$$

$$\hat{\kappa}_1^k := \begin{cases} 1 & \text{for } \bar{\kappa}^{k+1} > 0 \\ 0 & \text{for } \bar{\kappa}^{k+1} = 0 . \end{cases} \tag{3.163}$$

Proof of claim.:

- For $k = 1$, (3.152) holds due to (3.162), and, furthermore, (3.156) holds due to (3.163).
- For $k = 1$, (3.153) holds because of

$$\begin{aligned} \hat{\lambda}_1^k \rho_1 G_1 &\stackrel{(3.162)}{=} \rho_1 G_1 \\ &\stackrel{(3.161)}{=} \lambda_1^k \rho_1 G_1 + \lambda_p^k G_p^1 \\ &\stackrel{(3.77)}{=} G_p^2 \end{aligned}$$

and (3.157) holds because of

$$\begin{aligned} \hat{\kappa}_1^k \hat{\rho}_1 \hat{G}_1 &\stackrel{(3.163)}{=} \begin{cases} \rho_1 \hat{G}_1 & \text{for } \bar{\kappa}^{k+1} > 0 \\ 0 & \text{for } \bar{\kappa}^{k+1} = 0 \end{cases} \\ &\stackrel{(3.161)}{=} \kappa_1^k \hat{\rho}_1 \hat{G}_1 + \kappa_p^k \hat{G}_p^1 \\ &\stackrel{(3.82)}{=} \hat{G}_p^2 . \end{aligned}$$

- For $k = 1$, (3.154) holds because of

$$\begin{aligned} \hat{\lambda}_1^k g_1^k &\stackrel{(3.162)}{=} g_1^k \\ &\stackrel{(3.161)}{=} \lambda_1^k g_1^k + \lambda_p^k g_p^k \\ &\stackrel{(3.76)}{=} \hat{g}_p^1 \end{aligned}$$

and (3.158) holds because of

$$\begin{aligned} \hat{\kappa}_1^k \hat{g}_1^k &\stackrel{(3.163)}{=} \begin{cases} \hat{g}_1^k & \text{for } \bar{\kappa}^{k+1} > 0 \\ 0 & \text{for } \bar{\kappa}^{k+1} = 0 \end{cases} \\ &\stackrel{(3.161)}{=} \kappa_1^k \hat{g}_1^k + \kappa_p^k \hat{g}_p^k \\ &\stackrel{(3.81)}{=} \hat{\tilde{g}}_p^1 . \end{aligned}$$

- For $k = 1$, (3.155) holds because of

$$\begin{aligned}\hat{\lambda}_1^k s_1^k &\stackrel{(3.162)}{=} s_1^k \\ &\stackrel{(3.161)}{=} \lambda_1^k s_1^k + \lambda_p^k s_p^k \\ &\stackrel{(3.78)}{=} \tilde{s}_p^1\end{aligned}$$

and (3.159) holds because of

$$\begin{aligned}\hat{\kappa}_1^k \hat{s}_1^k &\stackrel{(3.163)}{=} \begin{cases} \hat{s}_1^k & \text{for } \bar{\kappa}^{k+1} > 0 \\ 0 & \text{for } \bar{\kappa}^{k+1} = 0 \end{cases} \\ &\stackrel{(3.161)}{=} \kappa_1^k \hat{s}_1^k + \kappa_p^k \hat{s}_p^k \\ &\stackrel{(3.83)}{=} \tilde{s}_p^1.\end{aligned}\quad \blacksquare$$

3. *Claim.*: Let the induction hypothesis be satisfied, i.e. (3.152), (3.153), (3.154), (3.155) hold as well as in the case $\bar{\kappa}^{k+1} > 0$ (3.156), (3.157), (3.158), (3.159) hold, then the inductive step $k \mapsto k + 1$ is true with

$$\hat{\lambda}_j^{k+1} := \lambda_j^{k+1} + \lambda_p^{k+1} \hat{\lambda}_j^k \quad \text{for } j = 1, \dots, k \quad (3.164)$$

$$\hat{\lambda}_{k+1}^{k+1} := \lambda_{k+1}^{k+1} \quad (3.165)$$

$$\hat{\kappa}_j^{k+1} := \begin{cases} \kappa_j^{k+1} + \kappa_p^{k+1} \hat{\kappa}_j^k & \text{for } \bar{\kappa}^{k+1} > 0 \\ \kappa_j^{k+1} & \text{for } \bar{\kappa}^{k+1} = 0 \end{cases} \quad \text{for } j = 1, \dots, k \quad (3.166)$$

$$\hat{\kappa}_{k+1}^{k+1} := \kappa_{k+1}^{k+1}, \quad (3.167)$$

where $(\lambda_j^{k+1}, \lambda_p^{k+1}, \mu_j^{k+1}, \mu_p^{k+1})$ is the solution of the dual problem (3.42) (including the aggregated terms) (this is well defined, as the inductive assumption tells us that Algorithm 3.20 has not stopped before iteration $k + 1$ and the search direction as well as the Lagrange multipliers have already been computed at step 3, but the termination criterion is first checked at step 5 (at each iteration)) and κ_j^{k+1} resp. κ_p^{k+1} are set according to (3.73) resp. (3.74).

Proof of claim.:

- In the case $\bar{\kappa}^{k+1} = 0$ we have

$$\bar{\kappa}^{k+1} = 0 \stackrel{(3.72), (3.73), (3.74)}{\iff} (\kappa_j^k = 0 \quad \text{for all } j = 1, \dots, k) \quad \wedge \quad \kappa_p^k = 0 \quad (3.168)$$

and therefore we obtain

$$\tilde{g}_p^k \stackrel{(3.81)}{=} \sum_{j \in J_k} \kappa_j^k \hat{g}_j^k + \kappa_p^k \hat{g}_p^k \stackrel{(3.168)}{=} 0 \quad (3.169)$$

$$\hat{G}_p^{k+1} \stackrel{(3.82)}{=} \sum_{j \in J_k} \kappa_j^k \hat{\rho}_j \hat{G}_j + \kappa_p^k \hat{G}_p^k \stackrel{(3.168)}{=} 0 \quad (3.170)$$

$$\hat{g}_p^{k+1} \stackrel{(3.108)}{=} \tilde{g}_p^k + \hat{G}_p^{k+1}(x_{k+1} - x_k) \stackrel{(3.169), (3.170)}{\stackrel{\downarrow}{=}} 0. \quad (3.171)$$

Hence, at iteration $k + 1$ in the QCQP (3.68) the aggregated constraint for F reads in the case $i_s \leq i_r$

$$F(x_{k+1}) - A_p^{k+1} + d^T \hat{g}_p^{k+1} + \frac{1}{2} d^T \bar{G}^{k+1} d \leq 0 \stackrel{(3.170), (3.171)}{\iff} F(x_{k+1}) - A_p^{k+1} \leq 0. \quad (3.172)$$

Since the second inequality from (3.172) is sharp

$$F(x_{k+1}) - \underbrace{A_p^{k+1}}_{\substack{(3.67) \\ \geq 0}} \leq F(x_{k+1}) \stackrel{(3.123)}{<} 0 ,$$

the aggregated constraint for F is inactive at iteration $k+1$. Since Lagrange multipliers for inactive constraints vanish, we obtain at iteration $k+1$ (Note that μ_p^{k+1} is the Lagrange multiplier corresponding to the aggregated constraint for F at iteration $k+1$ and note that $\bar{\kappa}^{k+2} > 0$ is the assumption for what we want to show by the inductive step $k \mapsto k+1$)

$$\begin{aligned} \mu_p^{k+1} &= 0 \\ \bar{\kappa}^{k+2} > 0, (3.74) \\ \iff \kappa_p^{k+1} &= 0 \end{aligned} \tag{3.173}$$

$$\begin{aligned} \bar{\kappa}^{k+2} > 0, (3.72), (3.73), (3.74) \\ \iff \sum_{j=1}^{k+1} \kappa_j^{k+1} &= 1 \quad \wedge \quad (\kappa_j^{k+1} \geq 0 \text{ for all } j = 1, \dots, k+1) . \end{aligned} \tag{3.174}$$

In the case $i_s > i_r$ (3.173) and (3.174) hold anyway, since then in the dual problem (3.42) for the QCQP (3.68) the aggregated constraints do not occur and therefore the corresponding Lagrange multiplier can be set to zero.

- We prove the inductive step for (3.152) by

$$\begin{aligned} \hat{\lambda}_j^{k+1} &\stackrel{(3.164)}{=} \underbrace{\lambda_j^{k+1}}_{\substack{(3.42) \\ \geq 0}} + \underbrace{\lambda_p^{k+1}}_{\substack{(3.42) \\ \geq 0}} \underbrace{\hat{\lambda}_j^k}_{\substack{(3.152) \\ \geq 0}} \geq 0 \quad \text{for } j = 1, \dots, k \\ \hat{\lambda}_{k+1}^{k+1} &\stackrel{(3.165)}{=} \lambda_{k+1}^{k+1} \stackrel{(3.42)}{\geq} 0 \end{aligned}$$

and

$$\begin{aligned} \sum_{j=1}^{k+1} \hat{\lambda}_j^{k+1} &\stackrel{(3.164), (3.165)}{=} \sum_{j=1}^k \lambda_j^{k+1} + \lambda_p^{k+1} \hat{\lambda}_j^k + \lambda_{k+1}^{k+1} \\ &\stackrel{(3.152)}{=} \sum_{j=1}^{k+1} \lambda_j^{k+1} + \lambda_p^{k+1} \\ &\stackrel{(3.42)}{=} 1 . \end{aligned}$$

In the case $\bar{\kappa}^{k+1} > 0$ we prove the inductive step for (3.156) by

$$\begin{aligned} \hat{\kappa}_j^{k+1} &\stackrel{(3.166)}{=} \underbrace{\kappa_j^{k+1}}_{\substack{(3.73), (3.72), (3.42) \\ \geq 0}} + \underbrace{\kappa_p^{k+1}}_{\substack{(3.74), (3.72), (3.42) \\ \geq 0}} \underbrace{\hat{\kappa}_j^k}_{\substack{(3.156) \\ \geq 0}} \geq 0 \\ \hat{\kappa}_{k+1}^{k+1} &\stackrel{(3.167)}{=} \kappa_{k+1}^{k+1} \stackrel{(3.73), (3.72), (3.42)}{\geq} 0 \end{aligned}$$

and (Note that we assumed that we consider the case $\bar{\kappa}^{k+1} > 0$ which implies that we can use the induction hypothesis for (3.156) and note that we have $\bar{\kappa}^{k+2} > 0$, since this is the

assumption for what we want to show by the inductive step $k \mapsto k+1$)

$$\begin{aligned}
\sum_{j=1}^{k+1} \hat{\kappa}_j^{k+1} &\stackrel{(3.166), (3.167)}{=} \sum_{j=1}^k \kappa_j^{k+1} + \kappa_p^{k+1} \hat{\kappa}_j^k + \kappa_{k+1}^{k+1} \\
&\stackrel{(3.156)}{=} \sum_{j=1}^{k+1} \kappa_j^{k+1} + \kappa_p^{k+1} \\
&\stackrel{(3.73), (3.74), (3.72)}{=} 1.
\end{aligned}$$

In the case $\bar{\kappa}^{k+1} = 0$ we prove the inductive step for (3.156) by

$$\begin{aligned}
\hat{\kappa}_j^{k+1} &\stackrel{(3.166)}{=} \kappa_j^{k+1} \stackrel{(3.174)}{\geq} 0 \\
\hat{\kappa}_{k+1}^{k+1} &\stackrel{(3.167)}{=} \kappa_{k+1}^{k+1} \stackrel{(3.174)}{\geq} 0
\end{aligned}$$

and

$$\begin{aligned}
\sum_{j=1}^{k+1} \hat{\kappa}_j^{k+1} &\stackrel{(3.166), (3.167)}{=} \sum_{j=1}^k \kappa_j^{k+1} + \kappa_{k+1}^{k+1} \\
&\stackrel{(3.174)}{=} 1.
\end{aligned}$$

- We prove the inductive step for (3.153) by

$$\begin{aligned}
G_p^{(k+1)+1} &\stackrel{(3.77)}{=} \sum_{j=1}^{k+1} \lambda_j^{k+1} \rho_j G_j + \lambda_p^{k+1} G_p^{k+1} \\
&\stackrel{(3.153)}{=} \sum_{j=1}^{k+1} \lambda_j^{k+1} \rho_j G_j + \lambda_p^{k+1} \left(\sum_{j=1}^k \hat{\lambda}_j^k \rho_j G_j \right) \\
&\stackrel{(3.165)}{=} \hat{\lambda}_{k+1}^{k+1} \rho_{k+1} G_{k+1} + \sum_{j=1}^k (\lambda_j^{k+1} + \lambda_p^{k+1} \hat{\lambda}_j^k) \rho_j G_j \\
&\stackrel{(3.164)}{=} \sum_{j=1}^{k+1} \hat{\lambda}_j^{k+1} \rho_j G_j.
\end{aligned}$$

In the case $\bar{\kappa}^{k+1} > 0$ we prove the inductive step for (3.157) by

$$\begin{aligned}
\hat{G}_p^{(k+1)+1} &\stackrel{(3.82)}{=} \sum_{j=1}^{k+1} \kappa_j^{k+1} \hat{\rho}_j \hat{G}_j + \kappa_p^{k+1} \hat{G}_p^{k+1} \\
&\stackrel{(3.157)}{=} \sum_{j=1}^{k+1} \kappa_j^{k+1} \hat{\rho}_j \hat{G}_j + \kappa_p^{k+1} \left(\sum_{j=1}^k \hat{\kappa}_j^k \hat{\rho}_j \hat{G}_j \right) \\
&\stackrel{(3.167)}{=} \hat{\kappa}_{k+1}^{k+1} \hat{\rho}_{k+1} \hat{G}_{k+1} + \sum_{j=1}^k (\kappa_j^{k+1} + \kappa_p^{k+1} \hat{\kappa}_j^k) \hat{\rho}_j \hat{G}_j \\
&\stackrel{(3.166)}{=} \sum_{j=1}^{k+1} \hat{\kappa}_j^{k+1} \hat{\rho}_j \hat{G}_j.
\end{aligned}$$

In the case $\bar{\kappa}^{k+1} = 0$ we prove the inductive step for (3.157) by

$$\begin{aligned}
 \hat{G}_p^{(k+1)+1} &\stackrel{(3.82)}{=} \sum_{j=1}^{k+1} \kappa_j^{k+1} \hat{\rho}_j \hat{G}_j + \kappa_p^{k+1} \hat{G}_p^{k+1} \\
 &\stackrel{(3.173)}{=} \sum_{j=1}^{k+1} \kappa_j^{k+1} \hat{\rho}_j \hat{G}_j \\
 &\stackrel{(3.166), (3.167)}{\Downarrow} \sum_{j=1}^{k+1} \hat{\kappa}_j^{k+1} \hat{\rho}_j \hat{G}_j .
 \end{aligned}$$

- We prove the inductive step for (3.154) by

$$\begin{aligned}
 \tilde{g}_p^{k+1} &\stackrel{(3.76)}{=} \sum_{j=1}^{k+1} \lambda_j^{k+1} g_j^{k+1} + \lambda_p^{k+1} g_p^{k+1} \\
 &\stackrel{(3.105)}{=} \sum_{j=1}^{k+1} \lambda_j^{k+1} g_j^{k+1} + \lambda_p^{k+1} (\tilde{g}_p^k + G_p^{k+1}(x_{k+1} - x_k)) \\
 &\stackrel{(3.153), (3.154)}{\Downarrow} \sum_{j=1}^{k+1} \lambda_j^{k+1} g_j^{k+1} + \lambda_p^{k+1} \left(\sum_{j=1}^k \hat{\lambda}_j^k g_j^k + \sum_{j=1}^k \hat{\lambda}_j^k \rho_j G_j(x_{k+1} - x_k) \right) \\
 &\stackrel{(3.103)}{=} \sum_{j=1}^{k+1} \lambda_j^{k+1} g_j^{k+1} + \sum_{j=1}^k \lambda_p^{k+1} \hat{\lambda}_j^k g_j^{k+1} \\
 &\stackrel{(3.165)}{=} \hat{\lambda}_{k+1}^{k+1} g_{k+1}^{k+1} + \sum_{j=1}^k (\lambda_j^{k+1} + \lambda_p^{k+1} \hat{\lambda}_j^k) g_j^{k+1} \\
 &\stackrel{(3.164)}{=} \sum_{j=1}^{k+1} \hat{\lambda}_j^{k+1} g_j^{k+1} .
 \end{aligned}$$

In the case $\bar{\kappa}^{k+1} > 0$ we prove the inductive step for (3.158) by

$$\begin{aligned}
 \tilde{\hat{g}}_p^{k+1} &\stackrel{(3.81)}{=} \sum_{j=1}^{k+1} \kappa_j^{k+1} \hat{g}_j^{k+1} + \kappa_p^{k+1} \hat{g}_p^{k+1} \\
 &\stackrel{(3.108)}{=} \sum_{j=1}^{k+1} \kappa_j^{k+1} \hat{g}_j^{k+1} + \kappa_p^{k+1} (\tilde{\hat{g}}_p^k + \hat{G}_p^{k+1}(x_{k+1} - x_k)) \\
 &\stackrel{(3.157), (3.158)}{\Downarrow} \sum_{j=1}^{k+1} \kappa_j^{k+1} \hat{g}_j^{k+1} + \kappa_p^{k+1} \left(\sum_{j=1}^k \hat{\kappa}_j^k \hat{g}_j^k + \sum_{j=1}^k \hat{\kappa}_j^k \hat{\rho}_j \hat{G}_j(x_{k+1} - x_k) \right) \\
 &\stackrel{(3.106)}{=} \sum_{j=1}^{k+1} \kappa_j^{k+1} \hat{g}_j^{k+1} + \sum_{j=1}^k \kappa_p^{k+1} \hat{\kappa}_j^k \hat{g}_j^{k+1} \\
 &\stackrel{(3.167)}{=} \hat{\kappa}_{k+1}^{k+1} \hat{g}_{k+1}^{k+1} + \sum_{j=1}^k (\kappa_j^{k+1} + \kappa_p^{k+1} \hat{\kappa}_j^k) \hat{g}_j^{k+1} \\
 &\stackrel{(3.166)}{=} \sum_{j=1}^{k+1} \hat{\kappa}_j^{k+1} \hat{g}_j^{k+1} .
 \end{aligned}$$

In the case $\bar{\kappa}^{k+1} = 0$ we prove the inductive step for (3.158) by

$$\begin{aligned}
\tilde{g}_p^{k+1} &\stackrel{(3.81)}{=} \sum_{j=1}^{k+1} \kappa_j^{k+1} \hat{g}_j^{k+1} + \kappa_p^{k+1} \hat{g}_p^{k+1} \\
&\stackrel{(3.173)}{=} \sum_{j=1}^{k+1} \kappa_j^{k+1} \hat{g}_j^{k+1} \\
&\stackrel{(3.166), (3.167)}{\Downarrow} \sum_{j=1}^{k+1} \hat{\kappa}_j^{k+1} \hat{g}_j^{k+1} .
\end{aligned}$$

- We prove the inductive step for (3.155) by

$$\begin{aligned}
\tilde{s}_p^{k+1} &\stackrel{(3.78)}{=} \sum_{j=1}^{k+1} \lambda_j^{k+1} s_j^{k+1} + \lambda_p^{k+1} s_p^{k+1} \\
&\stackrel{(3.95)}{=} \sum_{j=1}^{k+1} \lambda_j^{k+1} s_j^{k+1} + \lambda_p^{k+1} (\tilde{s}_p^k + |x_{k+1} - x_k|) \\
&\stackrel{(3.155), (3.152)}{\Downarrow} \sum_{j=1}^{k+1} \lambda_j^{k+1} s_j^{k+1} + \lambda_p^{k+1} \left(\sum_{j=1}^k \hat{\lambda}_j^k s_j^k + \sum_{j=1}^k \hat{\lambda}_j^k |x_{k+1} - x_k| \right) \\
&\stackrel{(3.93)}{=} \lambda_{k+1}^{k+1} s_{k+1}^{k+1} + \sum_{j=1}^k \lambda_j^{k+1} s_j^{k+1} + \sum_{j=1}^k \lambda_p^{k+1} \hat{\lambda}_j^k s_j^{k+1} \\
&\stackrel{(3.165)}{=} \hat{\lambda}_{k+1}^{k+1} s_{k+1}^{k+1} + \sum_{j=1}^k (\lambda_j^{k+1} + \lambda_p^{k+1} \hat{\lambda}_j^k) s_j^{k+1} \\
&\stackrel{(3.164)}{=} \sum_{j=1}^{k+1} \hat{\lambda}_j^{k+1} s_j^{k+1} .
\end{aligned}$$

In the case $\bar{\kappa}^{k+1} > 0$ we prove the inductive step for (3.159) by

$$\begin{aligned}
\tilde{s}_p^{k+1} &\stackrel{(3.83)}{=} \sum_{j=1}^{k+1} \kappa_j^{k+1} \hat{s}_j^{k+1} + \kappa_p^{k+1} \hat{s}_p^{k+1} \\
&\stackrel{(3.98)}{=} \sum_{j=1}^{k+1} \kappa_j^{k+1} \hat{s}_j^{k+1} + \kappa_p^{k+1} (\tilde{s}_p^k + |x_{k+1} - x_k|) \\
&\stackrel{(3.159), (3.156)}{\Downarrow} \sum_{j=1}^{k+1} \kappa_j^{k+1} \hat{s}_j^{k+1} + \kappa_p^{k+1} \left(\sum_{j=1}^k \hat{\kappa}_j^k \hat{s}_j^k + \sum_{j=1}^k \hat{\kappa}_j^k |x_{k+1} - x_k| \right) \\
&\stackrel{(3.96)}{=} \kappa_{k+1}^{k+1} \hat{s}_{k+1}^{k+1} + \sum_{j=1}^k \kappa_j^{k+1} \hat{s}_j^{k+1} + \sum_{j=1}^k \kappa_p^{k+1} \hat{\kappa}_j^k \hat{s}_j^{k+1} \\
&\stackrel{(3.167)}{=} \hat{\kappa}_{k+1}^{k+1} \hat{s}_{k+1}^{k+1} + \sum_{j=1}^k (\kappa_j^{k+1} + \kappa_p^{k+1} \hat{\kappa}_j^k) \hat{s}_j^{k+1} \\
&\stackrel{(3.166)}{=} \sum_{j=1}^{k+1} \hat{\kappa}_j^{k+1} \hat{s}_j^{k+1} .
\end{aligned}$$

In the case $\bar{\kappa}^{k+1} = 0$ we prove the inductive step for (3.159) by

$$\begin{aligned} \tilde{s}_p^{k+1} &\stackrel{(3.83)}{=} \sum_{j=1}^{k+1} \kappa_j^{k+1} \hat{s}_j^{k+1} + \kappa_p^{k+1} \hat{s}_p^{k+1} \\ &\stackrel{(3.173)}{=} \sum_{j=1}^{k+1} \kappa_j^{k+1} \hat{s}_j^{k+1} \\ &\stackrel{(3.166), (3.167)}{\downarrow} \sum_{j=1}^{k+1} \hat{\kappa}_j^{k+1} \hat{s}_j^{k+1} . \end{aligned} \quad \blacksquare$$

4. In the case $\bar{\kappa}^{k+1} = 0$ we obtain

$$\bar{\kappa}^{k+1} = 0 \stackrel{(3.73), (3.74), (3.72)}{\iff} \kappa_j^k = 0 \text{ for all } j = 1, \dots, k \quad \wedge \quad \kappa_p^k = 0 \quad (3.175)$$

and therefore (3.157) holds

$$\hat{G}_p^{k+1} \stackrel{(3.82), (3.175)}{\downarrow} 0 \stackrel{(3.160)}{=} \sum_{j=1}^k \hat{\kappa}_j^k \hat{\rho}_j \hat{G}_j ,$$

(3.158) holds

$$\tilde{g}_p^k \stackrel{(3.81), (3.175)}{\downarrow} 0 \stackrel{(3.160)}{=} \sum_{j=1}^k \hat{\kappa}_j^k \hat{g}_j^k$$

and (3.159) holds

$$\tilde{s}_p^k \stackrel{(3.83), (3.175)}{\downarrow} 0 \stackrel{(3.160)}{=} \sum_{j=1}^k \hat{\kappa}_j^k \hat{s}_j^k . \quad \square$$

Proposition 3.35. *If $\bar{x} \in \mathbb{R}^N$ and there exists $\bar{G}_j \in \mathbb{R}_{\text{sym}}^{N \times N}$, $\bar{q}, \bar{y}_j \in \mathbb{R}^n$, $\bar{g}_j \in \partial f(y_j)$, $\bar{s}_j, \bar{\lambda}_j \in \mathbb{R}$ for $j = 1, \dots, L$, where $L \geq 1$, with*

$$1 = \sum_{j=1}^L \bar{\lambda}_j , \quad \bar{\lambda}_j \geq 0 \quad (3.176)$$

$$\bar{q} = \sum_{j=1}^L (\bar{g}_j + \bar{G}_j(\bar{x} - \bar{y}_j)) \bar{\lambda}_j \quad (3.177)$$

$$0 = \sum_{j=1}^L \bar{s}_j \bar{\lambda}_j \quad (3.178)$$

$$|\bar{y}_j - \bar{x}| \leq \bar{s}_j \quad \text{for all } j = 1, \dots, L , \quad (3.179)$$

then

$$\bar{q} \in \partial f(\bar{x}) . \quad (3.180)$$

Proof. LUKŠAN & VLČEK [64, p. 381, Proof of Lemma 3.2]. □

Theorem 3.36. *If Algorithm 3.20 stops at iteration k , then there exists $\bar{\kappa}^{k+1} \geq 0$ with*

$$\begin{aligned} 0 &\in \partial f(x_k) + \bar{\kappa}^{k+1} \partial F(x_k) \\ \bar{\kappa}^{k+1} F(x_k) &= 0 \\ F(x_k) &\leq 0 , \end{aligned}$$

i.e. x_k is stationary for the optimization problem (2.35) according to (2.36).

Proof.

1. *Claim.:* If Algorithm 3.20 stops at iteration k , then

$$\tilde{s}_p^k = 0 \quad (3.181)$$

$$\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k = 0 \quad (3.182)$$

$$\begin{aligned} \bar{\kappa}^{k+1} F(x_k) &= 0 \\ \bar{\kappa}^{k+1} &= 0 \text{ or } (\bar{\kappa}^{k+1} > 0 \wedge \tilde{s}_p^k = 0) . \end{aligned} \quad (3.183)$$

Proof of claim.: If Algorithm 3.20 stops at iteration k , we have at step 5

$$\begin{aligned} w_k &\leq \varepsilon \stackrel{(3.151)}{=} 0 \stackrel{(3.130)}{\implies} w_k = 0 \\ &\stackrel{(3.86), (3.79), \bar{\kappa}^{k+1} \geq 0, F(x_k) \leq 0}{\iff} \frac{1}{2} |H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k)|^2 = 0 \quad \wedge \quad \tilde{\alpha}_p^k = 0 \\ &\quad \wedge \quad \bar{\kappa}^{k+1} \tilde{A}_p^k = 0 \quad \wedge \quad \bar{\kappa}^{k+1} (-F(x_k)) = 0 \\ &\stackrel{H_k \text{ regular}, (3.79), (3.78)}{\implies} \tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k = 0 \quad \wedge \quad \tilde{s}_p^k = 0 \\ &\quad \wedge \quad \bar{\kappa}^{k+1} \tilde{A}_p^k = 0 \quad \wedge \quad \bar{\kappa}^{k+1} F(x_k) = 0 . \end{aligned} \quad (3.184)$$

If $\bar{\kappa}^{k+1} > 0$, then

$$\bar{\kappa}^{k+1} > 0 \stackrel{(3.184)}{\implies} \tilde{A}_p^k = 0 \stackrel{(3.84), (3.83)}{\implies} \tilde{s}_p^k = 0 . \quad \blacksquare$$

2. *Claim.:* For

$$\begin{aligned} \bar{x} &:= x_k & \bar{G}_j &:= \rho_j G_j & \bar{y}_j &:= y_j & \bar{s}_j &:= s_j^k \\ L &:= k & \bar{q} &:= \tilde{g}_p^k & \bar{g}_j &:= g_j & \bar{\lambda}_j &:= \hat{\lambda}_j^k \end{aligned} \quad (3.185)$$

with $j = 1, \dots, L$ the assumptions of Proposition 3.35 are satisfied and therefore

$$\tilde{g}_p^k \in \partial f(x_k) . \quad (3.186)$$

Proof of claim.:

- (3.176) holds because of (3.152).
- Since the algorithm stops at iteration k , it has not stopped before iteration k and therefore we can apply Proposition 3.34

$$\bar{q} \stackrel{(3.177)}{=} \sum_{j=1}^L (\bar{g}_j + \bar{G}_j(\bar{x} - \bar{y}_j)) \bar{\lambda}_j \stackrel{(3.24)}{=} \sum_{j=1}^k g_j^k \hat{\lambda}_j^k \stackrel{(3.154)}{=} \tilde{g}_p^k ,$$

i.e. the assumption (3.177) holds.

- Since the algorithm stops at iteration k , it has not stopped before iteration k and therefore we can apply Proposition 3.34

$$\sum_{j=1}^L \bar{s}_j \bar{\lambda}_j \stackrel{(3.155)}{=} \tilde{s}_p^k \stackrel{(3.181)}{=} 0 ,$$

i.e. the assumption (3.178) holds.

- For $j = 1, \dots, k$ we have

$$|\bar{y}_j - \bar{x}| \stackrel{(3.33)}{\leq} s_j^k = \bar{s}_j ,$$

i.e. the assumption (3.179) holds.

- Since all assumptions of Proposition 3.35 are satisfied, we can apply this proposition and therefore we obtain

$$\tilde{g}_p^k \stackrel{(3.185)}{=} \bar{q} \stackrel{(3.180)}{\in} \partial f(\bar{x}) = \partial f(x_k) . \quad \blacksquare$$

3. *Claim.:* If $\bar{\kappa}^{k+1} > 0$ and

$$\begin{array}{llll} \bar{x}' := x_k & \bar{G}'_j := \hat{\rho}_j \hat{G}_j & \bar{y}'_j := y_j & \bar{s}'_j := \hat{s}_j^k \\ L' := k & \bar{q}' := \tilde{g}_p^k & \bar{g}'_j := \hat{g}_j & \bar{\lambda}'_j := \hat{\kappa}_j^k \end{array} \quad (3.187)$$

with $j = 1, \dots, L$, then the assumptions of Proposition 3.35 are satisfied and therefore

$$\tilde{g}_p^k \in \partial F(x_k) . \quad (3.188)$$

Proof of claim.:

- (3.176) holds because of (3.156) and the assumption $\bar{\kappa}^{k+1} > 0$.
- Since the algorithm stops at iteration k , it has not stopped before iteration k and therefore we can apply Proposition 3.34

$$\bar{q}' \stackrel{(3.177)}{=} \sum_{j=1}^L (\bar{g}'_j + \bar{G}'_j(\bar{x} - \bar{y}'_j)) \bar{\lambda}'_j \stackrel{(3.26)}{=} \sum_{j=1}^k \hat{g}_j^k \hat{\kappa}_j^k \stackrel{(3.158)}{=} \tilde{g}_p^k ,$$

i.e. the assumption (3.177) holds.

- Since the algorithm stops at iteration k , it has not stopped before iteration k and therefore we can apply Proposition 3.34

$$\sum_{j=1}^L \bar{s}'_j \bar{\lambda}'_j \stackrel{(3.159)}{=} \tilde{s}_p^k \stackrel{(3.183)}{=} 0 ,$$

i.e. the assumption (3.178) holds.

- For $j = 1, \dots, k$ we have

$$|\bar{y}'_j - \bar{x}| \stackrel{(3.33), (3.31), (3.30)}{\leq} \hat{s}_j^k = \bar{s}'_j ,$$

i.e. the assumption (3.179) holds.

- Since all assumptions of Proposition 3.35 are satisfied, we can apply this proposition and therefore we obtain

$$\tilde{g}_p^k \stackrel{(3.187)}{=} \bar{q}' \stackrel{(3.180)}{\in} \partial F(\bar{x}) = \partial F(x_k) . \quad \blacksquare$$

4. *Claim.:* It holds

$$0 \in \partial f(x_k) + \bar{\kappa}^{k+1} \partial F(x_k) .$$

Proof of claim.: We have

$$0 \stackrel{(3.182)}{=} \tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{\tilde{g}}_p^k \stackrel{(3.186), (3.188)}{\in} \partial f(x_k) + \bar{\kappa}^{k+1} \partial F(x_k) . \quad \blacksquare$$

□

Assumption 3.37. From now on, we demand that Algorithm 3.20 does not stop, i.e. according to step 5 of Algorithm 3.20 and (3.151) we have for all k

$$w_k > 0 . \quad (3.189)$$

Theorem 3.38 (Carathéodory). *For $S \subseteq \mathbb{R}^n$ each element of (the convex hull) $\text{ch}(S)$ can be expressed as a convex combination of at most $n + 1$ elements of S .*

Proof. A proof of this well-known theorem can be found in, e.g., NEUMAIER [82]. □

We summarize some properties of positive (semi)definite matrices.

Definition 3.39. Let $A, B \in \mathbb{R}_{\text{sym}}^{N \times N}$. We define

$$A \preceq B : \Longleftrightarrow B - A \text{ is positive semidefinite} . \quad (3.190)$$

Proposition 3.40. *Let $A, B \in \mathbb{R}_{\text{sym}}^{N \times N}$ with B positive semidefinite, then*

$$A \preceq A + B . \quad (3.191)$$

Proof. Since B is positive semidefinite due to assumption, we have for all $x \in \mathbb{R}^N$

$$x^T ((A + B) - A)x = x^T Bx \geq 0 ,$$

i.e. $(A + B) - A$ is positive semidefinite, which is equivalent to $A \preceq A + B$ due to (3.190). □

Proposition 3.41. *Let $A, B \in \mathbb{R}_{\text{sym}}^{N \times N}$ be positive definite.*

1. *It holds*

$$A^{-1} \text{ is positive definite} . \quad (3.192)$$

2. *We have*

$$A \preceq B \Longleftrightarrow B^{-1} \preceq A^{-1} . \quad (3.193)$$

3. *If $A \preceq B$ holds, then*

$$\lambda_i(A) \leq \lambda_i(B) \quad (3.194)$$

for all $i = 1, \dots, N$, where $\lambda_i(A)$ denotes the i th eigenvalue of A .

4. *If $A \preceq B$ holds, then*

$$|B^{-1}| \leq |A^{-1}| . \quad (3.195)$$

5. Then

$$|A^{\frac{1}{2}} - B^{\frac{1}{2}}| \leq \frac{1}{(\lambda_{\min}(A))^{\frac{1}{2}} + (\lambda_{\min}(B))^{\frac{1}{2}}} |A - B|. \quad (3.196)$$

Proof.

1. Since A is positive definite due to assumption, we have for all $y \in \mathbb{R}^N$ with $y \neq 0$

$$y^T A^{-1} y \stackrel{A \in \mathbb{R}_{\text{sym}}^{N \times N}}{=} (A^{-1} y)^T A (A^{-1} y) > 0,$$

i.e. A^{-1} is positive definite and hence (3.192) holds.

2. (3.193) holds due to HORN & JOHNSON [40, p. 471, Corollary 7.7.4(a)].
3. (3.194) holds due to HORN & JOHNSON [40, p. 471, Corollary 7.7.4(c)].
4. Since B is positive definite due to assumption, (3.192) implies that B^{-1} is positive definite. Since all eigenvalues of a positive definite matrix are positive (cf. Proposition 2.2), we have

$$\lambda_i(B^{-1}) > 0 \quad \text{for all } i = 1, \dots, N. \quad (3.197)$$

Therefore, we obtain (3.195)

$$\begin{aligned} \text{assumption} \\ A \preceq B &\iff B^{-1} \stackrel{(3.193)}{\preceq} A^{-1} \\ &\implies 0 < \lambda_i(B^{-1}) \stackrel{(3.194)}{\leq} \lambda_i(A^{-1}) \quad \text{for all } i = 1, \dots, N \\ &\implies 0 < \sqrt{\lambda_{\max}(B^{-1})} \leq \sqrt{\lambda_{\max}(A^{-1})} \\ &\stackrel{(2.2)}{\implies} |B^{-1}| \leq |A^{-1}|. \end{aligned}$$

5. (3.196) holds due to HIGHAM [39, p. 135, Theorem 6.2]. □

Proposition 3.42. Let $\{A_k\}$ be a sequence of positive definite matrices $A_k \in \mathbb{R}_{\text{sym}}^{N \times N}$. Then

$$\{A_k\} \text{ is bounded} \iff \{A_k^{\frac{1}{2}}\} \text{ is bounded} \quad (3.198)$$

and

$$\{A_k\} \text{ is uniformly positive definite} \iff \{A_k^{-1}\} \text{ is bounded}. \quad (3.199)$$

Proof.

- Since $A_k \in \mathbb{R}_{\text{sym}}^{N \times N}$ is positive definite due to assumption, there exists an eigenvalue decomposition of A_k

$$A_k = Q_k^T \Xi_k Q_k \quad (3.200)$$

with $Q_k \in \mathbb{R}^{N \times N}$ orthogonal and a diagonal matrix $\Xi_k \in \mathbb{R}^{N \times N}$ with positive diagonal elements (cf. Proposition 2.2) and we define

$$\mu_k := \lambda_{\max}(\Xi_k). \quad (3.201)$$

Then we have

$$|A_k| \stackrel{(3.200)}{=} |Q_k^T \Xi_k Q_k| \stackrel{(2.2)}{=} (\lambda_{\max}(\Xi_k))^{\frac{1}{2}} \stackrel{(3.201)}{=} \mu_k^{\frac{1}{2}}. \quad (3.202)$$

Furthermore, we obtain

$$A_k \stackrel{(3.200)}{=} Q_k^T \Xi_k Q_k = Q_k^T \Xi_k^{\frac{1}{2}} \Xi_k^{\frac{1}{2}} Q_k = (Q_k^T \Xi_k^{\frac{1}{2}} Q_k)(Q_k^T \Xi_k^{\frac{1}{2}} Q_k) \quad (3.203)$$

which implies

$$A_k^{\frac{1}{2}} \stackrel{(3.203)}{=} Q_k^T \Xi_k^{\frac{1}{2}} Q_k \quad (3.204)$$

and therefore we have

$$|A_k^{\frac{1}{2}}| \stackrel{(3.204)}{=} |Q_k^T \Xi_k^{\frac{1}{2}} Q_k| \stackrel{(2.2)}{=} (\lambda_{\max}(\Xi_k^{\frac{1}{2}}))^{\frac{1}{2}} \stackrel{(3.201)}{=} (\mu_k^{\frac{1}{2}})^{\frac{1}{2}} = \mu_k^{\frac{1}{4}}. \quad (3.205)$$

Now, let $\{A_k\}$ be bounded, then there exists $\psi_1 > 0$ with

$$\begin{aligned} \mu_k^{\frac{1}{2}} &\stackrel{(3.202)}{=} |A_k| \leq \psi_1 \quad \text{for all } k \geq 1 \\ \stackrel{\wedge \frac{1}{2}}{\iff} |A_k^{\frac{1}{2}}| &\stackrel{(3.205)}{=} \mu_k^{\frac{1}{4}} \leq \psi_1^{\frac{1}{2}} \quad \text{for all } k \geq 1, \end{aligned} \quad (3.206)$$

which yields (3.198), since (3.206) can be read in both directions.

- We denote the i th eigenvalue of A_k by $\lambda_i(A_k)$ for $i = 1, \dots, N$. Since $\{A_k\}$ is uniformly positive definite due to assumption, there exists $\psi_2 > 0$ with

$$\begin{aligned} \lambda_{\min}(A_k) \geq \psi_2 &\iff \lambda_i(A_k) \geq \psi_2 \quad \text{for all } i = 1, \dots, N \\ &\stackrel{\wedge -1}{\iff} (\lambda_i(A_k))^{-1} \leq \psi_2^{-1} \quad \text{for all } i = 1, \dots, N \\ &\iff \lambda_i(A_k^{-1}) \leq \psi_2^{-1} \quad \text{for all } i = 1, \dots, N \\ &\iff \lambda_{\max}(A_k^{-1}) \leq \psi_2^{-1} \\ &\stackrel{\wedge \frac{1}{2}}{\iff} (\lambda_{\max}(A_k^{-1}))^{\frac{1}{2}} \leq \psi_2^{-\frac{1}{2}} \\ &\stackrel{(2.2)}{\iff} |A_k^{-1}| \leq \psi_2^{-\frac{1}{2}}, \end{aligned}$$

which yields (3.199). □

Corollary 3.43.

- If $\{(\bar{W}_p^k)^{-\frac{1}{2}}\}$ is bounded, then $\{(\bar{W}_p^k)^{-1}\}$ and $\{H_k\}$ are bounded

$$|(\bar{W}_p^k)^{-1}| \leq C_0 \quad (3.207)$$

for all $k \geq 1$ with some positive constant $C_0 > 0$.

- If $\{\bar{\kappa}^{k+1}\}$ is bounded and $\{(\bar{W}_p^k)^{-\frac{1}{2}}\}$ is uniformly positive definite, then $\{H_k^{-1}\}$ is bounded and

$$|\bar{W}_p^k + \bar{\kappa}^{k+1} \bar{G}^k| \leq C_1, \quad (3.208)$$

for all $k \geq 1$ with some positive constant $C_1 > 0$.

Proof.

- Since

$$(\bar{W}_p^k)^{-\frac{1}{2}} = ((\bar{W}_p^k)^{-1})^{\frac{1}{2}}$$

is bounded due to assumption, $\{(\bar{W}_p^k)^{-1}\}$ is bounded due to (3.198) and therefore (3.207) holds with some positive constant $C_0 > 0$, which is equivalent to the uniform positive definiteness of $\{\bar{W}_p^k\}$ due to (3.199). Since

$$\bar{W}_p^k \preceq \bar{W}_p^k + \sum_{j \in J_k} \mu_j^k \bar{G}_j^k + \mu_p^k \bar{G}^k \stackrel{(3.71)}{=} H_k^{-2}$$

for all $\mu^k \geq 0$, we obtain

$$|H_k^2| \stackrel{(3.195)}{\leq} |(\bar{W}_p^k)^{-1}| \stackrel{(3.207)}{\leq} C_0 ,$$

which is equivalent that $\{H_k\}$ is bounded due to (3.198).

- Since $\{\bar{\kappa}^{k+1}\}$ is bounded due to assumption, there exists a positive constant $\chi_0 > 0$ with

$$\bar{\kappa}^{k+1} \leq \chi_0 \tag{3.209}$$

for all $k \geq 1$ (note that $\bar{\kappa}^{k+1} \stackrel{(3.72)}{\geq} 0$). Since $\{(\bar{W}_p^k)^{-\frac{1}{2}}\}$ is uniformly positive definite due to assumption, $\{(\bar{W}_p^k)^{\frac{1}{2}}\}$ is bounded due to (3.199), which is equivalent to $\{\bar{W}_p^k\}$ being bounded due to (3.198)

$$|\bar{W}_p^k| \leq \chi_1 \tag{3.210}$$

for some positive constant $\chi_1 > 0$ and for all $k \geq 1$. Therefore, we obtain the boundedness of

$$\begin{aligned} |H_k^{-2}| &\stackrel{(3.71)}{=} |\bar{W}_p^k + \sum_{j \in J_k} \mu_j^k \bar{G}_j^k + \mu_p^k \bar{G}^k| \\ &\stackrel{(3.73), (3.74), (3.72)}{\leq} |\bar{W}_p^k + \bar{\kappa}^{k+1} \left(\sum_{j \in J_k} \kappa_j^k \bar{G}_j^k + \kappa_p^k \bar{G}^k \right)| \\ &\leq |\bar{W}_p^k| + \bar{\kappa}^{k+1} \left(\sum_{j \in J_k} \kappa_j^k |\bar{G}_j^k| + \kappa_p^k |\bar{G}^k| \right) \\ &\stackrel{(3.210), (3.209), \text{initialization of Algorithm 3.20}}{\leq} \chi_1 + \chi_0 \left(\sum_{j \in J_k} \kappa_j^k \bar{C}_G + \kappa_p^k \bar{C}_G \right) \\ &\stackrel{(3.73), (3.74)}{\leq} \chi_1 + \chi_0 \bar{C}_G , \end{aligned}$$

which is equivalent to $\{H_k^{-1}\}$ being bounded due to (3.198). Furthermore, we estimate

$$\begin{aligned} |\bar{W}_p^k + \bar{\kappa}^{k+1} \bar{G}^k| &\leq |\bar{W}_p^k| + |\bar{\kappa}^{k+1}| \cdot |\bar{G}^k| \\ &\stackrel{(3.72)}{=} |\bar{W}_p^k| + \bar{\kappa}^{k+1} |\bar{G}^k| \\ &\stackrel{(3.210), (3.209), \text{initialization of Algorithm 3.20}}{\leq} \chi_1 + \chi_0 \bar{C}_G \end{aligned}$$

and now setting

$$C_1 := \chi_1 + \chi_0 \bar{C}_G$$

yields (3.208). $\square$

Proposition 3.44. *Let (3.189) be satisfied. If $\{(x_k, \bar{\kappa}^{k+1})\}$ is bounded and there exists $\bar{x} \in \mathbb{R}^N$ with*

$$\sigma(\bar{x}) = 0 ,$$

where $\sigma : \mathbb{R}^N \longrightarrow \mathbb{R}$

$$\sigma(x) := \liminf_{k \rightarrow \infty} \max(w_k, |x_k - x|) , \quad (3.211)$$

and if $\{(\bar{W}_p^k)^{-\frac{1}{2}}\}$ is uniformly positive definite, then there exists $\bar{\kappa} \in \mathbb{R}_{\geq 0}$ with

$$0 \in \partial f(\bar{x}) + \bar{\kappa} \partial F(\bar{x})$$

$$\bar{\kappa} F(\bar{x}) = 0$$

$$F(\bar{x}) \leq 0 ,$$

i.e. if the sequence of iteration points and (single) Lagrange multipliers is bounded and the sequence of iteration points has an accumulation point with $\sigma(\bar{x}) = 0$, then this accumulation point is stationary for the optimization problem (2.35).

Proof.

1. *Claim:* There exist (an infinite set) $K \subset \{1, 2, \dots\}$ and $\bar{\kappa} \in \mathbb{R}_{\geq 0}$ with

$$x_k \xrightarrow{K} \bar{x} \quad (3.212)$$

$$w_k \xrightarrow{K} 0 \quad (3.213)$$

$$\bar{\kappa}^{k+1} \xrightarrow{K} \bar{\kappa} , \quad (3.214)$$

where $F(\bar{x}) \leq 0$.

Proof of claim: We have

$$0 \stackrel{\text{assumption}}{\stackrel{\downarrow}{=}} \sigma(\bar{x}) \stackrel{(3.211)}{=} \liminf_{k \rightarrow \infty} \max(w_k, |x_k - \bar{x}|) .$$

Since $w_k \stackrel{(3.130)}{\geq} 0$ and $|\cdot| \geq 0$, there exist convergent subsequences of $\{w_k\}_{k \geq 1}$ and $\{x_k - \bar{x}\}_{k \geq 1}$, i.e. there exists (an infinite set) $\hat{K} \subset \{1, 2, \dots\}$ with

$$x_k \xrightarrow{\hat{K}} \bar{x} , \quad w_k \xrightarrow{\hat{K}} 0 . \quad (3.215)$$

Since $\{\bar{\kappa}^{k+1}\}_k$ is bounded by assumption, all its subsequences are also bounded. Therefore, in particular, its subsequence $\{\bar{\kappa}^{k+1}\}_{k \in \hat{K}}$ is bounded. Consequently, $\{\bar{\kappa}^{k+1}\}_{k \in \hat{K}}$ has an accumulation point, i.e. there exist (an infinite set) $K \subset \hat{K}$ and $\bar{\kappa} \in \mathbb{R}$ with

$$\bar{\kappa}^{k+1} \xrightarrow{K} \bar{\kappa} .$$

Since $\bar{\kappa}^{k+1} \stackrel{(3.72)}{\geq} 0$ for $k = 1, 2, \dots$, we have $\bar{\kappa} \in \mathbb{R}_{\geq 0}$. Since $K \subset \hat{K}$ and a sequence is convergent, if and only if all of its subsequences converge towards the same limit, (3.215) yields

$$x_k \xrightarrow{K} \bar{x} , \quad w_k \xrightarrow{K} 0 .$$

Consequently (3.212), (3.213) and (3.214) hold. Moreover, due to the continuity of F we obtain

$$0 \stackrel{(3.123)}{\geq} \lim_K F(x_k) = F(\lim_K x_k) \stackrel{(3.212)}{=} F(\bar{x}) . \quad \blacksquare$$

2. *Claim:* Let $I := \{1, 2, \dots, N+2\}$.

- For $i \in I$ and $k \geq 1$ there exist

$$(g^{k,i}, s^{k,i}) \in \{(g_j^k, s_j^k) : j = 1, \dots, k\} \subseteq \mathbb{R}^N \times \mathbb{R} = \mathbb{R}^{N+1} \quad (3.216)$$

and $\lambda^{k,i} \in \mathbb{R}$ with

$$1 = \sum_{i \in I} \lambda^{k,i}, \quad \lambda^{k,i} \geq 0 \quad (3.217)$$

$$\tilde{g}_p^k = \sum_{i \in I} \lambda^{k,i} g^{k,i} \quad (3.218)$$

$$\tilde{s}_p^k = \sum_{i \in I} \lambda^{k,i} s^{k,i} . \quad (3.219)$$

- For $i \in I$ and $k \geq 1$ there exist

$$(\hat{g}^{k,i}, \hat{s}^{k,i}) \in \{(\hat{g}_j^k, \hat{s}_j^k) : j = 1, \dots, k\} \subseteq \mathbb{R}^N \times \mathbb{R} = \mathbb{R}^{N+1} \quad (3.220)$$

and $\kappa^{k,i} \in \mathbb{R}$ with

$$(1 = \sum_{i \in I} \kappa^{k,i} \quad \wedge \quad \kappa^{k,i} \geq 0) \quad \text{or} \quad (\kappa^{k,i} = 0 \quad \text{for all } i \in I) \quad (3.221)$$

$$\tilde{\hat{g}}_p^k = \sum_{i \in I} \kappa^{k,i} \hat{g}^{k,i} \quad (3.222)$$

$$\tilde{\hat{s}}_p^k = \sum_{i \in I} \kappa^{k,i} \hat{s}^{k,i} . \quad (3.223)$$

Furthermore, we have

$$\sum_{i \in I} \kappa^{k,i} = \begin{cases} 1 & \text{if } \bar{\kappa}^{k+1} > 0 \\ 0 & \text{if } \bar{\kappa}^{k+1} = 0 . \end{cases} \quad (3.224)$$

Proof of claim:

- Let

$$S := \{(g_j^k, s_j^k) : j = 1, \dots, k\} \subseteq \mathbb{R}^{N+1} \stackrel{(3.189), (3.154), (3.155)}{\Downarrow} (\tilde{g}_p^k, \tilde{s}_p^k) \in \text{ch}(S) .$$

Due to Theorem 3.38 (Carathéodory), there exist $(g^{k,i}, s^{k,i}) \in S$, $\lambda^{k,i} \in \mathbb{R}$ for $i \in I$ (Note: $\text{card}(I) = n+2$) with

$$1 = \sum_{i \in I} \lambda^{k,i}, \quad \lambda^{k,i} \geq 0, \quad (\tilde{g}_p^k, \tilde{s}_p^k) = \sum_{i \in I} \lambda^{k,i} (g^{k,i}, s^{k,i}) .$$

- Let

$$\hat{S} := \{(\hat{g}_j^k, \hat{s}_j^k) : j = 1, \dots, k\} \subseteq \mathbb{R}^{N+1} \xrightarrow{(3.189), (3.158), (3.159)} \begin{cases} (\tilde{g}_p^k, \tilde{s}_p^k) \in \text{ch}(\hat{S}) & \text{for } \bar{\kappa}^{k+1} > 0 \\ (\tilde{g}_p^k, \tilde{s}_p^k) = 0 & \text{for } \bar{\kappa}^{k+1} = 0 \end{cases}.$$

In the case $\bar{\kappa}^{k+1} > 0$ there exist $(\hat{g}^{k,i}, \hat{s}^{k,i}) \in \hat{S}$, $\kappa^{k,i} \in \mathbb{R}$ for $i \in I$ (Note: $\text{card}(I) = n + 2$) with

$$1 = \sum_{i \in I} \kappa^{k,i}, \quad \kappa^{k,i} \geq 0, \quad (\tilde{g}_p^k, \tilde{s}_p^k) = \sum_{i \in I} \kappa^{k,i} (\hat{g}^{k,i}, \hat{s}^{k,i})$$

due to Theorem 3.38 (Carathéodory). In the case $\bar{\kappa}^{k+1} = 0$ choosing

$$\kappa^{k,i} := 0 \quad \text{for all } i \in I$$

yields

$$(\tilde{g}_p^k, \tilde{s}_p^k) = 0 = \sum_{i \in I} \kappa^{k,i} (\hat{g}^{k,i}, \hat{s}^{k,i}).$$

Furthermore, (3.224) follows immediately. ■

3. *Claim:* There exists (a function) $j : \{k \in \mathbb{N} : k \geq 1\} \times I \longrightarrow \{1, \dots, k\}$ with

$$g^{k,i} = g_{j(k,i)}^k \tag{3.225}$$

$$s^{k,i} = s_{j(k,i)}^k \tag{3.226}$$

$$\hat{g}^{k,i} = \hat{g}_{j(k,i)}^k \tag{3.227}$$

$$\hat{s}^{k,i} = \hat{s}_{j(k,i)}^k \tag{3.228}$$

(i.e. we can assign every $k \geq 1$ and every $i \in I$ an index $j(k, i)$, such that the above properties hold).

Proof of claim: For $i \in I$ and $k \geq 1$ we have

$$(g^{k,i}, s^{k,i}) \stackrel{(3.216)}{\in} \{(g_j^k, s_j^k) : j = 1, \dots, k\} \subseteq \mathbb{R}^N \times \mathbb{R} = \mathbb{R}^{N+1}$$

$$(\hat{g}^{k,i}, \hat{s}^{k,i}) \stackrel{(3.220)}{\in} \{(\hat{g}_j^k, \hat{s}_j^k) : j = 1, \dots, k\} \subseteq \mathbb{R}^N \times \mathbb{R} = \mathbb{R}^{N+1},$$

i.e. for all $i \in I$ and for all $k \geq 1$ there exists $j(k, i) \in \{1, \dots, k\}$ (i.e. a function $j : \{k \in \mathbb{N} : k \geq 1\} \times I \longrightarrow \{1, \dots, k\}$) with

$$(g^{k,i}, s^{k,i}) = (g_j^k, s_j^k), \quad (\hat{g}^{k,i}, \hat{s}^{k,i}) = (\hat{g}_j^k, \hat{s}_j^k). \quad \text{■}$$

4. *Claim:* We have

$$\{y_{j(k,i)}\}_{k \geq 1, i \in I} \text{ is bounded} \tag{3.229}$$

(Note: The index j of the members of the sequence depends on i and k).

Proof of claim: We have for all $i \in I$ and for all $k \geq 1$

$$|y_{j(k,i)}| - |x_{j(k,i)}| \leq |y_{j(k,i)} - x_{j(k,i)}| \leq C_S \implies |y_{j(k,i)}| \leq |x_{j(k,i)}| + C_S$$

and consequently — because of the assumption of the boundedness of $\{x_k\}$ — the sequence $\{y_{j(k,i)}\}$ is bounded. ■

5. *Claim:* For all $i \in I$ there exist $\bar{y}_i \in \mathbb{R}^N$ and (an infinite set) $K_1 \subset K$ with

$$y_{j(k,i)} \xrightarrow{K_1} \bar{y}_i . \quad (3.230)$$

Proof of claim: Since the sequence $\{y_{j(k,i)}\}$ is bounded due to (3.229), it has a convergent subsequence, i.e. for all $i \in I$ there exist $\bar{y}_i \in \mathbb{R}^N$ and (an infinite set) $K_1 \subset K$ with $y_{j(k,i)} \xrightarrow{K_1} \bar{y}_i$. ■

6. *Claim:* For all $i \in I$ there exist $\bar{g}_i \in \partial f(\bar{y}_i)$, $\bar{\bar{g}}_i \in \partial F(\bar{y}_i)$ and (an infinite set) $K_2 \subset K_1$ with

$$g_{j(k,i)} \xrightarrow{K_2} \bar{g}_i \quad (3.231)$$

$$\hat{g}_{j(k,i)} \xrightarrow{K_2} \bar{\bar{g}}_i . \quad (3.232)$$

Proof of claim:

- Since the sequence $\{y_{j(k,i)}\}$ is bounded according to (3.229), the local boundedness of ∂f resp. ∂F (cf. (2.32)) imply that the sets

$$B_1 := \{g \in \partial f(y_{j(k,i)}) : y_{j(k,i)} \in \mathbb{R}^N, k \geq 1, k \in K_1, i \in I\}$$

$$B_2 := \{\hat{g} \in \partial F(y_{j(k,i)}) : y_{j(k,i)} \in \mathbb{R}^N, k \geq 1, k \in K_1, i \in I\}$$

are bounded. Therefore, $B_1 \times B_2$ is bounded and consequently there exists a convergent subsequence $(g_{j(k,i)}, \hat{g}_{j(k,i)}) \in \partial f(y_{j(k,i)}) \times \partial F(y_{j(k,i)})$, i.e. there exists $(\bar{g}_i, \bar{\bar{g}}_i) \in \mathbb{R}^N \times \mathbb{R}^N$ and (an infinite set) $K_2 \subset K_1$ with

$$\begin{aligned} & (g_{j(k,i)}, \hat{g}_{j(k,i)}) \xrightarrow{K_2} (\bar{g}_i, \bar{\bar{g}}_i) \\ \iff & g_{j(k,i)} \xrightarrow{K_2} \bar{g}_i \quad \wedge \quad \hat{g}_{j(k,i)} \xrightarrow{K_2} \bar{\bar{g}}_i . \end{aligned}$$

The upper semicontinuity of ∂f resp. ∂F (cf. (2.33)) and (3.230) imply

$$\bar{g}_i \in \partial f(\bar{y}_i), \quad \bar{\bar{g}}_i \in \partial F(\bar{y}_i) . \quad \blacksquare$$

7. *Claim:* For all $i \in I$ there exist $\bar{G}_i, \bar{\bar{G}}_i \in \mathbb{R}^{N \times N}$, $\bar{\lambda}^{k,i}, \bar{\kappa}^{k,i} \in \mathbb{R}$ and (an infinite set) $K_3 \subset K_2$ with

$$\rho_{j(k,i)} G_{j(k,i)} \xrightarrow{K_3} \bar{G}_i \quad (3.233)$$

$$\lambda^{k,i} \xrightarrow{K_3} \bar{\lambda}_i \quad (3.234)$$

$$\hat{\rho}_{j(k,i)} \hat{G}_{j(k,i)} \xrightarrow{K_3} \bar{\bar{G}}_i \quad (3.235)$$

$$\kappa^{k,i} \xrightarrow{K_3} \bar{\kappa}_i . \quad (3.236)$$

Proof of claim: Since

$$\rho_{j(k,i)} \stackrel{(3.90)}{=} \min(1, \frac{C_G}{|G_{j(k,i)}|}) \xrightarrow{C_G > 0} \rho_{j(k,i)} \in (0, 1]$$

we obtain

$$\rho_{j(k,i)} \leq \frac{C_G}{|G_{j(k,i)}|} \iff \underbrace{\rho_{j(k,i)}|G_{j(k,i)}|}_{\in(0,1]} \leq C_G, \quad (3.237)$$

which yields the boundedness of the sequence $\{\rho_{j(k,i)}|G_{j(k,i)}|\}$.

Due to (3.217), the sequence $\{\lambda^{k,i}\}$ is bounded.

Since

$$\hat{\rho}_{j(k,i)} \stackrel{(3.91)}{=} \min(1, \frac{\hat{C}_G}{|\hat{G}_{j(k,i)}|}) \stackrel{\hat{C}_G \geq 0}{\implies} \hat{\rho}_{j(k,i)} \in (0, 1]$$

we obtain

$$\hat{\rho}_{j(k,i)} \leq \frac{\hat{C}_G}{|\hat{G}_{j(k,i)}|} \iff \underbrace{\hat{\rho}_{j(k,i)}|\hat{G}_{j(k,i)}|}_{\in(0,1]} \leq \hat{C}_G, \quad (3.238)$$

which yields the boundedness of the sequence $\{\hat{\rho}_{j(k,i)}|\hat{G}_{j(k,i)}|\}$.

Due to (3.221), the sequence $\{\kappa^{k,i}\}$ is bounded.

Therefore, the sequence $\{\rho_{j(k,i)}|G_{j(k,i)}|, \lambda^{k,i}, \hat{\rho}_{j(k,i)}|\hat{G}_{j(k,i)}|, \kappa^{k,i}\}$ is bounded. Consequently, there exists a convergent subsequence of $\{\rho_{j(k,i)}|G_{j(k,i)}|, \lambda^{k,i}, \hat{\rho}_{j(k,i)}|\hat{G}_{j(k,i)}|, \kappa^{k,i}\}$, i.e. there exist $\bar{G}_i, \bar{\hat{G}}_i \in \mathbb{R}^{N \times N}$, $\bar{\lambda}_i, \bar{\kappa}_i \in \mathbb{R}$ and (an infinite set) $K_3 \subset K_2$ with

$$\{\rho_{j(k,i)}|G_{j(k,i)}|, \lambda^{k,i}, \hat{\rho}_{j(k,i)}|\hat{G}_{j(k,i)}|, \kappa^{k,i}\} \xrightarrow{K_3} \{\bar{G}_i, \bar{\lambda}_i, \bar{\hat{G}}_i, \bar{\kappa}_i\}. \quad \blacksquare$$

8. *Claim:* We have

$$\tilde{g}_p^k \xrightarrow{K_3} \sum_{i \in I} \bar{\lambda}_i (\bar{g}_i + \bar{G}_i(\bar{x} - \bar{y}_i)) \quad (3.239)$$

$$\tilde{g}_p^k \xrightarrow{K_3} \sum_{i \in I} \bar{\kappa}_i (\bar{\hat{g}}_i + \bar{\hat{G}}_i(\bar{x} - \bar{y}_i)). \quad (3.240)$$

Proof of claim: We calculate

$$\begin{aligned} \lim_{K_3} \tilde{g}_p^k &\stackrel{(3.218)}{=} \lim_{K_3} \left(\sum_{i \in I} \lambda^{k,i} g_{j(k,i)}^k \right) \stackrel{(3.225)}{=} \lim_{K_3} \left(\sum_{i \in I} \lambda^{k,i} g_{j(k,i)}^k \right) \\ &\stackrel{(3.24)}{=} \lim_{K_3} \left(\sum_{i \in I} \lambda^{k,i} (g_{j(k,i)} + \rho_{j(k,i)} G_{j(k,i)} (x_k - y_{j(k,i)})) \right) \\ &\stackrel{(3.234), (3.231), (3.233), (3.212), (3.230)}{\stackrel{\perp}{=}} \sum_{i \in I} \bar{\lambda}_i (\bar{g}_i + \bar{G}_i(\bar{x} - \bar{y}_i)) \end{aligned}$$

and

$$\begin{aligned} \lim_{K_3} \tilde{g}_p^k &\stackrel{(3.222)}{=} \lim_{K_3} \left(\sum_{i \in I} \kappa^{k,i} \hat{g}_{j(k,i)}^k \right) \stackrel{(3.227)}{=} \lim_{K_3} \left(\sum_{i \in I} \kappa^{k,i} \hat{g}_{j(k,i)}^k \right) \\ &\stackrel{(3.26)}{=} \lim_{K_3} \left(\sum_{i \in I} \kappa^{k,i} (\hat{g}_{j(k,i)} + \hat{\rho}_{j(k,i)} \hat{G}_{j(k,i)} (x_k - y_{j(k,i)})) \right) \\ &\stackrel{(3.236), (3.232), (3.235), (3.212), (3.230)}{\stackrel{\perp}{=}} \sum_{i \in I} \bar{\kappa}_i (\bar{\hat{g}}_i + \bar{\hat{G}}_i(\bar{x} - \bar{y}_i)). \quad \blacksquare \end{aligned}$$

9. *Claim:* We have

$$\sum_{i \in I} \bar{\lambda}_i (\bar{g}_i + \bar{G}_i(\bar{x} - \bar{y}_i)) + \bar{\kappa} \sum_{i \in I} \bar{\kappa}_i (\bar{\hat{g}}_i + \bar{\hat{G}}_i(\bar{x} - \bar{y}_i)) = 0 \quad (3.241)$$

$$\lambda^{k,i} s^{k,i} \xrightarrow{K_3} 0 \quad (3.242)$$

$$\bar{\kappa} F(\bar{x}) = 0$$

$$\kappa^{k,i} \hat{s}^{k,i} \xrightarrow{K_3} 0 \quad \text{if } \bar{\kappa} > 0. \quad (3.243)$$

Proof of claim: Since $\{\bar{\kappa}^{k+1}\}$ is bounded and $\{(\bar{W}_p^k)^{-\frac{1}{2}}\}$ is uniformly positive definite (both due to assumption), Corollary 3.43 implies that

$$\{H_k^{-1}\} \text{ is bounded.} \quad (3.244)$$

Because of (3.213), (3.86) and (3.130), we have

$$\begin{aligned} & |H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{\hat{g}}_p^k)| \xrightarrow{K_3} 0 \\ \implies & |\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{\hat{g}}_p^k| \stackrel{H_k \text{ regular}}{\downarrow} |H_k^{-1} H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{\hat{g}}_p^k)| \leq \underbrace{|H_k^{-1}|}_{\substack{\text{bounded} \\ \text{due to (3.244)}}} \cdot |H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{\hat{g}}_p^k)| \xrightarrow{K_3} 0 \\ \iff & \tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{\hat{g}}_p^k \xrightarrow{K_3} 0 \\ \stackrel{(3.239), (3.240), (3.214), \text{limit unique}}{\iff} & \sum_{i \in I} \bar{\lambda}_i (\bar{g}_i + \bar{G}_i(\bar{x} - \bar{y}_i)) + \bar{\kappa} \sum_{i \in I} \bar{\kappa}_i (\bar{\hat{g}}_i + \bar{\hat{G}}_i(\bar{x} - \bar{y}_i)) = 0 \end{aligned}$$

and

$$\begin{aligned} & \tilde{\alpha}_p^k \xrightarrow{K_3} 0 \\ & \stackrel{(3.79)}{\implies} \tilde{s}_p^k \xrightarrow{K_3} 0 \\ & \stackrel{(3.219)}{\iff} \sum_{i \in I} \lambda^{k,i} s^{k,i} \xrightarrow{K_3} 0 \\ & \stackrel{(3.226), (3.217), (3.30)}{\downarrow} \lambda^{k,i} s^{k,i} \xrightarrow{K_3} 0 \end{aligned}$$

as well as

$$\bar{\kappa}^{k+1} F(x_k) \xrightarrow{K_3} 0 \quad (3.245)$$

$$\bar{\kappa}^{k+1} \tilde{A}_p^k \xrightarrow{K_3} 0. \quad (3.246)$$

Hence, it follows

$$\begin{aligned} 0 & \stackrel{(3.245)}{=} \lim_{K_3} (\bar{\kappa}^{k+1} F(x_k)) \\ & = (\lim_{K_3} \bar{\kappa}^{k+1}) (\lim_{K_3} F(x_k)) \\ & \stackrel{(3.214), F \text{ continuous}, (3.212)}{\downarrow} \bar{\kappa} F(\bar{x}). \end{aligned}$$

Furthermore, if $\bar{\kappa} > 0$, we obtain

$$\begin{aligned}
0 &\stackrel{(3.246)}{=} \lim_{K_3} (\bar{\kappa}^{k+1} \tilde{A}_p^k) \\
&\stackrel{(3.214)}{=} \bar{\kappa} (\lim_{K_3} \tilde{A}_p^k) \\
&\stackrel{\bar{\kappa} > 0}{\implies} \tilde{A}_p^k \xrightarrow{K_3} 0 \\
&\stackrel{(3.84)}{\implies} \tilde{s}_p^k \xrightarrow{K_3} 0 \\
&\stackrel{(3.223)}{\iff} \sum_{i \in I} \kappa^{k,i} \tilde{s}^{k,i} \xrightarrow{K_3} 0 \\
&\stackrel{(3.228), (3.221), (3.31)}{\implies} \kappa^{k,i} \tilde{s}^{k,i} \xrightarrow{K_3} 0 .
\end{aligned}$$

■

10. *Claim:*

- We have

$$\sum_{i \in I} \bar{\lambda}_i = 1 . \quad (3.247)$$

- If $\bar{\kappa} > 0$, then

$$\sum_{i \in I} \bar{\kappa}_i = 1 . \quad (3.248)$$

Proof of claim:

- (3.247) holds due to

$$\sum_{i \in I} \lambda^{k,i} \stackrel{(3.217)}{=} 1 \quad \text{for all } k \in K_3 \quad \stackrel{(3.234)}{\implies} \sum_{i \in I} \bar{\lambda}_i = 1 .$$

- We have

$$\lim_{K_3} \sum_{i \in I} \kappa^{k,i} = \sum_{i \in I} \lim_{K_3} \kappa^{k,i} \stackrel{(3.236)}{=} \sum_{i \in I} \bar{\kappa}_i . \quad (3.249)$$

If $\bar{\kappa} > 0$, then — because of (3.214) and since $K_3 (\subset K)$ is an infinite set — there exists $\hat{k} \in K_3$ such that

$$|\bar{\kappa}^{k+1} - \bar{\kappa}| < \frac{\bar{\kappa}}{2} \implies -\bar{\kappa}^{k+1} + \bar{\kappa} < \frac{\bar{\kappa}}{2} \implies 0 < \frac{\bar{\kappa}}{2} < \bar{\kappa}^{k+1} \quad (3.250)$$

for all $k \in \hat{K}_3$, where

$$\hat{K}_3 := \left\{ k \in K_3 : k \geq \hat{k} \right\} \subseteq K_3$$

is an infinite set. Therefore, we obtain for all $k \in \hat{K}_3$

$$\sum_{i \in I} \kappa^{k,i} \stackrel{(3.224)}{=} \begin{cases} 1 & \text{if } \bar{\kappa}^{k+1} > 0 \\ 0 & \text{if } \bar{\kappa}^{k+1} = 0 . \end{cases} \stackrel{(3.250)}{=} 1 ,$$

i.e. $\{\sum_{i \in I} \kappa^{k,i}\}_{k \in \hat{K}_3}$ is constant on $\hat{K}_3$ and hence we have

$$\lim_{\hat{K}_3} \sum_{i \in I} \kappa^{k,i} = 1 . \quad (3.251)$$

Since

- the sequence $\{\sum_{i \in I} \kappa^{k,i}\}_{k \in K_3}$ is convergent due to (3.249)
- the (infinite) subsequence $\{\sum_{i \in I} \kappa^{k,i}\}_{k \in \hat{K}_3}$ (of the sequence $\{\sum_{i \in I} \kappa^{k,i}\}_{k \in K_3}$) converges towards 1 due to (3.251)
- a sequence is convergent if and only if all its subsequences converge towards the same limit,

the limit of the sequence $\{\sum_{i \in I} \kappa^{k,i}\}_{k \in K_3}$ must be 1

$$1 = \lim_{K_3} \sum_{i \in I} \kappa^{k,i} \stackrel{(3.249)}{=} \sum_{i \in I} \bar{\kappa}_i ,$$

i.e. (3.248) holds. ■

11. *Claim:* If $\bar{\lambda}_i \neq 0$ resp. if $\bar{\kappa} > 0$ and $\bar{\kappa}_i \neq 0$, then

$$|\bar{x} - \bar{y}_i| = 0 . \quad (3.252)$$

Proof of claim:

- Due to (3.242) the sequence $\{\lambda^{k,i} s^{k,i}\}_{k \in K_3}$ is convergent and therefore necessarily bounded, i.e. there exists $C > 0$ with

$$\stackrel{(3.226),(3.30)}{\downarrow} 0 \leq \lambda^{k,i} s^{k,i} \leq C \iff 0 \leq s^{k,i} \leq \frac{C}{\lambda^{k,i}}$$

and therefore $\{s^{k,i}\}_{k \in K_3}$ is bounded due to (3.234) for $\bar{\lambda}_i \neq 0$, where at least one such $\bar{\lambda}_i$ exists due to (3.247). Since the locality measure is monotone due to (3.32), $\{s^{k,i}\}_{k \in K_3}$ is monotone. Consequently, $\{s^{k,i}\}_{k \in K_3}$ is convergent for $\bar{\lambda}_i \neq 0$, i.e.

$$\exists s_i := \lim_{K_3} s^{k,i} . \quad (3.253)$$

Therefore, it follows

$$0 \stackrel{(3.242)}{=} \lim_{K_3} (\lambda^{k,i} s^{k,i}) \stackrel{(3.234),(3.253)}{\downarrow} \bar{\lambda}_i s_i \stackrel{\bar{\lambda}_i \neq 0}{\implies} s_i = 0 . \quad (3.254)$$

Hence, we obtain (3.252) in the case $\bar{\lambda}_i \neq 0$ due to

$$\begin{aligned} 0 &\stackrel{(3.253),(3.254)}{\downarrow} \lim_{K_3} s^{k,i} \stackrel{(3.226)}{=} \lim_{K_3} s_{j(k,i)}^k \\ &\stackrel{(3.33)}{\geq} \lim_{K_3} |y_{j(k,i)} - x_k| \stackrel{(3.230),(3.212)}{\downarrow} |\bar{y}_i - \bar{x}| \\ \iff |\bar{y}_i - \bar{x}| &= 0 . \end{aligned}$$

- Due to the assumption $\bar{\kappa} > 0$ and due to (3.243) the sequence $\{\kappa^{k,i} s^{k,i}\}_{k \in K_3}$ is convergent and therefore necessarily bounded, i.e. there exists $\hat{C} > 0$ with

$$\stackrel{(3.228),(3.31)}{\downarrow} 0 \leq \kappa^{k,i} \hat{s}^{k,i} \leq \hat{C} \iff 0 \leq \hat{s}^{k,i} \leq \frac{\hat{C}}{\kappa^{k,i}}$$

and therefore $\{\hat{s}^{k,i}\}_{k \in K_3}$ is bounded due to (3.236) for $\bar{\kappa}_i \neq 0$, where at least one such $\bar{\kappa}_i$ exists due to (3.248). Since the locality measure is monotone due to (3.31) and (3.32), $\{\hat{s}^{k,i}\}_{k \in K_3}$ is monotone. Consequently, $\{\hat{s}^{k,i}\}_{k \in K_3}$ is convergent for $\bar{\kappa}_i \neq 0$, i.e.

$$\exists \hat{s}_i := \lim_{K_3} \hat{s}^{k,i} . \quad (3.255)$$

Therefore, it follows

$$0 \stackrel{(3.243)}{=} \lim_{K_3} (\kappa^{k,i} \hat{s}^{k,i}) \stackrel{(3.236), (3.255)}{=} \bar{\kappa}_i \hat{s}_i \stackrel{\bar{\kappa}_i \neq 0}{\implies} \hat{s}_i = 0 . \quad (3.256)$$

Hence, we obtain (3.252) in the case $\bar{\kappa} > 0$ and $\bar{\kappa}_i \neq 0$ due to

$$\begin{aligned} 0 &\stackrel{(3.255), (3.256)}{=} \lim_{K_3} \hat{s}^{k,i} \stackrel{(3.228)}{=} \lim_{K_3} \hat{s}_{j(k,i)}^k \\ &\stackrel{(3.31), (3.33)}{\geq} \lim_{K_3} |y_{j(k,i)} - x_k| \stackrel{(3.230), (3.212)}{=} |\bar{y}_i - \bar{x}| \\ &\iff |\bar{y}_i - \bar{x}| = 0 . \end{aligned} \quad \blacksquare$$

12. *Claim:*

- If we set

$$\bar{q} := \sum_{i \in I} \bar{\lambda}_i (\bar{g}_i + \bar{G}_i(\bar{x} - \bar{y}_i)) \quad (3.257)$$

$$\bar{s}_i := \begin{cases} |\bar{x} - \bar{y}_i| & \text{for } \bar{\lambda}_i = 0 \\ 0 & \text{for } \bar{\lambda}_i \neq 0 , \end{cases} \quad (3.258)$$

then

$$\sum_{i \in I} \bar{\lambda}_i (\bar{g}_i + \bar{G}_i(\bar{x} - \bar{y}_i)) \in \partial f(\bar{x}) . \quad (3.259)$$

- If $\bar{\kappa} > 0$, we set

$$\bar{q}' := \sum_{i \in I} \bar{\kappa}_i (\bar{g}_i + \bar{G}_i(\bar{x} - \bar{y}_i)) \quad (3.260)$$

$$\bar{s}'_i := \begin{cases} |\bar{x} - \bar{y}_i| & \text{for } \bar{\kappa}_i = 0 \\ 0 & \text{for } \bar{\kappa}_i \neq 0 . \end{cases} \quad (3.261)$$

Then we have for $\bar{\kappa} > 0$

$$\sum_{i \in I} \bar{\kappa}_i (\bar{g}_i + \bar{G}_i(\bar{x} - \bar{y}_i)) \in \partial F(\bar{x}) . \quad (3.262)$$

- If $\bar{\kappa} = 0$, then

$$\bar{\kappa} \partial F(\bar{x}) = \{0\} . \quad (3.263)$$

Proof of claim:

- For showing (3.259), we establish that the assumptions of Proposition 3.35 are satisfied:

- The non-negativity property of (3.176) holds, since

$$\lambda^{k,i} \stackrel{(3.217)}{\geq} 0 \quad \text{for all } k \in K_3 \stackrel{(3.234)}{\implies} \bar{\lambda}_i \geq 0 .$$

The convex combination property of (3.176) holds due to (3.247).

- (3.177) holds due to the definition of $\bar{q}$ in (3.257).
- (3.178) holds, since

$$\sum_{i \in I} \bar{s}_i \bar{\lambda}_i \stackrel{(3.258)}{=} \sum_{i \in I} \left\{ \begin{array}{ll} |\bar{x} - \bar{y}_i| & \text{for } \bar{\lambda}_i = 0 \\ 0 & \text{for } \bar{\lambda}_i \neq 0 , \end{array} \right\} \bar{\lambda}_i = 0 .$$

- (3.179) holds, since for $\bar{\lambda}_i = 0$ we have

$$|\bar{x} - \bar{y}_i| \stackrel{(3.258)}{=} \bar{s}_i \leq \bar{s}_i$$

and for $\bar{\lambda}_i \neq 0$ we have

$$|\bar{x} - \bar{y}_i| \stackrel{(3.252)}{=} 0 \stackrel{(3.258)}{=} \bar{s}_i \leq \bar{s}_i .$$

Therefore, it follows

$$\sum_{i \in I} \bar{\lambda}_i (\bar{g}_i + \bar{G}_i(\bar{x} - \bar{y}_i)) \stackrel{(3.257)}{=} \bar{q} \stackrel{(3.180)}{\in} \partial f(\bar{x}) .$$

- For showing (3.262), we establish that the assumptions of Proposition 3.35 are satisfied:
 - The non-negativity property of (3.176) holds, since

$$\kappa^{k,i} \stackrel{(3.221)}{\geq} 0 \quad \text{for all } k \in K_3 \stackrel{(3.236)}{\implies} \bar{\kappa}_i \geq 0 .$$

The convex combination property of (3.176) holds due to the assumption $\bar{\kappa} > 0$ and (3.248).

- (3.177) holds due to the definition of $\bar{q}'$ in (3.260).
- (3.178) holds, since

$$\sum_{i \in I} \bar{s}'_i \bar{\kappa}_i \stackrel{(3.261)}{=} \sum_{i \in I} \left\{ \begin{array}{ll} |\bar{x} - \bar{y}_i| & \text{for } \bar{\kappa}_i = 0 \\ 0 & \text{for } \bar{\kappa}_i \neq 0 , \end{array} \right\} \bar{\kappa}_i = 0 .$$

- (3.179) holds, since for $\bar{\kappa}_i = 0$ we have

$$|\bar{x} - \bar{y}_i| \stackrel{(3.261)}{=} \bar{s}'_i \leq \bar{s}'_i$$

and for $\bar{\kappa}_i \neq 0$ we have (note that $\bar{\kappa} > 0$ due to assumption)

$$|\bar{x} - \bar{y}_i| \stackrel{(3.252)}{=} 0 \stackrel{(3.261)}{=} \bar{s}'_i \leq \bar{s}'_i .$$

Therefore, it follows

$$\sum_{i \in I} \bar{\kappa}_i (\bar{g}_i + \bar{G}_i(\bar{x} - \bar{y}_i)) \stackrel{(3.260)}{=} \bar{q}' \stackrel{(3.180)}{\in} \partial F(\bar{x}) .$$

- Since F is locally Lipschitz continuous, $\partial F(\bar{x})$ is in particular bounded due to (2.30) and consequently (3.263) follows immediately. ■

13. *Claim:* We have

$$0 \in \partial f(\bar{x}) + \bar{\kappa} \partial F(\bar{x}) .$$

Proof of claim: We calculate

$$0 \stackrel{(3.241)}{=} \sum_{i \in I} \bar{\lambda}_i (\bar{g}_i + \bar{G}_i(\bar{x} - \bar{y}_i)) + \bar{\kappa} \sum_{i \in I} \bar{\kappa}_i (\bar{g}_i + \bar{G}_i(\bar{x} - \bar{y}_i)) \stackrel{(3.259), (3.262), (3.263)}{\in} \partial f(\bar{x}) + \bar{\kappa} \partial F(\bar{x}) . \quad \blacksquare$$

□

Proposition 3.45. *Let (3.189) be satisfied. If there exist $\bar{x} \in \mathbb{R}^N$ and $K \subset \{1, 2, \dots\}$ with $x \xrightarrow{K} \bar{x}$, then*

$$t_L^k v_k \xrightarrow{K} 0 . \quad (3.264)$$

Proof. LUKŠAN & VLČEK [64, p. 382, Proof of Lemma 3.5(ii)]. □

Proposition 3.46. *Let (3.189) be satisfied, let the sequence of (symmetric, positive definite matrices) $\{H_k\}$ be bounded and assume that there exist an infinite subset $K \subset \{1, 2, \dots\}$ and $\bar{x} \in \mathbb{R}^N$ with*

$$x_k \xrightarrow{K} \bar{x} . \quad (3.265)$$

1. *Then for all $i \geq 0$ we have*

$$x_{k+i} \xrightarrow{k \xrightarrow{K} \infty} \bar{x} . \quad (3.266)$$

2. *If additionally*

$$\sigma(\bar{x}) > 0$$

holds, then the following two properties are satisfied:

(a) *For all $i \geq 0$ we have*

$$t_L^{k+i} \xrightarrow{k \xrightarrow{K} \infty} 0 . \quad (3.267)$$

(b) *Let $\varepsilon_0 > 0$ be fixed. Then for all fixed $r \geq 0$ there exists $\tilde{k} \geq 0$ with*

$$w_{k+i} \geq \frac{\sigma(\bar{x})}{2} \quad (3.268)$$

$$t_L^{k+i} < \varepsilon_0 \quad (3.269)$$

for all $k > \tilde{k}$, $k \in K$ and $0 \leq i \leq r$.

Proof.

1. We show (3.266) by induction:

- *Claim:* The base case holds for $i = 0$.

Proof of claim: For $i = 0$, the assumption (3.265) implies

$$x_{k+i} = x_k \xrightarrow{k \xrightarrow{K} \infty} \bar{x} \quad \blacksquare$$

- *Claim:* Let the induction hypothesis be satisfied, i.e. (3.266) holds. Then the inductive step $i \mapsto i + 1$ is true.

Proof of claim: Because of

$$\begin{aligned}
 d_{k+i} &\stackrel{(3.69)}{=} H_{k+i}^2 \left(\sum_{j \in J_{k+i}} \lambda_j^{k+i} g_j^{k+i} + \lambda_p^{k+i} g_p^{k+i} + \sum_{j \in J_{k+i}} \mu_j^{k+i} g_j^{k+i} + \mu_p^{k+i} g_p^{k+i} \right) \\
 &\stackrel{(3.72), (3.73), (3.74)}{\leq} H_{k+i}^2 \left(\sum_{j \in J_{k+i}} \lambda_j^{k+i} g_j^{k+i} + \lambda_p^{k+i} g_p^{k+i} + \bar{\kappa}^{k+i+1} \left(\sum_{j \in J_{k+i}} \kappa_j^{k+i} g_j^{k+i} + \kappa_p^{k+i} g_p^{k+i} \right) \right) \\
 &\stackrel{(3.76), (3.81)}{\leq} H_{k+i}^2 (\tilde{g}_p^{k+i} + \bar{\kappa}^{k+i+1} \tilde{g}_p^{k+i}) \tag{3.270}
 \end{aligned}$$

and

$$\begin{aligned}
 \frac{1}{2} |H_{k+i}(\tilde{g}_p^{k+i} + \bar{\kappa}^{k+i+1} \tilde{g}_p^{k+i})|^2 &\stackrel{(3.109)}{=} \frac{1}{2} d_{k+i}^T (\bar{W}_p^{k+i} + \sum_{j \in J_{k+i}} \mu_j^{k+i} \bar{G}_j^{k+i} + \mu_p^{k+i} \bar{G}^{k+i}) d_{k+i} \\
 &= \frac{1}{2} \underbrace{d_{k+i}^T \bar{W}_p^{k+i} d_{k+i}}_{\geq 0} + \frac{1}{2} d_{k+i}^T \left(\sum_{j \in J_{k+i}} \mu_j^{k+i} \bar{G}_j^{k+i} + \mu_p^{k+i} \bar{G}^{k+i} \right) d_{k+i} \\
 &\leq d_{k+i}^T \bar{W}_p^{k+i} d_{k+i} + \frac{1}{2} d_{k+i}^T \left(\sum_{j \in J_{k+i}} \mu_j^{k+i} \bar{G}_j^{k+i} + \mu_p^{k+i} \bar{G}^{k+i} \right) d_{k+i} \tag{3.271}
 \end{aligned}$$

we obtain

$$\begin{aligned}
 &|x_{k+(i+1)} - x_{k+i}| \\
 &\stackrel{(3.87)}{=} |x_{k+i} + \underbrace{t_L^{k+i} d_{k+i}}_{\in [0,1]} - x_{k+i}| \\
 &\stackrel{(3.270)}{=} t_L^{k+i} |H_{k+i}^2 (\tilde{g}_p^{k+i} + \bar{\kappa}^{k+i+1} \tilde{g}_p^{k+i})| \\
 &\leq t_L^{k+i} |H_{k+i}| \cdot |H_{k+i}(\tilde{g}_p^{k+i} + \bar{\kappa}^{k+i+1} \tilde{g}_p^{k+i})| \\
 &= \sqrt{2} t_L^{k+i} |H_{k+i}| \sqrt{\frac{1}{2} |H_{k+i}(\tilde{g}_p^{k+i} + \bar{\kappa}^{k+i+1} \tilde{g}_p^{k+i})|^2} \\
 &\stackrel{(3.271)}{\leq} \sqrt{2} t_L^{k+i} |H_{k+i}| \left(d_{k+i}^T \bar{W}_p^{k+i} d_{k+i} + \frac{1}{2} d_{k+i}^T \left(\sum_{j \in J_{k+i}} \mu_j^{k+i} \bar{G}_j^{k+i} + \mu_p^{k+i} \bar{G}^{k+i} \right) d_{k+i} \right)^{\frac{1}{2}} \\
 &\leq \sqrt{2} t_L^{k+i} |H_{k+i}| \left(d_{k+i}^T \bar{W}_p^{k+i} d_{k+i} + \frac{1}{2} d_{k+i}^T \left(\sum_{j \in J_{k+i}} \mu_j^{k+i} \bar{G}_j^{k+i} + \mu_p^{k+i} \bar{G}^{k+i} \right) d_{k+i} \right. \\
 &\quad \left. + \underbrace{\alpha_p^{k+i}}_{\stackrel{(3.79)}{\geq 0}} + \underbrace{\bar{\kappa}^{k+i+1} A_p^{k+i}}_{\stackrel{(3.72)}{\geq 0}} + \underbrace{\bar{\kappa}^{k+i+1} (-F(x_{k+i}))}_{\stackrel{(3.84)}{\geq 0}} \right)^{\frac{1}{2}} \\
 &\stackrel{(3.85)}{=} \sqrt{2} t_L^{k+i} |H_{k+i}| \sqrt{-v_{k+i}} \\
 &= \sqrt{2} \underbrace{|H_{k+i}|}_{\text{bounded due to assumption}} \underbrace{\sqrt{t_L^{k+i}}}_{\in [0,1]} \underbrace{\sqrt{-t_L^{k+i} v_{k+i}}}_{\xrightarrow{K} 0 \text{ due to (3.264)}} \xrightarrow{K} 0, \tag{3.272}
 \end{aligned}$$

and therefore it follows

$$|x_{k+(i+1)} - \bar{x}| = |x_{k+(i+1)} - x_{k+i} + x_{k+i} - \bar{x}| \leq \underbrace{|x_{k+(i+1)} - x_{k+i}|}_{\xrightarrow{K} 0 \text{ due to (3.272)}} + \underbrace{|x_{k+i} - \bar{x}|}_{\xrightarrow{K} 0 \text{ due to induction hypothesis (3.266)}} \xrightarrow{K} 0. \blacksquare$$

2. (a) (By contradiction) Suppose (3.267) is false, i.e.

$$\exists i \geq 0, \bar{t} > 0, \bar{K} \subset K : t_L^{k+i} \geq \bar{t} \quad \text{for all } k \in \bar{K} \quad (3.273)$$

(If a sequence does not converge to 0, we can keep (the infinite part of) the sequence away from 0). Since

$$\begin{aligned} 0 &\leq \underbrace{\bar{t}}_{\substack{(3.273) \\ >0}} \underbrace{w_{k+i}}_{\substack{(3.130) \\ \geq 0}} \\ &\stackrel{(3.273)}{\leq} t_L^{k+i} w_{k+i} \\ &\stackrel{(3.86)}{=} t_L^{k+i} \left(\frac{1}{2} |H_{k+i}(\tilde{g}_p^{k+i} + \bar{\kappa}^{k+i+1} \tilde{g}_p^{k+i})|^2 + \tilde{\alpha}_p^{k+i} + \bar{\kappa}^{k+i+1} \tilde{A}_p^{k+i} + \bar{\kappa}^{k+i+1} (-F(x_{k+i})) \right) \\ &\stackrel{(3.109)}{=} \underbrace{t_L^{k+i}}_{\in [0,1]} \left(\frac{1}{2} d_{k+i}^T (\bar{W}_p^{k+i} + \sum_{j \in J_{k+i}} \mu_j^{k+i} \bar{G}_j^{k+i} + \mu_p^{k+i} \bar{G}^{k+i}) d_{k+i} \right. \\ &\quad \left. + \tilde{\alpha}_p^{k+i} + \bar{\kappa}^{k+i+1} \tilde{A}_p^{k+i} + \bar{\kappa}^{k+i+1} (-F(x_{k+i})) \right) \\ &\leq t_L^{k+i} \left(d_{k+i}^T \bar{W}_p^{k+i} d_{k+i} + \frac{1}{2} d_{k+i}^T \left(\sum_{j \in J_{k+i}} \mu_j^{k+i} \bar{G}_j^{k+i} + \mu_p^{k+i} \bar{G}^{k+i} \right) d_{k+i} \right. \\ &\quad \left. + \tilde{\alpha}_p^{k+i} + \bar{\kappa}^{k+i+1} \tilde{A}_p^{k+i} + \bar{\kappa}^{k+i+1} (-F(x_{k+i})) \right) \\ &\stackrel{(3.85)}{=} -t_L^{k+i} v_{k+i} \xrightarrow{\bar{K}} 0 \end{aligned}$$

due to (3.264), we have

$$w_{k+i} \xrightarrow{\bar{K}} 0$$

and therefore we obtain

$$\sigma(\bar{x}) \stackrel{(3.211)}{=} \liminf_{k \rightarrow \infty} \max (w_k, |x_k - \bar{x}|) \leq \lim_{\bar{K}} \max (\underbrace{w_{k+i}}_{\xrightarrow{\bar{K}} 0}, \underbrace{|x_{k+i} - \bar{x}|}_{\xrightarrow{\bar{K}} 0 \text{ due to (3.266)}}) = 0 ,$$

which is a contradiction to the assumption $\sigma(\bar{x}) > 0$.

(b) Let $r \geq 0$ be fixed and $0 \leq i \leq r$. Since

$$\begin{aligned} \frac{\sigma(\bar{x})}{2} &\stackrel{\sigma(\bar{x}) > 0 \text{ due to assumption}}{<} \downarrow \sigma(\bar{x}) \stackrel{(3.211)}{=} \liminf_{k \rightarrow \infty} \max (w_k, |x_k - \bar{x}|) \\ &\leq \lim_{\bar{K}} \max (w_{k+i}, |x_{k+i} - \bar{x}|) \stackrel{(3.266)}{=} \lim_{\bar{K}} w_{k+i} \end{aligned}$$

and (3.267) and because $\varepsilon_0 > 0$ is a fixed number by assumption, there exist $k_i \geq 0$ with

$$\frac{\sigma(\bar{x})}{2} \leq w_{k+i} , \quad t_L^{k+i} < \varepsilon_0$$

for all $k > k_i$ with $k \in K$. Now, setting

$$\tilde{k} := \max \{k_i : 0 \leq i \leq r\}$$

yields (3.268) and (3.269). □

Proposition 3.47. Let $p, g, \Delta \in \mathbb{R}^N$ and $c, u, w, \beta \in \mathbb{R}$, $m \in (0, 1)$, $\alpha \geq 0$ with

$$w = \frac{1}{2}|p|^2 + \alpha \quad (3.274)$$

$$v = -(|p|^2 + \alpha) \quad (3.275)$$

$$-\beta - g^T p \geq mv \quad (3.276)$$

$$c = \max(|g|, |p|, \sqrt{\alpha}) \quad (3.277)$$

and define $Q : \mathbb{R} \rightarrow \mathbb{R}$ by

$$Q(\nu) := \frac{1}{2}|\nu g + (1 - \nu)(p + \Delta)|^2 + \nu\beta + (1 - \nu)\alpha, \quad (3.278)$$

then

$$\min_{\nu \in [0, 1]} Q(\nu) \leq w - w^2 \frac{(1-m)^2}{8c^2} + 4c|\Delta| + \frac{1}{2}|\Delta|^2. \quad (3.279)$$

Proof. LUKŠAN & VLČEK [64, p. 381, Lemma 3.4]. □

We introduce the following notation (cf. MAGNUS & NEUDECKER [69, p. 31, Section 2 resp. p. 34, Section 4]).

Definition 3.48. Let $A, B \in \mathbb{R}^{N \times N}$ with

$$A := (A_{:,1} \dots A_{:,N}) = \begin{pmatrix} A_{11} & \dots & A_{1N} \\ \vdots & & \vdots \\ A_{N1} & \dots & A_{NN} \end{pmatrix}.$$

- We define the Kronecker product of A and B by

$$A \otimes B := \begin{pmatrix} A_{11}B & \dots & A_{1N}B \\ \vdots & & \vdots \\ A_{N1}B & \dots & A_{NN}B \end{pmatrix} \in \mathbb{R}^{N^2 \times N^2}. \quad (3.280)$$

- We define the vectorization of A by

$$A_{(\cdot)} := \begin{pmatrix} A_{:,1} \\ \vdots \\ A_{:,N} \end{pmatrix} \in \mathbb{R}^{N^2}.$$

- We define the Frobenius norm of A by

$$|A|_F := \left(\sum_{i,j=1}^N A_{ij}^2 \right)^{\frac{1}{2}}. \quad (3.281)$$

Proposition 3.49. Let $A, B, C \in \mathbb{R}^{N \times N}$. Then

$$|A| \leq |A|_F \leq \sqrt{N}|A| \quad (3.282)$$

$$(ABC)_{(\cdot)} = (C^T \otimes A)B_{(\cdot)} \quad (3.283)$$

$$|A \otimes A| \leq N|A|^2. \quad (3.284)$$

Proof.

- (3.282) holds due to GOLUB & VAN LOAN [27, p. 56, Section 2.3.2].

- (3.283) holds due to MAGNUS & NEUDECKER [69, p. 35, Theorem 2].
- (3.284) holds due to

$$\begin{aligned}
|A \otimes A|^2 &\stackrel{(3.282)}{\leq} |A \otimes A|_F^2 \\
&\stackrel{(3.280)}{=} \left| \begin{pmatrix} A_{11}A & \dots & A_{1N}A \\ \vdots & & \vdots \\ A_{N1}A & \dots & A_{NN}A \end{pmatrix} \right|_F^2 \stackrel{(3.281)}{=} \sum_{i,j=1}^N |A_{ij}A|_F^2 \\
&= \left(\sum_{i,j=1}^N A_{ij}^2 \right) |A|_F^2 \stackrel{(3.281)}{=} |A|_F^2 \cdot |A|_F^2 = |A|_F^4 \\
&\stackrel{(3.281)}{\leq} (\sqrt{N}|A|)^4 = N^2 |A|^4 \\
\stackrel{\sqrt{\cdot}}{\iff} |A \otimes A| &\leq N |A|^2 . \quad \square
\end{aligned}$$

Now, we introduce differentiability of matrix valued functions (cf. MAGNUS & NEUDECKER [69, p. 107, Definition 3]).

Definition 3.50. Let $A : \mathbb{R}^p \longrightarrow \mathbb{R}^{N \times N}$ and $\mu_0 \in \mathbb{R}^p$ be fixed. If there exists $B(\mu_0) \in \mathbb{R}^{N^2 \times p}$ with

$$A_{(\cdot)}(\mu_0 + \mu) = A_{(\cdot)}(\mu_0) + B(\mu_0)\mu + R_{(\cdot)}(\mu_0, \mu) \quad (3.285)$$

for all $\mu \in \mathbb{R}^p$ in a neighborhood of μ_0 and

$$\lim_{\mu \rightarrow 0} \frac{R_{(\cdot)}(\mu_0, \mu)}{|\mu|} = 0 ,$$

then A is said to be differentiable at μ_0 . Furthermore, the $N \times N$ -matrix $dA(\mu_0, \mu)$ defined by

$$dA_{(\cdot)}(\mu_0, \mu) := B(\mu_0)\mu \in \mathbb{R}^{N^2} \quad (3.286)$$

is called the (first) differential of A at μ_0 with increment μ and $B(\mu_0)$ is called the first derivative of A at μ_0 .

Proposition 3.51. *Let*

$$T := \{Y : Y \in \mathbb{R}^{N \times N}, \det Y \neq 0\}$$

be the set of non-singular $N \times N$ -matrices. If $A : \mathbb{R}^p \longrightarrow T$ is k times (continuously) differentiable, then so is $B : \mathbb{R}^p \longrightarrow T$ defined by

$$B(\mu) := A(\mu)^{-1}$$

and

$$dB(\mu_0, \mu) = -B(\mu_0)dA(\mu_0, \mu)B(\mu_0) . \quad (3.287)$$

Proof. MAGNUS & NEUDECKER [69, p. 156, Theorem 3]. □

Example 3.52. Consider $A : \mathbb{R} \longrightarrow \mathbb{R}$

$$A(x) := a(x - x_0) \quad (3.288)$$

with $a \in \mathbb{R} \setminus \{0\}$ and $x_0 \in \mathbb{R}$. Then

$$A'(x) \stackrel{(3.288)}{=} a \implies dA(x, h) \stackrel{(3.286)}{=} A'(x)h = ah . \quad (3.289)$$

Therefore, Proposition 3.51 implies the differentiability of

$$B(x) := A(x)^{-1} = (a(x - x_0))^{-1} = a^{-1}(x - x_0)^{-1} \quad (3.290)$$

and we calculate

$$B'(x) \stackrel{(3.290)}{=} -a^{-1}(x - x_0)^{-2} . \quad (3.291)$$

Furthermore, we obtain

$$\begin{aligned} dB(x, h) &\stackrel{(3.287)}{=} -B(x)dA(x, h)B(x) \\ &\stackrel{(3.290), (3.289)}{\stackrel{\downarrow}{=}} -(a^{-1}(x - x_0)^{-1})(ah)(a^{-1}(x - x_0)^{-1}) \\ &\stackrel{(3.291)}{=} B'(x)h . \end{aligned}$$

We recall the mean value theorem for vector valued functions.

Theorem 3.53. *Let $f : \Omega \subseteq \mathbb{R}^p \longrightarrow \mathbb{R}^q$ (with Ω open) be continuously differentiable.*

1. *Let $x, y \in \Omega$ with*

$$\begin{aligned} \Omega_{x,y} &:= \{(1-t)x + ty : t \in [0, 1]\} \subseteq \Omega \\ \omega_{x,y} &:= \sup_{z \in \Omega_{x,y}} |f'(z)| , \end{aligned}$$

where $f'(z) \in \mathbb{R}^{q \times p}$ denotes the derivative of f at z , then

$$\begin{aligned} f(y) - f(x) &= \left(\int_0^1 f'((1-t)x + ty) dt \right) (y - x) \\ |f(y) - f(x)| &\leq \omega_{x,y} |y - x| . \end{aligned}$$

2. *Let Ω be convex and*

$$\omega := \sup_{z \in \Omega} |f'(z)| < \infty ,$$

then

$$|f(y) - f(x)| \leq \omega |y - x| \quad (3.292)$$

for all $x, y \in \Omega$ (i.e. f is Lipschitz continuous on Ω due to Definition 2.21).

Proof.

1. HEUSER [38, p. 278, 167.4 Mittelwertsatz für vektorwertige Funktionen].

2. HEUSER [38, p. 280, Exercise 3]. □

Proposition 3.54. *Let $\{\bar{\kappa}^{k+1}\}$ be bounded and let $\{(\bar{W}_p^k)^{-\frac{1}{2}}\}$ be bounded and uniformly positive definite. For $k \geq 1$ we define $Z_k : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}^{N \times N}$*

$$Z_k(s) := (\bar{W}_p^k + s\bar{G}^k)^{-\frac{1}{2}}. \quad (3.293)$$

Then we have for all $k \geq 1$

$$|Z_k(\bar{\kappa}^{k+2}) - Z_k(\bar{\kappa}^{k+1})| \leq C_5 |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}| \quad (3.294)$$

$$0 < C_5 < \infty, \quad (3.295)$$

where

$$C_5 := C_2 C_4 \quad (3.296)$$

$$C_4 := N C_0^2 C_3 \quad (3.297)$$

$$C_3 := N^{\frac{1}{2}} \bar{C}_G \quad (3.298)$$

and C_2 is a positive constant.

Proof.

1. *Claim:* We define for all $k \geq 1$

$$Y_k(s) := (\bar{W}_p^k + s\bar{G}^k)^{-1}. \quad (3.299)$$

Then we have for all $k \geq 1$

$$|Y_k(\bar{\kappa}^{k+2})^{\frac{1}{2}} - Y_k(\bar{\kappa}^{k+1})^{\frac{1}{2}}| \leq C_2 |Y_k(\bar{\kappa}^{k+2}) - Y_k(\bar{\kappa}^{k+1})|. \quad (3.300)$$

with a positive constant C_2 .

Proof of claim: We have for all $k \geq 1$

$$\begin{aligned} |Y_k(\bar{\kappa}^{k+1})^{-1}| &\stackrel{(3.299)}{=} |\bar{W}_p^k + \bar{\kappa}^{k+1} \bar{G}^k| \\ &\stackrel{(3.208)}{\leq} C_1, \end{aligned}$$

which is equivalent to $\{Y_k(\bar{\kappa}^{k+1})\}$ being uniformly positive definite due to (3.199), i.e. there exists $\tilde{C}_2 > 0$ with

$$\lambda_{\min}(Y_k(\bar{\kappa}^{k+1})) \geq \tilde{C}_2 \stackrel{\wedge \frac{1}{2}}{\iff} (\lambda_{\min}(Y_k(\bar{\kappa}^{k+1})))^{\frac{1}{2}} \geq \tilde{C}_2^{\frac{1}{2}}. \quad (3.301)$$

Consequently, we obtain for all $k \geq 1$

$$\begin{aligned} &(\lambda_{\min}(Y_k(\bar{\kappa}^{k+2})))^{\frac{1}{2}} + (\lambda_{\min}(Y_k(\bar{\kappa}^{k+1})))^{\frac{1}{2}} \stackrel{(3.301)}{\geq} 2\tilde{C}_2^{\frac{1}{2}} \\ &\stackrel{\wedge -1}{\iff} \frac{1}{(\lambda_{\min}(Y_k(\bar{\kappa}^{k+2})))^{\frac{1}{2}} + (\lambda_{\min}(Y_k(\bar{\kappa}^{k+1})))^{\frac{1}{2}}} \leq \frac{1}{2} \tilde{C}_2^{-\frac{1}{2}} \end{aligned} \quad (3.302)$$

and hence we estimate for all $k \geq 1$

$$\begin{aligned} |Y_k(\bar{\kappa}^{k+2})^{\frac{1}{2}} - Y_k(\bar{\kappa}^{k+1})^{\frac{1}{2}}| &\stackrel{(3.196)}{\leq} \frac{1}{(\lambda_{\min}(Y_k(\bar{\kappa}^{k+2}))^{\frac{1}{2}} + (\lambda_{\min}(Y_k(\bar{\kappa}^{k+1}))^{\frac{1}{2}}))} |Y_k(\bar{\kappa}^{k+2}) - Y_k(\bar{\kappa}^{k+1})| \\ &\stackrel{(3.302)}{\leq} \frac{1}{2} \tilde{C}_2^{-\frac{1}{2}} |Y_k(\bar{\kappa}^{k+2}) - Y_k(\bar{\kappa}^{k+1})| , \end{aligned}$$

where setting

$$C_2 := \frac{1}{2} \tilde{C}_2^{-\frac{1}{2}} > 0$$

yields the desired result. ■

2. *Claim:* We define for $k \geq 1$

$$X_k(s) := \bar{W}_p^k + U_k(s) , \quad (3.303)$$

where

$$U_k(s) := s \bar{G}^k . \quad (3.304)$$

Then X_k is continuously differentiable, the differential of X_k at s is given by

$$dX_{k,(\cdot)}(s, t-s) = \hat{U}_k(t-s) \quad (3.305)$$

and we have

$$|\hat{U}_k| \leq C_3 , \quad (3.306)$$

where

$$\hat{U}_k := \bar{G}_{(\cdot)}^k \in \mathbb{R}^{N^2} . \quad (3.307)$$

Proof of claim: We calculate

$$\begin{aligned} U_{k,(\cdot)}(t-s) &\stackrel{(3.304)}{=} (\bar{G}^k(t-s))_{(\cdot)} \\ &\stackrel{(3.307)}{=} \hat{U}_k(t-s) . \end{aligned} \quad (3.308)$$

Therefore, we have for all $k \geq 1$ and for all $s, t \in \mathbb{R}$

$$\begin{aligned} X_k(t) &\stackrel{(3.303)}{=} \bar{W}_p^k + U_k(t) \stackrel{(3.304)}{=} \bar{W}_p^k + t \bar{G}^k \\ &\stackrel{(3.304)}{=} \bar{W}_p^k + U_k(s) + U_k(t-s) \stackrel{(3.303)}{=} X_k(s) + U_k(t-s) , \end{aligned} \quad (3.309)$$

which is equivalent to

$$X_{k,(\cdot)}(t) \stackrel{(3.309)}{=} X_{k,(\cdot)}(s) + U_{k,(\cdot)}(t-s) \stackrel{(3.308)}{=} X_{k,(\cdot)}(s) + \hat{U}_k(t-s) .$$

Consequently, (3.285) and (3.286) imply that the differential of X_k at s is given by

$$dX_{k,(\cdot)}(s, t-s) = \hat{U}_k(t-s)$$

(with $R_k(s, t-s) \equiv 0$) and that the derivative of X_k at s denoted by X'_k is constant

$$X'_k(s) = \hat{U}_k \quad \text{for all } s \in \mathbb{R} ,$$

which implies that X_k is continuously differentiable. Furthermore, we estimate

$$\begin{aligned}
|\hat{U}_k|^2 &\stackrel{(3.307)}{=} |\bar{G}_{(\cdot)}^k|^2 \stackrel{(3.281)}{=} |\bar{G}^k|_F^2 \stackrel{(3.282)}{\leq} (\sqrt{N}|\bar{G}^k|)^2 = N|\bar{G}^k|^2 \\
&\text{initialization of Algorithm 3.20} \\
&\downarrow \\
&\leq N\bar{C}_G^2 \\
\iff |\hat{U}_k| &\leq \sqrt{N}\bar{C}_G \stackrel{(3.298)}{=} C_3 .
\end{aligned}$$

■

3. *Claim:* We have for all $k \geq 1$ and for all $s \geq 0$

$$|Y_k(s)| \leq C_0 . \quad (3.310)$$

Proof of claim: Since $\bar{G}^k$ is symmetric and positive definite, we obtain that $U_k(s)$ is symmetric and positive semidefinite for all $s \geq 0$ (cf. (3.304)). Consequently, we have

$$\bar{W}_p^k \stackrel{(3.191)}{\preceq} \bar{W}_p^k + U_k(s) \stackrel{(3.303)}{=} X_k(s) .$$

Therefore, we estimate for all $k \geq 1$ and for all $s \geq 0$

$$|Y_k(s)| \stackrel{(3.299),(3.304),(3.303)}{\leq} |X_k(s)^{-1}| \stackrel{(3.195)}{\leq} |(\bar{W}_p^k)^{-1}| \stackrel{(3.207)}{\leq} C_0 .$$

■

4. *Claim:* For $k \geq 1$ Y_k is continuously differentiable, the derivative of Y_k at s is given by

$$V_k(s) := (Y_k(s) \otimes Y_k(s)) \hat{U}_k \in \mathbb{R}^{N^2} \quad (3.311)$$

and we have

$$\sup_{s \geq 0} |V_k(s)| \leq C_4 . \quad (3.312)$$

Proof of claim: Since X_k is continuously differentiable (cf. (3.305)), Proposition 3.51 yields the continuous differentiability of

$$Y_k(s) \stackrel{(3.299)}{=} (\bar{W}_p^k + \bar{G}^k)^{-1} \stackrel{(3.303)}{=} (\bar{W}_p^k + U_k(s))^{-1} \stackrel{(3.304)}{=} X_k(s)^{-1}$$

and

$$dY_k(s, t-s) \stackrel{(3.287)}{=} -Y_k(s) dX_k(s, t-s) Y_k(s) . \quad (3.313)$$

Hence, we obtain

$$\begin{aligned}
dY_{k,(\cdot)}(s, t-s) &\stackrel{(3.313)}{=} (Y_k(s) dX_k(s, t-s) Y_k(s))_{(\cdot)} \\
&\stackrel{(3.283)}{=} (Y_k(s)^T \otimes Y_k(s)) dX_{k,(\cdot)}(s, t-s) \\
&\stackrel{Y_k(s) \in \mathbb{R}_{\text{sym}}^{N \times N}, (3.305)}{\leq} (Y_k(s) \otimes Y_k(s)) \hat{U}_k(t-s) \\
&\stackrel{(3.311)}{=} V_k(s)(t-s)
\end{aligned}$$

and therefore (3.286) implies that $V_k(s)$ is the derivative of Y_k at s . Furthermore, we estimate

for all $k \geq 1$ and for all $s \geq 0$

$$|V_k(s)| \stackrel{(3.311)}{=} |(Y_k(s) \otimes Y_k(s))\hat{U}_k| \stackrel{(3.284),(3.306)}{\leq} N|Y_k(s)|^2 C_3 \stackrel{(3.310)}{\leq} NC_0^2 C_3 \stackrel{(3.297)}{=} C_4, \quad (3.314)$$

which is equivalent to the fact that for all $k \geq 1$

$$\sup_{s \geq 0} |V_k(s)| \stackrel{(3.314)}{\leq} C_4. \quad \blacksquare$$

5. *Claim:* We have for all $k \geq 1$ and for all $s, t \geq 0$

$$|Y_{k,(\cdot)}(t) - Y_{k,(\cdot)}(s)| \leq C_4 |t - s|. \quad (3.315)$$

Proof of claim: Since Y_k is continuously differentiable for all $s, t \in S$ defined by

$$S := \{\xi \in \mathbb{R} : \xi \geq 0\}$$

(Note that S is an interval and consequently convex) and since the derivative of Y_k at s is given by $V_k(s)$ (cf. (3.311)) and since the norm of the derivative $|V_k(s)|$ is bounded on S due to (3.312), we obtain

$$|Y_{k,(\cdot)}(t) - Y_{k,(\cdot)}(s)| \stackrel{(3.292)}{\leq} \left(\sup_{\xi \geq 0} |V_k(\xi)| \right) |t - s| \stackrel{(3.312)}{\leq} C_4 |t - s|$$

for all $s, t \in S$. \blacksquare

6. *Claim:* We have for all $k \geq 1$

$$|Z_k(\bar{\kappa}^{k+2}) - Z_k(\bar{\kappa}^{k+1})| \leq C_5 |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}|.$$

Proof of claim: We estimate for all $k \geq 1$

$$\begin{aligned} |Z_k(\bar{\kappa}^{k+2}) - Z_k(\bar{\kappa}^{k+1})| &\stackrel{(3.293),(3.299)}{\leq} |Y_k(\bar{\kappa}^{k+2})^{\frac{1}{2}} - Y_k(\bar{\kappa}^{k+1})^{\frac{1}{2}}| \\ &\stackrel{(3.300)}{\leq} C_2 |Y_k(\bar{\kappa}^{k+2}) - Y_k(\bar{\kappa}^{k+1})| \\ &\stackrel{(3.282)}{\leq} C_2 |Y_k(\bar{\kappa}^{k+2}) - Y_k(\bar{\kappa}^{k+1})|_F \\ &\stackrel{(3.281)}{=} C_2 |(Y_k(\bar{\kappa}^{k+2}) - Y_k(\bar{\kappa}^{k+1}))_{(\cdot)}| \\ &\stackrel{(3.315)}{\leq} C_2 C_4 |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}| \\ &\stackrel{(3.296)}{=} C_5 |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}|. \end{aligned}$$

Furthermore, we obtain

$$C_5 \stackrel{(3.296)}{=} C_2 C_4 \stackrel{(3.297)}{=} C_2 (NC_0^2 C_3) \stackrel{(3.298)}{=} C_2 (NC_0^2) (N^{\frac{1}{2}} \bar{\tilde{C}}_G) = C_2 N^{\frac{3}{2}} C_0^2 \bar{\tilde{C}}_G$$

and therefore the fact that C_2 is a positive constant due to (3.300), the fact that $N \geq 1$ is a fixed finite natural number, combining (3.207) with the positive definiteness of $\bar{W}_p^k$, and the initialization of Algorithm 3.20 yield (3.295). \blacksquare

□

Theorem 3.55. *Let (3.189) be satisfied. Furthermore, let the sequence $\{(x_k, \bar{\kappa}^{k+1})\}$ be bounded, let the sequence (of symmetric, positive definite matrices) $\{(\bar{W}_p^k)^{-\frac{1}{2}}\}$ be bounded as well as uniformly positive definite and let $\bar{x} \in \mathbb{R}^n$ be any accumulation point of $\{x_k\}$, i.e. there exists (an infinite set) $K \subset \{1, 2, \dots\}$ with*

$$x_k \xrightarrow{K} \bar{x} , \quad (3.316)$$

and demand

$$\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1} \xrightarrow{K} 0 \quad (3.317)$$

$$t_0^{\inf} := \inf_{k \geq 0} t_0^k > 0 \quad (3.318)$$

(cf. Remark 3.26). Then there exists $\bar{\kappa} \in \mathbb{R}_{\geq 0}$ with

$$\begin{aligned} 0 &\in \partial f(\bar{x}) + \bar{\kappa} \partial F(\bar{x}) \\ \bar{\kappa} F(\bar{x}) &= 0 \\ F(\bar{x}) &\leq 0 , \end{aligned}$$

i.e. each accumulation point of the sequence of iteration points $\{x_k\}$ is stationary for the optimization problem (2.35).

Proof.

1. *Statement:* We have

$$\sigma(\bar{x}) = 0 . \quad (3.319)$$

Proof of statement: (By contradiction) Suppose (3.319) is false, i.e. we have

$$\sigma(\bar{x}) \stackrel{(3.211)}{>} 0 \quad \text{or} \quad \sigma(\bar{x}) \stackrel{(3.211)}{=} \infty . \quad (3.320)$$

(a) *Claim:* We have

$$\{y_k\}, \{\rho_k G_k\}, \{\hat{\rho}_k \hat{G}_k\} \text{ and } \{g_k\} \text{ are bounded} . \quad (3.321)$$

Proof of claim: This statement was shown in the proof of Proposition 3.44 (cf. (3.229), (3.233), (3.235), (3.231)), where only the assumption of the boundedness of $\{x_k\}$ was used. Consequently, this statement is here also true. $\blacksquare$

(b) *Claim:* We have

$$\{H_k\} \text{ is bounded} . \quad (3.322)$$

Proof of claim: Since $\{(\bar{W}_p^k)^{-\frac{1}{2}}\}$ is bounded due to assumption, $\{H_k\}$ is bounded due to Corollary 3.43. $\blacksquare$

(c) *Claim:* We have

$$\{g_k^k\}, \{H_k g_k^k\} \text{ and } \{\alpha_k^k\} \text{ are bounded} . \quad (3.323)$$

Proof of claim:

• Because of

$$|g_k^k| \stackrel{(3.104)}{=} |g_k + \rho_k G_k(x_k - y_k)| \leq |g_k| + |\rho_k G_k|(|x_k| + |y_k|) ,$$

the assumption of the boundedness of $\{x_k\}$ and (3.321), $\{g_k^k\}$ is bounded. From this, (3.322) and $|H_k g_k^k| \leq |H_k| \cdot |g_k^k|$, we get the boundedness of $\{H_k g_k^k\}$.

- Because of

$$\begin{aligned}
 \alpha_k^k &\stackrel{(3.64)}{=} \max(|f_k^k - f(x_k)|, \gamma_1(s_k^k)^{\omega_1}) \\
 &\stackrel{(3.99), (3.94), (3.89)}{=} \max(|f(y_k) + g_k^T(x_k - y_k) + \frac{1}{2}\rho_k(x_k - y_k)^T G_k(x_k - y_k) - f(x_k)|, \\
 &\quad \gamma_1|x_k - y_k|^{\omega_1}) \\
 &\stackrel{(2.1)}{\leq} \max(|f(y_k) - f(x_k)| + |g_k| \cdot |x_k - y_k| + \frac{1}{2}|x_k - y_k| \cdot |\rho_k G_k| \cdot |x_k - y_k|, \\
 &\quad \gamma_1|x_k - y_k|^{\omega_1}) \\
 &\stackrel{\omega_1 \geq 1}{\leq} \max(|f(y_k)| + |f(x_k)| + |g_k|(|x_k| + |y_k|) + \frac{1}{2}|\rho_k G_k|(|x_k| + |y_k|)^2, \\
 &\quad \gamma_1(|x_k| + |y_k|)^{\omega_1}) ,
 \end{aligned}$$

the boundedness of $\{x_k\}$, (3.321), and the fact that f is continuous on (the whole) $\mathbb{R}^n$ (cf. Assumption 3.1), $\{\alpha_k^k\}$ is bounded. ■

(d) *Claim:* We have for $k \geq 1$

$$\hat{w}_k \leq \frac{1}{2}|H_k g_k^k|^2 + \alpha_k^k . \quad (3.324)$$

Proof of claim: For $k \geq 1$, $(\lambda, \lambda_p, \mu, \mu_p) \in \mathbb{R}^{2(|J_k|+1)}$ with

$$\begin{aligned}
 \lambda_j &:= \begin{cases} 1 & \text{for } j = k \\ 0 & \text{for } j \in J_k \setminus \{k\} \end{cases} \\
 \lambda_p &:= 0 \\
 \mu_j &:= 0 \quad \text{for all } j \in J_k \\
 \mu_p &:= 0
 \end{aligned} \quad (3.325)$$

is feasible for the (dual) problem (3.42) (Note: This problem is written as a minimization problem) and therefore we obtain (Note: $\hat{w}_k$ is the optimal function value of (3.42))

$$\begin{aligned}
 \hat{w}_k &\stackrel{(3.127)}{=} \frac{1}{2}|H_k(\hat{g}_p^k + \bar{\kappa}^{k+1}\hat{g}_p^k)|^2 + \hat{\alpha}_p^k + \bar{\kappa}^{k+1}\hat{A}_p^k + \bar{\kappa}^{k+1}(-F(x_k)) \\
 &\stackrel{(3.76), (3.125), (3.81), (3.126)}{=} \frac{1}{2}|H_k\left(\sum_{j \in J_k} \lambda_j^k g_j^k + \lambda_p^k g_p^k + \bar{\kappa}^{k+1}\left(\sum_{j \in J_k} \kappa_j^k \hat{g}_j^k + \kappa_p^k \hat{g}_p^k\right)\right)|^2 \\
 &\quad + \sum_{j \in J_k} \lambda_j^k \alpha_j^k + \lambda_p^k \alpha_p^k + \bar{\kappa}^{k+1}\left(\sum_{j \in J_k} \kappa_j^k A_j^k + \kappa_p^k A_p^k\right) \\
 &\stackrel{(3.73), (3.74), (3.72)}{=} \frac{1}{2}|H_k\left(\sum_{j \in J_k} \lambda_j^k g_j^k + \lambda_p^k g_p^k + \sum_{j \in J_k} \mu_j^k \hat{g}_j^k + \mu_p^k \hat{g}_p^k\right)|^2 \\
 &\quad + \sum_{j \in J_k} \lambda_j^k \alpha_j^k + \lambda_p^k \alpha_p^k + \sum_{j \in J_k} \mu_j^k A_j^k + \mu_p^k A_p^k \\
 &\stackrel{(\lambda, \lambda_p, \mu, \mu_p) \text{ from (3.325) feasible}}{\leq} \frac{1}{2}|H_k\left(\sum_{j \in J_k} \lambda_j^k g_j^k + \lambda_p^k g_p^k\right)|^2 + \sum_{j \in J_k} \lambda_j^k \alpha_j^k + \lambda_p^k \alpha_p^k \\
 &\stackrel{(3.325)}{=} \frac{1}{2}|H_k g_k^k|^2 + \alpha_k^k .
 \end{aligned} \quad \blacksquare$$

(e) *Claim:* We have

$$\{w_k\}, \{\tilde{g}_p^k\}, \{\tilde{g}_p^k\}, \{\tilde{\alpha}_p^k\}, \{\bar{\kappa}^{k+1}\tilde{A}_p^k\}, \{H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1}\tilde{g}_p^k)\} \text{ and } \{\tau_k\} \text{ are bounded ,} \quad (3.326)$$

where

$$\tau_k := \tilde{\alpha}_p^k + \bar{\kappa}^{k+1}\tilde{A}_p^k + \bar{\kappa}^{k+1}(-F(x_k)) \geq 0 . \quad (3.327)$$

Proof of claim:

- Due to (3.79), (3.72), (3.84), (3.72) and $F(x_k) < 0$, we have

$$\tau_k \stackrel{(3.327)}{=} \tilde{\alpha}_p^k + \bar{\kappa}^{k+1}\tilde{A}_p^k + \bar{\kappa}^{k+1}(-F(x_k)) \geq 0 . \quad (3.328)$$

Hence, since

$$\begin{aligned} 0 &\leq \frac{1}{2}|H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1}\tilde{g}_p^k)|^2 + \tilde{\alpha}_p^k + \bar{\kappa}^{k+1}\tilde{A}_p^k + \bar{\kappa}^{k+1}(-F(x_k)) \\ &\stackrel{(3.327)}{=} \frac{1}{2}|H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1}\tilde{g}_p^k)|^2 + \tau_k \stackrel{(3.86)}{=} w_k \stackrel{(3.130)}{\leq} \hat{w}_k \\ &\stackrel{(3.324)}{\leq} \frac{1}{2}|H_k g_k^k|^2 + \alpha_k^k \end{aligned}$$

and (3.323) as well as (3.328), we obtain that $\{w_k\}$, $\{\tilde{\alpha}_p^k\}$, $\{\bar{\kappa}^{k+1}\tilde{A}_p^k\}$, $\{H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1}\tilde{g}_p^k)\}$ and $\{\tau_k\}$ are bounded.

- Consider the proof of Proposition 3.44: There we only used the consequence (3.212) ($x_k \xrightarrow{K} \bar{x}$ and this property is also satisfied here due to (3.316)) of the assumption $\sigma(\bar{x}) = 0$ for showing (3.239) resp. (3.240) (convergence of $\tilde{g}_p^k$ resp. $\tilde{g}_p^k$ on a subsequence). Consequently, $\tilde{g}_p^k$ resp. $\tilde{g}_p^k$ are also bounded here.

The second property (3.213) ($w_k \xrightarrow{K} 0$) resulting from $\sigma(\bar{x}) = 0$ there, is first used directly after the proof of (3.241) (and therefore in particular after the proof of (3.239) resp. (3.240)). If this property was already used in the proof of (3.239) or (3.240), the above implication would be false, since then indeed $\sigma(\bar{x}) = 0$ (and not only $x_k \xrightarrow{K} \bar{x}$) would be used in the proof of (3.239) or (3.240), and the situation here is $\sigma(\bar{x}) \stackrel{(3.320)}{>} 0$. ■

(f) *Claim:* We have

$$\sigma(\bar{x}) \text{ is finite ,} \quad (3.329)$$

i.e. in (3.320) only the case $\sigma(\bar{x}) > 0$ occurs.

Proof of claim: Because of

$$\sigma(\bar{x}) \stackrel{(3.211)}{=} \liminf_{k \rightarrow \infty} \max(w_k, |x_k - \bar{x}|) ,$$

the assumption of the boundedness of $\{x_k\}$ and (3.326), $\sigma(\bar{x})$ is finite. ■

(g) *Claim:* We have

$$x_{k+1} - x_k \xrightarrow{K} 0 . \quad (3.330)$$

Proof of claim: The assumptions of Proposition 3.46 for applying (3.266) are satisfied:

- Due to (3.316), we have $x_k \xrightarrow{K} \bar{x}$.
- Due to (3.329), we have $\sigma(\bar{x}) > 0$.

Therefore, applying (3.266) for $i = 1$ and $i = 0$ yields

$$\lim_K (x_{k+1} - x_k) = \lim_K x_{k+1} - \lim_K x_k \stackrel{(3.266)}{=} \bar{x} - \bar{x} = 0 . \quad \blacksquare$$

(h) *Claim:* We have

$$s_p^{k+1} - \tilde{s}_p^k \xrightarrow{K} 0 \quad (3.331)$$

$$\hat{s}_p^{k+1} - \tilde{\hat{s}}_p^k \xrightarrow{K} 0 . \quad (3.332)$$

Proof of claim: We calculate

$$\lim_K (s_p^{k+1} - \tilde{s}_p^k) \stackrel{(3.95)}{=} \lim_K |x_{k+1} - x_k| \stackrel{(3.330)}{=} 0$$

and

$$\lim_K (\hat{s}_p^{k+1} - \tilde{\hat{s}}_p^k) \stackrel{(3.98)}{=} \lim_K |x_{k+1} - x_k| \stackrel{(3.330)}{=} 0 . \quad \blacksquare$$

(i) *Claim:* We have

$$f(x_{k+1}) - f(x_k) \xrightarrow{K} 0 \quad (3.333)$$

$$F(x_{k+1}) - F(x_k) \xrightarrow{K} 0 . \quad (3.334)$$

Proof of claim: This follows directly from (3.316), (3.330), and the continuity of f and F . $\blacksquare$

(j) *Claim:* We have

$$|G_p^{k+1}| \leq C_G \quad (3.335)$$

$$|\hat{G}_p^{k+1}| \leq \hat{C}_G . \quad (3.336)$$

Proof of claim: Due to (3.189) the assumptions of Proposition 3.34 are satisfied and therefore we can make use of it. We estimate

$$\begin{aligned} |G_p^{k+1}| &\stackrel{(3.153)}{=} \left| \sum_{j=1}^k \hat{\lambda}_j^k \rho_j G_j \right| \leq \sum_{j=1}^k |\hat{\lambda}_j^k \rho_j G_j| \\ &\stackrel{(3.152)}{=} \sum_{j=1}^k \hat{\lambda}_j^k |\rho_j G_j| \stackrel{(3.321)}{\leq} \sum_{j=1}^k \hat{\lambda}_j^k C_G \\ &\stackrel{(3.152), \text{compare (3.237)}}{\leq} C_G \end{aligned}$$

(Note that (3.237) was shown in the proof of Proposition 3.44, where we only used (3.90), and hence (3.237) holds here, too) and

$$\begin{aligned} |\hat{G}_p^{k+1}| &\stackrel{(3.157)}{=} \left| \sum_{j=1}^k \hat{\kappa}_j^k \hat{\rho}_j \hat{G}_j \right| \leq \sum_{j=1}^k |\hat{\kappa}_j^k \hat{\rho}_j \hat{G}_j| \\ &\stackrel{(3.156), (3.160)}{\leq} \sum_{j=1}^k \hat{\kappa}_j^k |\hat{\rho}_j \hat{G}_j| \stackrel{(3.321)}{\leq} \sum_{j=1}^k \hat{\kappa}_j^k \hat{C}_G \\ &\stackrel{(3.156), (3.160), \text{compare (3.238)}}{\leq} \hat{C}_G \end{aligned}$$

(Note that (3.238) was shown in the proof of Proposition 3.44, where we only used (3.91), and hence (3.238) holds here, too). $\blacksquare$

(k) *Claim:* We have

$$f_p^{k+1} - \tilde{f}_p^k \xrightarrow{K} 0 \quad (3.337)$$

$$F_p^{k+1} - \tilde{F}_p^k \xrightarrow{K} 0 \quad (3.338)$$

$$\Delta_k \xrightarrow{K} 0, \quad (3.339)$$

where

$$\Delta_k := H_{k+1}((g_p^{k+1} + \bar{\kappa}^{k+1}\hat{g}_p^{k+1}) - (\tilde{g}_p^k + \bar{\kappa}^{k+1}\tilde{g}_p^k)). \quad (3.340)$$

Proof of claim:

- We have

$$\begin{aligned} |f_p^{k+1} - \tilde{f}_p^k| &\stackrel{(3.100)}{=} |\tilde{g}_p^{kT}(x_{k+1} - x_k) + \frac{1}{2}(x_{k+1} - x_k)^T G_p^{k+1}(x_{k+1} - x_k)| \\ &\stackrel{(2.1)}{\leq} |\tilde{g}_p^k| \cdot |x_{k+1} - x_k| + \frac{1}{2}|x_{k+1} - x_k| \cdot |G_p^{k+1}| \cdot |x_{k+1} - x_k| \\ &\stackrel{(3.335)}{\leq} |\tilde{g}_p^k| \cdot |x_{k+1} - x_k| + C_G \frac{1}{2}|x_{k+1} - x_k|^2 \xrightarrow{K} 0 \end{aligned}$$

due to (3.326), (2.4) and (3.330).

- We have

$$\begin{aligned} |F_p^{k+1} - \tilde{F}_p^k| &\stackrel{(3.102)}{=} |\tilde{g}_p^{kT}(x_{k+1} - x_k) + \frac{1}{2}(x_{k+1} - x_k)^T \hat{G}_p^{k+1}(x_{k+1} - x_k)| \\ &\stackrel{(2.1)}{\leq} |\tilde{g}_p^k| \cdot |x_{k+1} - x_k| + \frac{1}{2}|x_{k+1} - x_k| \cdot |\hat{G}_p^{k+1}| \cdot |x_{k+1} - x_k| \\ &\stackrel{(3.336)}{\leq} |\tilde{g}_p^k| \cdot |x_{k+1} - x_k| + \hat{C}_G \frac{1}{2}|x_{k+1} - x_k|^2 \xrightarrow{K} 0 \end{aligned}$$

due to (3.326), (2.4) and (3.330).

- We have

$$\begin{aligned} |\Delta_k| &\stackrel{(3.340)}{=} |H_{k+1}((g_p^{k+1} + \bar{\kappa}^{k+1}\hat{g}_p^{k+1}) - (\tilde{g}_p^k + \bar{\kappa}^{k+1}\tilde{g}_p^k))| \\ &= |H_{k+1}(g_p^{k+1} - \tilde{g}_p^k + \bar{\kappa}^{k+1}(\hat{g}_p^{k+1} - \tilde{g}_p^k))| \\ &\stackrel{(3.105), (3.108)}{\stackrel{\downarrow}{=}} |H_{k+1}(G_p^{k+1}(x_{k+1} - x_k) + \underbrace{\bar{\kappa}^{k+1}\hat{G}_p^{k+1}}_{\stackrel{(3.72)}{\geq 0}}(x_{k+1} - x_k))| \\ &\leq |H_{k+1}|(|G_p^{k+1}| \cdot |x_{k+1} - x_k| + \bar{\kappa}^{k+1}|\hat{G}_p^{k+1}| \cdot |x_{k+1} - x_k|) \\ &= |H_{k+1}|(|G_p^{k+1}| + \bar{\kappa}^{k+1}|\hat{G}_p^{k+1}|)|x_{k+1} - x_k| \\ &\stackrel{(3.335), (3.336)}{\stackrel{\downarrow}{\leq}} \underbrace{|H_{k+1}|}_{\text{bounded due to (3.322)}}(C_G + \underbrace{\bar{\kappa}^{k+1}\hat{C}_G}_{\text{bounded due to assumption}})|x_{k+1} - x_k| \xrightarrow{K} 0 \end{aligned}$$

due to (3.330). $\blacksquare$

(1) *Claim:* There exist constants $c_L, \hat{c}_L > 0$ with

$$|(s_p^{k+1})^{\omega_1} - (\tilde{s}_p^k)^{\omega_1}| \leq c_L |s_p^{k+1} - \tilde{s}_p^k| \quad (3.341)$$

$$\bar{\kappa}^{k+1} |(\hat{s}_p^{k+1})^{\omega_2} - (\tilde{\hat{s}}_p^k)^{\omega_2}| \leq \hat{c}_L |\hat{s}_p^{k+1} - \tilde{\hat{s}}_p^k| \quad (3.342)$$

for all $k \geq 1$.

Proof of claim:

- We prove (3.341): Because of

$$\begin{aligned} & \tilde{\alpha}_p^k \stackrel{(3.79)}{=} \max(|\tilde{f}_p^k - f(x_k)|, \gamma_1(\tilde{s}_p^k)^{\omega_1}) \\ \implies & \gamma_1(\tilde{s}_p^k)^{\omega_1} \leq \tilde{\alpha}_p^k \\ \iff & \tilde{s}_p^k \leq \left(\frac{\tilde{\alpha}_p^k}{\gamma_1}\right)^{\frac{1}{\omega_1}} \end{aligned}$$

and the boundedness of $\{\tilde{\alpha}_p^k\}$ due to (3.326), $\tilde{s}_p^k \stackrel{(3.78)}{\geq} 0$ is bounded. Since the function $\xi \mapsto \xi^{\omega_1}$ with $\omega_1 \geq 1$ is Lipschitz continuous on every bounded subset of $\mathbb{R}_+$, there exists $c_L > 0$ with (cf. Definition 2.21)

$$|(s_p^{k+1})^{\omega_1} - (\tilde{s}_p^k)^{\omega_1}| \leq c_L |s_p^{k+1} - \tilde{s}_p^k|.$$

- We prove (3.342): In the case $\bar{\kappa}^{k+1} = 0$, we have

$$\begin{aligned} & \stackrel{(3.72), (3.73), (3.74), (3.83)}{\tilde{s}_p^k \downarrow} \tilde{s}_p^k \stackrel{\cdot \bar{\kappa}^{k+1}=0}{\implies} \bar{\kappa}^{k+1} \tilde{s}_p^k = 0. \end{aligned}$$

Now consider the case $\bar{\kappa}^{k+1} > 0$. Because of

$$\begin{aligned} & \tilde{A}_p^k \stackrel{(3.84)}{=} \max(|\tilde{F}_p^k - F(x_k)|, \gamma_2(\tilde{s}_p^k)^{\omega_2}) \\ \implies & \gamma_2(\tilde{s}_p^k)^{\omega_2} \leq \tilde{A}_p^k \\ \iff & \tilde{s}_p^k \leq \left(\frac{\tilde{A}_p^k}{\gamma_2}\right)^{\frac{1}{\omega_2}} \\ \stackrel{\cdot \bar{\kappa}^{k+1} > 0}{\iff} & \bar{\kappa}^{k+1} \tilde{s}_p^k \leq \bar{\kappa}^{k+1} \left(\frac{\tilde{A}_p^k}{\gamma_2}\right)^{\frac{1}{\omega_2}} \\ & = (\bar{\kappa}^{k+1})^{\omega_2} \left(\frac{\kappa^{k+1} \tilde{A}_p^k}{\gamma_2}\right)^{\frac{1}{\omega_2}} \end{aligned}$$

and the boundedness of $\{\bar{\kappa}^{k+1}\}$ due to assumption and the boundedness of $\{\bar{\kappa}^{k+1} \tilde{A}_p^k\}$ due to (3.326), $\bar{\kappa}^{k+1} \tilde{s}_p^k \stackrel{(3.83)}{\geq} 0$ is bounded.

Therefore, $\{\bar{\kappa}^{k+1} \tilde{s}_p^k\}$ is bounded for all $\bar{\kappa}^{k+1} \geq 0$. Since the function $\xi \mapsto \xi^{\omega_2}$ with $\omega_2 \geq 1$ is Lipschitz continuous on every bounded subset of $\mathbb{R}_+$, there exists $\bar{c}_L > 0$ with (cf. Definition 2.21)

$$\begin{aligned} |(\bar{\kappa}^{k+1} \hat{s}_p^{k+1})^{\omega_2} - (\bar{\kappa}^{k+1} \tilde{\hat{s}}_p^k)^{\omega_2}| & \leq \bar{c}_L |(\bar{\kappa}^{k+1} \hat{s}_p^{k+1}) - (\bar{\kappa}^{k+1} \tilde{\hat{s}}_p^k)| \\ & = \bar{c}_L \bar{\kappa}^{k+1} |\hat{s}_p^{k+1} - \tilde{\hat{s}}_p^k| \end{aligned} \quad (3.343)$$

and hence we obtain

$$\begin{aligned}
\bar{\kappa}^{k+1} |(\hat{s}_p^{k+1})^{\omega_2} - (\tilde{s}_p^k)^{\omega_2}| &= (\bar{\kappa}^{k+1})^{\frac{1}{\omega_2}} (\bar{\kappa}^{k+1})^{\omega_2} |(\hat{s}_p^{k+1})^{\omega_2} - (\tilde{s}_p^k)^{\omega_2}| \\
&= (\bar{\kappa}^{k+1})^{\frac{1}{\omega_2}} |(\bar{\kappa}^{k+1} \hat{s}_p^{k+1})^{\omega_2} - (\bar{\kappa}^{k+1} \tilde{s}_p^k)^{\omega_2}| \\
&\stackrel{(3.343)}{\leq} (\bar{\kappa}^{k+1})^{\frac{1}{\omega_2}} \bar{c}_L \bar{\kappa}^{k+1} |\hat{s}_p^{k+1} - \tilde{s}_p^k| \\
&= (\bar{\kappa}^{k+1})^{1+\frac{1}{\omega_2}} \bar{c}_L |\hat{s}_p^{k+1} - \tilde{s}_p^k| \\
&\leq \hat{c}_L |\hat{s}_p^{k+1} - \tilde{s}_p^k|
\end{aligned}$$

with

$$\hat{c}_L := \bar{c}_L \sup_{k \geq 1} (\bar{\kappa}^{k+1})^{1+\frac{1}{\omega_2}} < \infty$$

due to the assumption of the boundedness of $\{\bar{\kappa}^{k+1}\}$ and $\omega_2 \geq 1$. ■

(m) *Claim:* We have

$$|(\alpha_p^{k+1} - \tilde{\alpha}_p^k) + \bar{\kappa}^{k+1}(A_p^{k+1} - \tilde{A}_p^k)| \xrightarrow{K} 0. \quad (3.344)$$

Proof of claim: For the proof we remind of the formula

$$|\max(a, b) - \max(c, d)| \leq |a - c| + |b - d| \quad (3.345)$$

for all $a, b, c, d \in \mathbb{R}$. We have

$$\begin{aligned}
|\alpha_p^{k+1} - \tilde{\alpha}_p^k| &\stackrel{(3.65), (3.79)}{\stackrel{\downarrow}{\leq}} |\max(|f_p^{k+1} - f(x_{k+1})|, \gamma_1(s_p^{k+1})^{\omega_1}) - \max(|\tilde{f}_p^k - f(x_k)|, \gamma_1(\tilde{s}_p^k)^{\omega_1})| \\
&\stackrel{(3.345)}{\leq} ||f_p^{k+1} - f(x_{k+1})| - |\tilde{f}_p^k - f(x_k)|| + |\gamma_1(s_p^{k+1})^{\omega_1} - \gamma_1(\tilde{s}_p^k)^{\omega_1}| \\
&\leq |f_p^{k+1} - \tilde{f}_p^k| + |f(x_k) - f(x_{k+1})| + \gamma_1 |(s_p^{k+1})^{\omega_1} - (\tilde{s}_p^k)^{\omega_1}| \\
&\stackrel{(3.341)}{\leq} |f_p^{k+1} - \tilde{f}_p^k| + |f(x_k) - f(x_{k+1})| + \gamma c_L |s_p^{k+1} - \tilde{s}_p^k| \xrightarrow{K} 0
\end{aligned} \quad (3.346)$$

due to (3.337), (3.333) and (3.331). Furthermore, we obtain

$$\begin{aligned}
|A_p^{k+1} - \tilde{A}_p^k| &\stackrel{(3.67), (3.84)}{\stackrel{\downarrow}{\leq}} |\max(|F_p^{k+1} - F(x_{k+1})|, \gamma_2(\hat{s}_p^{k+1})^{\omega_2}) - \max(|\tilde{F}_p^k - F(x_k)|, \gamma_2(\tilde{s}_p^k)^{\omega_2})| \\
&\stackrel{(3.345)}{\leq} ||F_p^{k+1} - F(x_{k+1})| - |\tilde{F}_p^k - F(x_k)|| + |\gamma_2(\hat{s}_p^{k+1})^{\omega_2} - \gamma_2(\tilde{s}_p^k)^{\omega_2}| \\
&\leq |F_p^{k+1} - \tilde{F}_p^k| + |F(x_k) - F(x_{k+1})| + \gamma_2 |(\hat{s}_p^{k+1})^{\omega_2} - (\tilde{s}_p^k)^{\omega_2}|.
\end{aligned} \quad (3.347)$$

Multiplying (3.347) with $\bar{\kappa}^{k+1} \stackrel{(3.72)}{\geq} 0$ yields

$$\begin{aligned}
\bar{\kappa}^{k+1} |A_p^{k+1} - \tilde{A}_p^k| &\stackrel{(3.347)}{\leq} \bar{\kappa}^{k+1} (|F_p^{k+1} - \tilde{F}_p^k| + |F(x_k) - F(x_{k+1})| + \gamma_2 |(\hat{s}_p^{k+1})^{\omega_2} - (\tilde{s}_p^k)^{\omega_2}|) \\
&= \bar{\kappa}^{k+1} (|F_p^{k+1} - \tilde{F}_p^k| + |F(x_k) - F(x_{k+1})|) + \gamma_2 \bar{\kappa}^{k+1} |(\hat{s}_p^{k+1})^{\omega_2} - (\tilde{s}_p^k)^{\omega_2}| \\
&\stackrel{(3.342)}{\leq} \underbrace{\bar{\kappa}^{k+1}}_{\text{bounded due to assumption}} (|F_p^{k+1} - \tilde{F}_p^k| + |F(x_k) - F(x_{k+1})|) + \gamma \hat{c}_L |\hat{s}_p^{k+1} - \tilde{s}_p^k| \xrightarrow{K} 0
\end{aligned} \quad (3.348)$$

due to (3.338), (3.334) and (3.332). Therefore, we have

$$|(\alpha_p^{k+1} - \tilde{\alpha}_p^k) + \underbrace{\bar{\kappa}^{k+1}}_{\substack{(3.72) \\ \geq 0}}(A_p^{k+1} - \tilde{A}_p^k)| \leq \underbrace{|\alpha_p^{k+1} - \tilde{\alpha}_p^k|}_{\xrightarrow{K} 0 \text{ due to (3.346)}} + \underbrace{\bar{\kappa}^{k+1}|A_p^{k+1} - \tilde{A}_p^k|}_{\xrightarrow{K} 0 \text{ due to (3.348)}} \xrightarrow{K} 0$$

due to (2.4). ■

(n) *Claim:* We define

$$c := \sup_{k \geq 1} \left(|H_k g_{k+1}^{k+1}|, |H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k)|, \sqrt{\tau_k} \right) \quad (3.349)$$

$$\delta := \frac{\sigma(\bar{x})}{2} \quad (3.350)$$

$$\bar{c} := \delta^{\frac{1-m_R}{4c}} \quad (3.351)$$

and

$$\tilde{c} := \sup_{k \geq 1} (|g_{k+1}^{k+1}| + |\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k|) \quad (3.352)$$

$$C_6 := \tilde{c} C_5 \max(2c, 1, \frac{1}{2} \tilde{c} C_5). \quad (3.353)$$

Then there exists $\bar{k} \geq 0$ with

$$4c|\Delta_k| + \frac{|\Delta_k|^2}{2} + |(\alpha_p^{k+1} - \tilde{\alpha}_p^k) + \bar{\kappa}^{k+1}(A_p^{k+1} - \tilde{A}_p^k)| < \frac{1}{2} \bar{c}^2 \quad (3.354)$$

$$C_6(|\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}| + |\Delta_k| \cdot |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}| + |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}|^2) < \frac{1}{2} \bar{c}^2 \quad (3.355)$$

for all $k > \bar{k}$.

Proof of claim:

i. c is finite because of

$$c \stackrel{(3.349)}{=} \sup_{k \geq 1} \left(|H_k g_{k+1}^{k+1}|, |H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k)|, \sqrt{\tau_k} \right) < \infty \quad (3.356)$$

due to (3.322), (3.323) and (3.326).

ii. We have

$$c > 0, \quad (3.357)$$

since: (By contradiction) Suppose $c = 0$. Then we obtain

$$0 = c \stackrel{(3.349)}{=} \sup_{k \geq 1} \left(|H_k g_{k+1}^{k+1}|, |H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k)|, \sqrt{\tau_k} \right)$$

which is equivalent to having for all $k \geq 1$

$$\begin{aligned} |H_k g_{k+1}^{k+1}| &= 0 \quad \wedge \quad |H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k)| = 0 \quad \wedge \quad \sqrt{\tau_k} = 0 \\ \implies \frac{1}{2} |H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k)|^2 &= 0 \quad \wedge \quad \tilde{\alpha}_p^k + \bar{\kappa}^{k+1} \tilde{A}_p^k + \bar{\kappa}^{k+1} (-F(x_k)) \stackrel{(3.327)}{=} 0 \\ \iff w_k &\stackrel{(3.86)}{=} \frac{1}{2} |H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k)|^2 + \tilde{\alpha}_p^k + \bar{\kappa}^{k+1} \tilde{A}_p^k + \bar{\kappa}^{k+1} (-F(x_k)) \\ &= 0, \end{aligned}$$

which is a contradiction to assumption (3.189).

iii. We have (cf. the initialization of Algorithm 3.20)

$$\begin{aligned} m_L \in (0, \tfrac{1}{2}) \quad \wedge \quad m_R \in (m_L, 1) &\implies 0 < m_R < 1 \\ &\implies 1 - m_R > 0. \end{aligned} \quad (3.358)$$

iv. Due to (3.323), (3.326) and the assumption of the boundedness of $\{\bar{\kappa}^{k+1}\}$, $\tilde{c} \stackrel{(3.352)}{\geq} 0$ is bounded. Therefore, (3.353), (3.295) and (3.356) imply

$$0 \leq C_6 < \infty. \quad (3.359)$$

v. We have

$$\left(\frac{4c|\Delta_k| + \frac{|\Delta_k|^2}{2} + |(\alpha_p^{k+1} - \tilde{\alpha}_p^k) + \bar{\kappa}^{k+1}(A_p^{k+1} - \tilde{A}_p^k)|}{C_6(|\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}| + |\Delta_k| \cdot |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}| + |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}|^2)} \right) \xrightarrow{K} 0$$

due to (3.356), (3.339), (3.344), (3.359) and (3.317). Hence, there exists $\bar{k} \geq 0$ with

$$\begin{aligned} 4c|\Delta_k| + \frac{|\Delta_k|^2}{2} + |(\alpha_p^{k+1} - \tilde{\alpha}_p^k) + \bar{\kappa}^{k+1}(A_p^{k+1} - \tilde{A}_p^k)| &< \tfrac{1}{2}\bar{c}^2 \\ C_6(|\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}| + |\Delta_k| \cdot |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}| + |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}|^2) &< \tfrac{1}{2}\bar{c}^2 \end{aligned}$$

for all $k > \bar{k}$, where

$$\begin{aligned} \bar{c}^2 &\stackrel{(3.351)}{=} \left(\delta \frac{1-m_R}{4c} \right)^2 \stackrel{(3.350)}{=} \left(\frac{\sigma(\bar{x})}{2} \cdot \frac{1-m_R}{4c} \right)^2 \\ &\stackrel{\sqrt{\cdot}}{\iff} \bar{c} = \left| \frac{\sigma(\bar{x})}{2} \cdot \frac{1-m_R}{4c} \right| \stackrel{(3.320), (3.358), (3.357)}{=} \frac{\sigma(\bar{x})}{2} \cdot \frac{1-m_R}{4c}. \end{aligned} \quad (3.360)$$

In particular, this implies $0 < \bar{c} < \infty$ by (3.360), (3.320), (3.358), (3.357), and (3.329). ■

(o) *Claim:* We have for all $\nu \in [0, 1]$ and for all $k \geq 1$

$$\tfrac{1}{2}|\nu H_{k+1}g_{k+1}^{k+1} + (1-\nu)H_{k+1}(g_p^{k+1} + \bar{\kappa}^{k+1}\tilde{g}_p^{k+1})|^2 \leq \tfrac{1}{2}|\nu q_k + (1-\nu)(p_k + \Delta_k)|^2 + e_k, \quad (3.361)$$

where

$$q_k := H_k g_{k+1}^{k+1} \quad (3.362)$$

$$p_k := H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1}\tilde{g}_p^k) \quad (3.363)$$

$$e_k := (2c + |\Delta_k|)\tilde{c}|E_k| + \tfrac{1}{2}\tilde{c}^2|E_k|^2 \quad (3.364)$$

$$E_k := H_{k+1} - H_k. \quad (3.365)$$

Proof of claim: Setting

$$z_k := E_k g_{k+1}^{k+1} \quad (3.366)$$

we obtain

$$H_{k+1}g_{k+1}^{k+1} \stackrel{(3.365)}{=} (H_k + E_k)g_{k+1}^{k+1} \stackrel{(3.362), (3.366)}{\downarrow} q_k + z_k. \quad (3.367)$$

Setting

$$\hat{z}_k := E_k(\tilde{g}_p^k + \bar{\kappa}^{k+1}\tilde{g}_p^k) \quad (3.368)$$

we obtain

$$\begin{aligned}
 H_{k+1}(g_p^{k+1} + \bar{\kappa}^{k+1}\hat{g}_p^{k+1}) &\stackrel{(3.340)}{=} H_{k+1}(\tilde{g}_p^k + \bar{\kappa}^{k+1}\tilde{\hat{g}}_p^k) + \Delta_k \\
 &\stackrel{(3.365)}{=} (H_k + E_k)(\tilde{g}_p^k + \bar{\kappa}^{k+1}\tilde{\hat{g}}_p^k) + \Delta_k \\
 &\stackrel{(3.363), (3.368)}{\downarrow} p_k + \Delta_k + \hat{z}_k .
 \end{aligned} \tag{3.369}$$

Furthermore, we estimate for all $\nu \in [0, 1]$

$$\begin{aligned}
 &(\nu q_k + (1 - \nu)(p_k + \Delta_k))^T (\nu z_k + (1 - \nu)\hat{z}_k) \\
 &\stackrel{(2.1)}{\leq} |\nu q_k + (1 - \nu)(p_k + \Delta_k)| \cdot |\nu z_k + (1 - \nu)\hat{z}_k| \\
 &\leq \left(\underbrace{\nu}_{\in[0,1]} |q_k| + \underbrace{(1 - \nu)}_{\in[0,1]} |p_k + \Delta_k| \right) |\nu z_k + (1 - \nu)\hat{z}_k| \\
 &\leq (|q_k| + |p_k| + |\Delta_k|) |\nu z_k + (1 - \nu)\hat{z}_k| \\
 &\stackrel{(3.362), (3.363), (3.349)}{\downarrow} \leq (2c + |\Delta_k|) |\nu z_k + (1 - \nu)\hat{z}_k|
 \end{aligned} \tag{3.370}$$

as well as

$$\begin{aligned}
 |\nu z_k + (1 - \nu)\hat{z}_k| &\leq \underbrace{\nu}_{\nu \in [0,1]} |z_k| + \underbrace{(1 - \nu)}_{\nu \in [0,1]} |\hat{z}_k| \\
 &\leq |z_k| + |\hat{z}_k| \\
 &\stackrel{(3.366), (3.368)}{\downarrow} \leq |E_k g_{k+1}^{k+1}| + |E_k(\tilde{g}_p^k + \bar{\kappa}^{k+1}\tilde{\hat{g}}_p^k)| \\
 &\leq |E_k|(|g_{k+1}^{k+1}| + |\tilde{g}_p^k + \bar{\kappa}^{k+1}\tilde{\hat{g}}_p^k|) \\
 &\stackrel{(3.352)}{\leq} \tilde{c}|E_k| .
 \end{aligned} \tag{3.371}$$

Hence we obtain for all $\nu \in [0, 1]$

$$\begin{aligned}
 &\frac{1}{2} |\nu H_{k+1} g_{k+1}^{k+1} + (1 - \nu) H_{k+1}(g_p^{k+1} + \bar{\kappa}^{k+1}\hat{g}_p^{k+1})|^2 \\
 &\stackrel{(3.367), (3.369)}{\downarrow} \leq \frac{1}{2} |\nu(q_k + z_k) + (1 - \nu)(p_k + \Delta_k + \hat{z}_k)|^2 \\
 &= \frac{1}{2} |(\nu q_k + (1 - \nu)(p_k + \Delta_k)) + (\nu z_k + (1 - \nu)\hat{z}_k)|^2 \\
 &= \frac{1}{2} |\nu q_k + (1 - \nu)(p_k + \Delta_k)|^2 \\
 &\quad + (\nu q_k + (1 - \nu)(p_k + \Delta_k))^T (\nu z_k + (1 - \nu)\hat{z}_k) + \frac{1}{2} |\nu z_k + (1 - \nu)\hat{z}_k|^2 \\
 &\stackrel{(3.370)}{\leq} \frac{1}{2} |\nu q_k + (1 - \nu)(p_k + \Delta_k)|^2 + (2c + |\Delta_k|) |\nu z_k + (1 - \nu)\hat{z}_k| \\
 &\quad + \frac{1}{2} |\nu z_k + (1 - \nu)\hat{z}_k|^2 \\
 &\stackrel{(3.371)}{\leq} \frac{1}{2} |\nu q_k + (1 - \nu)(p_k + \Delta_k)|^2 + (2c + |\Delta_k|) \tilde{c}|E_k| + \frac{1}{2} \tilde{c}^2 |E_k|^2 \\
 &\stackrel{(3.364)}{=} \frac{1}{2} |\nu q_k + (1 - \nu)(p_k + \Delta_k)|^2 + e_k .
 \end{aligned} \tag{3.372}$$

■

(p) *Claim:* Let

$$\hat{r} := \frac{3}{2} \cdot \frac{c^2}{\bar{c}^2} + i_m \tag{3.372}$$

$$r := i_l + \hat{r} , \tag{3.373}$$

let $\tilde{k}$ be the number from Proposition 3.46 which belongs to r , let $k_0 \in K$ with $k_0 > \max(\tilde{k}, \bar{k})$ and set

$$k := k_0 + i_l + i \quad (3.374)$$

with any $i \in [i_m, \hat{r}] \cap \{0, 1, \dots\}$. Then we have after step 6 (line search) of Algorithm 3.20

$$w_k \geq \delta \quad (3.375)$$

$$t_L^k < t_0^k \quad (3.376)$$

$$i_n > i_l + i_m \quad (3.377)$$

(i.e. from now on we fix the index k which was constructed above).

Proof of claim:

i. We have

$$\begin{aligned} r &\stackrel{(3.373)}{=} i_l + \hat{r} \\ &\stackrel{(3.372)}{=} i_l + \underbrace{\left(\frac{3}{2}\left(\frac{c}{\bar{c}}\right)^2\right)}_{\geq 0} + i_m \\ &\geq i_l + i_m \geq i_l \geq 0. \end{aligned} \quad (3.378)$$

The inequalities (3.378) imply that $[i_l, r]$ is a well-defined interval and since $i_l \geq 0$ is a natural number (cf. Algorithm 3.20), we have

$$\begin{aligned} &[i_l, r] \cap \{0, 1, \dots\} \neq \emptyset \\ \implies &\exists i \in [i_l, r] \cap \{0, 1, \dots\} \subseteq [0, r]. \end{aligned} \quad (3.379)$$

Furthermore, (3.378) implies that $[i_m, \hat{r}]$ is a well-defined interval and since $i_m \geq 0$ is a natural number (cf. Algorithm 3.20), we have

$$\begin{aligned} &[i_m, \hat{r}] \cap \{0, 1, \dots\} \neq \emptyset \\ \implies &\exists i \in [i_m, \hat{r}] \cap \{0, 1, \dots\} \subseteq [0, \hat{r}]. \end{aligned} \quad (3.380)$$

ii. We have that

$$\text{the assumptions of Proposition 3.46 are satisfied,} \quad (3.381)$$

since:

- (3.189) holds due to assumption.
- The sequence (of symmetric, positive definite matrices $\{H_k\}$ is bounded due to (3.322).
- Due to (3.316), we have $x_k \xrightarrow{K} \bar{x}$.
- Due to (3.320) and (3.329), we have $\sigma(\bar{x}) > 0$.
- Due to (3.373) r is a fixed number and due to (3.378) we have $r \geq 0$.
- The choice

$$\varepsilon_0 := t_0^{\inf} \stackrel{(3.318)}{>} 0 \quad (3.382)$$

yields a fixed positive number ε_0 .

- iii. Because of (3.381) we can apply Proposition 3.46: For r defined in (3.373) there exists $\tilde{k} \geq 0$ with

$$\begin{aligned} w_{k+i} &\stackrel{(3.268)}{\geq} \frac{\sigma(\bar{x})}{2} \stackrel{(3.350)}{=} \delta \\ t_L^{k+i} &\stackrel{(3.269)}{<} \varepsilon_0 \stackrel{(3.382)}{=} t_0^{\inf} \end{aligned} \quad (3.383)$$

for all $k > \tilde{k}$, $k \in K$ and for all $0 \leq i \leq r$.

Since K is an infinite set due to (3.316) ($K \subset \{1, 2, \dots\}$), we can choose $k_0 \in K$ with

$$k_0 > \max(\tilde{k}, \bar{k}) \quad (3.384)$$

($\bar{k}$ was introduced in (3.354)). Because of

$$k_0 \stackrel{(3.384)}{>} \max(\tilde{k}, \bar{k}) \geq \tilde{k} ,$$

(3.383) holds in particular for all $k \geq k_0$ and hence for $k = k_0$

$$\begin{aligned} w_{k_0+i} &\stackrel{(3.383)}{\geq} \delta \\ t_L^{k_0+i} &\stackrel{(3.383)}{<} t_0^{\inf} \end{aligned} \quad (3.385)$$

for all $0 \leq i \leq r$. Because of

$$\begin{aligned} t_0^{\inf} &\stackrel{(3.318)}{=} \inf_{k=0,1,\dots} t_0^k \\ \iff t_0^{\inf} &\leq t_0^k \quad \text{for all } k = 0, 1, \dots \\ \implies t_L^{k_0+i} &\stackrel{(3.385)}{\leq} t_0^{\inf} \\ &\leq t_0^{k_0+i} \quad \text{for all } 0 \leq i \leq r , \end{aligned} \quad (3.386)$$

we obtain for (3.385)

$$\begin{aligned} w_{k_0+i} &\stackrel{(3.385)}{\geq} \delta \\ t_L^{k_0+i} &\stackrel{(3.385), (3.386)}{<} t_0^{k_0+i} \end{aligned} \quad (3.387)$$

for all $0 \leq i \leq r$. Because of (3.379), (3.387) holds in particular for all $i \in [i_l, r] \stackrel{(3.373)}{=} i_l + [0, \hat{r}]$ which yields

$$\begin{aligned} w_{k_0+i_l+i} &\stackrel{(3.387)}{\geq} \delta \\ t_L^{k_0+i_l+i} &\stackrel{(3.387)}{<} t_0^{k_0+i_l+i} \end{aligned} \quad (3.388)$$

with $i \in [0, \hat{r}]$. In particular, because of (3.380), (3.388) holds for all $i \in [i_m, \hat{r}] \cap$

$\{0, 1, \dots\}$ and now setting

$$k' := k_0 + i_l + i$$

yields the desired result.

iv. Because of (3.378), we have

$$t_L^{k_0+i} \stackrel{(3.387)}{<} t_0^{k_0+i}$$

in particular for all $0 \leq i \leq i_l + i_m$. Consequently, the case (3.92) always occurs for the $i_l + i_m + 1$ subsequent iterations $k_0 + 0, \dots, k_0 + i_l, \dots, k_0 + i_l + i_m$ (Remember: $i_n \geq 0$ denotes the number of subsequent short and null steps according to the initialization of Algorithm 3.20)

$i_n = i_n^0 \geq 0$	value of i_n at the beginning of iteration $k_0 + 0$
$i_n = i_n^0 + 1$	value of i_n at the end of iteration $k_0 + 0$
$i_n = i_n^0 + 1$	value of i_n at the beginning of iteration $k_0 + 1$
$i_n = i_n^0 + 2$	value of i_n at the end of iteration $k_0 + 1$
$i_n = i_n^0 + 2$	value of i_n at the beginning of iteration $k_0 + 2$
$i_n = i_n^0 + 3$	value of i_n at the end of iteration $k_0 + 2$
$\vdots$	$\vdots$
$i_n = i_n^0 + (i_l + i_m - 1)$	value of i_n at the beginning of iteration $k_0 + (i_l + i_m - 1)$
$i_n = i_n^0 + i_l + i_m$	value of i_n at the end of iteration $k_0 + (i_l + i_m - 1)$
$i_n = i_n^0 + i_l + i_m$	value of i_n at the beginning of iteration $k_0 + i_l + i_m$
$i_n = (i_n^0 + i_l + i_m) + 1$	value of i_n at the end of iteration $k_0 + i_l + i_m$

and therefore we have at the end of iteration $k_0 + i_l + i_m$ (even if the value $i_n^0 = 0$ at the beginning of iteration $k_0 + 0$)

$$i_n = \underbrace{i_n^0}_{\geq 0} + i_l + i_m + 1 \geq i_l + i_m + 1 > i_l + i_m . \quad \blacksquare$$

(q) *Claim:* For k defined in (3.374) we have

$$e_k < \frac{1}{2} \bar{c}^2 . \quad (3.389)$$

Proof of claim: Because of

$$i_n \stackrel{(3.377)}{>} i_l + i_m$$

and since, due to (3.92), i_n increases at most by one at each iteration, we have at iteration k at least

$$i_n \geq i_l + i_m ,$$

which implies that at iteration k either the case (3.62) or the case (3.63) occurs.

Because of

$$i_n \stackrel{(3.377)}{>} i_l + i_m$$

the cases (3.60) and (3.63) occur at iteration $k + 1$.

These facts yield

$$\begin{aligned}
E_k &\stackrel{(3.365)}{=} H_{k+1} - H_k \\
&\stackrel{(3.71)}{=} (\bar{W}_p^{k+1} + \sum_{j \in J_{k+1}} \mu_j^{k+1} \bar{G}_j^{k+1} + \mu_p^{k+1} \bar{G}^{k+1})^{-\frac{1}{2}} - (\bar{W}_p^k + \sum_{j \in J_k} \mu_j^k \bar{G}_j^k + \mu_p^k \bar{G}^k)^{-\frac{1}{2}} \\
(3.62) \text{ or } (3.63) \text{ for iteration } k &\stackrel{\downarrow}{=} (\bar{W}_p^{k+1} + \sum_{j \in J_{k+1}} \mu_j^{k+1} \bar{G}_j^{k+1} + \mu_p^{k+1} \bar{G}^{k+1})^{-\frac{1}{2}} - (\bar{W}_p^k + \sum_{j \in J_k} \mu_j^k \bar{G}_j^k + \mu_p^k \bar{G}^k)^{-\frac{1}{2}} \\
(3.60), (3.63) \text{ for iteration } k+1 &\stackrel{\downarrow}{=} (\bar{W}_p^k + \sum_{j \in J_{k+1}} \mu_j^{k+1} \bar{G}_j^k + \mu_p^{k+1} \bar{G}^k)^{-\frac{1}{2}} - (\bar{W}_p^k + \sum_{j \in J_k} \mu_j^k \bar{G}_j^k + \mu_p^k \bar{G}^k)^{-\frac{1}{2}} \\
&= (\bar{W}_p^k + (\sum_{j \in J_{k+1}} \mu_j^{k+1} + \mu_p^{k+1}) \bar{G}^k)^{-\frac{1}{2}} - (\bar{W}_p^k + (\sum_{j \in J_k} \mu_j^k + \mu_p^k) \bar{G}^k)^{-\frac{1}{2}} \\
&\stackrel{(3.72)}{=} (\bar{W}_p^k + \bar{\kappa}^{k+2} \bar{G}^k)^{-\frac{1}{2}} - (\bar{W}_p^k + \bar{\kappa}^{k+1} \bar{G}^k)^{-\frac{1}{2}} \\
&\stackrel{(3.293)}{=} Z_k(\bar{\kappa}^{k+2}) - Z_k(\bar{\kappa}^{k+1}) . \tag{3.390}
\end{aligned}$$

Since $\{\bar{\kappa}^{k+1}\}$ is bounded and $\{(\bar{W}_p^k)^{-\frac{1}{2}}\}$ is bounded as well as uniformly positive definite (by assumption), we can make use of Proposition 3.54 and hence we obtain

$$|E_k| \stackrel{(3.390)}{=} |Z_k(\bar{\kappa}^{k+2}) - Z_k(\bar{\kappa}^{k+1})| \stackrel{(3.294)}{\leq} C_5 |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}| . \tag{3.391}$$

Consequently, we estimate

$$\begin{aligned}
e_k &\stackrel{(3.364)}{=} (2c + |\Delta_k|) \tilde{c} |E_k| + \frac{1}{2} \tilde{c}^2 |E_k|^2 \\
&= 2c \tilde{c} |E_k| + \tilde{c} |\Delta_k| \cdot |E_k| + \frac{1}{2} \tilde{c}^2 |E_k|^2 \\
&\stackrel{(3.391)}{\leq} 2c \tilde{c} (C_5 |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}|) + \tilde{c} |\Delta_k| (C_5 |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}|) + \frac{1}{2} \tilde{c}^2 (C_5 |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}|)^2 \\
&= (2c \tilde{c} C_5) |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}| + (\tilde{c} C_5) |\Delta_k| \cdot |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}| + \frac{1}{2} (\tilde{c} C_5)^2 |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}|^2 \\
&\stackrel{(3.353)}{\leq} C_6 |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}| + C_6 |\Delta_k| \cdot |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}| + C_6 |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}|^2 \\
&= C_6 (|\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}| + |\Delta_k| \cdot |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}| + |\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1}|^2) \\
&\stackrel{(3.355)}{<} \frac{1}{2} \tilde{c}^2 . \quad \blacksquare
\end{aligned}$$

(r) *Claim:* For k defined in (3.374)

a short or null step which changes the model
of the objective function is executed (3.392)

and we have

$$\begin{aligned}
w_{k+1} &\leq |(\alpha_p^{k+1} - \tilde{\alpha}_p^k) + \bar{\kappa}^{k+1} (A_p^{k+1} - \tilde{A}_p^k)| + e_k \\
&\quad + \min_{\nu \in [0,1]} \frac{1}{2} |\nu q_k + (1 - \nu)(p_k + \Delta_k)|^2 + \nu \alpha_{k+1}^{k+1} + (1 - \nu) \tau_k . \tag{3.393}
\end{aligned}$$

Proof of claim: For $\nu \in [0, 1]$ we estimate

$$\begin{aligned}
& \nu \alpha_{k+1}^{k+1} + (1 - \nu) \alpha_p^{k+1} + (1 - \nu) \bar{\kappa}^{k+1} A_p^{k+1} + (1 - \nu) \bar{\kappa}^{k+1} (-F(x_k)) \\
&= \nu \alpha_{k+1}^{k+1} + (1 - \nu) (\tilde{\alpha}_p^k + (\alpha_p^{k+1} - \tilde{\alpha}_p^k)) + (1 - \nu) \bar{\kappa}^{k+1} (\tilde{A}_p^k + (A_p^{k+1} - \tilde{A}_p^k)) \\
&\quad + (1 - \nu) \bar{\kappa}^{k+1} (-F(x_k)) \\
&= \nu \alpha_{k+1}^{k+1} + (1 - \nu) (\tilde{\alpha}_p^k + \bar{\kappa}^{k+1} \tilde{A}_p^k + \bar{\kappa}^{k+1} (-F(x_k))) \\
&\quad + (1 - \nu) ((\alpha_p^{k+1} - \tilde{\alpha}_p^k) + \bar{\kappa}^{k+1} (A_p^{k+1} - \tilde{A}_p^k)) \\
&\stackrel{(3.327)}{=} \nu \alpha_{k+1}^{k+1} + (1 - \nu) \tau_k + (1 - \nu) ((\alpha_p^{k+1} - \tilde{\alpha}_p^k) + \bar{\kappa}^{k+1} (A_p^{k+1} - \tilde{A}_p^k)) \\
&\leq \nu \alpha_{k+1}^{k+1} + (1 - \nu) \tau_k + |1 - \nu| \cdot |(\alpha_p^{k+1} - \tilde{\alpha}_p^k) + \bar{\kappa}^{k+1} (A_p^{k+1} - \tilde{A}_p^k)| \\
&\stackrel{\nu \in [0, 1]}{\leq} \nu \alpha_{k+1}^{k+1} + (1 - \nu) \tau_k + |(\alpha_p^{k+1} - \tilde{\alpha}_p^k) + \bar{\kappa}^{k+1} (A_p^{k+1} - \tilde{A}_p^k)| \tag{3.394}
\end{aligned}$$

and

$$\begin{aligned}
& \frac{1}{2} |H_{k+1}(\nu g_{k+1}^{k+1} + (1 - \nu) g_p^{k+1} + (1 - \nu) \bar{\kappa}^{k+1} \hat{g}_p^{k+1})|^2 \\
&= \frac{1}{2} |\nu H_{k+1} g_{k+1}^{k+1} + (1 - \nu) H_{k+1} (g_p^{k+1} + \bar{\kappa}^{k+1} \hat{g}_p^{k+1})|^2 \\
&\stackrel{(3.361)}{\leq} \frac{1}{2} |\nu q_k + (1 - \nu) (p_k + \Delta_k)|^2 + e_k . \tag{3.395}
\end{aligned}$$

Because of (3.376), step 6 (line search) of Algorithm 3.20 gives us the following information about the iteration k :

- A short or null step is executed. Because of

$$i_n \stackrel{(3.377)}{>} i_l + \underbrace{i_m}_{\geq 0} \geq i_l ,$$

the line search (Algorithm 3.25) yields a change of the model of the objective function.

- i_s is unchanged (since no serious step is executed), i.e. $i_s \leq i_r$ (no bundle reset) still holds (If $i_s > i_r$, then we would have had a serious step at iteration k , as a bundle reset can only occur after a serious step).

Therefore, $(\lambda, \lambda_p, \mu, \mu_p) \in \mathbb{R}^{2(|J_{k+1}|+1)}$ with

$$\begin{aligned}
\lambda_j &:= \begin{cases} \nu & \text{for } j = k + 1 \\ 0 & \text{for } j \in J_{k+1} \setminus \{k + 1\} \end{cases} \\
\lambda_p &:= 1 - \nu \\
\mu_j &:= 0 \quad \text{for all } j \in J_{k+1} \\
\mu_p &:= (1 - \nu) \bar{\kappa}^{k+1} , \tag{3.396}
\end{aligned}$$

where $\nu \in [0, 1]$, is feasible for the $(k + 1)$ st (dual) problem (3.42) (Note: This problem is written as a minimization problem) and, hence, we obtain (Note: $\hat{w}_{k+1}$ in (3.127) is the

optimal function value of (3.42))

$$\begin{aligned}
w_{k+1} &\stackrel{(3.130)}{\leq} \hat{w}_{k+1} \\
&\stackrel{(3.127)}{=} \frac{1}{2} |H_{k+1}(\tilde{g}_p^{k+1} + \bar{\kappa}^{k+1} \tilde{g}_p^{k+1})|^2 + \hat{\alpha}_p^{k+1} + \bar{\kappa}^{k+1} \hat{A}_p^k + \bar{\kappa}^{k+1} (-F(x_{k+1})) \\
&\stackrel{(3.76), (3.125)}{\stackrel{\downarrow}{=}} \frac{1}{2} |H_{k+1} \left(\sum_{j \in J_{k+1}} \lambda_j^{k+1} g_j^{k+1} + \lambda_p^{k+1} g_p^{k+1} + \bar{\kappa}^{k+1} \left(\sum_{j \in J_{k+1}} \kappa_j^{k+1} \hat{g}_j^{k+1} + \kappa_p^{k+1} \hat{g}_p^{k+1} \right) \right)|^2 \\
&\quad + \sum_{j \in J_{k+1}} \lambda_j^{k+1} \alpha_j^{k+1} + \lambda_p^{k+1} \alpha_p^{k+1} + \bar{\kappa}^{k+1} \left(\sum_{j \in J_{k+1}} \kappa_j^{k+1} A_j^{k+1} + \kappa_p^{k+1} A_p^{k+1} \right) \\
&\quad + \bar{\kappa}^{k+1} (-F(x_{k+1})) \\
&\stackrel{(3.73), (3.74), (3.72)}{\stackrel{\downarrow}{=}} \frac{1}{2} |H_{k+1} \left(\sum_{j \in J_{k+1}} \lambda_j^{k+1} g_j^{k+1} + \lambda_p^{k+1} g_p^{k+1} + \sum_{j \in J_{k+1}} \mu_j^{k+1} \hat{g}_j^{k+1} + \mu_p^{k+1} \hat{g}_p^{k+1} \right)|^2 \\
&\quad + \sum_{j \in J_{k+1}} \lambda_j^{k+1} \alpha_j^{k+1} + \lambda_p^{k+1} \alpha_p^{k+1} + \sum_{j \in J_{k+1}} \mu_j^{k+1} A_j^{k+1} + \mu_p^{k+1} A_p^{k+1} \\
&\quad + \left(\sum_{j \in J_{k+1}} \mu_j^{k+1} + \mu_p^{k+1} \right) (-F(x_{k+1})) \\
&\stackrel{(\lambda, \lambda_p, \mu, \mu_p) \text{ from (3.396) feasible}}{\stackrel{\downarrow}{\leq}} \frac{1}{2} |H_{k+1}(\nu g_{k+1}^{k+1} + (1-\nu)g_p^{k+1} + (1-\nu)\bar{\kappa}^{k+1}\hat{g}_p^{k+1})|^2 \\
&\quad + \nu \alpha_{k+1}^{k+1} + (1-\nu)\alpha_p^{k+1} + (1-\nu)\bar{\kappa}^{k+1} A_p^{k+1} + (1-\nu)\bar{\kappa}^{k+1} (-F(x_k)) \\
&\stackrel{(3.394), (3.395)}{\stackrel{\downarrow}{\leq}} \frac{1}{2} |\nu q_k + (1-\nu)(p_k + \Delta_k)|^2 + e_k + \nu \alpha_{k+1}^{k+1} + (1-\nu)\tau_k \\
&\quad + |(\alpha_p^{k+1} - \tilde{\alpha}_p^k) + \bar{\kappa}^{k+1}(A_p^{k+1} - \tilde{A}_p^k)|. \tag{3.397}
\end{aligned}$$

Since $\nu \in [0, 1]$ is arbitrary, (3.393) follows directly from (3.397). $\blacksquare$

(s) *Claim:* For k defined in (3.374) we have

$$w_{k+1} < w_k - \bar{c}^2. \tag{3.398}$$

Proof of claim:

- For

$$p := p_k \tag{3.399}$$

$$g := q_k \tag{3.400}$$

$$\Delta := \Delta_k \tag{3.401}$$

$$v := v_k - \frac{1}{2} d_k^T \left(\sum_{j \in J_k} \mu_j^k \bar{G}_j^k + \mu_p^k \bar{G}^k \right) d_k \tag{3.402}$$

$$w := w_k \tag{3.403}$$

$$\beta := \alpha_{k+1}^{k+1} \tag{3.404}$$

$$m := m_R \tag{3.405}$$

$$\alpha := \tau_k \tag{3.406}$$

the assumptions of Proposition 3.47 are satisfied:

– According to an assumption of Algorithm 3.20 we have

$$m_L \in (0, \frac{1}{2}) \quad \wedge \quad m_R \in (m_L, 1) \implies 0 < m_R < 1 \stackrel{(3.405)}{\iff} m \in (0, 1) . \quad (3.407)$$

– We have

$$\alpha \stackrel{(3.406)}{=} \tau_k \stackrel{(3.327)}{\geq} 0 .$$

– (3.274) is satisfied because of

$$\begin{aligned} w &\stackrel{(3.403)}{=} w_k \\ &\stackrel{(3.86)}{=} \frac{1}{2} |H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k)|^2 + \tilde{\alpha}_p^k + \bar{\kappa}^{k+1} \tilde{A}_p^k + \bar{\kappa}^{k+1} (-F(x_k)) \\ &\stackrel{(3.363), (3.327)}{\downarrow} \stackrel{(3.399), (3.406)}{=} \frac{1}{2} |p_k|^2 + \tau_k \stackrel{(3.399), (3.406)}{=} \frac{1}{2} |p|^2 + \alpha , \end{aligned}$$

– (3.275) is satisfied because of

$$\begin{aligned} v &\stackrel{(3.402)}{=} v_k - \frac{1}{2} d_k^T \left(\sum_{j \in J_k} \mu_j^k \bar{G}_j^k + \mu_p^k \bar{G}^k \right) d_k \\ &\stackrel{(3.85)}{=} -d_k^T \bar{W}_p^k d_k - \frac{1}{2} d_k^T \left(\sum_{j \in J_k} \mu_j^k \bar{G}_j^k + \mu_p^k \bar{G}^k \right) d_k \\ &\quad - \tilde{\alpha}_p^k - \bar{\kappa}^{k+1} \tilde{A}_p^k - \bar{\kappa}^{k+1} (-F(x_k)) - \frac{1}{2} d_k^T \left(\sum_{j \in J_k} \mu_j^k \bar{G}_j^k + \mu_p^k \bar{G}^k \right) d_k \\ &\stackrel{(3.327)}{=} -d_k^T (\bar{W}_p^k + \sum_{j \in J_k} \mu_j^k \bar{G}_j^k + \mu_p^k \bar{G}^k) d_k - \tau_k \\ &\stackrel{(3.109)}{=} -|H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k)|^2 - \tau_k \stackrel{(3.363)}{=} -|p_k|^2 - \tau_k \stackrel{(3.399), (3.406)}{\downarrow} = -(|p|^2 + \alpha) , \end{aligned}$$

– Because of

$$\begin{aligned} p &\stackrel{(3.399)}{=} p_k \stackrel{(3.363)}{=} H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k) \\ &\stackrel{(3.76), (3.81)}{\downarrow} = H_k \left(\sum_{j \in J_k} \lambda_j^k g_j^k + \lambda_p^k g_p^k + \bar{\kappa}^{k+1} \left(\sum_{j \in J_k} \kappa_j^k \hat{g}_j^k + \kappa_p^k \hat{g}_p^k \right) \right) \\ &\stackrel{(3.72), (3.73), (3.74), (3.69)}{\downarrow} = H_k((-1)H_k^{-2}d_k) = -H_k^{-1}d_k \end{aligned} \quad (3.408)$$

we obtain that (3.276) is satisfied

$$\begin{aligned} -\beta - g^T p &\stackrel{(3.404), (3.408), (3.400)}{\downarrow} = -\alpha_{k+1}^{k+1} + (H_k^{-1}d_k)^T q_k \stackrel{(3.362)}{=} -\alpha_{k+1}^{k+1} + (H_k^{-1}d_k)^T (H_k g_{k+1}^{k+1}) \\ &\stackrel{H_k \in \mathbb{R}_{\text{sym}}^{N \times N}}{\downarrow} = -\alpha_{k+1}^{k+1} + d_k^T g_{k+1}^{k+1} \\ &\stackrel{(3.392), (3.140)}{\downarrow} \geq m_R v_k \\ &\stackrel{(3.405), (3.402)}{\downarrow} = mv + \underbrace{\underbrace{m}_{\substack{(3.407) \\ \in (0,1)}} \frac{1}{2} d_k^T \left(\sum_{j \in J_k} \mu_j^k \bar{G}_j^k + \mu_p^k \bar{G}^k \right) d_k}_{\geq 0} \\ &\geq mv . \end{aligned}$$

– (3.277) is satisfied because of

$$\begin{aligned}
 c &\stackrel{(3.349)}{=} \sup_{k \geq 1} \left(|H_k g_{k+1}^{k+1}|, |H_k(\tilde{g}_p^k + \bar{\kappa}^{k+1} \tilde{g}_p^k)|, \sqrt{\tau_k} \right) \\
 &\stackrel{(3.362), (3.363)}{\downarrow} \sup_{k \geq 1} \left(|q_k|, |p_k|, \sqrt{\tau_k} \right) \\
 &\stackrel{(3.400), (3.399), (3.406)}{\downarrow} \sup_{k \geq 1} (|g|, |p|, \sqrt{|\alpha|}) .
 \end{aligned}$$

• We have

$$\delta^2 \frac{(1-m_R)^2}{8c^2} \stackrel{(3.351)}{=} 2\bar{c}^2 . \quad (3.409)$$

• Now, applying (3.279) from Proposition 3.47 yields

$$\begin{aligned}
 w_{k+1} &\stackrel{(3.393)}{\leq} |(\alpha_p^{k+1} - \tilde{\alpha}_p^k) + \bar{\kappa}^{k+1}(A_p^{k+1} - \tilde{A}_p^k)| + e_k \\
 &\quad + \min_{\nu \in [0,1]} \frac{1}{2} |\nu q_k + (1-\nu)(p_k + \Delta_k)|^2 + \nu \alpha_{k+1}^{k+1} + (1-\nu)\tau_k \\
 &\stackrel{(3.400), (3.399), (3.401), (3.404), (3.406)}{\downarrow} |(\alpha_p^{k+1} - \tilde{\alpha}_p^k) + \bar{\kappa}^{k+1}(A_p^{k+1} - \tilde{A}_p^k)| + e_k \\
 &\quad + \min_{\nu \in [0,1]} \frac{1}{2} |\nu g + (1-\nu)(p + \Delta)|^2 + \nu \beta + (1-\nu)\alpha \\
 &\stackrel{(3.278)}{=} |(\alpha_p^{k+1} - \tilde{\alpha}_p^k) + \bar{\kappa}^{k+1}(A_p^{k+1} - \tilde{A}_p^k)| + e_k + \min_{\nu \in [0,1]} Q(\nu) \\
 &\stackrel{(3.279)}{\leq} |(\alpha_p^{k+1} - \tilde{\alpha}_p^k) + \bar{\kappa}^{k+1}(A_p^{k+1} - \tilde{A}_p^k)| + e_k + w - w^2 \frac{(1-m)^2}{8c^2} + 4c|\Delta| + \frac{1}{2}|\Delta|^2 \\
 &\stackrel{(3.403), (3.405), (3.401)}{\downarrow} |(\alpha_p^{k+1} - \tilde{\alpha}_p^k) + \bar{\kappa}^{k+1}(A_p^{k+1} - \tilde{A}_p^k)| + e_k + w_k - w_k^2 \frac{(1-m_R)^2}{8c^2} + 4c|\Delta_k| + \frac{1}{2}|\Delta_k|^2 \\
 &\stackrel{(3.354), (3.389)}{\downarrow} w_k - w_k^2 \frac{(1-m_R)^2}{8c^2} + \frac{1}{2}\bar{c}^2 + \frac{1}{2}\bar{c}^2 \\
 &\stackrel{(3.375)}{\leq} w_k - \delta^2 \frac{(1-m_R)^2}{8c^2} + \bar{c}^2 \\
 &\stackrel{(3.409)}{=} w_k - \bar{c}^2 . \quad \blacksquare
 \end{aligned}$$

(t) *Claim:* For the largest $n \leq \hat{r}$ we have

$$w_{k_0+n+1} < 0 ,$$

which is a contradiction to

$$w_k \stackrel{(3.130)}{\geq} 0 \text{ for all } k \geq 1 .$$

Proof of claim: Let $n \in \mathbb{N}$ be the largest number lower equal $\hat{r} \stackrel{(3.372)}{>} 0$

$$n := \max_{\substack{z \leq \hat{r} \\ z \in \{0,1,\dots\}}} z ,$$

then

$$\begin{aligned}
 n+1 > \hat{r} &\stackrel{-i_m}{\iff} n+1-i_m > \hat{r}-i_m \\
 &\stackrel{(-\bar{c}^2)}{\iff} -\bar{c}^2(n+1-i_m) < -\bar{c}^2(\hat{r}-i_m) \\
 &\stackrel{(3.372)}{=} -\frac{3}{2}\bar{c}^2 . \quad (3.410)
 \end{aligned}$$

Therefore, applying (3.398) $(n - i_m) + 1$ times yields

$$\begin{aligned}
w_{k_0+n+1} &\stackrel{(3.398)}{<} w_{k_0+n} - \bar{c}^2 \stackrel{(3.398)}{<} w_{k_0+n-1} - \bar{c}^2 - \bar{c}^2 = w_{k_0+n-1} - 2\bar{c}^2 \\
&\stackrel{(3.398)}{<} w_{k_0+n-2} - \bar{c}^2 - 2\bar{c}^2 = w_{k_0+n-2} - 3\bar{c}^2 \\
&\vdots \\
&\stackrel{(3.398)}{<} w_{k_0+n-(n-i_m)} - ((n - i_m) + 1)\bar{c}^2 = w_{k_0+i_m} - (n + 1 - i_m)\bar{c}^2 \\
&\stackrel{(3.410)}{<} w_{k_0+i_m} - \frac{3}{2}c^2 \\
&\stackrel{(3.86)}{=} \frac{1}{2}|H_{k_0+i_m}(\tilde{g}_p^{k_0+i_m} + \bar{\kappa}^{k_0+i_m+1}\tilde{g}_p^{k_0+i_m})|^2 \\
&\quad + \tilde{\alpha}_p^{k_0+i_m} + \bar{\kappa}^{k_0+i_m+1}\tilde{A}_p^{k_0+i_m} + \bar{\kappa}^{k_0+i_m+1}(-F(x_{k_0+i_m})) - \frac{3}{2}c^2 \\
&\stackrel{(3.327)}{=} \frac{1}{2}|H_{k_0+i_m}(\tilde{g}_p^{k_0+i_m} + \bar{\kappa}^{k_0+i_m+1}\tilde{g}_p^{k_0+i_m})|^2 + (\sqrt{\tau_{k_0+i_m}})^2 - \frac{3}{2}c^2 \\
&\stackrel{(3.349)}{\leq} \frac{1}{2}c^2 + c^2 - \frac{3}{2}c^2 \\
&= 0 .
\end{aligned}$$

■

☒

2. *Statement:* There exists $\bar{\kappa} \in \mathbb{R}_{\geq 0}$ with

$$\begin{aligned}
0 &\in \partial f(\bar{x}) + \bar{\kappa} \partial F(\bar{x}) \\
\bar{\kappa} F(\bar{x}) &= 0 \\
F(\bar{x}) &\leq 0 .
\end{aligned}$$

Proof of statement: Since $\{(x_k, \bar{\kappa}^{k+1})\}$ is bounded as well as $\{(\bar{W}_p^k)^{-\frac{1}{2}}\}$ is uniformly positive definite (both due to assumption) and because of (3.319), the statement follows from Proposition 3.44. ☒

□

Remark 3.56.

- In examples that do not satisfy the nonsmooth constraint qualification (2.34), $\bar{\kappa}^{k+1}$ became very large in Algorithm 3.20 (Note that Theorem 3.55 has in particular the assumption that $\bar{\kappa}^{k+1}$ is bounded).
- The assumption (3.317)

$$\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1} \xrightarrow{K} 0$$

of Theorem 3.55 was satisfied in all our numerical examples of Chapter 6 in which the termination criterion of Algorithm 3.20 was satisfied.

- If t_0^k is only modified in, e.g., finitely many iterations of Algorithm 3.20, then (3.318) is satisfied (cf. Remark 3.26).
- For an unconstrained optimization problem we obtain in the proof of (3.389)

$$e_k \stackrel{(3.389)}{<} \frac{1}{2}\bar{c}^2$$

that

$$E_k = 0 \implies e_k \stackrel{(3.364)}{=} 0 .$$

Therefore, (3.389) is trivially satisfied in the unconstrained case, since $\bar{c}$ from 3.351 is positive.

Chapter 4

Implementation issues

In the following chapter we present some issues that arise when using a convex QCQP for the computation of the search direction problem like the reduction of its dimension and the existence of a strictly feasible starting point for its SOCP-reformulation. Moreover, we give a numerical justification of the approach of determining the search direction by solving a QCQP by comparing the results of some well-known solvers for our search direction problem.

4.1 Introduction

The most time-consuming part of the bundle-Newton method for nonsmooth unconstrained minimization by LUKŠAN & VLČEK [64] is solving a (convex) QP. This QP is solved by the FORTRAN solver PLQDF1 described in LUKŠAN [60] which exploits the special structure of the QP. Analogously, the most time-consuming part of Algorithm 3.20 is solving the (convex) QCQP (3.68).

For solving the QCQP (3.68), Algorithm 3.20 can use one of the following solvers

1. MOSEK by ANDERSEN [2], ANDERSEN et al. [3] (which is written in C and available as commercial software resp. as a trial version without any limitations of the problem size that may be used by an academic institution for 90 days)
2. IPOPT by WÄCHTER [108], WÄCHTER & BIEGLER [109] (which is written in C++ and freely available),

where the ordering represents the performance of the solvers according to the tests in MITTELMANN [78].

For solving the SOCP-reformulation of the QCQP (3.68) (cf. Subsection 4.3.2), Algorithm 3.20 can use one of the following solvers

1. MOSEK
2. SEDUMI by PÓLIK [86], STURM [102] (which is written in MATLAB and freely available)

3. SDPT3 by TOH et al. [104] (which is written in MATLAB and freely available)
4. socp by LOBO et al. [57] (which is written in C and freely available).

Again, the ordering represents the performance of the solvers according to the tests in MITTELMANN [79], except for socp which was not tested there.

The comparisons in MITTELMANN [78, 79] coincide with our own observations (cf. Subsection 4.4.1).

4.2 Second-order cone programming

Assumption 4.1. In this section we demand

- $L \in \mathbb{N}$, $N \in \mathbb{N}^L$
- $x \in \mathbb{R}^n$ and $z_i \in \mathbb{R}^{N_i-1}$, $w_i \in \mathbb{R}$ for $i = 1, \dots, L$
- $f \in \mathbb{R}^n$ and $A_i \in \mathbb{R}^{(N_i-1) \times n}$, $b_i \in \mathbb{R}^{N_i-1}$, $c_i \in \mathbb{R}^n$, $d_i \in \mathbb{R}$ for $i = 1, \dots, L$.

Definition 4.2. An SOCP (Second-order cone program) is an optimization problem of the form

$$\begin{aligned} \min_x \quad & f^T x \\ \text{s.t.} \quad & |A_i x + b_i|_2 \leq c_i^T x + d_i \quad \text{for } i = 1, \dots, L. \end{aligned} \tag{4.1}$$

Proposition 4.3. The dual problem of the SOCP (4.1) is given by (the SOCP)

$$\begin{aligned} \max_{w, z} \quad & - \sum_{i=1}^L (b_i^T z_i + d_i w_i) \\ \text{s.t.} \quad & \sum_{i=1}^L A_i^T z_i + c_i w_i = f \\ & |z_i|_2 \leq w_i \quad \text{for } i = 1, \dots, L. \end{aligned} \tag{4.2}$$

Proof. BOYD & VANDENBERGHE [13, p. 287, Exercise 5.43]. □

Proposition 4.4. Let

- p^* be the optimal function value of the primal SOCP (4.1) (with $p^* = \infty$, if the primal SOCP is infeasible)
- and d^* be the optimal function value of the dual SOCP (4.2) (with $d^* = -\infty$, if the dual SOCP is infeasible),

then we have:

1. If both the primal and dual SOCP are strictly feasible, then there exists a primal-dual feasible pair with

$$p^* = d^*. \tag{4.3}$$

2. Strong duality (i.e. $p^* = d^*$) holds, if and only if the conic complementarity conditions

$$\begin{aligned} \text{If } |u_i|_2 < t_i &\implies w_i = |z_i|_2 = 0 \\ \text{If } |z_i|_2 < w_i &\implies t_i = |u_i|_2 = 0 \\ \text{If } |u_i|_2 = t_i, |z_i|_2 = w_i &\implies w_i u_i = -t_i z_i \end{aligned} \quad (4.4)$$

are satisfied, where

$$u_i := A_i x + b_i \quad (4.5)$$

$$t_i := c_i^T x + d_i. \quad (4.6)$$

Proof. NESTEROV & NEMIROVSKY [80, §4.2.2]. □

Proposition 4.5. *If x solves the primal SOCP (4.1) and both the primal and dual SOCP are strictly feasible, then we have*

$$\begin{aligned} w_i &\geq 0 \\ |A_i x + b_i|_2 - (c_i^T x + d_i) &\leq 0 \\ w_i (|A_i x + b_i|_2 - (c_i^T x + d_i)) &= 0. \end{aligned} \quad (4.7)$$

Proof.

- Since x solves the primal SOCP (4.1) due to assumption, x is in particular feasible for the primal SOCP

$$|A_i x + b_i|_2 \leq c_i^T x + d_i.$$

- Since both the primal and dual SOCP are strictly feasible due to assumption, strong duality holds due to (4.3), which is equivalent to (4.4).
- For the case

$$\begin{aligned} |A_i x + b_i|_2 &\stackrel{(4.5)}{=} |u_i|_2 \\ &< t_i \stackrel{(4.6)}{=} c_i^T x + d_i \\ &\stackrel{(4.4)}{\implies} w_i = |z_i|_2 = 0 \end{aligned}$$

of (4.4), (4.7) is trivially satisfied.

- Because of

$$\begin{aligned} 0 &\leq |A_i x + b_i|_2 \\ &\leq c_i^T x + d_i \end{aligned} \quad (4.8)$$

we obtain for the case

$$|u_i|_2 = t_i, \quad |z_i|_2 = w_i \quad (4.9)$$

of (4.4)

$$\begin{aligned}
& \underbrace{w_i}_{\substack{(4.9) \\ \geq 0}} u_i \stackrel{(4.4)}{=} -t_i z_i \\
& \stackrel{|\cdot|_2}{\iff} w_i |u_i|_2 \stackrel{(4.9)}{=} |t_i| w_i \\
& \iff 0 \stackrel{(4.5),(4.6)}{\leq} w_i (|A_i x + b_i|_2 - |c_i^T x + d_i|) \\
& \stackrel{(4.8)}{=} w_i (|A_i x + b_i|_2 - (c_i^T x + d_i)) . \quad \square
\end{aligned}$$

4.3 The reduced problem

4.3.1 Reduction of problem size

Idea 4.6. We want to reduce the problem size of the QCQP (3.68). For this purpose we choose

$$\begin{aligned}
\bar{G}^k &:= \text{“positive definite modification of } \hat{G}_p^k \text{”} \\
\bar{G}_j^k &:= \bar{G}^k \quad \text{for all } j \in J_k ,
\end{aligned} \tag{4.10}$$

i.e. we choose all matrices for the constraint approximation equal to a positive definite modification of an aggregated Hessian of the constraint (i.e. similar to the choice of $\bar{W}_p^k$ in the bundle-Newton method for nonsmooth unconstrained minimization by LUKŠAN & VLČEK [64]). For the implementation, we will extract linear constraints

$$Bx \leq b$$

with $B \in \mathbb{R}^{\bar{m} \times n}$ and $b \in \mathbb{R}^{\bar{m}}$ that may occur in the single nonsmooth function $F : \mathbb{R}^n \rightarrow \mathbb{R}$ (via a max-function of the rows $B_i: x - b_i \leq 0$ for all $i = 1, \dots, \bar{m}$) in the nonsmooth constrained optimization problem (2.35) and put them directly into the search direction problem (this is the usual way of handling linear constraints in bundle methods). For easiness of exposition, we drop the p-constraints. These facts altogether yield the (convex) QCQP

$$\begin{aligned}
& \min_{d, \hat{v}} \hat{v} + \frac{1}{2} d^T \bar{W}_p^k d \\
& \text{s.t.} \quad -\alpha_j^k + d^T \hat{g}_j^k \leq \hat{v} \quad \text{for } j \in J_k \\
& \quad F(x_k) - A_j^k + d^T \hat{g}_j^k + \frac{1}{2} d^T \bar{G}^k d \leq 0 \quad \text{for } j \in J_k \\
& \quad B_i: (x_k + d) \leq b_i \quad \text{for } i = 1, \dots, \bar{m} .
\end{aligned} \tag{4.11}$$

Furthermore, we consider the following modification of the QCQP (4.11)

$$\begin{aligned}
& \min_{d, \hat{v}, \hat{u}} \hat{v} + \frac{1}{2} d^T \bar{W}_p^k d \\
& \text{s.t.} \quad -\alpha_j^k + d^T \hat{g}_j^k \leq \hat{v} \quad \text{for } j \in J_k \\
& \quad F(x_k) - A_j^k + d^T \hat{g}_j^k + \hat{u} \leq 0 \quad \text{for } j \in J_k \\
& \quad \frac{1}{2} d^T \bar{G}^k d \leq \hat{u} \\
& \quad B_i: (x_k + d) \leq b_i \quad \text{for } i = 1, \dots, \bar{m} ,
\end{aligned} \tag{4.12}$$

which is a (convex) QCQP with only one quadratic constraint.

Proposition 4.7. *If $(d_k, \hat{v}_k, \hat{u}_k)$ solves the QCQP (4.12), then there exists $(\lambda^k, \mu^k, \nu^k)$ with*

$$\begin{aligned}
 & \left(\bar{W}_p^k + \left(\sum_{j \in J_k} \mu_j^k \right) \bar{G}^k \right) d_k + \sum_{j \in J_k} \lambda_j^k g_j^k + \sum_{j \in J_k} \mu_j^k \hat{g}_j^k + \sum_{i=1}^m \nu_i^k B_{i:}^T = 0 \\
 & \sum_{j \in J_k} \lambda_j^k = 1 \\
 & -\alpha_j^k + d_k^T g_j^k - \hat{v}_k = 0 \quad \wedge \quad \lambda_j^k \geq 0 \\
 & -\alpha_j^k + d_k^T g_j^k - \hat{v}_k < 0 \quad \wedge \quad \lambda_j^k = 0 \\
 & F(x_k) - A_j^k + d_k^T \hat{g}_j^k + \hat{u}_k = 0 \quad \wedge \quad \mu_j^k \geq 0 \\
 & F(x_k) - A_j^k + d_k^T \hat{g}_j^k + \hat{u}_k < 0 \quad \wedge \quad \mu_j^k = 0 \\
 & \frac{1}{2} d_k^T \bar{G}^k d_k - \hat{u}_k = 0 \quad \wedge \quad \sum_{j \in J_k} \mu_j^k \geq 0 \\
 & \frac{1}{2} d_k^T \bar{G}^k d_k - \hat{u}_k < 0 \quad \wedge \quad \sum_{j \in J_k} \mu_j^k = 0 \\
 & B_{i:}(x_k + d_k) - b_i = 0 \quad \wedge \quad \nu_i^k \geq 0 \\
 & B_{i:}(x_k + d_k) - b_i < 0 \quad \wedge \quad \nu_i^k = 0 .
 \end{aligned} \tag{4.13}$$

Proof. Insert (4.12) into (2.10), where we have to take the following facts under consideration: If we denote the Lagrange multiplier of the quadratic constraint of (4.12)

$$\frac{1}{2} d_k^T \bar{G}^k d_k - \hat{u} \leq 0 \tag{4.14}$$

by $\bar{\nu}_k \geq 0$, then the complementarity condition of (2.10) for (4.14) reads

$$\begin{aligned}
 & \frac{1}{2} d_k^T \bar{G}^k d_k - \hat{u} = 0 \quad \wedge \quad \bar{\nu}_k \geq 0 \\
 & \frac{1}{2} d_k^T \bar{G}^k d_k - \hat{u} < 0 \quad \wedge \quad \bar{\nu}_k = 0 .
 \end{aligned} \tag{4.15}$$

The component of (2.10) with respect to the $\hat{u}_k$ -derivative reads

$$0 \stackrel{(2.10)}{=} \sum_{j \in J_k} \mu_j^k - \bar{\nu}^k \iff \bar{\nu}^k = \sum_{j \in J_k} \mu_j^k \geq 0 \tag{4.16}$$

and therefore we have

$$(4.15) \iff \begin{cases} \frac{1}{2} d_k^T \bar{G}^k d_k - \hat{u} = 0 & \wedge & \sum_{j \in J_k} \mu_j^k \geq 0 \\ \frac{1}{2} d_k^T \bar{G}^k d_k - \hat{u} < 0 & \wedge & \sum_{j \in J_k} \mu_j^k = 0 . \end{cases}$$

Inserting (4.16) into the component of (2.10) with respect to the d_k -derivative yields

$$\begin{aligned}
 0 & \stackrel{(2.10)}{=} \bar{W}_p^k d_k + \sum_{j \in J_k} \lambda_j^k g_j^k + \sum_{j \in J_k} \mu_j^k \hat{g}_j^k + \bar{\nu}^k \bar{G}^k d_k + \sum_{i=1}^m \nu_i^k B_{i:}^T + \sum_{i=1}^{\hat{m}} \hat{\nu}_i^k \hat{B}_{i:}^T \\
 & \stackrel{(4.16)}{=} \left(\bar{W}_p^k + \left(\sum_{j \in J_k} \mu_j^k \right) \bar{G}^k \right) d_k + \sum_{j \in J_k} \lambda_j^k g_j^k + \sum_{j \in J_k} \mu_j^k \hat{g}_j^k + \sum_{i=1}^m \nu_i^k B_{i:}^T + \sum_{i=1}^{\hat{m}} \hat{\nu}_i^k \hat{B}_{i:}^T .
 \end{aligned}$$

The component of (2.10) with respect to the $\hat{v}_k$ -derivative reads

$$0 \stackrel{(2.10)}{=} 1 - \sum_{j \in J_k} \lambda_j^k . \quad \square$$

Proposition 4.8.

- If $(d_k, \hat{v}_k, \lambda_k, \mu_k)$ solves the QCQP (4.11), then

$$(\tilde{d}_k, \tilde{v}_k, \tilde{u}_k, \tilde{\lambda}_k, \tilde{\mu}_k, \tilde{\nu}_k) := (d_k, \hat{v}_k, \frac{1}{2} d_k^T \bar{G}^k d_k, \lambda_k, \mu_k, \nu_k) \quad (4.17)$$

solves the reduced QCQP (4.12).

- If $(\tilde{d}_k, \tilde{v}_k, \tilde{u}_k, \tilde{\lambda}_k, \tilde{\mu}_k)$ solves the reduced QCQP (4.12), then

$$(d_k, \hat{v}_k, \lambda_k, \mu_k, \nu_k) := (\tilde{d}_k, \tilde{v}_k, \tilde{\lambda}_k, \tilde{\mu}_k, \tilde{\nu}_k) \quad (4.18)$$

solves the QCQP (4.11).

Proof. Because of (4.10), comparing the optimality conditions (3.41) of the QCQP (4.11) (extended by linear constraints which is no problem) with the optimality conditions (4.13) of the reduced QCQP (4.12), which are both sufficient due to the convexity of both QCQPs and Theorem 2.10, implies that we only need to check the terms concerning the approximation of F .

- (4.17) satisfies (4.13) because:

– We obtain

$$F(x_k) - A_j^k + \tilde{d}_k^T \hat{g}_j^k + \tilde{u}_k \stackrel{(4.17)}{=} F(x_k) - A_j^k + d_k^T \hat{g}_j^k + \frac{1}{2} d_k^T \bar{G}^k d_k$$

$$\begin{cases} \stackrel{(3.41)}{<} 0 & \text{for } \tilde{\mu}_j^k \stackrel{(4.17)}{=} \mu_j^k \stackrel{(3.41)}{=} 0 \\ \stackrel{(3.41)}{=} 0 & \text{for } \tilde{\mu}_j^k \stackrel{(4.17)}{=} \mu_j^k \stackrel{(3.41)}{\geq} 0 . \end{cases}$$

– Furthermore, we have

$$\frac{1}{2} \tilde{d}_k^T \bar{G}^k \tilde{d}_k - \tilde{u}_k \stackrel{(4.17)}{=} \frac{1}{2} d_k^T \bar{G}^k d_k - \frac{1}{2} d_k^T \bar{G}^k d_k = 0$$

and

$$\tilde{\mu}_j^k \stackrel{(4.17)}{=} \mu_j^k \stackrel{(3.41)}{\geq} 0 .$$

- (4.18) satisfies (3.41) because:

– In the case

$$\frac{1}{2} d_k^T \bar{G}^k d_k - \tilde{u}_k = 0 \stackrel{(4.13)}{\implies} \sum_{j \in J_k} \tilde{\mu}_j^k \geq 0 \quad (4.19)$$

we obtain

$$F(x_k) - A_j^k + d_k^T \hat{g}_j^k + \frac{1}{2} d_k^T \bar{G}^k d_k \stackrel{(4.19)}{=} F(x_k) - A_j^k + \tilde{d}_k^T \hat{g}_j^k + \tilde{u}_k$$

$$\begin{cases} \stackrel{(4.13)}{<} 0 & \text{for } \mu_j^k \stackrel{(4.18)}{=} \tilde{\mu}_j^k \stackrel{(4.13)}{=} 0 \\ \stackrel{(4.13)}{=} 0 & \text{for } \mu_j^k \stackrel{(4.18)}{=} \tilde{\mu}_j^k \stackrel{(4.13)}{\geq} 0 . \end{cases}$$

– In the case

$$\begin{aligned} \frac{1}{2} \tilde{d}_k^T \bar{G}^k \tilde{d}_k - \tilde{u}_k < 0 &\stackrel{(4.13)}{\iff} \sum_{j \in J_k} \tilde{\mu}_j^k = 0 \\ &\stackrel{(4.13)}{\iff} \tilde{\mu}_j^k = 0 \quad \text{for all } j \in J_k \\ &\stackrel{(4.18)}{\iff} \mu_j^k = 0 \quad \text{for all } j \in J_k \end{aligned} \tag{4.20}$$

we obtain

$$F(x_k) - A_j^k + d_k^T \hat{g}_j^k + \frac{1}{2} d_k^T \bar{G}^k d_k \stackrel{(4.20)}{<} F(x_k) - A_j^k + \tilde{d}_k^T \hat{g}_j^k + \tilde{u}_k \stackrel{(4.13)}{\leq} 0. \quad \square$$

Remark 4.9. If we make the choice (4.10) in Algorithm 3.20, then we can solve the reduced QCQP (4.12) instead of the QCQP (4.11). This yields the following observations:

- We expect that the reduced QCQP (4.12) should be solved much faster than the QCQP (4.11) because:
 - An interior point method for solving QPs/QCQPs solves a linear system (called the KKT-system) at each iteration which is the most time consuming operation, i.e. the bigger the KKT-system is, the longer the interior point method will need to solve the problem.
 - If we solved a QP to determine the search direction (we do not do this because of Remark 3.11), we would obtain $|J_k| + 1$ linear constraints for approximating F which increases the size of the KKT-system by $|J_k| + 1$ rows compared to the unconstrained case (i.e. without F).
 - If we solve the QCQP (4.11) to determine the search direction, we will obtain — in addition to the $|J_k| + 1$ rows which are due to the linear terms — $|J_k| + 1$ many $n \times n$ -blocks (i.e. $(|J_k| + 1)n$ rows) which are due to the $|J_k| + 1$ quadratic terms. Since J_k is bounded by the maximal bundle dimension M and if we choose, e.g., $M = n + 3$ (this is the recommended default value for M in the bundle-Newton method by LUKŠAN & VLČEK [64] for nonsmooth unconstrained minimization), then the KKT-system can become big quite fast.
 - If we solve the reduced QCQP (4.12) to determine the search direction, we will obtain — in addition to the $|J_k| + 1$ rows which are due to the linear terms — only one $n \times n$ -block (i.e. n rows) since we only have one quadratic term. Therefore, if n is not too big, we expect that solving the reduced QCQP should not take significantly more time than solving the corresponding QP at least for a good interior point method and this turns out to be true indeed (cf. the comparisons in Subsection 4.4.1).

So the big advantage of the reduced QCQP (4.12) is that it has a size similar to that of the corresponding QP (i.e. its size is much smaller than that of the QCQP (4.11)), but it still uses quadratic information to deal with the nonlinearity of F .

- We do not need to compute a positive definite modification $\bar{\tilde{G}}_j^k$ of $\hat{G}_j$ in (3.61).
- We can replace the model change condition in (3.121) by

$$F(x_k + t_L d_k) - \hat{\beta} + d_k^T (\hat{g} + \hat{\rho}(t_L - t) \hat{G} d_k) \geq m_F \cdot (-\hat{u}_k)$$

and therefore we do not need to compute a positive definite modification $\bar{\tilde{G}}$ of $\hat{G}$ in (3.120).

4.3.2 SOCP-reformulation

Motivation 4.10. We want to reformulate the QCQP (4.12) as an SOCP (cf. ALIZADEH & GOLDFARB [1, p. 8, 2.1. QPs and QCQPs] and LOBO et al. [58, p. 199, 2.3 Problems with hyperbolic constraints, (8)]).

Proposition 4.11. *We have for all $u \in \mathbb{R}$*

$$4u = (1 + u)^2 - (1 - u)^2 . \quad (4.21)$$

Proof. (4.21) follows from an easy calculation. $\square$

Notation 4.12. We denote the linear space of upper triangular $n \times n$ -matrices by

$$\mathbb{R}_{\text{triu}}^{n \times n} := \{A \in \mathbb{R}^{n \times n} : A_{ij} = 0 \text{ for } i > j\} . \quad (4.22)$$

Proposition 4.13.

- We have

$$-\alpha_j^k + d^T g_j^k \leq \hat{v} \iff \begin{pmatrix} -g_j^k \\ 1 \\ 0 \\ 0 \end{pmatrix}^T \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + \alpha_j^k \geq 0 . \quad (4.23)$$

- We have

$$F(x_k) - A_j^k + d^T \hat{g}_j^k + \hat{u} \leq 0 \iff \begin{pmatrix} -\hat{g}_j^k \\ 0 \\ 0 \\ -1 \end{pmatrix}^T \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} - (F(x_k) - A_j^k) \geq 0 . \quad (4.24)$$

- We have

$$\frac{1}{2}d^T \bar{G}^k d \leq \hat{u} \iff \left| \begin{pmatrix} \sqrt{2}\hat{R}^k & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right|_2 \leq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}^T \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + 1 , \quad (4.25)$$

where

$$\bar{G}^k = \hat{R}^{kT} \hat{R}^k \quad (4.26)$$

is the Cholesky factorization of $\bar{G}^k$ with $\hat{R}^k \in \mathbb{R}_{\text{triu}}^{n \times n}$.

- We have

$$B_{i:}(x_k + d) \leq b_i \iff (-B_{i:} \ 0 \ 0 \ 0) \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + (b_i - B_{i:}x_k) \geq 0 . \quad (4.27)$$

- Let $v \in \mathbb{R}$. Then we have

$$\hat{v} + \frac{1}{2}d^T \bar{W}_p^k d \leq v \iff \left| \begin{pmatrix} \sqrt{2}R_p^k & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right|_2 \leq \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}^T \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + 1 , \quad (4.28)$$

where

$$\bar{W}_p^k = R_p^{kT} R_p^k \quad (4.29)$$

is the Cholesky factorization of $\bar{W}_p^k$ with $R_p^k \in \mathbb{R}_{\text{triu}}^{n \times n}$.

Proof.

- (4.23) and (4.24) hold due to an easy calculation.
- Since $\bar{G}^k$ is positive definite, we have

$$\begin{aligned} \frac{1}{2}d^T \bar{G}^k d &\leq \hat{u} \\ \Leftrightarrow^{+1} 1 + \hat{u} &\geq 1 + \underbrace{\frac{1}{2}d^T \bar{G}^k d}_{\geq 0} \geq 0 \end{aligned} \quad (4.30)$$

and therefore we obtain (4.25)

$$\begin{aligned} \frac{1}{2}d^T \bar{G}^k d &\leq \hat{u} \\ \Leftrightarrow 0 &\geq 2d^T \bar{G}^k d - 4\hat{u} \\ &\stackrel{(4.21)}{=} 2d^T \bar{G}^k d - ((1 + \hat{u})^2 - (1 - \hat{u})^2) \\ &\stackrel{(4.26)}{=} 2d^T R_p^{kT} \hat{R}^k d - (1 + \hat{u})^2 + (1 - \hat{u})^2 \\ &= \left| \begin{pmatrix} \sqrt{2}\hat{R}^k d \\ 1 - \hat{u} \end{pmatrix} \right|_2^2 - (1 + \hat{u})^2 \\ \Leftrightarrow \left| \begin{pmatrix} \sqrt{2}\hat{R}^k d \\ 1 - \hat{u} \end{pmatrix} \right|_2 &\leq |1 + \hat{u}| \stackrel{(4.30)}{=} 1 + \hat{u} \\ \Leftrightarrow \left| \begin{pmatrix} \sqrt{2}\hat{R}^k & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right| &\leq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}^T \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + 1 . \end{aligned}$$

- (4.27) holds due to an easy calculation.
- Since $\bar{W}_p^k$ is positive definite, we have

$$\begin{aligned} \hat{v} + \frac{1}{2}d^T \bar{W}_p^k d &\leq v \\ \Leftrightarrow \frac{1}{2}d^T \bar{W}_p^k d &\leq -(\hat{v} - v) \\ \Leftrightarrow^{+1} 1 + \underbrace{\frac{1}{2}d^T \bar{W}_p^k d}_{\geq 0} &\leq 1 - (\hat{v} - v) \\ \Rightarrow 0 &\leq 1 - (\hat{v} - v) \end{aligned} \quad (4.31)$$

and therefore we obtain (4.28)

$$\begin{aligned} \hat{v} + \frac{1}{2}d^T \bar{W}_p^k d &\leq v \\ \stackrel{-v}{\Leftrightarrow} \hat{v} - v + \frac{1}{2}d^T \bar{W}_p^k d &\leq 0 \\ \stackrel{4}{\Leftrightarrow} 0 &\geq 4(\hat{v} - v) + 2d^T \bar{W}_p^k d \\ &\stackrel{(4.21)}{=} 2d^T \bar{W}_p^k d + (1 + (\hat{v} - v))^2 - (1 - (\hat{v} - v))^2 \\ &\stackrel{(4.29)}{=} 2d^T R_p^{kT} R_p^k d + (1 + (\hat{v} - v))^2 - (1 - (\hat{v} - v))^2 \\ &= \left| \begin{pmatrix} \sqrt{2}R_p^k d \\ 1 + (\hat{v} - v) \end{pmatrix} \right|_2^2 - (1 - (\hat{v} - v))^2 \\ \stackrel{\sqrt{}}{\Leftrightarrow} \left| \begin{pmatrix} \sqrt{2}R_p^k d \\ 1 + \hat{v} - v \end{pmatrix} \right|_2 &\leq |1 - (\hat{v} - v)| \stackrel{(4.31)}{=} 1 - \hat{v} + v \\ \Leftrightarrow \left| \begin{pmatrix} \sqrt{2}R_p^k & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right|_2 &\leq \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}^T \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + 1 . \end{aligned}$$

□

Proposition 4.14. *The primal SOCP-formulation of the QCQP (4.12) reads*

$$\begin{aligned}
& \min_{d, \hat{v}, v, \hat{u}} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}^T \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} \\
& \text{s.t.} \quad \begin{pmatrix} -g_j^k \\ 1 \\ 0 \\ 0 \end{pmatrix}^T \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + \alpha_j^k \geq 0 \\
& \quad \begin{pmatrix} -\hat{g}_j^k \\ 0 \\ 0 \\ -1 \end{pmatrix}^T \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} - (F(x_k) - A_j^k) \geq 0 \\
& \quad \left| \begin{pmatrix} \sqrt{2}\hat{R}^k & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right|_2 \leq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}^T \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + 1 \\
& \quad (-B_i: 0 \ 0 \ 0) \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + (b_i - B_i x_k) \geq 0 \\
& \quad \left| \begin{pmatrix} \sqrt{2}R_p^k & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right|_2 \leq \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}^T \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + 1 .
\end{aligned} \tag{4.32}$$

Proof. Insert (4.23), (4.24), (4.25), (4.27) and (4.28) into (4.12). $\square$

4.3.3 Strict feasibility of the SOCP

Proposition 4.15. *We have for all $u < 0$*

$$|1 + u| < 1 - u . \tag{4.33}$$

Proof. Because of

$$-1 < 1 \xLeftrightarrow{u < 0} -u > u \xLeftrightarrow{+1} 1 - u > 1 + u \tag{4.34}$$

and

$$-1 < 1 \xLeftrightarrow{-u} -(1 + u) < 1 - u , \tag{4.35}$$

we obtain

$$(4.34) \wedge (4.35) \iff |1 + u| < 1 - u . \quad \square$$

Proposition 4.16. *If $x_k \in \mathbb{R}^n$ is strictly feasible for the general linear inequality constraints and $F(x_k) < 0$, then every point $(d, \hat{v}, v, \hat{u})$ with*

$$\begin{aligned}
& d = 0 \\
& -\min_{j \in J_k} \alpha_j^k < \hat{v} < v \\
& 0 < \hat{u} < -F(x_k) + \min_{j \in J_k} A_j^k
\end{aligned} \tag{4.36}$$

is strictly primal feasible for the SOCP (4.32).

Proof.

1. Inserting (4.36) into the left side of the linear bundle constraints (with respect to the approximation of the objective function) of (4.32) yields

$$\hat{v} + \alpha_j^k \geq \hat{v} + \min_{i \in J_k} \alpha_i^k \stackrel{(4.36)}{>} 0 .$$

2. Due to

$$-F(x_k) + \min_{i \in J_k} A_i^k \stackrel{(3.29)}{\geq} \underbrace{-F(x_k)}_{<0} > 0 ,$$

the choice of $\hat{u}$ according to (4.36) is always possible. Therefore, inserting (4.36) into the left side of the linear bundle constraints (with respect to the approximation of the constraint) of (4.32) yields

$$-\hat{u} - (F(x_k) - A_j^k) \geq -\hat{u} - F(x_k) + \min_{i \in J_k} A_i^k \stackrel{(4.36)}{>} 0 .$$

3. Inserting (4.36) into the quadratic constraint (with respect to the constraint) of (4.32) yields

$$|-\hat{u} + 1| \leq \hat{u} + 1 \iff |1 + (-\hat{u})| \leq 1 - (-\hat{u}) .$$

This inequality holds because of

$$-\hat{u} \stackrel{(4.36)}{<} 0$$

and (4.33), and it is sharp, in particular.

4. Inserting (4.36) into the left side of the general linear inequality constraints of (4.32) yields

$$b_i - B_{i:}x_k > 0 ,$$

where this inequality holds because of the assumption of the strict feasibility of x_k for the general linear inequality constraints.

5. Inserting (4.36) into the constraint of (4.32) that corresponds with the objective function yields

$$|\hat{v} - v + 1| \leq -\hat{v} + v + 1 \iff |1 + (\hat{v} - v)| \leq 1 - (\hat{v} - v) .$$

This inequality holds because of

$$\hat{v} - v \stackrel{(4.36)}{<} 0$$

and (4.33), and it is sharp, in particular. □

Proposition 4.17. *If $y \in \mathbb{R}^{n+1}$ with*

$$y_{n+1} := \frac{|y_{1:n}|^2 - (\tilde{y} - \hat{y})^2}{2(\tilde{y} - \hat{y})} , \tag{4.37}$$

where $\tilde{y} > 0$ and $\hat{y} \in (0, \tilde{y})$, then

$$|y| + \hat{y} = y_{n+1} + \tilde{y} . \tag{4.38}$$

Proof. Due to assumption we obtain

$$0 < \hat{y} < \tilde{y} \xrightarrow{-\hat{y}} 0 < \tilde{y} - \hat{y} . \tag{4.39}$$

Therefore, we estimate

$$y_{n+1} + \tilde{y} - \hat{y} \stackrel{(4.37)}{=} \frac{|y_{1:n}|^2 - (\tilde{y} - \hat{y})^2}{2(\tilde{y} - \hat{y})} + \tilde{y} - \hat{y} = \frac{|y_{1:n}|^2 + (\tilde{y} - \hat{y})^2}{2(\tilde{y} - \hat{y})} \stackrel{(4.39)}{>} 0 , \tag{4.40}$$

which implies

$$\begin{aligned}
|y|^2 &= |y_{1:n}|^2 + y_{n+1}^2 \\
&\stackrel{(4.37)}{=} 2y_{n+1}(\check{y} - \hat{y}) + (\check{y} - \hat{y})^2 + y_{n+1}^2 \\
&= (y_{n+1} + \check{y} - \hat{y})^2 \\
&\stackrel{\sqrt{\cdot}}{\iff} |y| \stackrel{(4.40)}{=} y_{n+1} + \check{y} - \hat{y} \\
&\stackrel{+\hat{y}}{\iff} |y| + \hat{y} = y_{n+1} + \check{y} . \quad \square
\end{aligned}$$

Proposition 4.18. *Let the following requirements be satisfied:*

1. Let

$$w_j > 0 \text{ for all } j = 1, \dots, m . \quad (4.41)$$

2. Let $z_{m+1} \in \mathbb{R}^{n+1}$ with

$$z_{m+1}^{n+1} := \frac{|z_{m+1}^{1:n}|^2 - (\sum_{j=1}^m w_j - \delta_1)^2}{2(\sum_{j=1}^m w_j - \delta_1)} , \quad (4.42)$$

where

$$\delta_1 \in (0, \sum_{j=1}^m w_j) \quad (4.43)$$

is a fixed number, and set

$$w_{m+1} := \sum_{j=1}^m w_j + z_{m+1}^{n+1} . \quad (4.44)$$

3. Let

$$w_{m+2+j} > 0 \text{ for all } j = 1, \dots, m \quad (4.45)$$

such that

$$\sum_{j=1}^m w_{m+2+j} = 1 . \quad (4.46)$$

4. Let

$$w_{m+2+m+i} > 0 \text{ for all } i = 1, \dots, \bar{m} . \quad (4.47)$$

5. We set

$$r := \sum_{j=1}^m \hat{g}_j^k w_j + \sum_{j=1}^m g_j^k w_{m+2+j} + \sum_{i=1}^{\bar{m}} B_{i\cdot}^T w_{m+2+m+i} - \sqrt{2} \hat{R}^{k^T} z_{m+1}^{1:n} . \quad (4.48)$$

Furthermore, let $z_{m+2}^{1:n} \in \mathbb{R}^n$ be the unique solution of the (upper triangular) linear system

$$\sqrt{2} R_p^{k^T} z_{m+2}^{1:n} = r \quad (4.49)$$

and set

$$z_{m+2}^{n+1} := \frac{|z_{m+2}^{1:n}|^2 - (1 - \delta_2)^2}{2(1 - \delta_2)} , \quad (4.50)$$

where

$$\delta_2 \in (0, 1) \quad (4.51)$$

is a fixed number, as well as

$$w_{m+2} := z_{m+2}^{n+1} + 1 . \quad (4.52)$$

Then the point that is obtained by this construction is strictly dual feasible for the SOCP (4.32) and the following estimation holds

$$|z_{m+i}| < |z_{m+i}| + \delta_i = w_{m+i} \quad \text{for } i = 1, 2 . \quad (4.53)$$

Proof.

- The linear system in (4.49) is uniquely solvable, since R_p^k is a Cholesky factor of the positive definite (and therefore regular) matrix $\bar{W}_p^k$ due to (4.29), and consequently R_p^k is also regular.
- Our point is strictly feasible for the norm-constraint of (4.2) for (4.32), because of (4.41), (4.45), (4.47) and because of the following two facts:

1. For

$$\begin{aligned} y &:= z_{m+1} \\ \check{y} &:= \sum_{j=1}^m w_j \\ \hat{y} &:= \delta_1 \end{aligned} \quad (4.54)$$

the assumptions of Proposition 4.17 are satisfied due to

$$\check{y} \stackrel{(4.54)}{=} \sum_{j=1}^m w_j \stackrel{(4.41)}{>} 0$$

and

$$\hat{y} \stackrel{(4.54)}{=} \delta_1 \stackrel{(4.43)}{\in} (0, \sum_{j=1}^m w_j) \stackrel{(4.54)}{=} (0, \check{y})$$

as well as

$$y_{n+1} \stackrel{(4.54)}{=} z_{m+1}^{n+1} \stackrel{(4.42)}{=} \frac{|z_{m+1}^{1:n}|^2 - (\sum_{j=1}^m w_j - \delta_1)^2}{2(\sum_{j=1}^m w_j - \delta_1)} \stackrel{(4.54)}{=} \frac{|y_{1:n}|^2 - (\check{y} - \hat{y})^2}{2(\check{y} - \hat{y})} ,$$

and therefore we obtain

$$|z_{m+1}| \stackrel{(4.43)}{<} |z_{m+1}| + \delta_1 \stackrel{(4.54)}{=} |y| + \hat{y} \stackrel{(4.38)}{=} y_{n+1} + \check{y} \stackrel{(4.54)}{=} z_{m+1}^{n+1} + \sum_{j=1}^m w_j \stackrel{(4.44)}{=} w_{m+1} .$$

2. For

$$\begin{aligned} y &:= z_{m+2} \\ \check{y} &:= 1 \\ \hat{y} &:= \delta_2 \end{aligned} \quad (4.55)$$

the assumptions of Proposition 4.17 are satisfied due to

$$\check{y} \stackrel{(4.55)}{>} 0$$

and

$$\hat{y} \stackrel{(4.55)}{=} \delta_2 \stackrel{(4.51)}{\in} (0, 1) \stackrel{(4.55)}{=} (0, \tilde{y})$$

as well as

$$y_{n+1} \stackrel{(4.55)}{=} z_{m+2}^{n+1} \stackrel{(4.50)}{=} \frac{|z_{m+2}^{1:n}|^2 - (1 - \delta_2)^2}{2(1 - \delta_2)} \stackrel{(4.55)}{=} \frac{|y_{1:n}|^2 - (\tilde{y} - \hat{y})^2}{2(\tilde{y} - \hat{y})},$$

and therefore we obtain

$$|z_{m+2}| \stackrel{(4.51)}{<} |z_{m+2}| + \delta_2 \stackrel{(4.55)}{=} |y| + \hat{y} \stackrel{(4.38)}{=} y_{n+1} + \tilde{y} \stackrel{(4.55)}{=} z_{m+2}^{n+1} + 1 \stackrel{(4.52)}{=} w_{m+2}.$$

- Inserting our point into the left side of the linear equality constraint of (4.2) for (4.32) yields the following equation

$$\begin{aligned} & \sum_{i=1}^L A_i^T z_i + c_i w_i \\ & \stackrel{(4.1), (4.32)}{\downarrow} \stackrel{=}{=} \sum_{j=1}^m c_j w_j + A_{m+1}^T z_{m+1} + c_{m+1} w_{m+1} + A_{m+2}^T z_{m+2} + c_{m+2} w_{m+2} \\ & \quad + \sum_{j=1}^m c_{m+2+j} w_{m+2+j} + \sum_{i=1}^{\tilde{m}} c_{m+2+m+i} w_{m+2+m+i} \\ & \stackrel{(4.32)}{=} \sum_{j=1}^m \begin{pmatrix} -\hat{g}_j^k \\ 0 \\ 0 \\ -1 \end{pmatrix} w_j + \begin{pmatrix} \sqrt{2}\hat{R}^{k^T} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & -1 \end{pmatrix} z_{m+1} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} w_{m+1} \\ & \quad + \begin{pmatrix} \sqrt{2}\hat{R}_p^{k^T} & 0 \\ 0 & 0 \\ 0 & -1 \end{pmatrix} z_{m+2} + \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} w_{m+2} + \sum_{j=1}^m \begin{pmatrix} -g_j^k \\ 1 \\ 0 \\ 0 \end{pmatrix} w_{m+2+j} + \sum_{i=1}^{\tilde{m}} \begin{pmatrix} -B_{i:}^T \\ 0 \\ 0 \\ 0 \end{pmatrix} w_{m+2+m+i} \\ & = \begin{pmatrix} -\sum_{j=1}^m \hat{g}_j^k w_j + \sqrt{2}\hat{R}^{k^T} z_{m+1}^{1:n} + \sqrt{2}\hat{R}_p^{k^T} z_{m+2}^{1:n} - \sum_{j=1}^m g_j^k w_{m+2+j} - \sum_{i=1}^{\tilde{m}} B_{i:}^T w_{m+2+m+i} \\ z_{m+2}^{n+1} - w_{m+2} + \sum_{j=1}^m w_{m+2+j} \\ -z_{m+2}^{n+1} + w_{m+2} \\ -\sum_{j=1}^m w_j - z_{m+1}^{n+1} + w_{m+1} \end{pmatrix} \\ & =: \hat{f} \in \mathbb{R}^{n+3}. \end{aligned} \tag{4.56}$$

From (4.56) we obtain

$$\begin{aligned} & \hat{f}^{1:n} \stackrel{(4.56)}{=} -\sum_{j=1}^m \hat{g}_j^k w_j + \sqrt{2}\hat{R}^{k^T} z_{m+1}^{1:n} + \sqrt{2}\hat{R}_p^{k^T} z_{m+2}^{1:n} - \sum_{j=1}^m g_j^k w_{m+2+j} - \sum_{i=1}^{\tilde{m}} B_{i:}^T w_{m+2+m+i} \\ & \stackrel{(4.48)}{=} \sqrt{2}\hat{R}_p^{k^T} z_{m+2}^{1:n} - r \stackrel{(4.1), (4.32)}{=} 0 \stackrel{(4.49)}{\downarrow} \stackrel{=}{=} f^{1:n} \end{aligned}$$

and

$$\begin{aligned} & \hat{f}^{n+1} \stackrel{(4.56)}{=} z_{m+2}^{n+1} - w_{m+2} + \sum_{j=1}^m w_{m+2+j} \\ & \stackrel{(4.52), (4.46)}{\downarrow} \stackrel{=}{=} z_{m+2}^{n+1} - (z_{m+2}^{n+1} + 1) + 1 = 0 \stackrel{(4.1), (4.32)}{\downarrow} \stackrel{=}{=} f^{n+1} \end{aligned}$$

as well as

$$\hat{f}^{n+2} \stackrel{(4.56)}{=} -z_{m+2}^{n+1} + w_{m+2} \stackrel{(4.52)}{=} 1 \stackrel{(4.1),(4.32)}{\downarrow} f^{n+2}$$

and

$$\hat{f}^{n+3} \stackrel{(4.56)}{=} -\sum_{j=1}^m w_j - z_{m+1}^{n+1} + w_{m+1} \stackrel{(4.44)}{=} 0 \stackrel{(4.1),(4.32)}{\downarrow} f^{n+3} . \quad \square$$

Remark 4.19.

- The estimate (4.53) allows to control the distance to the boundary for the dual starting point.
- When using `socp` by LOBO et al. [57] for the computation of the search direction in Algorithm 3.20, we chose the values

$$\begin{aligned} \hat{v} &:= \frac{1}{10} \\ v &:= 1 \\ \hat{u} &:= -\frac{1}{2}F(x_k) \end{aligned}$$

for the primal starting point and

$$\begin{aligned} w_j &:= 1 && \text{for all } j = 1, \dots, m \\ z_{m+1}^{1:n} &:= 0 \\ \delta_1 &:= m - \frac{1}{5} \\ w_{m+2+j} &:= \frac{1}{m} && \text{for all } j = 1, \dots, m \\ w_{m+2+m+i} &:= 1 && \text{for all } i = 1, \dots, \bar{m} \\ \delta_2 &:= \frac{1}{2} \end{aligned} \tag{4.57}$$

for the dual starting point, where for these choices all conditions of (4.36) and Proposition 4.18 are satisfied, which follows from an easy examination (for the numerical behavior of `socp` in Algorithm 3.20 cf. Remark 6.6).

4.3.4 Connection between the Lagrange multipliers

Motivation 4.20. We want to work out the connection between the Lagrange multiplier $(\lambda, \mu, \nu) \in \mathbb{R}_{\geq 0}^{m+m+\bar{m}}$ of the QCQP (4.12) and the Lagrange multipliers $(z_i, w_i) \in \mathbb{R}^{N_i-1} \times \mathbb{R}_{\geq 0}$ for all $i = 1, \dots, m+1+m+\bar{m}$ of the SOCP (4.32).

Proposition 4.21. *We have*

$$\begin{aligned} \left(\bar{W}_p^k + \left(\sum_{j=1}^m w_j \right) \hat{G}^k \right) d + \sum_{j=1}^m g_j^k w_{m+2+j} + \sum_{j=1}^m \hat{g}_j^k w_j + \sum_{i=1}^{\bar{m}} B_i^T w_{m+2+m+i} &= 0 \\ \sum_{j=1}^m w_{m+2+j} &= 1 . \end{aligned} \tag{4.58}$$

Proof.

- In our notation we have the correspondence

$$x \longleftrightarrow (d^T, \hat{v}, v, \hat{u})^T. \quad (4.59)$$

- We have (Note: d and d_{m+1} have nothing to do with each other here)

$$\begin{aligned} t_{m+1} &\stackrel{(4.6)}{=} c_{m+1}^T x + d_{m+1} \\ &\stackrel{(4.1), (4.32), (4.59)}{\downarrow} \stackrel{=}{=} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}^T \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + 1 \\ &= \hat{u} + 1 \\ \iff \hat{u} - 1 - t_{m+1} &= -2 \end{aligned} \quad (4.60)$$

and therefore we obtain

$$\begin{aligned} A_{m+1} \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + b_{m+1} &\stackrel{(4.1), (4.32)}{\downarrow} \stackrel{=}{=} \begin{pmatrix} \sqrt{2}\hat{R}^k & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \sqrt{2}\hat{R}^k d \\ -\hat{u} + 1 \end{pmatrix} \\ \xrightarrow{A_{m+1}^T} A_{m+1}^T (A_{m+1} \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + b_{m+1}) &\stackrel{(4.26)}{=} \begin{pmatrix} 2\hat{G}^k d \\ 0 \\ 0 \\ \hat{u} - 1 \end{pmatrix} \\ \xrightarrow{-t_{m+1}c_{m+1}} A_{m+1}^T (A_{m+1} \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + b_{m+1}) - t_{m+1}c_{m+1} &\stackrel{(4.1), (4.32)}{\downarrow} \stackrel{=}{=} \begin{pmatrix} 2\hat{G}^k d \\ 0 \\ 0 \\ \hat{u} - 1 \end{pmatrix} - t_{m+1} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \\ &\stackrel{(4.60)}{=} \begin{pmatrix} 2\hat{G}^k d \\ 0 \\ 0 \\ -2 \end{pmatrix}. \end{aligned} \quad (4.61)$$

- We have (Note: d and d_{m+2} have nothing to do with each other here)

$$\begin{aligned} t_{m+2} &\stackrel{(4.6)}{=} c_{m+2}^T x + d_{m+2} \\ &\stackrel{(4.1), (4.32), (4.59)}{\downarrow} \stackrel{=}{=} \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}^T \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + 1 \\ &= -\hat{v} + v + 1 \\ \iff \hat{v} - v + 1 + t_{m+2} &= 2 \end{aligned} \quad (4.62)$$

and therefore we obtain

$$\begin{aligned} A_{m+2} \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + b_{m+2} &\stackrel{(4.1), (4.32)}{\downarrow} \stackrel{=}{=} \begin{pmatrix} \sqrt{2}R_p^k & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} \sqrt{2}R_p^k d \\ \hat{v} - v + 1 \end{pmatrix} \\ \xrightarrow{A_{m+2}^T} A_{m+2}^T (A_{m+2} \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + b_{m+2}) &\stackrel{(4.29)}{=} \begin{pmatrix} 2\bar{W}_p^k d \\ \hat{v} - v + 1 \\ -(\hat{v} - v + 1) \\ 0 \end{pmatrix} \\ \xrightarrow{-t_{m+2}c_{m+2}} A_{m+2}^T (A_{m+2} \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + b_{m+2}) - t_{m+2}c_{m+2} &\stackrel{(4.1), (4.32)}{\downarrow} \stackrel{=}{=} \begin{pmatrix} 2\bar{W}_p^k d \\ \hat{v} - v + 1 \\ -(\hat{v} - v + 1) \\ 0 \end{pmatrix} - t_{m+2} \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} \\ &\stackrel{(4.62)}{=} \begin{pmatrix} 2\bar{W}_p^k d \\ 2 \\ -2 \\ 0 \end{pmatrix}. \end{aligned} \quad (4.63)$$

- We have

$$z_j \stackrel{(4.4)}{=} -\frac{w_j}{t_j} u_j \stackrel{(4.5), (4.59)}{\downarrow} -\frac{w_j}{t_j} \left(A_j \begin{pmatrix} d \\ \hat{v} \\ \hat{v} \\ \hat{u} \end{pmatrix} + b_j \right), \quad (4.64)$$

and therefore we obtain

$$\begin{aligned} \begin{pmatrix} 0 \\ 0 \\ -1 \\ 0 \end{pmatrix} &\stackrel{(4.1), (4.32)}{\downarrow} -f \\ &\stackrel{(4.2)}{=} -\left(\sum_{i=1}^L A_i^T z_i + c_i w_i \right) \\ &= -\left(\sum_{j=1}^m c_j w_j \right) - \left(A_{m+1}^T z_{m+1} + c_{m+1} w_{m+1} \right) - \left(A_{m+2}^T z_{m+2} + c_{m+2} w_{m+2} \right) \\ &\quad - \left(\sum_{j=1}^m c_{m+2+j} w_{m+2+j} \right) - \left(\sum_{i=1}^{\tilde{m}} c_{m+2+m+i} w_{m+2+m+i} \right) \\ &\stackrel{(4.64)}{=} -\left(\sum_{j=1}^m c_j w_j \right) - \left(A_{m+1}^T \left(-\frac{w_{m+1}}{t_{m+1}} \left(A_{m+1} \begin{pmatrix} d \\ \hat{v} \\ \hat{v} \\ \hat{u} \end{pmatrix} + b_{m+1} \right) \right) + c_{m+1} w_{m+1} \right) \\ &\quad - \left(A_{m+2}^T \left(-\frac{w_{m+2}}{t_{m+2}} \left(A_{m+2} \begin{pmatrix} d \\ \hat{v} \\ \hat{v} \\ \hat{u} \end{pmatrix} + b_{m+2} \right) \right) + c_{m+2} w_{m+2} \right) \\ &\quad - \left(\sum_{j=1}^m c_{m+2+j} w_{m+2+j} \right) - \left(\sum_{i=1}^{\tilde{m}} c_{m+2+m+i} w_{m+2+m+i} \right) \\ &\stackrel{(4.1), (4.32)}{\downarrow} -\sum_{j=1}^m \begin{pmatrix} -\hat{g}_j^k \\ 0 \\ 0 \\ -1 \end{pmatrix} w_j - \left(-\frac{w_{m+1}}{t_{m+1}} \left(A_{m+1}^T \left(A_{m+1} \begin{pmatrix} d \\ \hat{v} \\ \hat{v} \\ \hat{u} \end{pmatrix} + b_{m+1} \right) - t_{m+1} c_{m+1} \right) \right) \\ &\quad - \left(-\frac{w_{m+2}}{t_{m+2}} \left(A_{m+2}^T \left(A_{m+2} \begin{pmatrix} d \\ \hat{v} \\ \hat{v} \\ \hat{u} \end{pmatrix} + b_{m+2} \right) - t_{m+2} c_{m+2} \right) \right) \\ &\quad - \sum_{j=1}^m \begin{pmatrix} -g_j^k \\ 1 \\ 0 \\ 0 \end{pmatrix} w_{m+2+j} - \sum_{i=1}^{\tilde{m}} \begin{pmatrix} -B_{i:}^T \\ 0 \\ 0 \\ 0 \end{pmatrix} w_{m+2+m+i} \\ &\stackrel{(4.61), (4.63)}{\downarrow} \sum_{j=1}^m \begin{pmatrix} \hat{g}_j^k w_j \\ 0 \\ 0 \\ w_j \end{pmatrix} + \frac{w_{m+1}}{t_{m+1}} \begin{pmatrix} 2\hat{G}^k d \\ 0 \\ 0 \\ -2 \end{pmatrix} + \frac{w_{m+2}}{t_{m+2}} \begin{pmatrix} 2\bar{W}_p^k d \\ 2 \\ -2 \\ 0 \end{pmatrix} \\ &\quad + \sum_{j=1}^m \begin{pmatrix} g_j^k w_{m+2+j} \\ -w_{m+2+j} \\ 0 \\ 0 \end{pmatrix} + \sum_{i=1}^{\tilde{m}} \begin{pmatrix} B_{i:}^T w_{m+2+m+i} \\ 0 \\ 0 \\ 0 \end{pmatrix}. \end{aligned} \quad (4.65)$$

The last but first equation of (4.65) reads

$$1 = 2 \frac{w_{m+2}}{t_{m+2}} \quad (4.66)$$

and hence we obtain for the last but second equation of (4.65)

$$0 = \frac{w_{m+2}}{t_{m+2}} \cdot 2 + \sum_{j=1}^m -w_{m+2+j} \stackrel{(4.66)}{=} 1 - \sum_{j=1}^m w_{m+2+j}$$

Since we have for the last equation of (4.65)

$$0 = \sum_{j=1}^m w_j - 2 \frac{w_{m+1}}{t_{m+1}} \iff 2 \frac{w_{m+1}}{t_{m+1}} = \sum_{j=1}^m w_j, \quad (4.67)$$

we obtain for the first n equations of (4.65)

$$\begin{aligned}
0 &= \sum_{j=1}^m \hat{g}_j^k w_j + \frac{w_{m+1}}{t_{m+1}} 2\hat{G}^k d + \frac{w_{m+2}}{t_{m+2}} 2\bar{W}^k d + \sum_{j=1}^m g_j^k w_{m+2+j} + \sum_{i=1}^{\bar{m}} B_{i:}^T w_{m+2+m+i} \\
&\stackrel{(4.66), (4.67)}{=} \sum_{j=1}^m \hat{g}_j^k w_j + \left(\sum_{j=1}^m w_j \right) \hat{G}^k d + \bar{W}_p^k d + \sum_{j=1}^m g_j^k w_{m+2+j} + \sum_{i=1}^{\bar{m}} B_{i:}^T w_{m+2+m+i} . \quad \square
\end{aligned}$$

Proposition 4.22. *The Lagrange multiplier $(\lambda, \mu, \nu) \in \mathbb{R}_{\geq 0}^{m+m+\bar{m}}$ of the QCQP (4.12) is exactly the part*

$$\begin{aligned}
\mu_j &= w_j && \text{for all } j = 1, \dots, m \\
\lambda_j &= w_{m+2+j} && \text{for all } j = 1, \dots, m \\
\nu_i &= w_{m+2+m+i} && \text{for all } i = 1, \dots, \bar{m} .
\end{aligned} \tag{4.68}$$

of the Lagrange multiplier of the SOCP (4.32).

Proof.

- Due to (4.36) resp. Proposition 4.18, the SOCP (4.32) is strictly primal and strictly dual feasible.
- Therefore (4.7) — under consideration of (4.23), (4.24), (4.25), (4.27) and (4.28) — yields the complementarity condition of the optimality conditions (4.13) of the QCQP (4.12).
- Inserting (4.68) into (4.58) yields the gradient-equations of the optimality conditions (4.13) of the QCQP (4.12). $\square$

4.4 Third party software

4.4.1 Comparison of the QCQP-solvers

All tests were performed on an Intel Pentium IV with 3 GHz and 1 GB RAM running Microsoft Windows XP and MATLAB R2010a.

We are comparing the time for solving 50 randomly generated problems of the following types

$$\begin{aligned}
\text{L(inear)} &:= \text{“QP obtained by setting } \tilde{G}_j^k = \tilde{G}^k = 0 \text{ in QCQP (4.11)”} \\
\text{D(ifferent)} &:= \text{“QCQP (4.11)”} \\
\text{E(qual)} &:= \text{“QCQP (4.11) with } \tilde{G}_j^k = \tilde{G}^k \text{”} \\
\text{R(educd)} &:= \text{“Reduced QCQP (4.12)”} ,
\end{aligned}$$

where we set $m := |J_k|$ and we choose $\bar{m} = 0$. For obtaining a first insight, how long the computation of the search direction will take, we compare the plots (based on the data from Table A.1 in Appendix A) of the median solving times (in milliseconds) for the following solvers

- MOSEK QCQP-solver (blue)
- MOSEK SOCP-solver (red)
- SEDUMI (yellow)
- SDPT3 (green),

where we use the colors to distinguish the results of the different solvers (since the only purpose of this subsection is to obtain a rough estimation of the solving times of the different types of search direction problems, we only tested these solvers here because the MATLAB tools CVX by GRANT & BOYD [29] resp. YALMIP by LÖFBERG [59] offer an excellent interface for easily generating the input data of the different search direction problems for these different solvers; the performance of `socp` resp. IPOPT is discussed in Remark 6.6 within the framework of using one of these two algorithms as the (QC)QP-solver in Algorithm 3.20):

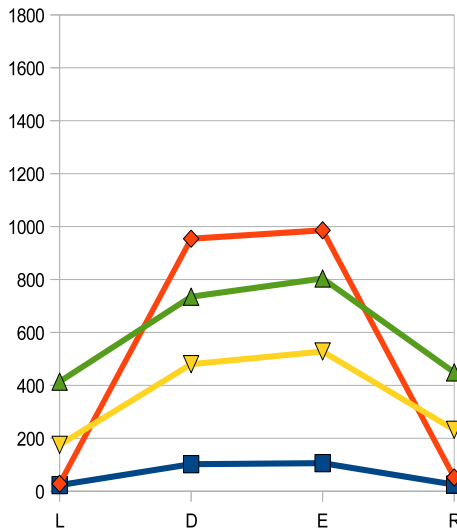


FIGURE 4.1: Median solving time for $n = 50$ and $m = 25$

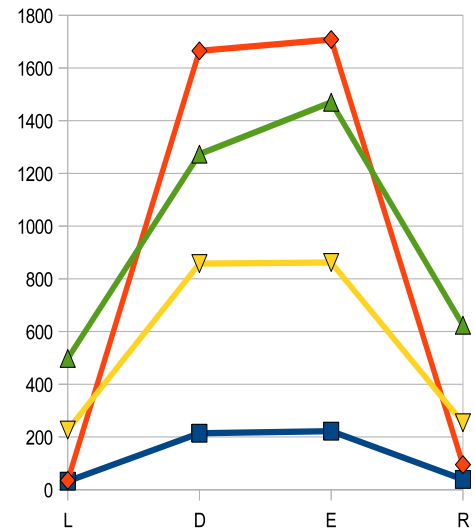


FIGURE 4.2: Median solving time for $n = 50$ and $m = 50$

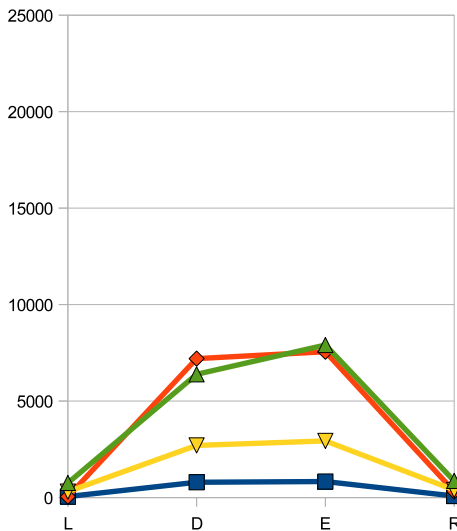


FIGURE 4.3: Median solving time for $n = 100$ and $m = 50$

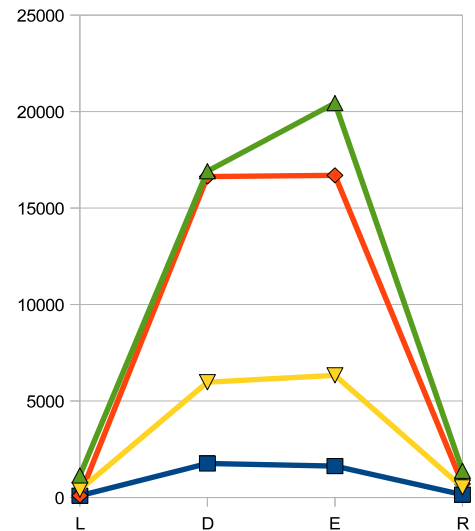


FIGURE 4.4: Median solving time for $n = 100$ and $m = 100$

By magnifying the results of

- L (dashed line)
- R (solid line),

where we use the two different line types for a better distinction of the comparisons only in Figure 4.5, from Figure 4.1, Figure 4.2, Figure 4.3 and Figure 4.4, we obtain the following plot:

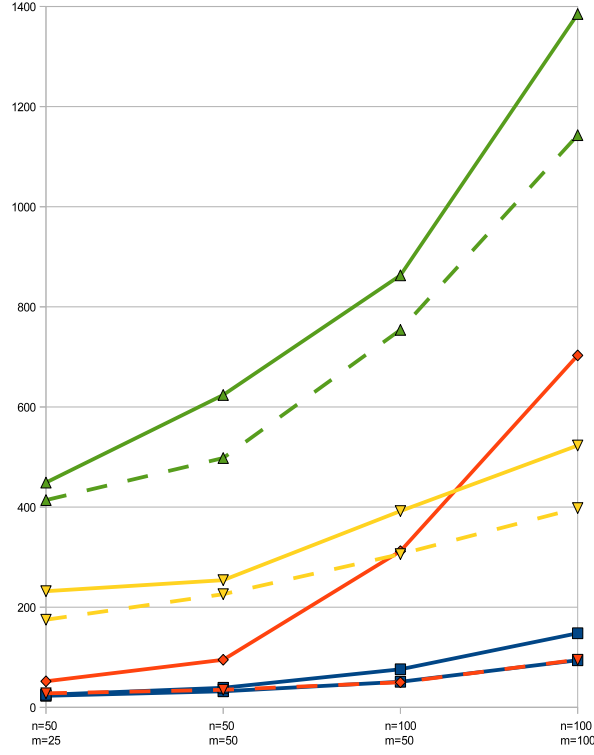


FIGURE 4.5: Magnification of the median solving time for L and R

Remark 4.23. Although ANDERSEN [2, p. 131, Section 7.2 and 7.2.1] recommends to rather use the MOSEK SOCP-solver than the MOSEK QCQP-solver for solving convex QCQPs, this does not coincide with the above results in which the MOSEK QCQP-solver has a significantly better performance than the MOSEK SOCP-solver for solving a QCQP of our shape.

The results from Figures 4.1–4.5 suggest that we will only test the MOSEK QCQP-solver on the reduced QCQP (4.12) in higher dimensions as this is the only combination that does not significantly exceed the shortest duration for solving the corresponding QP (which is always achieved by the MOSEK QP-solver). Therefore, we plot in Figure 4.6 (based on the data from Table A.2 in Appendix A) the

- minimal & maximal (lower and upper end of the vertical line)
- median (horizontal line)

solving times (in milliseconds) obtained by MOSEK for

- L (violet)
- R (orange)

(from $n = m = 400$ on, our computer started to swap and, consequently, we did not test higher

dimensional problems).

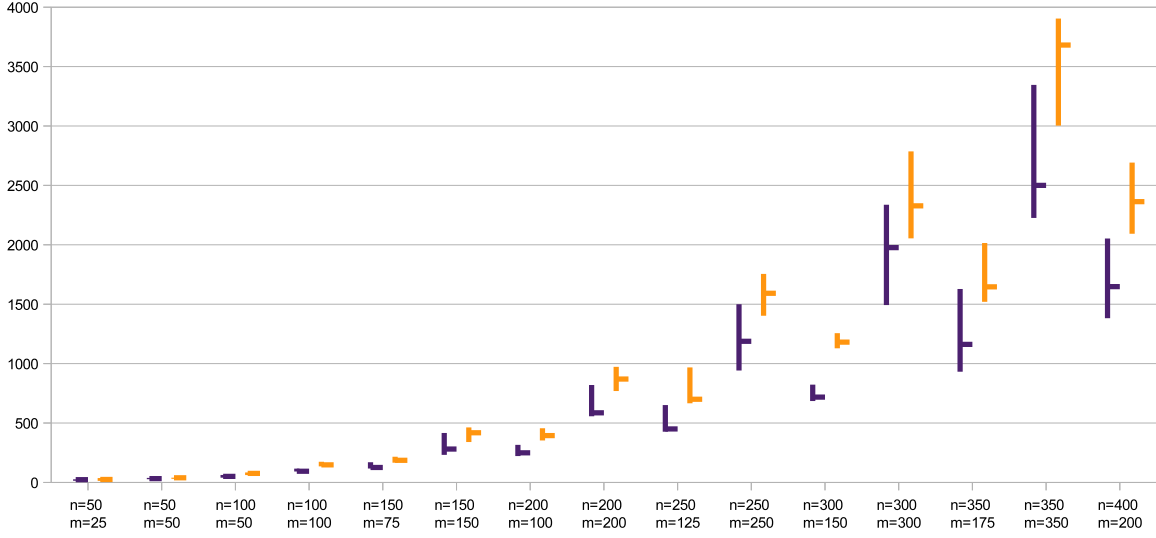


FIGURE 4.6: Minimal, median and maximal solving time

These results justify that we will mainly concentrate on the reduced QCQP (4.12) in the implementation as it is the only QCQP for which the solving time is competitive to that of the corresponding QP. Therefore, we will only give detailed implementation formulations for the reduced QCQP (4.12) in Subsections 4.4.2–4.4.4.

4.4.2 socp

Motivation 4.24. The interior-point solver `socp` by LOBO et al. [57], which is written in C, computes a solution of the primal SOCP (4.1) and its dual problem (4.2) under the assumption that a strictly primal feasible and a strictly dual feasible point are available. If no such points are available, there exists a MATLAB-extension of `socp` which can treat infeasible starting points by a big-M procedure (i.e. instead of the original SOCP, a sequence of SOCPs for which strictly feasible starting points are available easily is solved). Nevertheless, this feature is not implemented in C.

The input data for `socp` must be declared in the following form

$$A := \begin{pmatrix} A_1 \\ c_1^T \\ \vdots \\ A_L \\ c_L^T \end{pmatrix} \in \mathbb{R}^{\hat{L} \times n}, \quad b := \begin{pmatrix} b_1 \\ d_1 \\ \vdots \\ b_L \\ d_L \end{pmatrix} \in \mathbb{R}^{\hat{L}}, \quad z := \begin{pmatrix} z_1 \\ w_1 \\ \vdots \\ z_L \\ w_L \end{pmatrix} \in \mathbb{R}^{\hat{L}}, \quad (4.69)$$

where

$$\hat{L} := \sum_{i=1}^L N_i.$$

Proposition 4.25. *The input data of the SOCP (4.32) for `socp` is given by*

$$N = \begin{pmatrix} 1 \\ \frac{(n+1)+1}{(n+1)+1} \\ \frac{1}{1} \\ -1 \end{pmatrix}, \quad L = (m+1) + 1 + m + \hat{m},$$

and

$$f = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad A = \begin{pmatrix} -\hat{g}_1^k & 0 & 0 & -1 \\ \vdots & \vdots & \vdots & \vdots \\ -\hat{g}_m^k & 0 & 0 & -1 \\ \hline \sqrt{2}\hat{R}^k & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ \hline 0 & 0 & 0 & 1 \\ \hline \sqrt{2}\hat{R}_p^k & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ \hline 0 & -1 & 1 & 0 \\ \hline -\hat{g}_1^k & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ -\hat{g}_m^k & 1 & 0 & 0 \\ \hline -B_{1:} & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ -B_{\bar{m}:} & 0 & 0 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} -(F(x_k) - A_1^k) \\ \vdots \\ -(F(x_k) - A_m^k) \\ \hline 0 \\ 1 \\ \hline 1 \\ \hline 0 \\ 1 \\ \hline 1 \\ \hline \alpha_1^k \\ \vdots \\ \alpha_m^k \\ \hline b_1 - B_{1:}x_k \\ \vdots \\ b_{\bar{m}} - B_{\bar{m}:}x_k \end{pmatrix}.$$

Proof. Insert (4.32) into (4.69). □

Remark 4.26. To stabilize `socp` for our purpose, we add an additional termination criterion to `socp` as it is used in SEDUMI by PÓLIK [86], STURM [102], if SEDUMI cannot achieve the desired accuracy for the duality gap (the additional termination criterion is referred to as `pars.bigeps` in SEDUMI): If the current duality gap is smaller than `bigeps` := 10^{-2} and differs at most by 10^{-5} from the duality gap of the last iteration, then we accept the current point as a solution.

4.4.3 IPOPT

Motivation 4.27. IPOPT by WÄCHTER & BIEGLER [109] is an open source interior-point optimizer, which is written in C++, for solving high-dimensional non-linear optimization problems of the form (cf. WÄCHTER [108, p. 3, top])

$$\begin{aligned} \min_{x \in \mathbb{R}^n} & f(x) \\ \text{s.t. } & \underline{g} \leq g(x) \leq \bar{g} \\ & \underline{x} \leq x \leq \bar{x}, \end{aligned}$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ C^2 , $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ C^2 , $-\infty \leq \underline{g} \leq \bar{g} \leq \infty$ (in particular, the choice $\underline{g} = \bar{g}$ is allowed, i.e. IPOPT can handle equality constraints) and $-\infty \leq \underline{x} \leq \bar{x} \leq \infty$. The starting point for IPOPT can be feasible or infeasible.

Remark 4.28. In IPOPT first and second derivatives can be declared explicitly (in a special sparse format), which should have a positive impact for solving QCQPs, since for QCQPs all occurring functions coincide with the corresponding quadratic approximation.

Proposition 4.29. *If we identify*

$$\begin{pmatrix} d \\ \hat{v} \\ \hat{u} \end{pmatrix} \longleftrightarrow x = \begin{pmatrix} x_{1:n} \\ x_{n+1} \\ x_{n+2} \end{pmatrix} \in \mathbb{R}^{n+2}$$

in the QCQP (4.12), then the evaluation of the objective function and its gradient is specified by

$$\begin{aligned} \text{eval_f} &= x_{n+1} + \frac{1}{2} x_{1:n}^T \bar{W}_p^k x_{1:n} \\ \text{eval_grad_f} &= \begin{pmatrix} \bar{W}_p^k x_{1:n} \\ 1 \\ 0 \end{pmatrix} \in \mathbb{R}^{n+2} \end{aligned}$$

in IPOPT, the evaluation of the constraints and its Jacobian is specified by

$$\begin{aligned} \text{eval_g} &= \begin{pmatrix} Sx \\ s(x) \end{pmatrix} \in \mathbb{R}^{m+\hat{m}+m+1} \\ \text{eval_jac_g} &= \begin{pmatrix} S \\ x_{1:n}^T \bar{G}^k, 0, -1 \end{pmatrix} \in \mathbb{R}^{(m+\hat{m}+m+1) \times (n+2)}, \end{aligned}$$

where

$$\begin{aligned} S &:= \begin{pmatrix} g_1^k & -1 & 0 \\ \vdots & \vdots & \vdots \\ g_m^k & -1 & 0 \\ \hline B_{1:} & 0 & 0 \\ \vdots & \vdots & \vdots \\ B_{\hat{m}:} & 0 & 0 \\ \hline \hat{g}_1^k & 0 & 1 \\ \vdots & \vdots & \vdots \\ \hat{g}_m^k & 0 & 1 \end{pmatrix} \in \mathbb{R}^{(m+\hat{m}+m) \times (n+2)} \\ s(x) &:= -x_{n+2} + \frac{1}{2} x_{1:n}^T \bar{G}^k x_{1:n}, \end{aligned}$$

and the evaluation of the Hessian of the Lagrangian is specified by

$$\text{eval_h} = \begin{pmatrix} \sigma_f \bar{W}_p^k + \sigma_g \bar{G}^k & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{(n+2) \times (n+2)},$$

where $\sigma_f \geq 0$ (cf. WÄCHTER [108, p. 20]) and σ_g denotes the Lagrange multiplier of the $s(x)$ -constraint. Furthermore, the bounds are given by

$$\begin{aligned} [\underline{x}, \bar{x}] &= [-\infty, \infty] \\ \underline{g} &= -\infty \\ \bar{g} &= \begin{pmatrix} \alpha_1^k \\ \vdots \\ \alpha_m^k \\ \hline b_1 - B_{1:} x_k \\ \vdots \\ b_{\hat{m}} - B_{\hat{m}:} x_k \\ \hline A_1^k - F(x_k) \\ \vdots \\ A_m^k - F(x_k) \\ \hline 0 \end{pmatrix} \in \mathbb{R}^{m+\hat{m}+m+1}. \end{aligned}$$

Proof. Clear (Compute the gradient and the Hessian of the quadratic functions). □

Remark 4.30. For `eval_jac_g` and `eval_h` only the lower triangular part, which is entered in the IPOPT-sparse-format, is relevant.

4.4.4 MOSEK

Motivation 4.31. MOSEK by ANDERSEN et al. [3] is a commercial solver (that is available for academic institutions via a liberal trial license), which is written in C. It can handle both feasible and infeasible starting points, though declaring a starting point is not necessary as MOSEK can automatically create a (good) starting point. MOSEK is in particular specialized for solving QCQPs and SOCPs:

- MOSEK solves (in particular) QCQPs of the form (cf. ANDERSEN [2, p. 37, bottom, ff])

$$\begin{aligned}
 & \min_{x \in \mathbb{R}^n} \frac{1}{2} x^T Q^o x + c^T x \\
 & \text{s.t. } \frac{1}{2} x^T Q^k x + a_k^T x \leq u_k^c \quad \text{for all } k = 0, \dots, m-1 \\
 & \quad l^x \leq x \leq u^x,
 \end{aligned} \tag{4.70}$$

where $Q^o, Q^k \in \mathcal{S}(n)$ are positive semidefinite, from which only the lower triangular part is entered in the MOSEK sparse-format, and $a_k \in \mathbb{R}^n$. For entering the problem data of (4.70) in C, one has to obey the following facts:

- Use `MSK_putcj` for declaring the linear terms $c_k^T x$ of the objective function.
 - Use `MSK_putavec(task, MSK_ACC_CON, ...)` for declaring the linear terms $a_k^T x$ of the constraints (for row-wise input and this is what we want; for column-wise input we would have to use `MSK_putavec(task, MSK_ACC_VAR, ...)`).
 - Use `MSK_putqobj` for declaring the matrix Q^o of the quadratic term of the objective function.
 - Use `MSK_putqconk` for declaring the matrices Q^k of the quadratic terms of the constraints.
 - Use `MSK_putbound` to declare the bounds for the variables (Note: A bound on a variable has to be declared explicitly, even if the bound is $\pm\infty$).
- MOSEK solves (in particular) SOCPs of the form (cf. ANDERSEN [2, p. 49])

$$\begin{aligned}
 & \min_{x \in \mathbb{R}^n} c^T x \\
 & \text{s.t. } l^c \leq Ax \leq u^c \\
 & \quad l^x \leq x \leq u^x \\
 & \quad x \in \mathcal{C},
 \end{aligned} \tag{4.71}$$

where

$$\begin{aligned}
 \mathcal{C} &:= \prod_{t=0}^{p-1} C_t \\
 C_t &:= \{y \in \mathbb{R}^{n_t} : y_{n_t} \geq |y_{1:(n_t-1)}|_2 \quad \text{for all } t = 0, \dots, p-1\},
 \end{aligned}$$

and a variable may occur in one cone at most. E.g., the conic constraint

$$x_4 \geq |(x_0, x_2)|_2$$

is declared in C by

```
MSKidx csub [3];          /* Indices of those components */
csub[0]=4;                /* of x which are in C_t */
csub[1]=0;
csub[2]=2;
MSK_appendcone(task,
                MSK_CT_QUAD,
                0.0,      /* For future versions of MOSEK */
                3,        /* Length of C_t */
                csub);
```

Proposition 4.32. *The formulation of the QCQP (4.12) for MOSEK (cf. (4.70)) reads*

$$\begin{aligned} \min & \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}^T \begin{pmatrix} d \\ \hat{v} \\ \hat{u} \end{pmatrix} + \frac{1}{2} \begin{pmatrix} d \\ \hat{v} \\ \hat{u} \end{pmatrix}^T \begin{pmatrix} \bar{W}_p^k & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} d \\ \hat{v} \\ \hat{u} \end{pmatrix} \\ \text{s.t.} & \begin{pmatrix} \hat{g}_1^{kT} & 0 & 1 \\ \vdots & \vdots & \vdots \\ \hat{g}_m^{kT} & 0 & 1 \\ \hline g_1^{kT} & -1 & 0 \\ \vdots & \vdots & \vdots \\ g_m^{kT} & -1 & 0 \\ \hline 0 & 0 & -1 \\ \hline B_{1:} & 0 & 0 \\ \vdots & \vdots & \vdots \\ B_{\bar{m}:} & 0 & 0 \end{pmatrix} \begin{pmatrix} d \\ \hat{v} \\ \hat{u} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \frac{1}{2} \begin{pmatrix} d \\ \hat{v} \\ \hat{u} \end{pmatrix}^T \begin{pmatrix} \hat{G}^k & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} d \\ \hat{v} \\ \hat{u} \end{pmatrix} \\ 0 \end{pmatrix} \leq \begin{pmatrix} A_1^k - F(x_k) \\ \vdots \\ A_m^k - F(x_k) \\ \hline \alpha_1^k \\ \vdots \\ \alpha_m^k \\ \hline 0 \\ \hline b_1 - B_{1:}x_k \\ \vdots \\ b_{\bar{m}} - B_{\bar{m}:}x_k \end{pmatrix}. \end{aligned}$$

Proof. Apply trivial, algebraic transformations to the QCQP (4.12). □

Proposition 4.33. *Let $y_1, y_3 \in \mathbb{R}^{n+1}$ and $y_2, y_4 \in \mathbb{R}$ and identity*

$$z \longleftrightarrow \begin{pmatrix} d^T & \hat{v} & \hat{u} & y_1^T & y_2 & y_3^T & y_4 \end{pmatrix}^T,$$

then the formulation of the SOCP (4.32) for MOSEK (cf. (4.71)) reads

$$\begin{aligned} \min & \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix} z \\ \text{s.t.} & \begin{pmatrix} \hat{g}_j^k & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} z \leq A_j^k - F(x_k) \\ & \begin{pmatrix} g_j^k & -1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} z \leq \alpha_j^k \\ & \begin{pmatrix} B_{i:} & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} z \leq b_i - B_{i:}x_k \\ & \begin{pmatrix} \sqrt{2}\hat{R}^k & 0 & 0 & 0 & -I & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \\ \hline \sqrt{2}\hat{R}_p^k & 0 & 0 & 0 & 0 & 0 & -I & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & -1 \end{pmatrix} z = \begin{pmatrix} 0 \\ -1 \\ -1 \\ 0 \\ -1 \\ -1 \end{pmatrix} \\ & |z_{(n+3)+(1:(n+1))}|_2 \leq z_{(n+3)+(n+2)} \\ & |z_{(n+3)+(n+2)+(1:(n+1))}|_2 \leq z_{(n+3)+2(n+2)}. \end{aligned}$$

Proof. Since a variable is only allowed to occur in one cone at most in MOSEK, we introduce new variables $y_1 \in \mathbb{R}^{n+1}$, $y_2 \in \mathbb{R}$ for the conic constraint with respect to the constraint-approximation in the SOCP (4.32)

$$\begin{aligned}
& \left| \begin{pmatrix} \sqrt{2}\hat{R}^k & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right|_2 \leq \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}^T \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + 1 \\
& \iff \begin{cases} |y_1|_2 \leq y_2 \\ y_1 = \begin{pmatrix} \sqrt{2}\hat{R}^k & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ y_2 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}^T \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + 1 \end{cases} \\
& \iff \begin{cases} |y_1|_2 \leq y_2 \\ \begin{pmatrix} \sqrt{2}\hat{R}^k & 0 & 0 & 0 & -I & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \\ y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix} \end{cases}
\end{aligned}$$

and we introduce new variables $y_3 \in \mathbb{R}^{n+1}$, $y_4 \in \mathbb{R}$ for the constraint with respect to the objective function in the SOCP (4.32)

$$\begin{aligned}
& \left| \begin{pmatrix} \sqrt{2}R_p^k & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right|_2 \leq \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}^T \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + 1 \\
& \iff \begin{cases} |y_3|_2 \leq y_4 \\ y_3 = \begin{pmatrix} \sqrt{2}R_p^k & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \end{pmatrix} \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ y_4 = \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}^T \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \end{pmatrix} + 1 \end{cases} \\
& \iff \begin{cases} |y_3|_2 \leq y_4 \\ \begin{pmatrix} \sqrt{2}R_p^k & 0 & 0 & 0 & 0 & 0 & -I & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} d \\ \hat{v} \\ v \\ \hat{u} \\ y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix} \end{cases} .
\end{aligned}$$

The shape of the objective function and the remaining constraints follows by trivial algebraic transformations of the corresponding constraints of the SOCP (4.32) and adding zero-vectors of the right size for y_i for $i = 1, \dots, 4$. $\square$

Chapter 5

Certificates of infeasibility via nonsmooth optimization

After summarizing the most basic facts of interval analysis, we construct the certificate of infeasibility in this chapter. Furthermore, we discuss how a nonsmooth solver can use this certificate to obtain an exclusion box in a CSP, we discuss the resulting optimization problems to which we apply the nonsmooth solver for finding an exclusion box and how we obtain a starting point for them.

5.1 Interval arithmetic

We recall some basic facts on interval arithmetic from, e.g., NEUMAIER [82].

Definition 5.1.

1. We denote a box (also called interval vector) by

$$\mathbf{x} := [\underline{x}, \bar{x}] .$$

2. We denote the set of all boxes by

$$\mathbb{IR} := \{\mathbf{x} : \mathbf{x} = [\underline{x}, \bar{x}], \underline{x} \leq \bar{x}\} .$$

3. For $S \subseteq \mathbb{R}$ bounded, we define the hull of S by

$$\square S := [\inf S, \sup S] .$$

4. We extend the arithmetic operations and functions $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ to boxes by defining

$$\begin{aligned} \mathbf{x} \circ \mathbf{y} &:= \square\{x \circ y : x \in \mathbf{x}, y \in \mathbf{y}\} \\ \varphi(\mathbf{x}) &:= \square\{\varphi(x) : x \in \mathbf{x}\} . \end{aligned} \tag{5.1}$$

Example 5.2. For $\mathbf{x} \in \mathbb{R}^n$, $c_1 \in \mathbb{R}$ and $c_2 \geq 0$ we have

$$\inf \sum_{k=1}^n \mathbf{x}_k = \sum_{k=1}^n \inf \mathbf{x}_k \quad (5.2)$$

$$\inf (c_1 \mathbf{x}) = \begin{cases} c_1 \cdot \inf \mathbf{x} & \text{for } c_2 \geq 0 \\ c_1 \cdot \sup \mathbf{x} & \text{for } c_2 < 0 \end{cases} \quad (5.3)$$

$$\sup (c_2 \mathbf{x}) = c_2 \cdot \sup \mathbf{x} . \quad (5.4)$$

The fundamental theorem of interval arithmetic states

Theorem 5.3. Let $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$ and let Φ be an expression of φ which is a composition of arithmetic operations and elementary functions, then

$$[\inf_{\mathbf{x} \in \mathbf{x}} \varphi(\mathbf{x}), \sup_{\mathbf{x} \in \mathbf{x}} \varphi(\mathbf{x})] \subseteq \Phi(\mathbf{x}) . \quad (5.5)$$

Another important, well-known theorem is

Theorem 5.4. Let $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be Lipschitz continuous, then for all $x, z \in \mathbb{R}^n$ there exists a slope matrix $\varphi[z, x] \in \mathbb{R}^{n \times n}$ with

$$\varphi(x) - \varphi(z) = \varphi[z, x](x - z) . \quad (5.6)$$

5.2 Certificate of infeasibility

Definition 5.5. Let $k = 1, \dots, m$, $c_k \in \mathbb{R}^n$ and $C_k : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$. We define $F_k : \mathbb{R}^n \rightarrow \mathbb{R}$ and $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ by

$$F_k(x) := c_k^T x + x^T C_k(x) x \quad (5.7)$$

$$F(x) := (F_k(x))_{k=1}^m . \quad (5.8)$$

Motivation 5.6. We want to construct a certificate of infeasibility for the CSP

$$\begin{aligned} F(x) &\in \mathbf{F} \\ x &\in \mathbf{x} . \end{aligned} \quad (5.9)$$

Remark 5.7. The definition of F_k in (5.7) is sufficient for our purpose, since: If we consider $x_0 \in \mathbb{R}^n$ and the CSP

$$\begin{aligned} \hat{F}_k(x) &\in \hat{\mathbf{F}}_k \text{ for } k = 1, \dots, m \\ x &\in \hat{\mathbf{x}} \end{aligned}$$

with

$$\hat{F}_k(x) := \gamma_k + c_k^T(x - x_0) + (x - x_0)^T \hat{C}_k(x)(x - x_0) , \quad (5.10)$$

we obtain

$$\begin{aligned}
\hat{F}_k(x) \in \hat{\mathbf{F}}_k \quad \wedge \quad x \in \hat{\mathbf{x}} &\stackrel{(5.10)}{\iff} \gamma_k + c_k^T(x - x_0) + (x - x_0)^T \hat{C}_k(x)(x - x_0) \in \hat{\mathbf{F}}_k \quad \wedge \quad x \in \hat{\mathbf{x}} \\
&\iff c_k^T y + y^T \underbrace{\hat{C}_k(y + x_0)}_{=: C_k(y)} y \in \hat{\mathbf{F}}_k - \gamma_k =: \mathbf{F}_k \quad \wedge \quad y \in \hat{\mathbf{x}} - x_0 =: \mathbf{x} \\
&\stackrel{(5.7)}{\iff} F_k(y) \in \mathbf{F}_k \quad \wedge \quad y \in \mathbf{x} .
\end{aligned}$$

Proposition 5.8. For fixed $y \in \mathbb{R}^m$ let the restriction $C(y, \cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ of $C : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$

$$C(y, x) := \sum_{k=1}^m C_k(x) y_k \quad (5.11)$$

be Lipschitz continuous, then there exists a slope matrix $C^i(y, [z, x]) \in \mathbb{R}^{n \times n}$ for (the columns) $C_{:i}(y, \cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ (for $i = 1, \dots, n$) such that we have for all $x, z \in \mathbb{R}^n$

$$C_{:i}(y, x) - C_{:i}(y, z) = C^i(y, [z, x])(x - z) . \quad (5.12)$$

Proof. Apply (5.6) to (the columns) $C_{:i}(y, \cdot) \in \mathbb{R}^n$ of $C(y, \cdot) \in \mathbb{R}^{n \times n}$. □

Proposition 5.9. We have

$$\begin{aligned}
\mathbb{R}_{\text{triu}}^{n \times n} &\cong \mathbb{R}^{n_1} \\
\mathbb{R}_{\text{striu}}^{n \times n} &\cong \mathbb{R}^{n_0} \\
n_0 + n_1 &= n^2 ,
\end{aligned} \quad (5.13)$$

where

$$\mathbb{R}_{\text{striu}}^{n \times n} := \{A \in \mathbb{R}^{n \times n} : A_{ij} = 0 \text{ for } i \geq j\}$$

denotes the linear space of strictly upper triangular $n \times n$ -matrices and

$$n_1 := \frac{n(n+1)}{2} , \quad n_0 := \frac{(n-1)n}{2} .$$

Proof. Clear (use the definition of $\mathbb{R}_{\text{triu}}^{n \times n}$ (cf. (4.22)) and $\mathbb{R}_{\text{striu}}^{n \times n}$). □

Definition 5.10. Let $R \in \mathbb{R}_{\text{triu}}^{n \times n}$ and $S \in \mathbb{R}_{\text{striu}}^{n \times n}$. We define $B, D, E : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$, $A : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^{n_1} \times \mathbb{R}^{n_0} \rightarrow \mathbb{R}^{n \times n}$ and $c : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ by

$$B(y, x, z) := \sum_{j=1}^n z_j \begin{pmatrix} C^1(y, [z, x])_{j:} \\ \vdots \\ C^m(y, [z, x])_{j:} \end{pmatrix} \quad (5.14)$$

$$D(y, x, z) := \sum_{i=1}^n z_i C^i(y, [z, x]) \quad (5.15)$$

$$E(y, x, z) := B(y, x, z) + D(y, x, z) \quad (5.16)$$

$$A(y, x, z, R, S) := C(y, x) + R^T R + S^T - S + E(y, x, z) \quad (5.17)$$

$$c(y, x, z) := \sum_{k=1}^m c_k y_k + (C(y, z) + C(y, z)^T + D(y, x, z)^T) z . \quad (5.18)$$

In the following we make a few technical preparations.

Proposition 5.11. *Let $U \in \mathbb{R}^{n \times n}$ be skew-symmetric (i.e. $U^T = -U$), then we have for all $x \in \mathbb{R}^n$*

$$x^T U x = 0 . \quad (5.19)$$

Proof. An easy calculation yields the desired result. $\square$

Proposition 5.12. *Let $U^i \in \mathbb{R}^{n \times n}$ for $i = 1, \dots, n$, then we have for all $x, z \in \mathbb{R}^n$*

$$x^T \left((U^i(x-z))_{i=1}^n \right) z = \left(x^T \left(\sum_{i=1}^n z_i U^i \right) \right) (x-z) \quad (5.20)$$

$$z^T (U^i(x-z))_{i=1}^n = \left(\left(\sum_{j=1}^n z_j \begin{pmatrix} U_{j:}^1 \\ \vdots \\ U_{j:}^n \end{pmatrix} \right) (x-z) \right)^T . \quad (5.21)$$

Proof.

- The following calculation yields (5.20)

$$\begin{aligned} x^T \left((U^i(x-z))_{i=1}^n \right) z &= x^T \left(U^1(x-z) \mid \dots \mid U^n(x-z) \right) z \\ &= x^T \left(\sum_{i=1}^n (U^i(x-z)) z_i \right) \\ &= \left(x^T \left(\sum_{i=1}^n z_i U^i \right) \right) (x-z) . \end{aligned}$$

- For obtaining (5.21), we calculate

$$\begin{aligned} z^T (U^i(x-z))_{i=1}^n &= z^T \left(U^1(x-z) \mid \dots \mid U^n(x-z) \right) \\ &= \sum_{j=1}^n z_j \left(U_{j:}^1(x-z) \mid \dots \mid U_{j:}^n(x-z) \right) \\ &= \sum_{j=1}^n z_j \begin{pmatrix} U_{j:}^1(x-z) \\ \vdots \\ U_{j:}^n(x-z) \end{pmatrix}^T \\ &= \left(\left(\sum_{j=1}^n z_j \begin{pmatrix} U_{j:}^1 \\ \vdots \\ U_{j:}^n \end{pmatrix} \right) (x-z) \right)^T . \end{aligned} \quad \square$$

Proposition 5.13.

- Denote

$$\hat{c}(y, z) := \sum_{k=1}^m c_k y_k + (C(y, z) + C(y, z)^T) z \quad (5.22)$$

$$h(y, x, z) := \sum_{k=1}^m y_k \left(x^T (C_k(x) - C_k(z)) z \right) , \quad (5.23)$$

then

$$\begin{aligned} c(y, x, z)^T (x-z) + (x-z)^T (C(y, x) + E(y, x, z)) (x-z) \\ = \hat{c}(y, z)^T (x-z) + (x-z)^T (C(y, x) + B(y, x, z)) (x-z) + h(y, x, z) . \end{aligned} \quad (5.24)$$

- Let $U \in \mathbb{R}^{n \times n}$. Then the following equalities hold

$$x^T U x = (U x)^T x \quad (5.25)$$

$$(B(y, x, z)(x - z))^T = z^T (C(y, x) - C(y, z)) \quad (5.26)$$

$$(x - z)^T C(y, x) + z^T (C(y, x) - C(y, z)) = x^T C(y, x) - z^T C(y, z) \quad (5.27)$$

$$\hat{c}(y, z)^T + x^T C(y, x) - z^T C(y, z) = \sum_{k=1}^m c_k^T y_k + z^T C(y, z)^T + x^T C(y, x) \quad (5.28)$$

$$x^T C(y, x) + z^T C(y, z)^T = \sum_{k=1}^m y_k (x^T C_k(x) + z^T C_k(z)^T) . \quad (5.29)$$

- In addition we have

$$F_k(x) - F_k(z) = (c_k^T + x^T C_k(x) + z^T C_k(z)^T)(x - z) + x^T (C_k(x) - C_k(z))z . \quad (5.30)$$

Proof.

- Since

$$\begin{aligned} c(y, x, z) &\stackrel{(5.18)}{=} \sum_{k=1}^m c_k y_k + (C(y, z) + C(y, z)^T + D(y, x, z)^T)z \\ &\stackrel{(5.22)}{=} \hat{c}(y, z) + D(y, x, z)^T z \end{aligned} \quad (5.31)$$

and

$$\begin{aligned} &(x - z)^T (C(y, x) + E(y, x, z))(x - z) \\ &\stackrel{(5.16)}{=} (x - z)^T (C(y, x) + B(y, x, z) + D(y, x, z))(x - z) \\ &= (x - z)^T (C(y, x) + B(y, x, z))(x - z) + x^T D(y, x, z)(x - z) - z^T D(y, x, z)(x - z) \end{aligned} \quad (5.32)$$

as well as

$$\begin{aligned} h(y, x, z) &\stackrel{(5.23)}{=} \sum_{k=1}^m y_k (x^T (C_k(x) - C_k(z))z) \\ &\stackrel{(5.11)}{=} x^T (C(y, x) - C(y, z))z \\ &\stackrel{(5.12)}{=} x^T \left((C^i(y, [z, x])(x - z))_{i=1}^n \right) z \\ &\stackrel{(5.20)}{=} \left(x^T \left(\sum_{i=1}^n z_i C^i(y, [z, x]) \right) \right) (x - z) \\ &\stackrel{(5.15)}{=} (x^T D(y, x, z))(x - z) , \end{aligned} \quad (5.33)$$

we obtain (5.24)

$$\begin{aligned} &c(y, x, z)^T (x - z) + (x - z)^T (C(y, x) + E(y, x, z))(x - z) \\ &\stackrel{(5.31)}{=} (\hat{c}(y, z) + D(y, x, z)^T z)^T (x - z) + (x - z)^T (C(y, x) + E(y, x, z))(x - z) \\ &\stackrel{(5.32)}{=} \hat{c}(y, z)^T (x - z) + z^T D(y, x, z)(x - z) + (x - z)^T (C(y, x) + B(y, x, z))(x - z) \\ &\quad + x^T D(y, x, z)(x - z) - z^T D(y, x, z)(x - z) \\ &\stackrel{(5.33)}{=} \hat{c}(y, z)^T (x - z) + (x - z)^T (C(y, x) + B(y, x, z))(x - z) + h(y, x, z) . \end{aligned}$$

- (5.25) holds due to

$$x^T U x = (x^T U x)^T = (U x)^T x .$$

- The following calculation yields (5.26)

$$\begin{aligned} z^T (C(y, x) - C(y, z)) &= z^T \left((C_{:i}(y, x) - C_{:i}(y, z))_{i=1}^n \right) \\ &\stackrel{(5.12)}{=} z^T \left((C^i(y, [z, x])(x - z))_{i=1}^n \right) \\ &\stackrel{(5.21)}{=} \left(\left(\sum_{j=1}^n z_j \begin{pmatrix} C^1(y, [z, x])_{j:} \\ \vdots \\ C^n(y, [z, x])_{j:} \end{pmatrix} \right) (x - z) \right)^T \\ &\stackrel{(5.14)}{=} (B(y, x, z)(x - z))^T . \end{aligned}$$

- (5.27) holds due to

$$\begin{aligned} (x - z)^T C(y, x) + z^T (C(y, x) - C(y, z)) &= x^T C(y, x) - z^T C(y, x) + z^T C(y, x) - z^T C(y, z) \\ &= x^T C(y, x) - z^T C(y, z) . \end{aligned}$$

- (5.28) holds due to

$$\begin{aligned} \hat{c}(y, z)^T + x^T C(y, x) - z^T C(y, z) \\ &\stackrel{(5.22)}{=} \left(\sum_{k=1}^m c_k y_k + (C(y, z) + C(y, z)^T) z \right)^T + x^T C(y, x) - z^T C(y, z) \\ &= \sum_{k=1}^m c_k^T y_k + z^T C(y, z)^T + x^T C(y, x) . \end{aligned}$$

- (5.29) holds due to

$$\begin{aligned} x^T C(y, x) + z^T C(y, z)^T &\stackrel{(5.11)}{=} x^T \left(\sum_{k=1}^m C_k(x) y_k \right) + z^T \left(\sum_{k=1}^m C_k(z) y_k \right)^T \\ &= \sum_{k=1}^m y_k (x^T C_k(x) + z^T C_k(z)^T) . \end{aligned}$$

- Since

$$\begin{aligned} x^T C_k(x) x - z^T C_k(z) z &= x^T C_k(x) (x - z) + x^T (C_k(x) - C_k(z)) z + (x - z)^T C_k(z) z \\ &= x^T C_k(x) (x - z) + ((x - z)^T C_k(z) z)^T + x^T (C_k(x) - C_k(z)) z \\ &= x^T C_k(x) (x - z) + z^T C_k(z)^T (x - z) + x^T (C_k(x) - C_k(z)) z \\ &= (x^T C_k(x) + z^T C_k(z)^T) (x - z) + x^T (C_k(x) - C_k(z)) z , \end{aligned} \quad (5.34)$$

we obtain (5.30)

$$\begin{aligned} F_k(x) - F_k(z) &\stackrel{(5.7)}{=} c_k^T x + x^T C_k(x) x - c_k^T z - z^T C_k(z) z \\ &= c_k^T (x - z) + x^T C_k(x) x - z^T C_k(z) z \\ &\stackrel{(5.34)}{=} (c_k^T + x^T C_k(x) + z^T C_k(z)^T) (x - z) + x^T (C_k(x) - C_k(z)) z . \quad \square \end{aligned}$$

Definition 5.14. We define $Y : \mathbb{R}^m \times \mathbb{R}^n \longrightarrow \mathbb{R}$ and $Z, f : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n_1} \times \mathbb{R}^{n_0} \times \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ by

$$Y(y, z) := \inf y^T (\mathbf{F} - F(z)) \quad (5.35)$$

$$Z(y, z, R, S, \underline{x}, \bar{x}) := \sup (c(y, \mathbf{x}, z)^T + (\mathbf{x} - z)^T A(y, \mathbf{x}, z, R, S))(\mathbf{x} - z) \quad (5.36)$$

$$f(y, z, R, S, \underline{x}, \bar{x}) := \frac{Z(y, z, R, S, \underline{x}, \bar{x}) - \max(0, Y(y, z))}{T(y, z, R, S, \underline{x}, \bar{x})}, \quad (5.37)$$

where

$$\mathbf{x} = [\underline{x}, \bar{x}] \in \mathbb{IR}^n, \quad \mathbf{F} = [\underline{F}, \bar{F}] \in \mathbb{IR}^m$$

and $T : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n_1} \times \mathbb{R}^{n_0} \times \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}_{>0}$ is positive, continuous and differentiable almost everywhere.

Remark 5.15. f from (5.37) depends on

$$N := m + n + n_1 + n_0 + n + n \stackrel{(5.13)}{=} m + 3n + n^2$$

variables and is not differentiable everywhere and not convex (in general).

Proposition 5.16. *It holds for all $z \in \mathbf{x}$*

$$Z(y, z, R, S, \underline{x}, \bar{x}) \geq 0. \quad (5.38)$$

Proof. (5.38) holds, since we have for all $z \in \mathbf{x}$

$$\begin{aligned} Z(y, z, R, S, \underline{x}, \bar{x}) &\stackrel{(5.36)}{=} \sup (c(y, \mathbf{x}, z)^T + (\mathbf{x} - z)^T A(y, \mathbf{x}, z, R, S))(\mathbf{x} - z) \\ &\stackrel{(5.5)}{\geq} \sup_{\mathbf{x} \in \mathbf{x}} (c(y, \mathbf{x}, z)^T + (\mathbf{x} - z)^T A(y, \mathbf{x}, z, R, S))(\mathbf{x} - z) \\ &\stackrel{z \in \mathbf{x}}{\downarrow} \geq \sup_{\mathbf{x} \in \{z\}} (c(y, \mathbf{x}, z)^T + (\mathbf{x} - z)^T A(y, \mathbf{x}, z, R, S))(\mathbf{x} - z) \\ &= 0. \end{aligned} \quad \square$$

Now we state the main theorem for our application.

Theorem 5.17. *If there exist $y \in \mathbb{R}^m$, $\underline{x} \leq z \leq \bar{x} \in \mathbb{R}^n$, $R \in \mathbb{R}^{n \times n}$ and $S \in \mathbb{R}^{n \times n}$ with*

$$f(y, z, R, S, \underline{x}, \bar{x}) < 0,$$

then for all $x \in \mathbf{x}$ there exists $k \in \{1, \dots, m\}$ with

$$F_k(x) \notin \mathbf{F}_k,$$

i.e. there is no $x \in \mathbf{x}$ with $F(x) \in \mathbf{F}$, i.e. there is no feasible point.

Proof. (By contradiction) Let the assumption of the theorem be satisfied and suppose that there exists $\hat{x} \in \mathbf{x} := [\underline{x}, \bar{x}]$ with $F_k(\hat{x}) \in \mathbf{F}_k$ for all $k \in \{1, \dots, m\}$, i.e. with $F(\hat{x}) \in \mathbf{F}$.

By assumption there exist $y \in \mathbb{R}^m$, $z \in \mathbf{x} \subseteq \mathbb{R}^n$, $R \in \mathbb{R}^{n \times n}$ and $S \in \mathbb{R}^{n \times n}$ with

$$\begin{aligned}
 & f(y, z, R, S, \underline{x}, \bar{x}) < 0 \\
 \stackrel{(5.37), T(y, z, R, S, \underline{x}, \bar{x}) > 0}{\iff} & Z(y, z, R, S, \underline{x}, \bar{x}) - \max(0, Y(y, z)) < 0 \\
 \iff & Z(y, z, R, S, \underline{x}, \bar{x}) < 0 \vee Z(y, z, R, S, \underline{x}, \bar{x}) < Y(y, z) \\
 \stackrel{z \in \mathbf{x}, (5.38)}{\iff} & Z(y, z, R, S, \underline{x}, \bar{x}) < Y(y, z) .
 \end{aligned} \tag{5.39}$$

Because of the skew-symmetry of $S^T - S$, we obtain

$$\begin{aligned}
 \inf_{\tilde{F} \in \mathbf{F}} y^T (\tilde{F} - F(z)) & \stackrel{(5.1)}{=} \inf y^T (\mathbf{F} - F(z)) \stackrel{(5.35)}{=} Y(y, z) \stackrel{(5.39)}{>} Z(y, z, R, S) \\
 & \stackrel{(5.36)}{=} \sup (c(y, \mathbf{x}, z)^T + (\mathbf{x} - z)^T A(y, \mathbf{x}, z, R, S))(\mathbf{x} - z) \\
 & \stackrel{(5.5)}{\geq} \sup_{x \in \mathbf{x}} (c(y, x, z)^T + (x - z)^T A(y, x, z, R, S))(x - z) \\
 & \stackrel{(5.17)}{=} \sup_{x \in \mathbf{x}} (c(y, x, z)^T + (x - z)^T (C(y, x) + R^T R + S^T - S + E(y, x, z)))(x - z) \\
 & \stackrel{(5.19)}{=} \sup_{x \in \mathbf{x}} (c(y, x, z)^T + (x - z)^T (C(y, x) + R^T R + E(y, x, z)))(x - z) ,
 \end{aligned} \tag{5.40}$$

which yields

$$Z(y, z, R, S) \stackrel{(5.40)}{<} \inf_{\tilde{F} \in \mathbf{F}} y^T (\tilde{F} - F(z)) \leq y^T (\tilde{F} - F(z)) \quad \text{for all } x \in \mathbf{x}, \tilde{F} \in \mathbf{F} \tag{5.41}$$

and

$$\begin{aligned}
 Z(y, z, R, S) & \stackrel{(5.40)}{\geq} \sup_{x \in \mathbf{x}} (c(y, x, z)^T + (x - z)^T (C(y, x) + R^T R + E(y, x, z)))(x - z) \\
 & \geq c(y, x, z)^T (x - z) + (x - z)^T (C(y, x) + E(y, x, z))(x - z) \quad \text{for all } x \in \mathbf{x}, \tilde{F} \in \mathbf{F} .
 \end{aligned} \tag{5.42}$$

Hence, we obtain for all $x \in \mathbf{x}$ and for all $\tilde{F} \in \mathbf{F}$

$$\begin{aligned}
 y^T (\tilde{F} - F(z)) & \stackrel{(5.41), (5.42)}{>} c(y, x, z)^T (x - z) + (x - z)^T (C(y, x) + E(y, x, z))(x - z) \\
 & \stackrel{(5.24)}{=} \hat{c}(y, z)^T (x - z) + (x - z)^T (C(y, x) + B(y, x, z))(x - z) + h(y, x, z) \\
 & \stackrel{(5.25)}{=} \hat{c}(y, z)^T (x - z) + (x - z)^T C(y, x)(x - z) + (B(y, x, z)(x - z))^T (x - z) + h(y, x, z) \\
 & \stackrel{(5.26)}{=} \hat{c}(y, z)^T (x - z) + (x - z)^T C(y, x)(x - z) + z^T (C(y, x) - C(y, z))(x - z) + h(y, x, z) \\
 & \stackrel{(5.27)}{=} (\hat{c}(y, z)^T + x^T C(y, x) - z^T C(y, z))(x - z) + h(y, x, z) \\
 & \stackrel{(5.28)}{=} \left(\sum_{k=1}^m y_k c_k^T + x^T C(y, x) + z^T C(y, z)^T \right) (x - z) + h(y, x, z) \\
 & \stackrel{(5.29)}{=} \left(\sum_{k=1}^m y_k c_k^T + \sum_{k=1}^m y_k (x^T C_k(x) + z^T C_k(z)^T) \right) (x - z) + h(y, x, z) \\
 & \stackrel{(5.23)}{=} \sum_{k=1}^m y_k \left((c_k^T + x^T C_k(x) + z^T C_k(z)) (x - z) \right) + \sum_{k=1}^m y_k \left(x^T (C_k(x) - C_k(z)) z \right) \\
 & \stackrel{(5.30), (5.8)}{\stackrel{\downarrow}{=}} y^T (F(x) - F(z)) .
 \end{aligned} \tag{5.43}$$

Choosing $x = \hat{x} \in \mathbf{x}$ and $\tilde{F} = F(\hat{x}) \in \mathbf{F}$ in (5.43) yields $y^T F(\hat{x}) > y^T \tilde{F}$, which is a contradiction. $\square$

The following proposition gives in particular a hint how the y -component of a starting point should be chosen (cf. (5.97), (5.98) and (5.99)).

Proposition 5.18.

- We have for all $y \in \mathbb{R}^m$ and $z \in \mathbb{R}^n$

$$Y(y, z) = \sum_{k=1}^m \begin{cases} y_k (\underline{F}_k - F_k(z)) & \text{for } y_k \geq 0 \\ y_k (\overline{F}_k - F_k(z)) & \text{for } y_k < 0 \end{cases} \quad (5.44)$$

- Let $I, J \subseteq \{1, \dots, m\}$ satisfy

$$\begin{aligned} I \neq \emptyset \quad \vee \quad J \neq \emptyset \\ \underline{F}_i = -\infty \wedge y_i > 0 \quad \text{for all } i \in I \\ \overline{F}_j = \infty \wedge y_j < 0 \quad \text{for all } j \in J, \end{aligned}$$

then we have for all $z \in \mathbb{R}^n$

$$Y(y, z) = -\infty. \quad (5.45)$$

Proof.

- (5.44) holds because of

$$\begin{aligned} Y(y, z) &\stackrel{(5.35)}{=} \inf y^T (\mathbf{F} - F(z)) \\ &= \inf \sum_{k=1}^m y_k (\mathbf{F}_k - F_k(z)) \\ &\stackrel{(5.2)}{=} \sum_{k=1}^m \inf y_k (\mathbf{F}_k - F_k(z)) \\ &= \sum_{k=1}^m \inf y_k ([\underline{F}_k, \overline{F}_k] - F_k(z)) \\ &\stackrel{(5.3)}{=} \sum_{k=1}^m \begin{cases} y_k (\underline{F}_k - F_k(z)) & \text{for } y_k \geq 0 \\ y_k (\overline{F}_k - F_k(z)) & \text{for } y_k < 0 \end{cases} \end{aligned}$$

- Without loss of generality consider $\underline{F}_1 = -\infty \wedge y_1 > 0$ and $\overline{F}_2 = \infty \wedge y_2 < 0$, then we obtain (5.45)

$$\begin{aligned} Y(y, z) &\stackrel{(5.44)}{=} \sum_{k=1}^m \begin{cases} y_k (\underline{F}_k - F_k(z)) & \text{for } y_k \geq 0 \\ y_k (\overline{F}_k - F_k(z)) & \text{for } y_k < 0 \end{cases} \\ &= \underbrace{y_1}_{>0} \underbrace{(\underline{F}_1 - F_1(z))}_{=-\infty} + \underbrace{y_2}_{<0} \underbrace{(\overline{F}_2 - F_2(z))}_{=\infty} + \sum_{k=2}^m \begin{cases} y_k (\underline{F}_k - F_k(z)) & \text{for } y_k \geq 0 \\ y_k (\overline{F}_k - F_k(z)) & \text{for } y_k < 0 \end{cases} \\ &= -\infty - \infty + \sum_{k=2}^m \begin{cases} y_k (\underline{F}_k - F_k(z)) & \text{for } y_k \geq 0 \\ y_k (\overline{F}_k - F_k(z)) & \text{for } y_k < 0 \end{cases} \\ &= -\infty. \end{aligned} \quad \square$$

5.3 Properties of the certificate for quadratic F

Now we consider the special case of f with quadratic F , which turns out to be a variant of the original, unpublished certificate of infeasibility proposed by Arnold Neumaier.

Proposition 5.19. *If $C_k : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n}$ is constant*

$$C_k(x) := C_k \quad (5.46)$$

with $C_k \in \mathbb{R}^{n \times n}$ and $k = 1, \dots, m$, then the slopes $C^i(y, [z, x])$ vanish and the terms occurring in (5.37) are of the form

$$F_k(x) = c_k^T x + x^T C_k x \quad (5.47)$$

$$C(y) := \sum_{k=1}^m C_k y_k \quad (5.48)$$

$$= C(y, x) \quad \text{for all } x \in \mathbf{x}$$

$$c(y, z) := \sum_{k=1}^m c_k y_k + (C(y) + C(y)^T)z \quad (5.49)$$

$$= c(y, x, z) \quad \text{for all } x \in \mathbf{x}$$

$$A(y, R, S) := C(y) + R^T R + S^T - S \quad (5.50)$$

$$= A(y, x, z, R, S) \quad \text{for all } x, z \in \mathbf{x}$$

$$Z(y, z, R, S, \underline{x}, \bar{x}) = \sup (c(y, z)^T + (\underline{x} - z)^T A(y, R, S))(\underline{x} - z) . \quad (5.51)$$

Proof.

- (5.47) holds because of

$$F_k(x) \stackrel{(5.7)}{=} c_k^T x + x^T C_k(x) x \stackrel{(5.46)}{=} c_k^T x + x^T C_k x .$$

- (5.48) holds because of

$$C(y, x) \stackrel{(5.11)}{=} \sum_{k=1}^m C_k(x) y_k \stackrel{(5.46)}{=} \sum_{k=1}^m C_k y_k \stackrel{(5.48)}{=} C(y) . \quad (5.52)$$

- Since

$$C^i(y, [z, x])(x - z) \stackrel{(5.12)}{=} C_{:,i}(y, x) - C_{:,i}(y, z) \stackrel{(5.52)}{=} C_{:,i}(y) - C_{:,i}(y) = 0$$

must hold for all $x, z \in \mathbb{R}^n$ and for all $i = 1, \dots, n$, we obtain for all $i = 1, \dots, n$

$$C^i(y, [z, x]) = 0 . \quad (5.53)$$

- We have

$$B(y, x, z) \stackrel{(5.14)}{=} \sum_{j=1}^n z_j \begin{pmatrix} C^1(y, [z, x])_{j,:} \\ \vdots \\ C^n(y, [z, x])_{j,:} \end{pmatrix} \stackrel{(5.53)}{=} 0 . \quad (5.54)$$

- We have

$$D(y, x, z) \stackrel{(5.15)}{=} \sum_{i=1}^n z_i C^i(y, [z, x]) \stackrel{(5.53)}{=} 0 . \quad (5.55)$$

- We have

$$E(y, x, z) \stackrel{(5.16)}{=} B(y, x, z) + D(y, x, z) \stackrel{(5.54), (5.55)}{\downarrow} 0 . \quad (5.56)$$

- (5.49) holds because of

$$\begin{aligned} c(y, x, z) &\stackrel{(5.18)}{=} \sum_{k=1}^m c_k y_k + (C(y, z) + C(y, z)^T + D(y, x, z)^T) z \\ &\stackrel{(5.52), (5.55)}{\downarrow} \sum_{k=1}^m c_k y_k + (C(y) + C(y)^T) z \\ &\stackrel{(5.49)}{=} c(y, z) . \end{aligned} \quad (5.57)$$

- (5.50) holds because of

$$\begin{aligned} A(y, x, z, R, S) &\stackrel{(5.17)}{=} C(y, x) + R^T R + S^T - S + E(y, x, z) \\ &\stackrel{(5.52), (5.56)}{\downarrow} C(y) + R^T R + S^T - S \\ &\stackrel{(5.50)}{=} A(y, R, S) . \end{aligned} \quad (5.58)$$

- (5.51) holds because of

$$\begin{aligned} Z(y, z, R, S, \underline{x}, \bar{x}) &\stackrel{(5.36)}{=} \sup (c(y, \underline{x}, z)^T + (\underline{x} - z)^T A(y, \underline{x}, z, R, S)) (\underline{x} - z) \\ &\stackrel{(5.57), (5.58)}{\downarrow} \sup (c(y, z)^T + (\underline{x} - z)^T A(y, R, S)) (\underline{x} - z) . \end{aligned} \quad \square$$

Proposition 5.20. *Let (5.46) be satisfied. Then we have for all $p \in \mathbb{R}^m$ and $\alpha \in \mathbb{R}$*

$$C(y + \alpha p) = C(y) + \alpha C(p) \quad (5.59)$$

$$c(y + \alpha p, z) = c(y, z) + \alpha c(p, z) \quad (5.60)$$

and furthermore we have for all $\kappa \geq 0$

$$A(\kappa^2 y, \kappa R, \kappa^2 S) = \kappa^2 A(y, R, S) . \quad (5.61)$$

Proof.

- (5.59) holds because of

$$C(y + \alpha p) \stackrel{(5.48)}{=} \sum_{k=1}^m C_k(y_k + \alpha_k p_k) \stackrel{(5.48)}{=} C(y) + \alpha C(p) .$$

- (5.60) holds because of

$$\begin{aligned} c(y + \alpha p) &\stackrel{(5.49)}{=} \sum_{k=1}^m c_k(y_k + \alpha p_k) + (C(y + \alpha p) + C(y + \alpha p)^T) z \\ &\stackrel{(5.59)}{=} \sum_{k=1}^m c_k(y_k + \alpha p_k) + (C(y) + \alpha C(p) + C(y)^T + \alpha C(p)^T) z \\ &\stackrel{(5.49)}{=} c(y, z) + \alpha c(p, z) . \end{aligned}$$

- (5.61) holds because of

$$\begin{aligned}
A(\kappa^2 y, \kappa R, \kappa^2 S) &\stackrel{(5.50)}{=} C(\kappa^2 y) + (\kappa R)^T (\kappa R) + (\kappa^2 S)^T - (\kappa^2 S) \\
&\stackrel{(5.59)}{=} \kappa^2 C(y) + \kappa^2 R^T R + \kappa^2 (S^T - S) \\
&\stackrel{(5.50)}{=} \kappa^2 A(y, R, S) .
\end{aligned} \quad \square$$

Proposition 5.21. *Let (5.46) be satisfied. If the positive function $T : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}^{n_1} \times \mathbb{R}^{n_0} \times \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}_{>0}$ satisfies the (partial) homogeneity condition*

$$T(\kappa^2 y, z, \kappa R, \kappa^2 S, \underline{x}, \bar{x}) = \kappa^2 T(y, z, R, S, \underline{x}, \bar{x}) \quad (5.62)$$

for all $\kappa > 0$, $y \in \mathbb{R}^m$, $z \in [\underline{x}, \bar{x}] \in \mathbb{R}^n$, $R \in \mathbb{R}_{\text{triu}}^{n \times n}$ and $S \in \mathbb{R}_{\text{striu}}^{n \times n}$, then the certificate f from (5.37) is (partially) homogeneous

$$f(\kappa^2 y, z, \kappa R, \kappa^2 S, \underline{x}, \bar{x}) = f(y, z, R, S, \underline{x}, \bar{x}) . \quad (5.63)$$

Proof. Since (5.46) is satisfied by assumption, the statements of Proposition 5.19 hold. We calculate

$$\begin{aligned}
c(\kappa^2 y, z)^T + (\mathbf{x} - z)^T A(\kappa^2 y, \kappa R, \kappa^2 S) &\stackrel{(5.60)}{=} \kappa^2 c(y, z)^T + (\mathbf{x} - z)^T A(\kappa^2 y, \kappa R, \kappa^2 S) \\
&\stackrel{(5.61)}{=} \kappa^2 (c(y, z)^T + (\mathbf{x} - z)^T A(y, R, S))
\end{aligned} \quad (5.64)$$

and therefore it follows

$$\begin{aligned}
Z(\kappa^2 y, z, \kappa R, \kappa^2 S, \underline{x}, \bar{x}) &\stackrel{(5.51)}{=} \sup \left((c(\kappa^2 y, z)^T + (\mathbf{x} - z)^T A(\kappa^2 y, \kappa R, \kappa^2 S)) (\mathbf{x} - z) \right) \\
&\stackrel{(5.64)}{=} \sup \left(\left(\kappa^2 (c(y, z)^T + (\mathbf{x} - z)^T A(y, R, S)) \right) (\mathbf{x} - z) \right) \\
&\stackrel{(5.4)}{=} \kappa^2 \sup \left((c(y, z)^T + (\mathbf{x} - z)^T A(y, R, S)) (\mathbf{x} - z) \right) \\
&\stackrel{(5.51)}{=} \kappa^2 Z(y, z, R, S, \underline{x}, \bar{x}) .
\end{aligned} \quad (5.65)$$

Furthermore, we have

$$Y(\kappa^2 y, z) \stackrel{(5.35)}{=} \inf (\kappa y^2)^T (\mathbf{F} - F(z)) \stackrel{(5.3)}{=} \kappa^2 \inf y^T (\mathbf{F} - F(z)) \stackrel{(5.35)}{=} \kappa^2 Y(y, z) \quad (5.66)$$

and hence we obtain

$$\max(0, Y(\kappa^2 y, z)) \stackrel{(5.66)}{=} \max(0, \kappa^2 Y(y, z)) \stackrel{\kappa > 0}{=} \kappa^2 \max(0, Y(y, z)) . \quad (5.67)$$

Consequently, (5.63) holds due to

$$\begin{aligned}
f(\kappa^2 y, z, \kappa R, \kappa^2 S, \underline{x}, \bar{x}) &\stackrel{(5.37)}{=} \frac{Z(\kappa^2 y, z, \kappa R, \kappa^2 S, \underline{x}, \bar{x}) - \max(0, Y(\kappa^2 y, z))}{T(\kappa^2 y, z, \kappa R, \kappa^2 S, \underline{x}, \bar{x})} \\
&\stackrel{(5.65), (5.67)}{=} \frac{\kappa^2 Z(y, z, R, S, \underline{x}, \bar{x}) - \kappa^2 \max(0, Y(y, z))}{T(\kappa^2 y, z, \kappa R, \kappa^2 S, \underline{x}, \bar{x})} \\
&\stackrel{(5.62)}{=} \frac{\kappa^2 Z(y, z, R, S, \underline{x}, \bar{x}) - \kappa^2 \max(0, Y(y, z))}{\kappa^2 T(y, z, R, S, \underline{x}, \bar{x})} \\
&\stackrel{(5.37)}{=} f(y, z, R, S, \underline{x}, \bar{x}) .
\end{aligned} \quad \square$$

Remark 5.22.

- The intention of (5.63) is to reduce the scale dependence of the unbounded variables y , R and S of f .
- If we go through the proof of Proposition 5.21 again, we notice that we use the scaling property (5.61) of A for showing (5.64). From the proof of (5.61) we notice that this proof only holds, if y , R and S are treated as variables and none of them is treated as a constant (since factoring κ^2 out of a constant, yields an additional factor κ^{-2} to the constant). Nevertheless, if one of the variables y , R resp. S is treated as a constant and we set the corresponding value to $y = 0$, $R = 0$ resp. $S = 0$, then the proof still holds.

Example 5.23.

1. Consider the variables R and S as constants and set $R = S = 0$. Then

$$T(y, z, R, S, \underline{x}, \bar{x}) := 1 \quad (5.68)$$

does not satisfy (5.62), while

$$T(y, z, R, S, \underline{x}, \bar{x}) := |y|_2 \quad (5.69)$$

does (cf. Remark 5.22). Note that T from (5.69) violates the requirement of positivity, as demanded in Definition 5.14, for $y = 0$, and hence in this case f is only defined outside the zero set of T , which is given by

$$\{(y, z, R, S, \underline{x}, \bar{x}) : T(y, z, R, S, \underline{x}, \bar{x}) = 0\} .$$

Nevertheless, since the zero set of T from (5.69) has measure zero, it is numerically very unlikely to end up at a point of this zero set and therefore we will also consider this choice of T due to the important fact of the reduction of the scale dependence of f as mentioned in Remark 5.22 (also cf. Idea 5.24 and Example 5.27).

2. Choose

$$n = 2, \quad m = 2, \quad c_1 = \begin{pmatrix} 1 \\ -3 \end{pmatrix}, \quad c_2 = \begin{pmatrix} 4 \\ 2 \end{pmatrix}, \quad C_1 = \begin{pmatrix} 2 & 0 \\ 3 & 4 \end{pmatrix}, \quad C_2 = \begin{pmatrix} -1 & 0 \\ -2 & 7 \end{pmatrix},$$

which yields

$$\begin{aligned} F_1(x) &\stackrel{(5.7)}{=} 2x_1^2 + x_1 + 3x_1x_2 - 3x_2 + 4x_2^2 \\ F_2(x) &\stackrel{(5.7)}{=} -x_1^2 + 4x_1 - 2x_1x_2 + 2x_2 + 7x_2^2, \end{aligned}$$

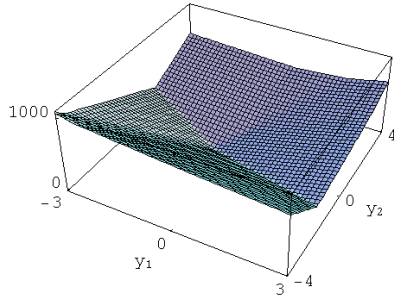
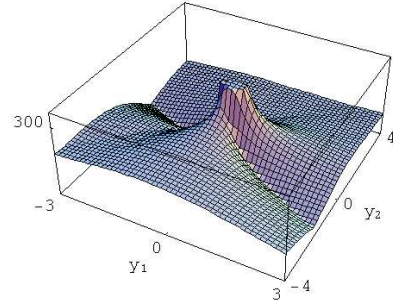
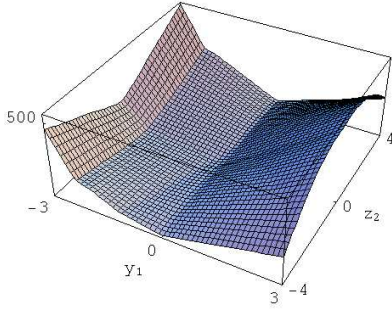
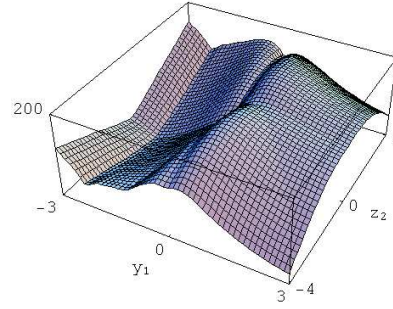
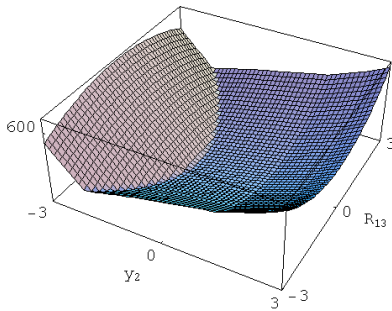
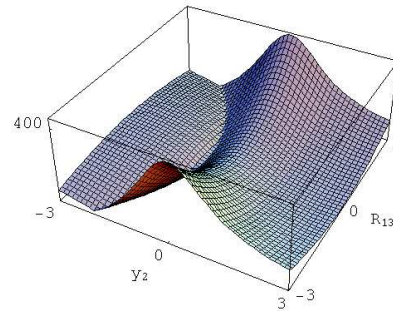
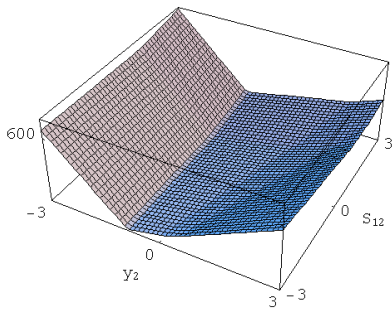
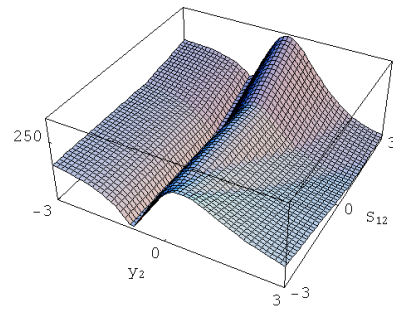
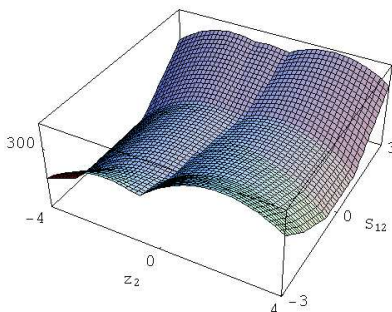
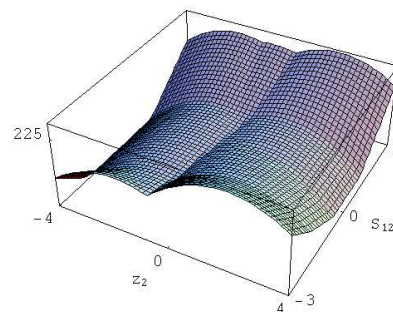
and

$$\mathbf{x} = [-3, 3] \times [-4, 4], \quad \mathbf{F} = [-1, 7] \times [-2, 0],$$

then we can illustrate certificate f from (5.37) with different T from (5.68) and (5.69) given by

$$T_1(y, z) := 1, \quad T_2(y, z) := |y|_2$$

by the following plots

FIGURE 5.1: $f(y_1, y_2)$ for T_1 FIGURE 5.2: $f(y_1, y_2)$ for T_2 FIGURE 5.3: $f(y_1, z_2)$ for T_1 FIGURE 5.4: $f(y_1, z_2)$ for T_2 FIGURE 5.5: $f(y_2, R_{13})$ for T_1 FIGURE 5.6: $f(y_2, R_{13})$ for T_2 FIGURE 5.7: $f(y_2, S_{12})$ for T_1 FIGURE 5.8: $f(y_2, S_{12})$ for T_2 FIGURE 5.9: $f(z_2, S_{12})$ for T_1 FIGURE 5.10: $f(z_2, S_{12})$ for T_2

5.4 Exclusion boxes for constraint satisfaction problems

Idea 5.24. Now we explain in detail how to use Theorem 5.17 for finding exclusion boxes for the CSP (5.9) by using a nonsmooth solver (Note: The certificate f from (5.37) is not differentiable everywhere due to Remark 5.15), although the convergence theory of many solvers (cf., e.g., Assumption 3.1 for Algorithm 3.20) requires that all occurring functions are defined on the whole $\mathbb{R}^N$, which might be violated for certain choices of T (cf. Example 5.23):

1. If we apply a solver for linearly constrained nonsmooth optimization to

$$\begin{aligned}
 & \min f(y, z, R, S, u, v) \\
 & \text{s.t. } u, v \in \mathbf{x} \\
 & \quad u + r \leq v \\
 & \quad z \in [u, v] \\
 & \quad y \in \mathbb{R}^m, u, v, z \in \mathbb{R}^n, R \in \mathbb{R}_{\text{triu}}^{n \times n}, S \in \mathbb{R}_{\text{striu}}^{n \times n}
 \end{aligned} \tag{5.70}$$

with a fixed

$$r \in [0, \bar{x} - \underline{x}] \tag{5.71}$$

and if there occurs a function value smaller than zero (during the optimization process), then there is no feasible point in $[u, v]$ according to Theorem 5.17 and consequently we can reduce the box $\mathbf{x}$ to the set

$$\mathbf{x} \setminus [u, v]$$

in the CSP (5.9).

If $[u, v] = \mathbf{x}$ (i.e. u and v are fixed and therefore no variables), then we can reduce the box $\mathbf{x}$ to

$$\mathbf{x} \setminus [u, v] = \mathbf{x} \setminus \mathbf{x} = \emptyset,$$

i.e. the reduction of a box to the empty set is equivalent to removing the box.

The constant r determines the size of the box $[u, v]$, which should be excluded: The closer r is to 0, the smaller the box $[u, v]$ can become (if $r = 0$, $[u, v]$ can become thin, what we want to prevent, since we want to remove a preferably large box $[u, v]$ out of $\mathbf{x}$, as then the remaining set $\mathbf{x} \setminus [u, v]$ is preferably small).

2. If $\mathbf{y} \subseteq \mathbf{x}$ and we remove $\mathbf{y}$ from $\mathbf{x}$, then the remaining set

$$\mathbf{x} \setminus \mathbf{y}$$

is not closed. Nevertheless, if we just remove $\mathbf{y}^\circ \subset \mathbf{y}$, then the remaining set

$$\mathbf{x} \setminus \mathbf{y}^\circ \supset \mathbf{x} \setminus \mathbf{y}$$

is closed (i.e. we remove a smaller box and therefore the remaining set is a bit larger, since it contains the boundary of $\mathbf{y}$). Furthermore, the set $\mathbf{x} \setminus \mathbf{y}$ can be represented as a union of at most $2n$ n -dimensional boxes, i.e. in particular the number of boxes obtained by this splitting process is linear in n .

3. We make the assumption that the certificate of infeasibility from (5.37) of the box $[\hat{u}, \hat{v}]$ satisfies

$$f(\hat{y}, \hat{z}, \hat{R}, \hat{S}, \hat{u}, \hat{v}) =: \hat{\delta} < 0 ,$$

i.e. $[\hat{u}, \hat{v}]$ is an exclusion box according to Theorem 5.17 (e.g., $\hat{\delta} = -10^{-2}$). For $\delta \in [\hat{\delta}, 0)$ (e.g., $\delta = -10^{-4}$) and a box $\mathbf{x}$ with $[\hat{u}, \hat{v}] \subseteq \mathbf{x}$, we can try to apply a solver for nonlinearly constrained nonsmooth optimization to

$$\begin{aligned}
 & \max -b(y, z, R, S, u, v) \\
 & \text{s.t. } f(y, z, R, S, u, v) \leq \delta \\
 & \quad u, v \in \mathbf{x} \\
 & \quad u \leq \hat{u} \\
 & \quad \hat{v} \leq v \\
 & \quad z \in [u, v] \\
 & \quad y \in \mathbb{R}^m, R \in \mathbb{R}_{\text{triu}}^{n \times n}, S \in \mathbb{R}_{\text{striu}}^{n \times n} \\
 & \stackrel{(2.3)}{=} -\min b(y, z, R, S, u, v) \\
 & \text{s.t. } f(y, z, R, S, u, v) - \delta \leq 0 \\
 & \quad u, v \in \mathbf{x} \\
 & \quad u \leq \hat{u} \\
 & \quad \hat{v} \leq v \\
 & \quad z \in [u, v] \\
 & \quad y \in \mathbb{R}^m, R \in \mathbb{R}_{\text{triu}}^{n \times n}, S \in \mathbb{R}_{\text{striu}}^{n \times n}
 \end{aligned} \tag{5.72}$$

to enlarge the exclusion box $[u, v]$ in $\mathbf{x}$, where $b : \mathbb{R}^N \rightarrow \mathbb{R}$ is a measure for the box $[u, v]$ in the following sense:

- If $b : \mathbb{R}^N \rightarrow \mathbb{R}_{\leq 0}$, then the following conditions must hold:
 - If $[u, v]$ is small, then $b(., u, v)$ is close to 0.
 - If $[u, v]$ is large, then $b(., u, v)$ is negative and large.

This means: The larger the box $[u, v]$ is, the more negative $b(., u, v)$ must be. For examples of this type of box measure cf. (5.75).

- If $b : \mathbb{R}^N \rightarrow \mathbb{R}_{\geq 0}$, then the following condition must hold: If $[u, v] \subseteq \mathbf{x}$ is close to $\mathbf{x}$, then $b(., u, v)$ is close to 0. For examples of this type of box measure cf. (5.76).

Notice that in (5.72) the constraint containing f is nondifferentiable (and possibly the objective function, too).

Remark 5.25.

- In opposite to (5.70), where the linear constraint

$$u + r \leq v$$

occurs, we use in (5.72) the bound constraints

$$u \leq \hat{u} , \quad \hat{v} \leq v .$$

- If during the optimization a point $z \in \mathbf{x}$ is found with $F(z) \in \mathbf{F}$, then we have found a feasible point and therefore we can stop the optimization, since then we cannot verify infeasibility for the box $\mathbf{x}$.
- We make the following observations for the global optimization problem

$$\begin{aligned}
& \min_x F_{\text{obj}}(x) \\
& \text{s.t. } F_i(x) \in \mathbf{F}_i \quad \text{for all } i = 2, \dots, m \\
& x \in \mathbf{x},
\end{aligned} \tag{5.73}$$

where $F_{\text{obj}} : \mathbb{R}^n \rightarrow \mathbb{R}$:

- The certificate f from (5.37) can be used for finding exclusion boxes in the global optimization problem (5.73) with an arbitrary objective function F_{obj} , since the certificate f only depends on the constraint data F , $\mathbf{F}$ and $\mathbf{x}$ (cf. the CSP (5.9)) and since a solution of an optimization problem is necessarily feasible.
- We consider the global optimization problem (5.73) with

$$\begin{aligned}
F_{\text{obj}}(x) &:= F_1(x) \\
&\stackrel{(5.7)}{=} c_1^T x + x^T C_1(x) x
\end{aligned}$$

and we denote the current lowest known function value of this optimization problem by $F_{\text{obj}}^{\text{cur}}$. Now, if we can find a box $[u, v] \subseteq x$ for which the certificate f from (5.37) with

$$\mathbf{F}_1 := [-\infty, F_{\text{obj}}^{\text{cur}}] \tag{5.74}$$

has a negative value, then Theorem 5.17 implies that for all $x \in [u, v]$

$$\begin{aligned}
& \exists k \in \{1, \dots, m\} : F_k(x) \notin \mathbf{F}_k \\
& \iff F_1(x) \notin \mathbf{F}_1 \quad \text{or} \quad \exists i \in \{2, \dots, m\} : F_i(x) \notin \mathbf{F}_i \\
& \stackrel{(5.74)}{\iff} F_1(x) < -\infty \quad \text{or} \quad F_{\text{obj}}^{\text{cur}} < F_1(x) \quad \text{or} \quad \exists i \in \{2, \dots, m\} : F_i(x) \notin \mathbf{F}_i \\
& \iff F_{\text{obj}}^{\text{cur}} < F_1(x) \quad \text{or} \quad \exists i \in \{2, \dots, m\} : F_i(x) \notin \mathbf{F}_i,
\end{aligned}$$

i.e. any point in the box $[u, v]$ has an objective function value which is higher than the current lowest known function value $F_{\text{obj}}^{\text{cur}}$ or is infeasible. Consequently, the box $[u, v]$ cannot contain a feasible point with function value lower equal $F_{\text{obj}}^{\text{cur}}$, and hence the box $[u, v]$ cannot contain a global minimizer of the global optimization problem (5.73). Therefore we can exclude the box $[u, v]$ from further consideration.

Example 5.26. For measuring the box $[u, v]$ with $b : \mathbb{R}^N \rightarrow \mathbb{R}_{\leq 0}$, we can use any negative p -norm with $p \in [1, \infty]$ as well as variants of them

$$\begin{aligned}
b_-^1(y, z, R, S, u, v) &:= -|v - u|_1 \\
b_-^2(y, z, R, S, u, v) &:= -\frac{1}{2}|v - u|_2^2 \\
b_-^\infty(y, z, R, S, u, v) &:= -|v - u|_\infty.
\end{aligned} \tag{5.75}$$

For measuring the box $[u, v]$ with $b : \mathbb{R}^N \rightarrow \mathbb{R}_{\geq 0}$, we can use any p -norm with $p \in [1, \infty]$ as well as variants of them

$$\begin{aligned} b_+^1(y, z, R, S, u, v) &:= \left| \begin{pmatrix} u-x \\ v-x \end{pmatrix} \right|_1 \\ b_+^2(y, z, R, S, u, v) &:= \frac{1}{2} \left| \begin{pmatrix} u-x \\ v-x \end{pmatrix} \right|_2^2 \\ b_+^\infty(y, z, R, S, u, v) &:= \left| \begin{pmatrix} u-x \\ v-x \end{pmatrix} \right|_\infty . \end{aligned} \quad (5.76)$$

b_- is concave, while b_+ is convex. b_- can be used for unbounded $\mathbf{x}$, while this is not possible for b_+ . b^2 is smooth, while b^1 and b^∞ are not differentiable.

b^1 has an equal growing rate for all components. The growing rate of b^2 depends on $\text{sgn}(\frac{1}{2}|\cdot|_2^2 - 1)$. b^∞ already grows, if the absolute value of the largest components grows.

Example 5.27. Choose

$$n = m = 1, \quad R = S = 0, \quad c_1 = 1, \quad C_1 = \frac{1}{2},$$

which yields

$$F(x) \stackrel{(5.47)}{=} \frac{1}{2}x^2 + x,$$

as well as

$$\mathbf{x} = [-1, 2]$$

and consider two CSPs (5.9) with

$$\mathbf{F}^1 = [-2, 1], \quad \mathbf{F}^2 = [-2, -1]$$

which yield the following two graphics

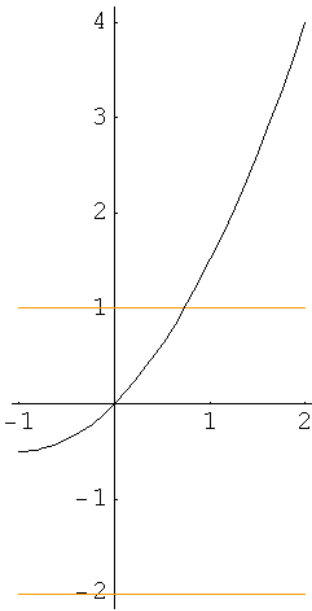


FIGURE 5.11: F for $\mathbf{F}^1$

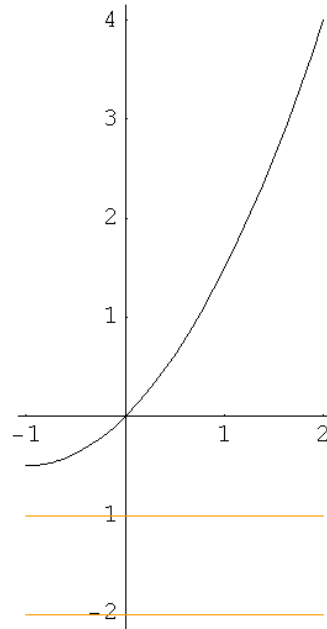


FIGURE 5.12: F for $\mathbf{F}^2$

from which we can see that the CSP has feasible points for $\mathbf{F}^1$, while it is infeasible for $\mathbf{F}^2$. The

corresponding certificates f from (5.37), where we only consider the variables $y \in \mathbb{R}$ and $z \in \mathbf{x}$ as well as different T from (5.68) and (5.69) given by

$$T_1(y, z) := 1, \quad T_2(y, z) := |y|,$$

and where we denote the function value of a local minimizer of the optimization problem (5.70) by $\hat{f}$, can be illustrated by the following plots

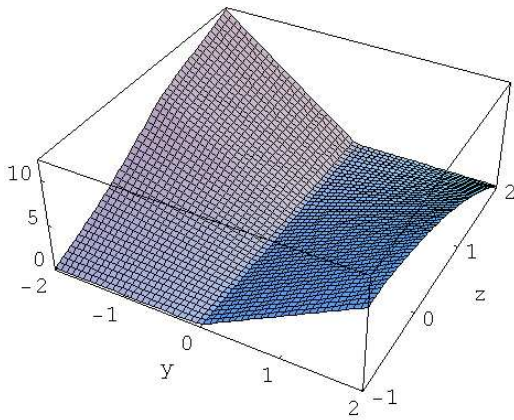


FIGURE 5.13: f for $(\mathbf{F}^1, T_1) \Rightarrow \hat{f} = 0$

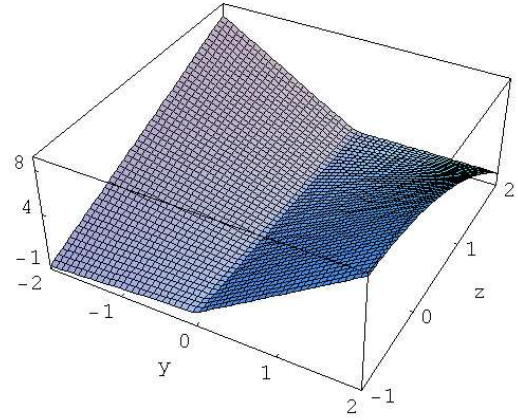


FIGURE 5.14: f for $(\mathbf{F}^2, T_1) \Rightarrow \hat{f} = -1$

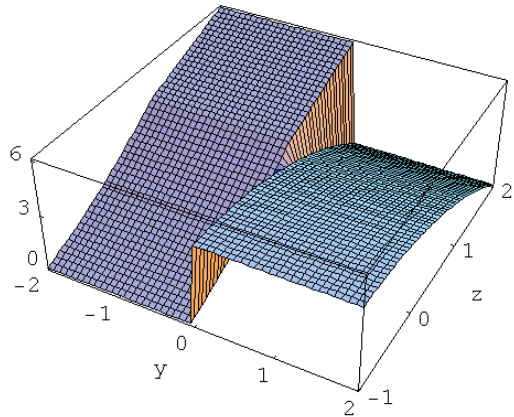


FIGURE 5.15: f for $(\mathbf{F}^1, T_2) \Rightarrow \hat{f} = 0$

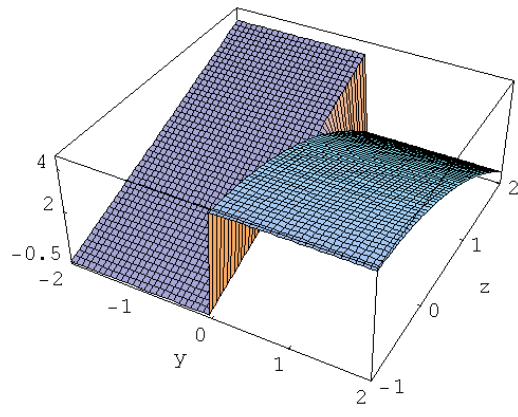


FIGURE 5.16: f for $(\mathbf{F}^2, T_2) \Rightarrow \hat{f} = -\frac{1}{2}$.

We see from Figure 5.15 and Figure 5.16 that the certificate f is not defined for $y = 0$ due to the definition of T_2 in (5.69) (cf. Example 5.23).

5.5 Implementation issues for finding exclusion boxes

5.5.1 Basics

Assumption 5.28. We implemented the suggestions from Idea 5.24 in GloptLab by DOMES [18] which is a configurable MATLAB framework for computing the global solution of a quadratic CSP (5.9), i.e. with F_k from (5.47). The matrices $C_k \in \mathbb{R}^{n \times n}$ are lower triangular in GloptLab

$$C_k \in \mathbb{R}_{\text{tril}}^{n \times n} := \{A \in \mathbb{R}^{n \times n} : A_{ij} = 0 \text{ for } i < j\}. \quad (5.77)$$

For running GloptLab the MATLAB toolbox INTLAB by RUMP [87], `lp_solve` by BERKELAAR et al. [9], SEDUMI by PÓLIK [86], STURM [102] as well as SDPT3 by TOH et al. [104] were installed for using all features of GloptLab.

Definition 5.29. Let $x \in \{y, z, R, S, u, v\}$, then we define

$$\text{opt}(x) := \begin{cases} 1 & \text{if } x \text{ is optimized} \\ 0 & \text{if } x \text{ is not optimized} . \end{cases}$$

5.5.2 Linearly constrained case

Assumption 5.30. Without loss of generality, we make the following assumptions for the optimization problem (5.70):

1. If u is not minimized, then we set

$$u \equiv u^0 := \underline{x} . \quad (5.78)$$

2. If v is not minimized, then we set

$$v \equiv v^0 := \bar{x} . \quad (5.79)$$

3. The starting point is assumed to satisfy

$$z^0 \in [u^0, v^0] \quad (5.80)$$

(which is satisfied for the choice (5.96)).

Remark 5.31.

- The choices (5.78) resp. (5.79) are justified, since otherwise $\mathbf{x}$ can be replaced by $[u, \bar{x}]$, $[\underline{x}, v]$ resp. $[u, v]$ anyway.
- The choice (5.80) ensures that the starting point component z^0 for the z -component is feasible for optimization problem (5.70).

Proposition 5.32. For the variables of the optimization problem (5.70) the following properties hold:

- The variables y, R, S are unbounded, if they are optimized.
- The constraints on the variables z, u, v satisfy

$\text{opt}(z)$	$\text{opt}(u)$	$\text{opt}(v)$	Bound constraints	Linear constraints
0	0	0	none	none
0	1	0	$u \in [\underline{x}, \min(z, \bar{x} - r)]$	none
0	0	1	$v \in [\max(\underline{x} + r, z), \bar{x}]$	none
0	1	1	$u \in [\underline{x}, z], v \in [z, \bar{x}]$	$B_2 \begin{pmatrix} u \\ v \end{pmatrix} \leq -r$
1	0	0	$z \in \mathbf{x}$	none
1	1	0	$z \in \mathbf{x}, u \in [\underline{x}, \bar{x} - r]$	$B_1 \begin{pmatrix} z \\ u \end{pmatrix} \leq 0$
1	0	1	$z \in \mathbf{x}, v \in [\underline{x} + r, \bar{x}]$	$B_2 \begin{pmatrix} z \\ v \end{pmatrix} \leq 0$
1	1	1	$u, v \in \mathbf{x}$	$\begin{pmatrix} B_3 \\ 0 \ B_2 \end{pmatrix} \begin{pmatrix} z \\ u \\ v \end{pmatrix} \leq \begin{pmatrix} 0 \\ -r \end{pmatrix} ,$

where

$$B_1 := (-I \ I) \quad (5.81)$$

$$B_2 := (I \ -I) \quad (5.82)$$

$$B_3 := \begin{pmatrix} -I & I & 0 \\ I & 0 & -I \end{pmatrix} . \quad (5.83)$$

In particular, the constraints with $r_i = 0$ can be dropped.

Proof. We consider the constraints of optimization problem (5.70)

$$u, v \in \mathbf{x} \quad , \quad u + r \leq v \quad , \quad z \in [u, v]$$

with $r \in [0, \bar{x} - \underline{x}]$ which are equivalent to

$$\underline{x} \leq u \leq z \leq v \leq \bar{x} \quad (5.84)$$

$$u + r \leq v \quad , \quad (5.85)$$

and feasible because of (5.80). In particular, we can drop the constraints in (5.85) with $r_i = 0$, since these constraints then are already satisfied by (5.84). Now we want to simplify these constraints (if available) which yields the following observations:

- The variables y, R, S are unbounded due to (5.84), if they are optimized.
- If — apart from y, R or S — only the variable u is optimized, then (5.84) and (5.85) imply because of (5.79)

$$\begin{aligned} u &\in [\underline{x}, z] \\ u + r \leq \bar{x} &\iff u \leq \bar{x} - r \quad , \end{aligned}$$

which is equivalent to

$$u \in [\underline{x}, \min(z, \bar{x} - r)] .$$

- If — apart from y, R or S — only the variable v is optimized, then (5.84) and (5.85) imply because of (5.78)

$$\begin{aligned} v &\in [z, \bar{x}] \\ \underline{x} + r \leq v &\quad , \end{aligned}$$

which is equivalent to

$$v \in [\max(z, \underline{x} + r), \bar{x}] .$$

- If — apart from y, R or S — exactly the variables u and v are optimized, then (5.84) is equivalent to

$$\begin{aligned} \underline{x} \leq u \leq z \quad \wedge \quad z \leq v \leq \bar{x} \\ \iff u \in [\underline{x}, z] \quad \wedge \quad v \in [z, \bar{x}] \end{aligned}$$

and (5.85) is equivalent to

$$u + r \leq v \stackrel{(5.82)}{\iff} B_2 \left(\begin{smallmatrix} u \\ v \end{smallmatrix} \right) \leq -r .$$

- If — apart from y , R or S — only the variable z is optimized, then (5.84) implies because of (5.78) and (5.79) that $z \in \mathbf{x}$.
- If — apart from y , R and S — exactly the variables u and z are optimized, then (5.84) and (5.85) imply because of (5.79)

$$\begin{aligned} \underline{x} &\leq u \leq z \leq \bar{x} \\ u + r &\leq \bar{x} \iff u \leq \bar{x} - r , \end{aligned}$$

which is equivalent to

$$\begin{aligned} u &\in [\underline{x}, \min(\bar{x}, \bar{x} - r)] \stackrel{(5.71)}{=} [\underline{x}, \bar{x} - r] \\ z &\in \mathbf{x} \end{aligned}$$

and

$$u \leq z \stackrel{(5.81)}{\iff} B_1 \left(\begin{smallmatrix} z \\ u \end{smallmatrix} \right) \leq 0 .$$

- If — apart from y , R or S — exactly the variables v and z are optimized, then (5.84) and (5.85) imply because of (5.78)

$$\begin{aligned} \underline{x} &\leq z \leq v \leq \bar{x} \\ \underline{x} + r &\leq v , \end{aligned}$$

which is equivalent to

$$\begin{aligned} z &\in \mathbf{x} \\ v &\in [\max(\underline{x}, \underline{x} + r), \bar{x}] \stackrel{(5.71)}{=} [\underline{x} + r, \bar{x}] \end{aligned}$$

and

$$z \leq v \stackrel{(5.82)}{\iff} B_2 \left(\begin{smallmatrix} z \\ v \end{smallmatrix} \right) \leq 0 .$$

- If — apart from y , R or S — exactly the the variables u , v and z are optimized, then (5.84) and (5.85) are equivalent to

$$u, v \in \mathbf{x} \quad \wedge \quad u \leq z \leq v \quad \wedge \quad u + r \leq v , \tag{5.86}$$

where the inequalities of (5.86) are equivalent to

$$u \leq z \quad \wedge \quad z \leq v \quad \wedge \quad u - v \leq r . \tag{5.87}$$

The inequalities (5.87) are equivalent to

$$\begin{pmatrix} -I & I & 0 \\ I & 0 & -I \\ 0 & I & -I \end{pmatrix} \begin{pmatrix} z \\ u \\ v \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \\ -r \end{pmatrix} \stackrel{(5.83), (5.82)}{\iff} \begin{pmatrix} B_3 \\ 0 & B_2 \end{pmatrix} \begin{pmatrix} z \\ u \\ v \end{pmatrix} \leq \begin{pmatrix} 0 \\ 0 \\ -r \end{pmatrix} .$$

□

If we want to apply Algorithm 3.20 for solving optimization problem (5.70) and we want to use a solver for computing the search direction in Algorithm 3.20 which needs a strictly feasible starting point (e.g., socp by LOBO et al. [57]), we have to take special care for the initial choices of u and v , which we will discuss next.

Definition 5.33. For $[a, b] \in \mathbb{IR}^n$ we define $\omega_{[a,b]} : (0, 1) \longrightarrow \mathbb{R}^n$ by

$$\omega_{[a,b]}(t) := a + t(b - a) . \quad (5.88)$$

Remark 5.34.

- If the box $[a, b]$ is given, then ω assigns every $t \in (0, 1)$ a point in (a, b) , i.e. an interior point.
- If t is close to 0, then $\omega_{[a,b]}(t)$ is close to a .
- If t is close to 1, then $\omega_{[a,b]}(t)$ is close to b .

Proposition 5.35. Let $r \in (0, \bar{x} - \underline{x})$ and $t_0, t_1 \in (0, 1)$ with $t_0 < t_1$. Then

$$u := \omega_{[\underline{x}, \bar{x}-r]}(t_0) \quad (5.89)$$

$$v := \omega_{[\underline{x}+r, \bar{x}]}(t_1) , \quad (5.90)$$

satisfy

$$u, v \in (\underline{x}, \bar{x}) , \quad u + r < v .$$

Proof.

- Because of

$$\begin{array}{c} \text{assumption} \\ \downarrow \\ r \in (0, \bar{x} - \underline{x}) \end{array} \iff 0 < r < \bar{x} - \underline{x} \implies -(\bar{x} - \underline{x}) < -r \quad (5.91)$$

we obtain

$$\underline{x} = \bar{x} - (\bar{x} - \underline{x}) \stackrel{(5.91)}{<} \bar{x} - r ,$$

i.e. $[\underline{x}, \bar{x} - r]$ is a well defined box which is not thin (and in particular $(\underline{x}, \bar{x} - r) \neq \emptyset$). Therefore we have for $t_0 \in (0, 1)$

$$u \stackrel{(5.89)}{=} \omega_{[\underline{x}, \bar{x}-r]}(t_0) \stackrel{(5.88)}{\in} (\underline{x}, \bar{x} - r) \subset [\underline{x}, \bar{x}] = \mathbf{x} .$$

- Because of

$$\begin{array}{c} \text{assumption} \\ \downarrow \\ r \in (0, \bar{x} - \underline{x}) \end{array} \implies r < \bar{x} - \underline{x} \iff \underline{x} + r < \bar{x} ,$$

$[\underline{x} + r, \bar{x}]$ is a well defined box which is not thin (and in particular $(\underline{x} + r, \bar{x}) \neq \emptyset$). Therefore we have for $t_1 \in (0, 1)$

$$v \stackrel{(5.90)}{=} \omega_{[\underline{x}+r, \bar{x}]}(t_1) \stackrel{(5.88)}{\in} (\underline{x} + r, \bar{x}) \subset [\underline{x}, \bar{x}] = \mathbf{x} .$$

- Since

$$\begin{array}{c} \text{assumption} \\ \downarrow \\ r \in (0, \bar{x} - \underline{x}) \end{array} \iff 0 < r < \bar{x} - \underline{x} \implies 0 < \bar{x} - \underline{x} - r \quad (5.92)$$

yields

$$\begin{array}{c} \text{assumption} \\ \downarrow \\ t_0 < t_1 \end{array} \implies t_0(\bar{x} - \underline{x} - r) \stackrel{(5.92)}{<} t_1(\bar{x} - \underline{x} - r) ,$$

we obtain

$$\begin{aligned} u + r &\stackrel{(5.89)}{=} \omega_{[\underline{x}, \bar{x}-r]}(t_0) + r \stackrel{(5.88)}{=} \underline{x} + t_0((\bar{x} - r) - \underline{x}) + r \\ &= (\underline{x} + r) + t_0(\bar{x} - \underline{x} - r) \stackrel{(5.5.2)}{<} (\underline{x} + r) + t_1(\bar{x} - \underline{x} - r) \stackrel{(5.88)}{=} \omega_{[\underline{x}+r, \bar{x}]}(t_1) \stackrel{(5.90)}{=} v . \quad \square \end{aligned}$$

5.5.3 Nonlinearly constrained case

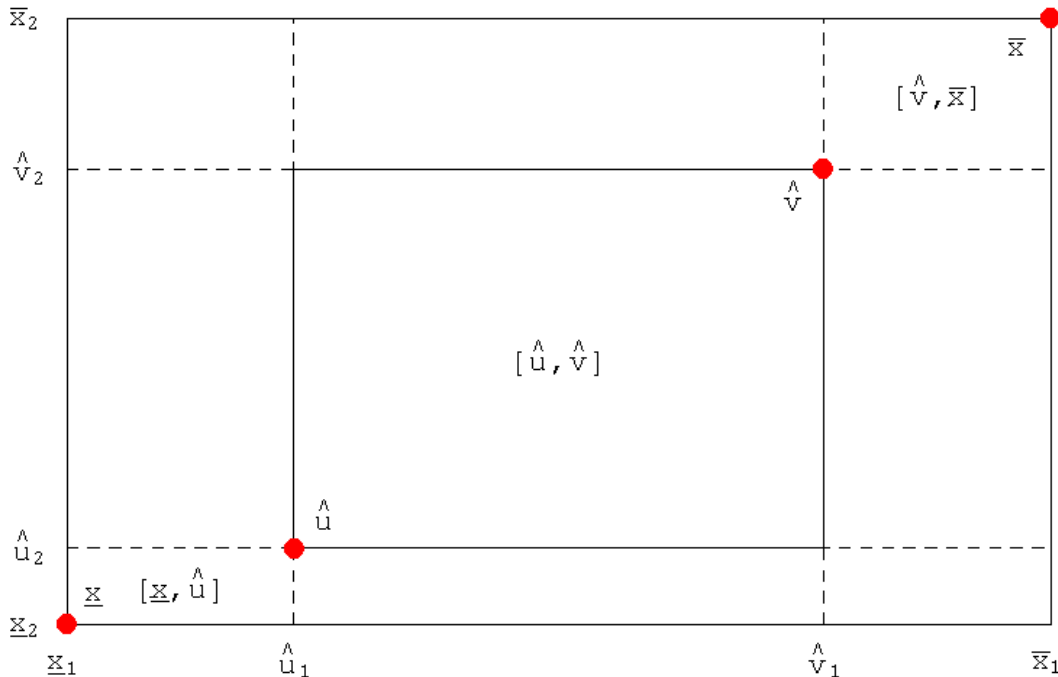
Proposition 5.36. *The bound constraints of the optimization problem (5.72) satisfy*

$$\left. \begin{array}{l} u, v \in \mathbf{x} \\ u \leq \hat{u} \\ \hat{v} \leq v \end{array} \right\} \iff \left\{ \begin{array}{l} u \in [\underline{x}, \hat{u}] \\ v \in [\hat{v}, \bar{x}] . \end{array} \right. \quad (5.93)$$

Proof. We have

$$\begin{aligned} \left. \begin{array}{l} u, v \in \mathbf{x} \\ u \leq \hat{u} \\ \hat{v} \leq v \end{array} \right\} &\iff \left\{ \begin{array}{l} (\underline{x} \leq u \wedge u \leq \bar{x}) \wedge (\underline{x} \leq v \wedge v \leq \bar{x}) \\ u \leq \hat{u} \\ \hat{v} \leq v \end{array} \right. \\ &\iff \left\{ \begin{array}{l} \underline{x} \leq u \wedge (u \leq \hat{u} \wedge u \leq \bar{x}) \\ (\underline{x} \leq v \wedge \hat{v} \leq v) \wedge v \leq \bar{x} \end{array} \right. \\ &\stackrel{\hat{u}, \hat{v} \in \mathbf{x}}{\iff} \left\{ \begin{array}{l} u \in [\underline{x}, \hat{u}] \\ v \in [\hat{v}, \bar{x}] . \end{array} \right. \quad \square \end{aligned}$$

Remark 5.37. A graphical illustration of (5.93) is given by



Proposition 5.38. *The constraints on the variables z, u, v of the optimization problem (5.72) satisfy*

$\text{opt}(z)$	$\text{opt}(u)$	$\text{opt}(v)$	Bound constraints	Linear constraints
0	1	0	$u \in [\underline{x}, \hat{u}]$	none
0	0	1	$v \in [\hat{v}, \bar{x}]$	none
0	1	1	$u \in [\underline{x}, \hat{u}], v \in [\hat{v}, \bar{x}]$	none
1	1	0	$z \leq \bar{x}, u \in [\underline{x}, \hat{u}]$	$B_1 \left(\begin{smallmatrix} z \\ u \end{smallmatrix} \right) \leq 0$
1	0	1	$\underline{x} \leq z, v \in [\hat{v}, \bar{x}]$	$B_2 \left(\begin{smallmatrix} z \\ v \end{smallmatrix} \right) \leq 0$
1	1	1	$u \in [\underline{x}, \hat{u}], v \in [\hat{v}, \bar{x}]$	$B_3 \left(\begin{smallmatrix} z \\ u \\ v \end{smallmatrix} \right) \leq 0$

(5.94)

Proof.

- If $\text{opt}(z) = 0$ (i.e. z is not optimized), then the bound constraint for z in (5.72) does not occur.
 - If $\text{opt}(u) = 1 \wedge \text{opt}(v) = 0$, then in (5.72) the bound constraint

$$u \in [\underline{x}, \hat{u}]$$

remains due to (5.93).

- If $\text{opt}(u) = 0 \wedge \text{opt}(v) = 1$, then in (5.72) the bound constraint

$$v \in [\hat{v}, \bar{x}]$$

remains due to (5.93).

- If $\text{opt}(u) = 1 \wedge \text{opt}(v) = 1$, then in (5.72) the bound constraints

$$u \in [\underline{x}, \hat{u}] , \quad v \in [\hat{v}, \bar{x}]$$

remain due to (5.93).

- If $\text{opt}(z) = 1$ (i.e. z is optimized), then the bound constraints of (5.72) read:
 - If $\text{opt}(u) = 1 \wedge \text{opt}(v) = 0$, then in (5.72) the bound constraints

$$\begin{aligned}
 &u \in [\underline{x}, \hat{u}] \\
 &\quad \text{opt}(v)=0 \\
 &z \in [u, v] \xLeftrightarrow{\downarrow} z \in [u, \bar{x}] \\
 &\quad \xLeftrightarrow{(5.81)} B_1 \left(\begin{smallmatrix} z \\ u \end{smallmatrix} \right) \leq 0 \wedge z \leq \bar{x}
 \end{aligned}$$

remain due to (5.93).

- If $\text{opt}(u) = 0 \wedge \text{opt}(v) = 1$, then in (5.72) the bound constraints

$$\begin{aligned}
 &v \in [\hat{v}, \bar{x}] \\
 &\quad \text{opt}(u)=0 \\
 &z \in [u, v] \xLeftrightarrow{\downarrow} z \in [\underline{x}, v] \\
 &\quad \xLeftrightarrow{(5.82)} \underline{x} \leq z \wedge B_2 \left(\begin{smallmatrix} z \\ v \end{smallmatrix} \right) \leq 0
 \end{aligned}$$

remain due to (5.93).

- If $\text{opt}(u) = 1 \wedge \text{opt}(v) = 1$, then all bound constraints remain in (5.72) and hence we obtain by using (5.93)

$$\begin{aligned} u &\in [\underline{x}, \hat{u}] \\ v &\in [\hat{v}, \bar{x}] \\ z &\in [u, v] \stackrel{(5.83)}{\iff} B_3 \begin{pmatrix} z \\ u \\ v \end{pmatrix} \leq 0. \end{aligned} \quad \square$$

Remark 5.39.

- If the search direction in Algorithm 3.20 is computed by a solver that needs a strictly feasible starting point (e.g., `socp` by LOBO et al. [57]), then we must slightly perturb the bound constraints for u and v in (5.72): Choose $\varepsilon \in (0, \frac{1}{2}(\hat{v} - \hat{u}))$ near 0 and consider instead of (5.72) (under consideration of (5.93)) the optimization problem

$$\begin{aligned} & - \min b(y, z, R, S, u, v) \\ & \text{s.t. } f(y, z, R, S, u, v) - \delta \leq 0 \\ & \quad u \in [\underline{x}, \hat{u} + \varepsilon] \\ & \quad v \in [\hat{v} - \varepsilon, \bar{x}] \\ & \quad z \in [u, v] \\ & \quad y \in \mathbb{R}^m, R \in \mathbb{R}_{\text{triu}}^{n \times n}, S \in \mathbb{R}_{\text{striu}}^{n \times n} \end{aligned} \quad (5.95)$$

for which $\hat{u}$ and $\hat{v}$ are strictly feasible. If we want to use (5.95) in Algorithm 3.20, then we have to replace $\hat{u}$ by $\hat{u} + \varepsilon$ resp. $\hat{v}$ by $\hat{v} - \varepsilon$ in (5.94).

- $[\hat{u} + \varepsilon, \hat{v} - \varepsilon]$ is a box, since

$$\begin{aligned} \varepsilon \in (0, \tfrac{1}{2}(\hat{v} - \hat{u})) & \iff 0 < \varepsilon \wedge \varepsilon < \tfrac{1}{2}(\hat{v} - \hat{u}) \\ & \stackrel{\cdot 2}{\iff} 0 < \varepsilon \wedge 2\varepsilon < \hat{v} - \hat{u} \\ & \stackrel{+\hat{u}-\varepsilon}{\iff} 0 < \varepsilon \wedge \hat{u} + \varepsilon < \hat{v} - \varepsilon. \end{aligned}$$

- A possible choice for $\varepsilon \in (0, \frac{1}{2}(\hat{v} - \hat{u}))$ is

$$\varepsilon_i := \begin{cases} \hat{\varepsilon} & \text{for } \hat{\varepsilon} < \frac{1}{2}(\hat{v}_i - \hat{u}_i) \\ \frac{1}{2}(\hat{v}_i - \hat{u}_i)\hat{\varepsilon} & \text{for } \hat{\varepsilon} \geq \frac{1}{2}(\hat{v}_i - \hat{u}_i) \end{cases}$$

for $i = 1, \dots, n$ with $\hat{\varepsilon} := 10^{-5}$.

5.5.4 Starting point

Idea 5.40. We want to find a point (y, z, R, S, u, v) which is feasible for the linearly constrained optimization problem (5.70) with

$$f(y, z, R, S, u, v) \stackrel{(5.37)}{=} \frac{Z(y, z, R, S, u, v) - \max(0, Y(y, z))}{T(y, z, R, S, u, v)} < 0$$

quickly. For this we need a good starting point $(y^0, z^0, R^0, S^0, u^0, v^0)$ and therefore we must take the following observations into account:

- y^0 and z^0 should be chosen so, that $Y(y^0, z^0)$ is positive.
- $(y^0, z^0, R^0, S^0, u^0, v^0)$ should be chosen so, that $Z(y^0, z^0, R^0, S^0, u^0, v^0) \stackrel{(5.38)}{\geq} 0$ is near zero.

These facts lead to the following suggestions for choosing a starting point $(y^0, z^0, R^0, S^0, u^0, v^0)$:

1. If the solver in use can only handle strictly feasible bound/linear constraints (e.g., Algorithm 3.20 with using `socp` by LOBO et al. [57] for computing the search direction), then we choose u^0 resp. v^0 according to (5.89) resp. (5.90). Otherwise (e.g., SolvOpt by KAPPEL & KUNTSEVICH [41] or Algorithm 3.20 with using MOSEK for computing the search direction) we take the endpoints of $\mathbf{x}$ for u^0 and v^0 .
2. The natural choice for the starting value of $z \in [u, v] \subseteq \mathbf{x}$ is the midpoint

$$z^0 := \frac{1}{2}(u^0 + v^0) \quad (5.96)$$

of the box $[u^0, v^0]$.

3. To get the term $\max(0, Y(y, z))$ in the certificate f from (5.37) as large as possible, we make the following choices:

- (a) If $\underline{F}_k = -\infty$, then we choose

$$y_k^0 := \begin{cases} -1 & \text{if } \overline{F}_k < F_k(z^0) \\ 0 & \text{else} \end{cases} \quad (5.97)$$

and if $\overline{F}_k = \infty$, then we choose

$$y_k^0 := \begin{cases} 1 & \text{if } F_k(z^0) < \underline{F}_k \\ 0 & \text{else} \end{cases} \quad (5.98)$$

both due to (5.44) and (5.45).

- (b) If both $\underline{F}_k$ and $\overline{F}_k$ are finite, then we choose

$$y_k^0 := \begin{cases} 1 & \text{if } F_k(z^0) < \underline{F}_k \\ -1 & \text{if } \overline{F}_k < F_k(z^0) \\ 0 & \text{else} \end{cases} \quad (5.99)$$

due to (5.44).

4. For the choices of R and S we refer to Proposition 5.42.

Remark 5.41. If T is chosen according to (5.69), then

$$T(y, z, R, S, u, v) = 0 \stackrel{(5.69)}{\iff} y = 0 .$$

Therefore, if $y^0 = 0$ occurs as starting point, then we have a feasible point $F(z) \in \mathbf{F}$ due to (5.97), (5.98) and (5.99). Furthermore, we can expect that no solver should have difficulties with this choice of T because of the small size of the zero set of T due to Example 5.23.

In the following we will make use of the MATLAB operators `diag`, `tril` and `triu`.

Proposition 5.42. *Let (5.46) and (5.77) be satisfied. Choose any $y \in \mathbb{R}^m$ and consider the modified Cholesky factorization*

$$\hat{A} = \hat{R}^T \hat{R} - D \quad (5.100)$$

of $\hat{A}$ (with $\hat{R} \in \mathbb{R}_{\text{triu}}^{n \times n}$ and the non-negative diagonal matrix $D \in \mathbb{R}^{n \times n}$), where

$$\hat{A} := C(y) + S^T - S \in \mathbb{R}_{\text{sym}}^{n \times n} \quad (5.101)$$

$$S := -\frac{1}{2} \text{triu}(C(y)^T, 1) \in \mathbb{R}_{\text{striu}}^{n \times n}. \quad (5.102)$$

1. Then

$$\hat{A} = C(y) - \frac{1}{2} \text{tril}(C(y), -1) + \frac{1}{2} \text{triu}(C(y)^T, 1) \quad (5.103)$$

$$\text{diag}(\hat{A}) = \text{diag}(C(y)). \quad (5.104)$$

2. If we set

$$R := D^{\frac{1}{2}}, \quad (5.105)$$

then

$$A(y, R, S) = \hat{R}^T \hat{R} \in \mathbb{R}_{\text{sym}}^{n \times n},$$

i.e. $A(y, R, S)$ is positive semidefinite.

Proof.

- Since (5.46) is satisfied by assumption, the statements of Proposition 5.19 hold.
- Since $C(y)$ is lower triangular due to (5.77) and (5.48) we obtain

$$S \stackrel{(5.102)}{=} -\frac{1}{2} \text{triu}(C(y)^T, 1) \stackrel{(5.77), (5.48)}{\downarrow} \in \mathbb{R}_{\text{striu}}^{n \times n}$$

and

$$\begin{aligned} \hat{A} &\stackrel{(5.101)}{=} C(y) + S^T - S \\ &\stackrel{(5.102)}{=} C(y) + \left(-\frac{1}{2} \text{triu}(C(y)^T, 1) \right)^T - \left(-\frac{1}{2} \text{triu}(C(y)^T, 1) \right) \\ &= C(y) - \frac{1}{2} \left(\text{triu}(C(y)^T, 1) \right)^T + \frac{1}{2} \text{triu}(C(y)^T, 1). \end{aligned} \quad (5.106)$$

- We obtain (5.103)

$$\begin{aligned} \hat{A} &\stackrel{(5.106)}{=} C(y) - \frac{1}{2} \left(\text{triu}(C(y)^T, 1) \right)^T + \frac{1}{2} \text{triu}(C(y)^T, 1) \\ &= C(y) - \frac{1}{2} \text{tril}(C(y), -1) + \frac{1}{2} \text{triu}(C(y)^T, 1) \end{aligned}$$

and (5.104)

$$\begin{aligned} \text{diag}(\hat{A}) &\stackrel{(5.103)}{=} \text{diag}\left(C(y) - \frac{1}{2} \text{tril}(C(y), -1) + \frac{1}{2} \text{triu}(C(y)^T, 1)\right) \\ &= \text{diag}(C(y)) - \frac{1}{2} \text{diag}\left(\text{tril}(C(y), -1)\right) + \frac{1}{2} \text{diag}\left(\text{triu}(C(y)^T, 1)\right) \\ &= \text{diag}(C(y)). \end{aligned}$$

- We calculate

$$\begin{aligned} C(y)^T - \frac{1}{2}\text{triu}(C(y)^T, 1) &= \left(\text{diag}(C(y)^T) + \text{triu}(C(y)^T, 1) \right) - \frac{1}{2}\text{triu}(C(y)^T, 1) \\ &= \text{diag}(C(y)) + \frac{1}{2}\text{triu}(C(y)^T, 1). \end{aligned} \quad (5.107)$$

Because of

$$\begin{aligned} \frac{1}{2}\left(\text{triu}(C(y)^T, 1)\right)^T &= \left(\text{triu}(C(y)^T, 1)\right)^T - \frac{1}{2}\left(\text{triu}(C(y)^T, 1)\right)^T \\ &= \text{tril}(C(y), -1) - \frac{1}{2}\left(\text{triu}(C(y)^T, 1)\right)^T \end{aligned} \quad (5.108)$$

we have

$$\text{diag}(C(y)) + \frac{1}{2}\left(\text{triu}(C(y)^T, 1)\right)^T \stackrel{(5.108)}{=} C(y) - \frac{1}{2}\left(\text{triu}(C(y)^T, 1)\right)^T. \quad (5.109)$$

Therefore, we obtain

$$\begin{aligned} \hat{A}^T &\stackrel{(5.106)}{=} \left(C(y) - \frac{1}{2}\left(\text{triu}(C(y)^T, 1)\right)^T + \frac{1}{2}\text{triu}(C(y)^T, 1) \right)^T \\ &= C(y)^T - \frac{1}{2}\text{triu}(C(y)^T, 1) + \frac{1}{2}\left(\text{triu}(C(y)^T, 1)\right)^T \\ &\stackrel{(5.107)}{=} \text{diag}(C(y)) + \frac{1}{2}\text{triu}(C(y)^T, 1) + \frac{1}{2}\left(\text{triu}(C(y)^T, 1)\right)^T \\ &\stackrel{(5.109)}{=} C(y) - \frac{1}{2}\left(\text{triu}(C(y)^T, 1)\right)^T + \frac{1}{2}\text{triu}(C(y)^T, 1) \\ &\stackrel{(5.106)}{=} \hat{A}. \end{aligned} \quad (5.110)$$

- Since $\hat{A} \in \mathbb{R}_{\text{sym}}^{n \times n}$ due to (5.110), there exists a modified Cholesky factorization of $\hat{A}$ of the form (5.100). Hence, we can choose R according to (5.105) and evaluating A at (y, R, S) with R from (5.105) and S from (5.102) yields

$$\begin{aligned} A(y, R, S) &\stackrel{(5.50)}{=} C(y) + S^T - S + R^T R \stackrel{(5.101), (5.105)}{\downarrow} \hat{A} + (D^{\frac{1}{2}})^T D^{\frac{1}{2}} \\ &\stackrel{(5.100)}{=} (\hat{R}^T \hat{R} - D) + D = \hat{R}^T \hat{R}. \end{aligned} \quad \square$$

Remark 5.43.

- If $\hat{A}$ is positive semidefinite, then $D = 0$ due to (5.100).
- If $C(y)$ is a diagonal matrix, then $S = 0$ due to (5.102).
- Due to (5.103) and (5.104), we can construct $\hat{A}$ by setting $\hat{A}$ equal to $C(y)$, then multiplying the lower triangular part of $\hat{A}$ by $\frac{1}{2}$ and finally copying the resulting lower triangular part of $\hat{A}$ to the upper triangular part of $\hat{A}$.

Now we combine the facts that we presented in this subsection (i.e (5.96), (5.97), (5.98), (5.99), and Proposition 5.42) to the Algorithm 5.44, which we will use for creating a starting point for the optimization problem (5.70).

Algorithm 5.44.

if the solver can only handle strictly feasible bound/linear constraints

Choose $0 < t_0 < t_1 < 1$ (e.g., $t_0 = 0.1$ and $t_1 = 0.9$)

$$u = (1 - t_0)\underline{x} + t_0(\bar{x} - r)$$

$$v = (1 - t_1)(\underline{x} + r) + t_1\bar{x}$$

else

$$u = \underline{x}$$

$$v = \bar{x}$$

end

$$z = \frac{1}{2}(u + v)$$

$$F = F(z)$$

if $F \in \mathbf{F}$

stop (found feasible point => cannot verify infeasibility)

end if

for $k = 1 : m$

if $\underline{F}_k = -\infty$

if $\bar{F}_k < F_k$

$$y_k = -1$$

else

$$y_k = 0$$

end if

else if $\bar{F}_k = \infty$

if $F_k < \underline{F}(k)$

$$y_k = 1$$

else

$$y_k = 0$$

end if

else

if $F_k < \underline{F}_k$

$$y_k = 1$$

else if $\bar{F}_k < F_k$

$$y_k = -1$$

else

$$y_k = 0$$

end if

end if

end for

Compute $C(y)$

$$S = -\frac{1}{2}\text{triu}(C(y)^T, 1)$$

$$[\hat{R}, D] = \text{modified_cholesky_factorization}(C(y) + S^T - S)$$

$$R = \text{sqrt}(D)$$

Remark 5.45. Infeasible constrained solvers (e.g., SolvOpt by KAPPEL & KUNTSEVICH [41]) can be applied directly to the nonlinearly constrained optimization problems (5.72). In this case the starting point created by Algorithm 5.44 can be used at once without solving optimization problem (5.70) first as it is necessary for Algorithm 3.20. Therefore, the bound constraints

$$u \leq \hat{u} \ , \quad \hat{v} \leq v$$

do not occur in this situation and optimization problem (5.95) does not need to be considered at all. Nevertheless, it is useful in this case to add the linear constraint

$$u + r \leq v$$

(with a fixed $r > 0$) from optimization problem (5.70) to the constrained problem for preventing the box $[u, v]$ from becoming too small.

Chapter 6

Numerical results

In the following chapter we compare the numerical results of the second order bundle algorithm, MPBNGC by MÄKELÄ [70] and SolvOpt by KAPPEL & KUNTSEVICH [41] for some examples of the Hock-Schittkowski collection by SCHITTKOWSKI [95, 96], for custom examples that arise in the context of finding exclusion boxes for a quadratic CSP in GloptLab by DOMES [18], and for higher dimensional piecewise quadratic examples.

6.1 Introduction

The following implementations of Algorithm 3.20 are available:

- A pure MATLAB version (for easy understanding, modifying and testing new ideas concerning the algorithm).
- A MATLAB version in which the main parts of the algorithm are split into several subroutines, where every subroutine can either be called as pure MATLAB code or via a C `mex`-file (this is useful for partially speeding up the algorithm, but still keeping it simple enough for modifying and testing many examples of the modified code).
- A pure C version (for performance), which is used throughout all the tests.

The C `mex`-files and the C version require a BLAS/LAPACK implementation (e.g., ATLAS by WHALEY & PETITET [110], GotoBLAS by GOTO & VAN DE GEIJN [28], or the Netlib BLAS reference implementation by BLACKFORD et al. [10]). In the unconstrained case, all three versions produce the same results as the original FORTRAN bundle-Newton method.

Although there exist some test collections for nonsmooth unconstrained optimization (e.g., LUKŠAN & VLČEK [68]; also cf. KARMITSA et al. [44] for an extensive comparison of numerical results) and nonsmooth linearly constrained optimization (e.g., LUKŠAN & VLČEK [67]), I do not know a standardized, prominent test collection for nonsmooth constrained optimization. Therefore, a common

way for testing nonsmooth constrained solvers is to take a test collection for smooth constrained optimization (e.g., the Hock-Schittkowski collection from SCHITTKOWSKI [95, 96]) and to treat the smooth constraints as one nonsmooth constraint (by using a max-function, cf. Remark 2.27).

We will make tests for

- Algorithm 3.20 (with optimality tolerance $\varepsilon := 10^{-5}$), where we refer to the linearly constrained version as “BNLC”, to the version with the QCQP (4.11) as “Full Alg(orithm)”, and to the version with the reduced QCQP (4.12) as “Red(uced) Alg(orithm)”
- MPBNGC by MÄKELÄ [70] (with the standard termination criterions; although MPBNGC supports the handling of multiple nonsmooth constraints, we do not use this feature, since we are interested here, how well the different solvers handle the nonsmoothness of a constraint, i.e. without exploiting the knowledge of the structure of a max-function; since MPBNGC turned out to be very fast with respect to pure solving time for the low dimensional examples in the case of successful termination with a stationary point, the number of iterations and function evaluations was chosen in a way that in the other case the solving times of the different algorithms have approximately at least the same magnitude)
- SolvOpt by KAPPEL & KUNTSEVICH [41] (with the standard termination criterions, which are described in KUNTSEVICH & KAPPEL [54])

(we choose MPBNGC and SolvOpt for our comparisons, since both are written in a compiled programming language, both are publicly available, and both support nonconvex constraints), where we will modify the termination criterions slightly only in Section 6.4, on the following examples (the corresponding result tables can be found in Appendix A):

- Example 2.11 (denoted by E1 resp. E2), Example 3.12 (denoted by E3) and the Hock-Schittkowski collection (in the above sense; no problems which contain nonlinear equality constraints; linear constraints are inserted into the search direction problem in Algorithm 3.20; feasible starting point). This yields 58 test problems (cf. Table A.3), which we will discuss in Section 6.2.
- Optimization problems as described in Idea 5.24 (for finding exclusion boxes for CSPs; cf. Tables A.4–A.7), which we will discuss in Section 6.3.
- Higher dimensional piecewise quadratic examples with up to 100 variables (cf. Tables A.8–A.11), which we will discuss in Section 6.4.

All test examples will be sorted with respect to the problem dimension (beginning with the smallest). Furthermore, we use analytic derivative information for all occurring functions (Note: Implementing analytic derivative information for certificate f from (5.37) effectively, is a nontrivial task) and we perform all tests on the same machine as in Subsection 4.4.1.

Notation 6.1. We introduce the following notation for the record of the solution process of an algorithm (which is used in this chapter as well as in Appendix A):

- We denote

$N :=$ “Dimension of the optimization problem”

$Nit :=$ “Number of performed iterations” .

- We denote the final number of evaluations of function dependent data by

$$\begin{aligned} N_a &:= \text{“Number of calls to } (f, g, G, F, \hat{g}, \hat{G}) \text{” (Algorithm 3.20)} \\ N_b &:= \text{“Number of calls to } (f, g, F, \hat{g}) \text{” (MPBNGC)} \\ N_c &:= \text{“Number of calls to } (f, F) \text{” (SolvOpt)} \\ N_g &:= \text{“Number of calls to } g \text{” (SolvOpt)} \\ N_{\hat{g}} &:= \text{“Number of calls to } \hat{g} \text{” (SolvOpt)} . \end{aligned} \tag{6.1}$$

- We denote the duration of the solution process by

$$\begin{aligned} t_1 &:= \text{“Time in milliseconds”} \\ t_2 &:= \text{“Time in milliseconds (without (QC)QP)” (only relevant for Algorithm 3.20)} . \end{aligned} \tag{6.2}$$

- We denote the additional algorithmic information by

$$\begin{aligned} R &:= \text{“Remark” (e.g., if } t_0^k \text{ is modified in Algorithm 3.20,} \\ &\quad \text{additional SolvOpt termination information,} \\ &\quad \text{supplementary problem dependent facts, . . .)} \\ nt &:= \text{“No termination” (within the given number of Nit, . . .)} \\ wm &:= \text{“Wrong minimum”} . \end{aligned}$$

Remark 6.2. The percentage of the time spent in the (QC)QP in Algorithm 3.20 is given by

$$p_1 := \frac{t_1(\text{Algorithm 3.20}) - t_2(\text{Algorithm 3.20})}{t_1(\text{Algorithm 3.20})} \tag{6.3}$$

due to (6.2).

For comparing the cost of evaluating function dependent data (like, e.g., function values, subgradients, . . .) in a preferably fair way (especially for solvers that use different function dependent data), we will make use of the following realistic “credit point system” that an optimal implementation of algorithmic differentiation in backward mode suggests (cf. GRIEWANK & CORLISS [30] and SCHICHL [91, 92, 93]).

Definition 6.3. Let f_A , g_A and G_A resp. F_A , $\hat{g}_A$ and $\hat{G}_A$ be the number of function values, subgradients and (substitutes of) Hessians of the objective function resp. the constraint that an algorithm A used for solving a nonsmooth optimization problem which may have linear constraints and at most

one single nonsmooth nonlinear constraint. Then we define the cost of these evaluations by

$$c(A) := f_A + 3g_A + 3N \cdot G_A + \text{nlc} \cdot (F_A + 3\hat{g}_A + 3N \cdot \hat{G}_A) , \quad (6.4)$$

where

$$\text{nlc} := \begin{cases} 1 & \text{if the optimization problem has a nonsmooth nonlinear constraint} \\ 0 & \text{else .} \end{cases}$$

Now we give explicit formulas for the cost c of the algorithms for which we will make performance comparisons.

Proposition 6.4. *We have*

$$\begin{aligned} c(\text{Algorithm 3.20}) &= (1 + \text{nlc}) \cdot \text{Na} \cdot (1 + 3 + 3N) \\ c(\text{MPBNGC}) &= (1 + \text{nlc}) \cdot \text{Nb} \cdot (1 + 3) \\ c(\text{SolvOpt}) &= (1 + \text{nlc}) \cdot \text{Nc} + 3(\text{Ng} + \text{nlc} \cdot \text{N}\hat{g}) . \end{aligned}$$

Proof.

- Since Algorithm 3.20 evaluates f, g, G and $F, \hat{g}, \hat{G}$ at every call that computes function dependent data, we obtain for $A = \text{Algorithm 3.20}$

$$\begin{aligned} c(A) &\stackrel{(6.4)}{=} f_A + 3g_A + 3N \cdot G_A + \text{nlc} \cdot (F_A + 3\hat{g}_A + 3N \cdot \hat{G}_A) \\ &\stackrel{(6.1)}{=} \text{Na} + 3\text{Na} + 3N \cdot \text{Na} + \text{nlc} \cdot (\text{Na} + 3\text{Na} + 3N \cdot \text{Na}) \\ &= (1 + \text{nlc}) \cdot \text{Na} \cdot (1 + 3 + 3N) . \end{aligned}$$

- Since MPBNGC evaluates f, g and $F, \hat{g}$ at every call that computes function dependent data (cf. MÄKELÄ [70]), the only difference to Algorithm 3.20 with respect to c from (6.4) is that MPBNGC uses no information of Hessians and hence we obtain for $A = \text{MPBNGC}$

$$c(A) \stackrel{(6.1)}{=} (1 + \text{nlc}) \cdot \text{Nb} \cdot (1 + 3) .$$

- Since SolvOpt evaluates f and F at every call that computes function dependent data and only sometimes g or $\hat{g}$ (cf. KUNTSEVICH & KAPPEL [54]), we obtain for $A = \text{SolvOpt}$

$$\begin{aligned} c(A) &\stackrel{(6.4)}{=} f_A + 3g_A + 3N \cdot G_A + \text{nlc} \cdot (F_A + 3\hat{g}_A + 3N \cdot \hat{G}_A) \\ &\stackrel{(6.1)}{=} \text{Nc} + 3\text{Ng} + \text{nlc} \cdot (\text{Nc} + 3\text{N}\hat{g}) \\ &= (1 + \text{nlc}) \cdot \text{Nc} + 3(\text{Ng} + \text{nlc} \cdot \text{N}\hat{g}) . \end{aligned} \quad \square$$

Motivation 6.5. We will visualize the performance of two algorithms A and B for $s \in \{c, \text{Nit}\}$ in Section 6.2 and Section 6.3 the by following record-plot: In this plot the abscissa is labeled by the

name of the test example and the value of the ordinate is given by

$$\text{rp}(s) := s(B) - s(A)$$

(i.e. if $\text{rp}(s) > 0$, then $\text{rp}(s)$ tells us how much algorithm A is better than algorithm B with respect to s for the considered example by absolute numbers; if $\text{rp}(s) < 0$, then $\text{rp}(s)$ quantifies the advantage of algorithm B in comparison to algorithm A ; if $\text{rp}(s) = 0$, then both algorithms are equally good with respect to s). The scaling of the plots is chosen in a way that plots that contain the same test examples are comparable (although the plots may have been generated by results from different algorithms).

Remark 6.6.

1. All results for Algorithm 3.20 that are given in the tables of Appendix A were obtained by using MOSEK by ANDERSEN et al. [3] for determining the search direction, where we used the MOSEK QCQP-solver which turned out to be much faster than the MOSEK SOCP-solver again (as we already noticed in Remark 4.23). We emphasize that in our tests there occurred no search direction problem which MOSEK was not able to solve.
2. The results for computing the search direction in Algorithm 3.20 with IPOPT by WÄCHTER & BIEGLER [109] are practically the same with respect to Nit and Na. Furthermore, IPOPT was as robust and reliable as MOSEK. Nevertheless, IPOPT was slower than MOSEK with respect to the solving time which we expected as IPOPT is designed for general non-linear optimization problems (cf. Subsection 4.4.3), while MOSEK is specialized in particular for QCQPs (cf. Subsection 4.4.4).
3. When using `socp` by LOBO et al. [57] for the computation of the search direction in Algorithm 3.20, the results are also practically the same with respect to Nit and Na — as long as `socp` did not fail to solve the search direction problem:
 - The most successful effort of stabilizing `socp` was achieved by adding a `bigeps`-termination criterion as it is used in SEDUMI by PÓLIK [86], STURM [102] (cf. Remark 4.26).
 - In our empirical experiments `socp` tended to be more reliable, when we chose δ_2 from (4.51) not too close to zero (e.g., δ_2 from (4.57) was a rather good choice).
 - We were not able to make `socp` more robust by improving the strict feasibility of the starting point by solving various linear programs that are obtained from the primal SOCP (4.1) and the dual SOCP (4.2) by exploiting the fact that $|x|_2 \leq |x|_1$ for all $x \in \mathbb{R}^n$ (`lp_solve` by BERKELAAR et al. [9], which is based on the revised simplex method and which we used for computing a solution of these linear programs, solved all of them easily).

At least when we used the variant of `socp` which was best for our purposes (i.e. `socp` with a `bigeps`-termination criterion and δ_2 from (4.57)) in Algorithm 3.20, then we were able to solve all examples that we took from the Hock-Schittkowski collection, while we were not able to achieve this for the other variants of `socp`. Furthermore, many examples of the nonlinearly constrained optimization problem (5.72) were not solvable by Algorithm 3.20 when using `socp` for the computation of the search direction (even when we used the best variant of `socp`).

6.2 Hock-Schittkowski Test-set

From Table A.3 in Appendix A, in which the results for the Hock-Schittkowski collection can be found and which is the basis for all plots in this section, we draw the following conclusions:

- To compare the solving time t_1 for the reduced algorithm (with MOSEK as (QC)QP-solver) and MPBNGC, we consider

	$t_1(\text{Red Alg})$	$t_2(\text{Red Alg})$	p_1	$t_1(\text{MPBNGC})$
HS	1198	961	0.80	1386
HS (*)	902	751	0.83	154

where we make use of (6.3) and in (*) we consider only those examples for which MPBNGC satisfied one of its termination criterions (cf. Subsection 6.3.5). Hence, for those examples of the Hock-Schittkowski collection for which MPBNGC was able to terminate successfully, MPBNGC is faster than the reduced algorithm. Furthermore, we notice that the reduced algorithm spent at least 80% of its time in the QCQP-solver, which is mostly overhead time in particular for the examples with lower dimension (which most examples are) as MOSEK has to, e.g., set up sparse matrix structures.

- The reduced algorithm needs approximately 65% of the solving time t_1 of the full algorithm. Nevertheless, SolvOpt only needs approximately 23% resp. 36% of the solving time t_1 of the full algorithm resp. the reduced algorithm. Not surprisingly, the full algorithm spent 80% of the time for solving the QCQPs (like the reduced algorithm did). Since SolvOpt terminated for the higher dimensional examples (i.e. the 15-dimensional examples 284, 285 and 384) with points that are not stationary, while both the full and the reduced algorithm were able to solve them, and since the reduced algorithm needs significantly less pure solving time than the full algorithm for these examples

ex	$t_1(\text{Full Alg})$	$t_1(\text{Red Alg})$	p_2
284	92	46	0.50
285	796	140	0.18
384	589	125	0.21

where

$$p_2 := \frac{t_1(\text{Red Alg})}{t_1(\text{Full Alg})},$$

we may expect that for more difficult examples the performance of the reduced algorithm increases with respect to t_1 (cf. Subsection 6.3.2 and Section 6.4).

Therefore, we will concentrate our comparison of Algorithm 3.20 (full and reduced version), MPBNGC and SolvOpt on the following qualitative aspects

- the cost c of the evaluations (solid line)
- the number of iterations Nit (dashed line; this comparison is only meaningful for the comparison between the full algorithm and the reduced algorithm),

where we use the two different line types for a better distinction of the comparisons in Figure 6.1, in this section, where before making detailed comparisons of our 58 examples, we give a short overview of them as a reason of clarity of the presentation: This yields the following summary table consisting of the number of examples for which the reduced algorithm is better than the full algorithm, MPBNGC resp. SolvOpt (and vice versa)

	no termi- nation	significantly better	better	a bit better	nearly equal	a bit better	better	significantly better
(Color code: Turquoise)	Full Alg					Red Alg		
Nit	0	2	1	3	51	0	1	0
<i>c</i>	0	4	1	7	40	1	1	4
(Color code: Violet)	MPBNGC					Red Alg		
<i>c</i>	5	2	3	10	16	8	7	7
(Color code: Orange)	SolvOpt					Red Alg		
<i>c</i>	3	3	4	3	1	31	9	4

that is visualized in Figure 6.1

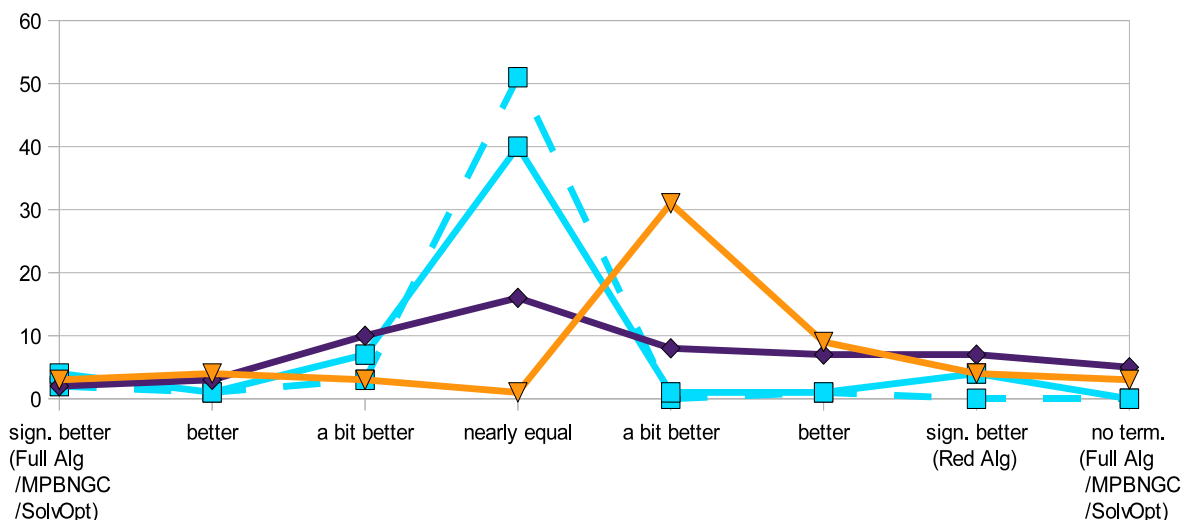


FIGURE 6.1: Hock-Schittkowski collection (summary)

and that let us draw the following conclusions:

- The performances of the full algorithm and the reduced algorithm are quite similar.
- The reduced algorithm is superior to MPBNGC in one third of the examples, for a further third of the examples one of these two solvers has only small advantages over the other, the performance differences between the two algorithms considered can be completely neglected for one quarter of the examples, and for the remaining ten percent of the examples MPBNGC beats the reduced algorithm clearly.
- The reduced algorithm is superior to SolvOpt in about one quarter of the examples, for sixty percent of the examples one of these two solvers has only small advantages over the other (in most cases the reduced algorithm is the slightly more successful one), and in the remaining twelve percent of the examples SolvOpt beats the reduced algorithm clearly.
- Furthermore, only the full algorithm and the reduced algorithm solved all examples successfully.

6.2.1 Reduced algorithm vs. Full algorithm

First of all, in the full algorithm t_0^k is only modified in 11 examples (34, 43, 66, 83, 100, 113, 227, 230, 264, 285, 384), while in the reduced algorithm this happens in 14 examples (the additional examples are 284, 330, 341). In all these examples t_0^k is only modified a few times and a modification only occurs at very early iterations of the optimization process (cf. Remark 3.26).

From Figure 6.2 and Figure 6.3 we conclude that the full and the reduced algorithm produce in most of the 58 examples approximately the same results — exceptions from this observation are in view of iterations the following 7 examples: The reduced algorithm is

- better in 1 example (33)

in comparison with the full algorithm, while the full algorithm is

- significantly better in 2 examples (34, 343)
- better in 1 example (66)
- a bit better in 3 examples (20, 113, 227)

in comparison with the reduced algorithm.

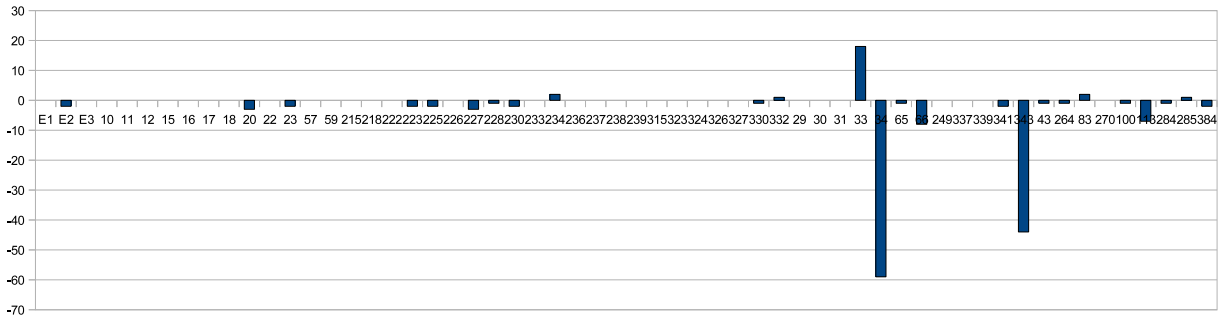


FIGURE 6.2: rp(Nit) for Red Alg & Full Alg

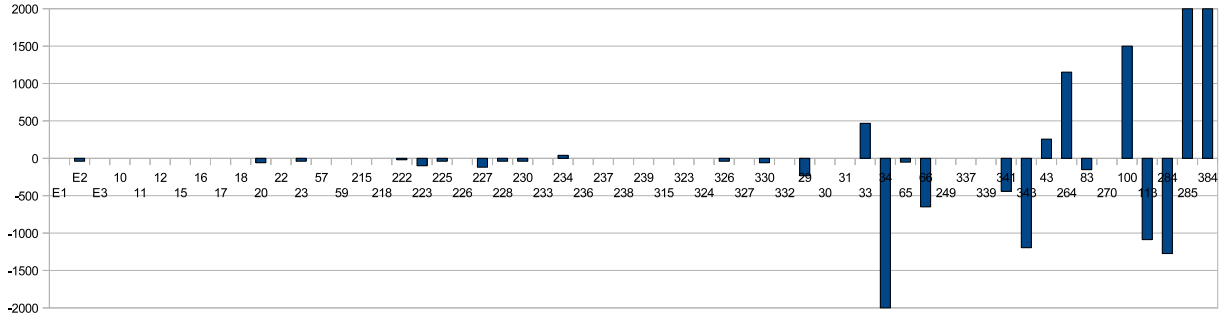
In view of costs the exceptions are given by the following 18 examples: The reduced algorithm is

- significantly better in 4 examples (100, 264, 285, 384)
- better in 1 example (33)
- a bit better in 1 example (43)

in comparison with the full algorithm, while the full algorithm is

- significantly better in 4 examples (34, 113, 284, 343)
- better in 1 example (66)
- a bit better in 7 examples (20, 29, 83, 223, 227, 330, 341)

in comparison with the reduced algorithm.

FIGURE 6.3: $rp(c)$ for Red Alg & Full Alg

6.2.2 Reduced algorithm vs. MPBNGC

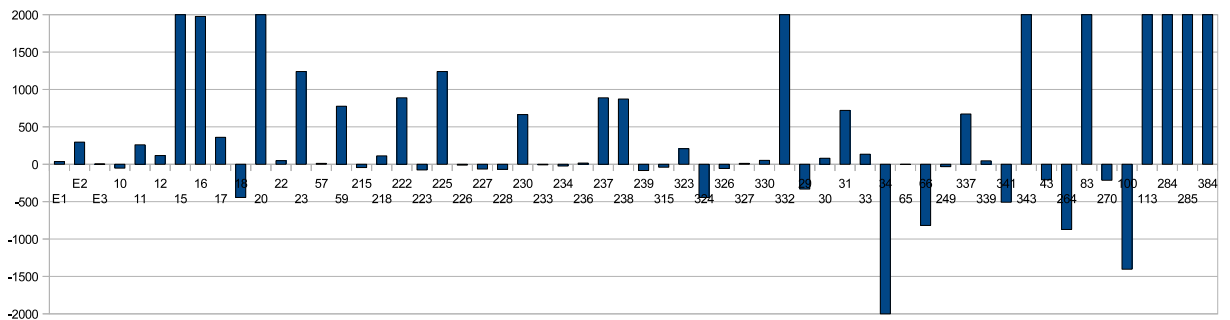
MPBNGC does not satisfy any of its termination criterions for five examples (15, 20, 83, 285, 384) within the given number of iterations and function evaluations. For the other 53 examples from Figure 6.4 we emphasize the following ones: The reduced algorithm is

- significantly better in 7 examples (16, 23, 113, 225, 284, 332, 343)
- better in 7 examples (31, 59, 222, 230, 237:238, 337)
- a bit better in 8 examples (E2, 11:12, 17, 30, 33, 218, 323)

in comparison with MPBNGC, while MPBNGC is

- significantly better in 2 examples (34, 100)
- better in 3 examples (66, 264, 341)
- a bit better in 10 examples (18, 29, 43, 223, 227:228, 239, 270, 324, 326)

in comparison with the reduced algorithm. In the remaining 16 examples the cost of the reduced algorithm and MPBNGC is practically the same.

FIGURE 6.4: $rp(c)$ for Red Alg & MPBNGC

6.2.3 Reduced algorithm vs. SolvOpt

SolvOpt terminates for the three 15-dimensional examples 284, 285 and 384 with points that are not stationary. For the other 55 examples from Figure 6.5 we emphasize the following ones: The reduced

algorithm is

- significantly better in 4 examples (18, 33, 270, 324)
- better in 9 examples (E3, 16, 215, 218, 223, 234, 237:238, 332)

in comparison with SolvOpt, while SolvOpt is

- significantly better in 3 examples (34, 66, 100)
- better in 4 examples (43, 83, 264, 343)
- a bit better in 3 examples (29, 113, 341)

in comparison with the reduced algorithm. Except for example 233 in which the cost of the reduced algorithm and SolvOpt are practically the same, in all 31 remaining examples the reduced algorithm is a bit better than SolvOpt.

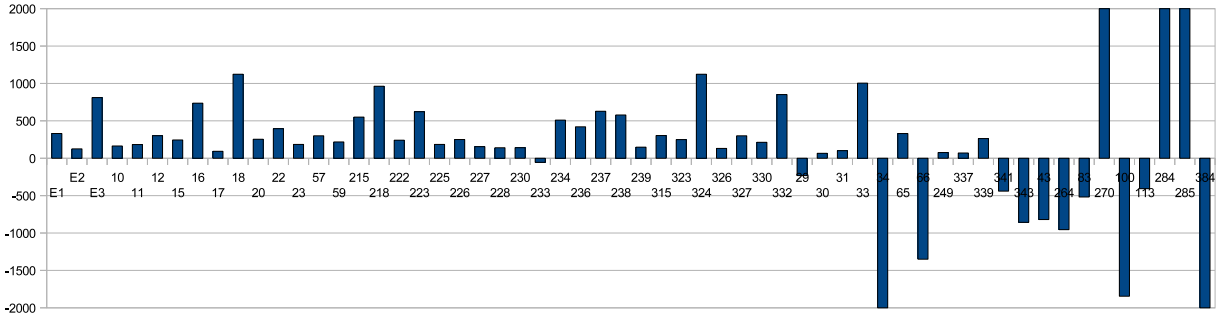


FIGURE 6.5: rp(c) for Red Alg & SolvOpt

6.3 Exclusion boxes

6.3.1 Notation

We will make the following tests concerning exclusion boxes:

- We give results for the linearly constrained optimization problem (5.70) with a fixed box (i.e. without optimizing u and v) for dimensions between 4 and 11 in Subsection 6.3.3.
- We give results for the linearly constrained optimization problem (5.70) with a variable box (i.e. with optimizing u and v) for dimensions between 8 and 21 in Subsection 6.3.4.
- We give results for the nonlinearly constrained optimization problem (5.72) for dimension 8 in Subsection 6.3.5, where we use b_+^1 from (5.76) as the objective function.

The underlying data for these nonsmooth optimization problems was extracted from real CSPs that occur in GloptLab by DOMES [18]. Apart from u and v , we will concentrate on the optimization of the variables y and z due to the large number of tested examples (cf. Subsubsection 6.3.2), and since the additional optimization of R and S did not have much impact on the quality of the results which

was discovered in additional empirical observations, where a detailed analysis of these observations goes beyond the scope of this thesis. Furthermore, we will make our tests for the two different choices of the function T from (5.68) and (5.69), which occurs in the denominator of the certificate f from (5.37), where for the latter one f is only defined outside of the zero set of T which has measure zero (cf. Example 5.23).

We represent the information on the feasibility of example i that is obtained by GloptLab in the following way (cf. Appendix A):

ex	Status of example i as identified by GloptLab
$\textcolor{red}{i}$	infeasible
$\textcolor{blue}{i}$	feasible
i	No statement on (in)feasibility can be made
$\textcolor{green}{i}$	A local minimizer is found
$\textcolor{brown}{i}$	Maximum number of iterations is exceeded

We represent the status of a point concerning certificate f from (5.37) after performing Nit iterations of a nonsmooth solver in the linearly constrained case (cf. Subsection 6.3.3 and Subsection 6.3.4) by

Nit	Status after Nit iterations (linearly constrained case)
$\textcolor{red}{k}$	A point with $f < 0$ is found
k	A stationary point with $f \geq 0$ is found
$\textcolor{blue}{k}$	A feasible point is found (of course with $f \geq 0$)
$\textcolor{brown}{k}$	Maximum number of iterations is exceeded (without (in)feasibility statement)

and in the nonlinearly constrained case (cf. Subsection 6.3.5) by

Nit	Status after Nit iterations (nonlinearly constrained case)
$\textcolor{red}{k}$	A point which yields a sufficiently large exclusion box is found
k	A stationary point is found where the exclusion box is not large enough
$\textcolor{brown}{k}$	Maximum number of iterations is exceeded (without finding a sufficiently large exclusion box or being able to make a stationarity statement)

where an exclusion box is sufficiently large in the nonlinearly constrained case, if the objective function satisfies

$$b_+^1(y, z, R, S, u, v) \leq 10^{-6} .$$

Remark 6.7. Because SolvOpt cannot distinguish between linear and nonlinear constraints (cf. KUNTSEVICH & KAPPEL [54, p. 15]), the linear constraints of the optimization problems (5.70) and (5.72) must be formulated as nonlinear constraints in SolvOpt. Nevertheless, we will not include the number of these evaluations in the computation of the cost c from (6.4) for the mentioned optimization problems in Subsection 6.3.3 and Subsection 6.3.4, since these evaluations may be considered as easy in comparison to the evaluation of the certificate f from (5.37) which is the objective function in these optimization problems.

6.3.2 Overview of the results

We compare the total time t_1 of the solution process, where we used the reduced algorithm (with MOSEK as the (QC)QP-solver) in the constrained case: From Tables A.4–A.7 (s. Appendix A) we obtain

	$t_1(\text{Red Alg})$	$t_2(\text{Red Alg})$	p_1	$t_1(\text{MPBNGC})$	$t_1(\text{SolvOpt})$
$T = 1$					
Linearly constrained (fixed box)	1477	215	0.85	231	2754
Linearly constrained (variable box)	782	60	0.92	30	1546
Nonlinearly constrained	25420	4885	0.81	21860	38761
Nonlinearly constrained (*)	19053	3723	0.80	2067	30312
$T = y _2$					
Linearly constrained (fixed box)	1316	129	0.90	15	1508
Linearly constrained (variable box)	797	45	0.94	30	2263
Nonlinearly constrained	24055	4284	0.82	25383	16909
Nonlinearly constrained (*)	18038	3112	0.83	3719	12635

where we make use of (6.3) and in (*) we consider only those examples for which MPBNGC satisfied one of its termination criterions (cf. Subsection 6.3.5).

For the linearly constrained problems MPBNGC was the fastest of the tested algorithms, followed by BNLC and SolvOpt. If we consider only those nonlinearly constrained examples for which MPBNGC was able to terminate successfully, MPBNGC was the fastest algorithm again. Considering the competitors, for the nonlinearly constrained problems with $T = 1$ the reduced algorithm is 13.3 seconds resp. 11.3 seconds faster than SolvOpt, while for the nonlinearly constrained problems with $T = |y|_2$ SolvOpt is 7.1 seconds resp. 5.4 seconds faster than the reduced algorithm.

Again (cf. Section 6.2), taking a closer look at p_1 yields the observation that at least 85% of the time is consumed by solving the QP (in the linearly constrained case) resp. at least 80% of the time is consumed by solving the QCQP (in the nonlinearly constrained case), which implies that the difference in the percentage between the QP and the QCQP is small in particular (an investigation of the behavior of the solving time t_1 for higher dimensional problems can be found in Section 6.4).

Therefore, we will concentrate in Subsection 6.3.3, Subsection 6.3.4 and Subsection 6.3.5 on the comparison of qualitative aspects between Algorithm 3.20, MPBNGC and SolvOpt (like, e.g., the cost c of the evaluations), where before making these detailed comparisons, we give a short overview of them as a reason of clarity of the presentation: In both cases

- $T = 1$ (solid line)
- $T = |y|_2$ (dashed line),

where we use the two different line types for a better distinction of the comparisons in Figures 6.6, 6.7, and 6.8, we tested

- 128 linearly constrained examples with a fixed box
- 117 linearly constrained examples with a variable box
- 201 nonlinearly constrained examples,

which yields the following two summary tables consisting of the number of examples for which Algorithm 3.20 (BNLC resp. the reduced algorithm) is better than MPBNGC resp. SolvOpt (and vice versa) with respect to the cost c of the evaluations

(Color code: Violet)	MPBNGC					BNLC/Red Alg		
	no termination	significantly better	better	a bit better	nearly equal	a bit better	better	significantly better
	$T = 1$							
Linearly constrained (fixed box)	0	2	5	12	106	2	0	1
Linearly constrained (variable box)	0	0	0	1	116	0	0	0
Nonlinearly constrained	32	6	28	89	31	10	2	3
	$T = y _2$							
Linearly constrained (fixed box)	0	2	5	30	91	0	0	0
Linearly constrained (variable box)	0	0	0	5	112	0	0	0
Nonlinearly constrained	43	4	28	59	30	15	14	8

(Color code: Orange)	SolvOpt					BNLC/Red Alg		
	no termination	significantly better	better	a bit better	nearly equal	a bit better	better	significantly better
	$T = 1$							
Linearly constrained (fixed box)	0	1	3	0	61	25	13	25
Linearly constrained (variable box)	0	0	0	0	48	37	24	8
Nonlinearly constrained	0	0	14	20	21	76	20	50
	$T = y _2$							
Linearly constrained (fixed box)	0	1	2	1	32	34	49	9
Linearly constrained (variable box)	0	0	0	5	41	32	19	20
Nonlinearly constrained	0	2	24	26	31	61	45	12

that are visualized in Figures 6.6, 6.7, and 6.8

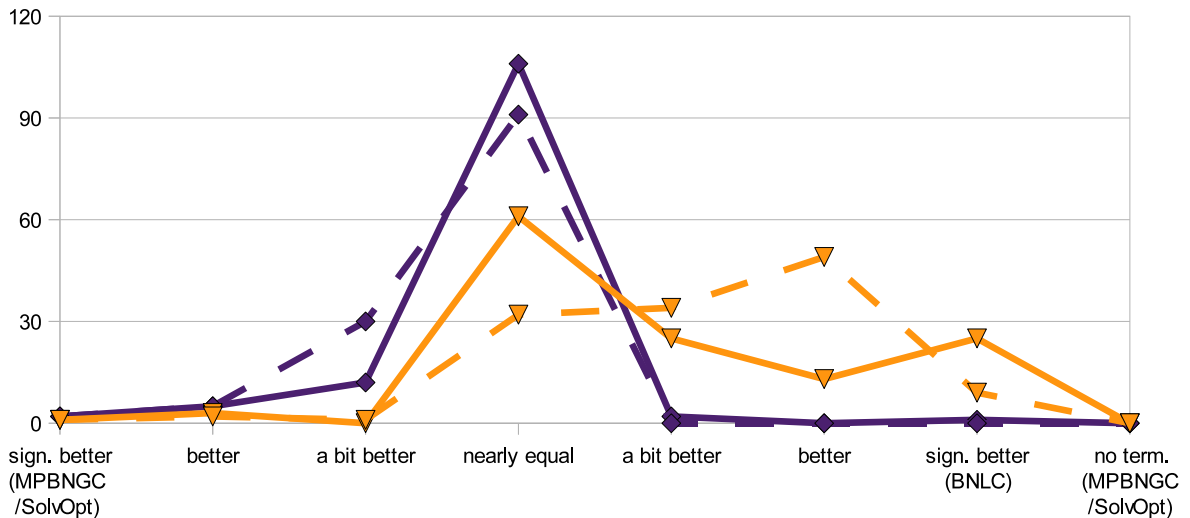


FIGURE 6.6: Linearly constrained — fixed box (summary)

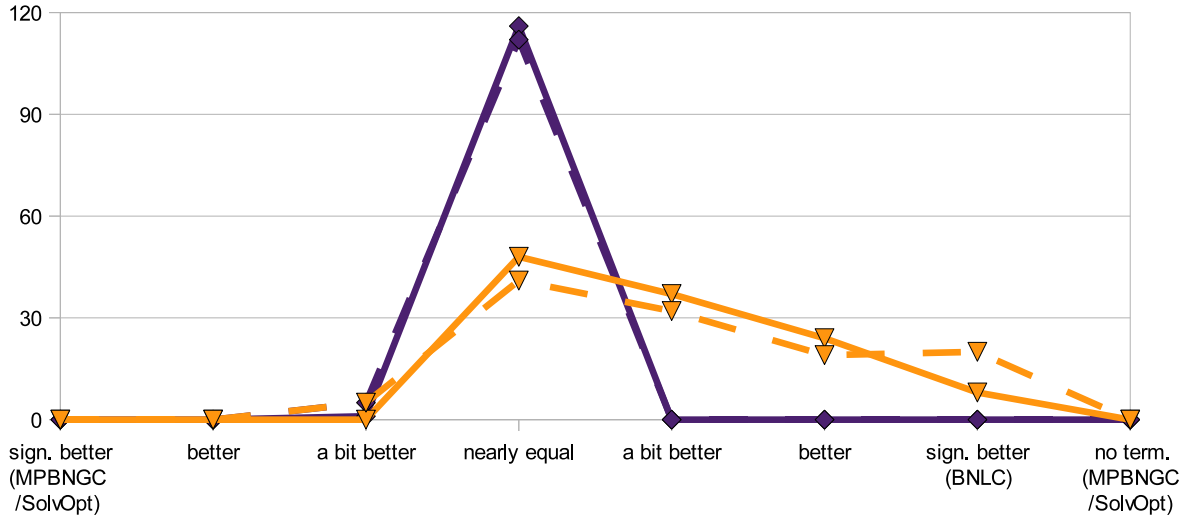


FIGURE 6.7: Linearly constrained — variable box (summary)

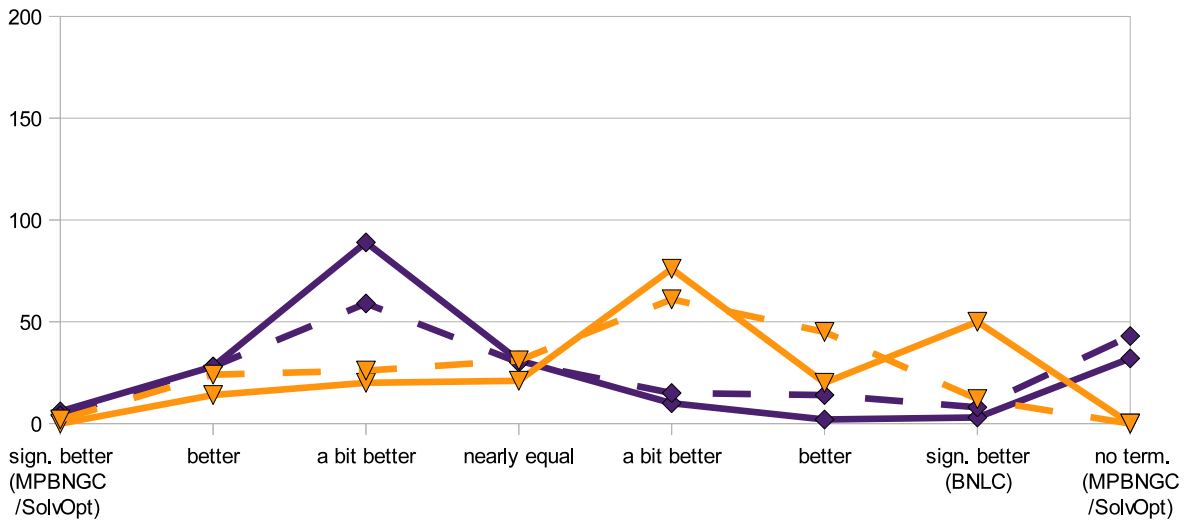


FIGURE 6.8: Nonlinearly constrained (summary)

and that let us draw the following conclusions:

- Reduced algorithm & MPBNGC:
 - The performance differences between BNLC and MPBNGC can be neglected for the largest part of the linearly constrained examples (with small advantages for MPBNGC in about ten percent of these examples).
 - For the nonlinearly constrained examples the reduced algorithm is superior to MPBNGC in one quarter of the examples, for forty percent of the examples one of these two solvers has small advantages over the other (in most cases MPBNGC is the slightly more successful one), the performance differences between the two algorithms considered can be completely neglected for fifteen percent of the examples, and for further fifteen percent of the examples MPBNGC beats the reduced algorithm clearly.
- Reduced algorithm & SolvOpt:
 - For the linearly constrained examples BNLC is superior to SolvOpt in one third of the

examples, for one quarter of the examples one of these two solvers has small advantages over the other (in nearly all cases BNLC is the slightly more successful one), the performance differences between the two algorithms considered can be completely neglected for forty percent of the examples, and in only one percent of the examples SolvOpt beats the reduced algorithm clearly.

- For the nonlinearly constrained examples the reduced algorithm is superior to SolvOpt in one third of the examples, for 45 percent of the examples one of these two solvers has small advantages over the other (the reduced algorithm is often the slightly more successful one), the performance differences between the considered two algorithms can be completely neglected for ten percent of the examples, and in the remaining ten percent of the examples SolvOpt beats the reduced algorithm clearly.
- In contrast to the linearly constrained case, in which all three solvers terminated successfully for all examples, only the reduced algorithm and SolvOpt were able to attain this goal in the nonlinearly constrained case, too.

6.3.3 Linearly constrained case (fixed box)

We took 310 examples from real CSPs that occur in GloptLab. We observe that

- for 79 examples (1:20, 23, 25:56, 63:66, 69, 75:90, 95, 119, 167, 169, 206) the starting point is feasible for the CSP and
- for 103 examples (105, 107, 111, 124:127, 131:133, 136:137, 141, 143:145, 148:150, 153, 155, 158, 174:188, 194, 196, 199, 202, 211:213, 216:217, 221:224, 227:232, 235:237, 241:243, 245:254, 256:258, 262:266, 272:277, 283, 287:288, 292, 296, 298:309) the evaluation of the certificate at the starting point identifies the box as infeasible

and hence there remain 128 test problems.

6.3.3.1 BNLC vs. MPBNGC

In the case $T = 1$ we conclude from Figure 6.9 that BNLC is

- significantly better in 1 example (168)
- a bit better in 2 examples (97, 280)

in comparison with MPBNGC, while MPBNGC is

- significantly better in 2 examples (121, 291)
- better in 5 examples (157, 238, 281, 285, 293)
- a bit better in 12 examples (100, 103, 110, 128, 156, 165, 189:190, 192, 198, 203, 310)

in comparison with BNLC. In the 106 remaining examples the costs of BNLC and MPBNGC are

practically the same.

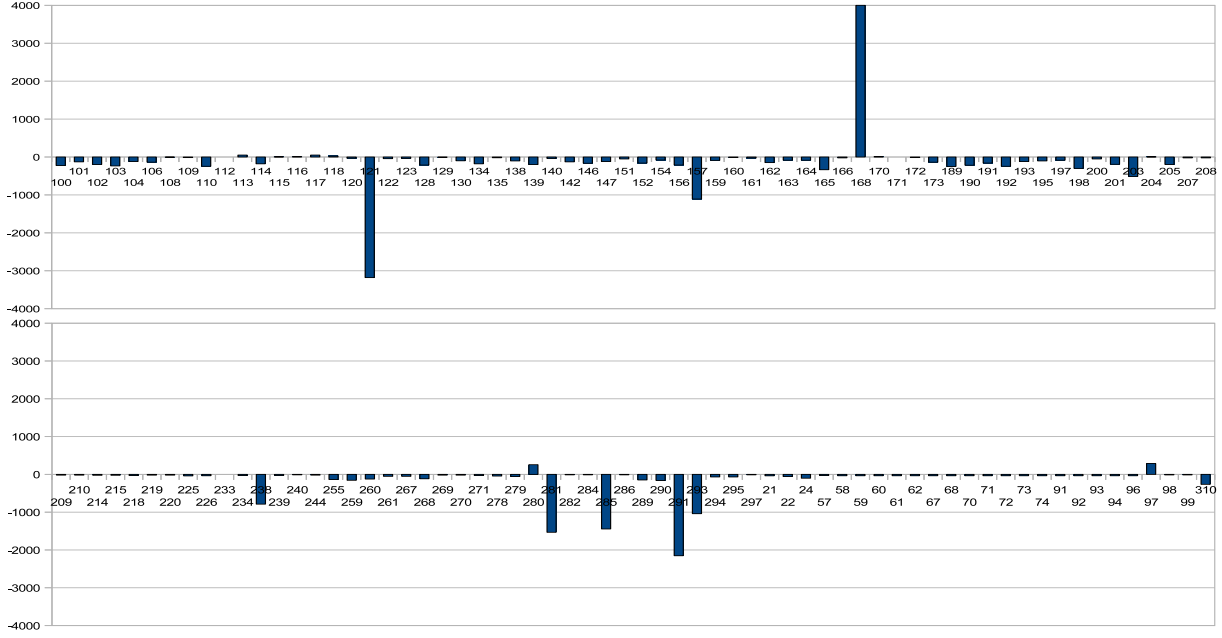


FIGURE 6.9: $\text{rp}(c)$ for BNLC & MPBNGC ($T = 1$)

In the case $T = |y|_2$ it follows from Figure 6.10 that MPBNGC is

- significantly better in 2 examples (122, 140)
- better in 5 examples (160, 171, 225, 293, 310)
- a bit better in 30 examples (100:103, 106, 115, 121, 123, 128, 134, 147, 152, 154, 157, 159, 161, 166, 168, 170, 172, 189, 195, 205, 209:210, 281:282, 284:285, 291)

in comparison with BNLC. In the 91 remaining examples the costs of BNLC and MPBNGC are practically the same.

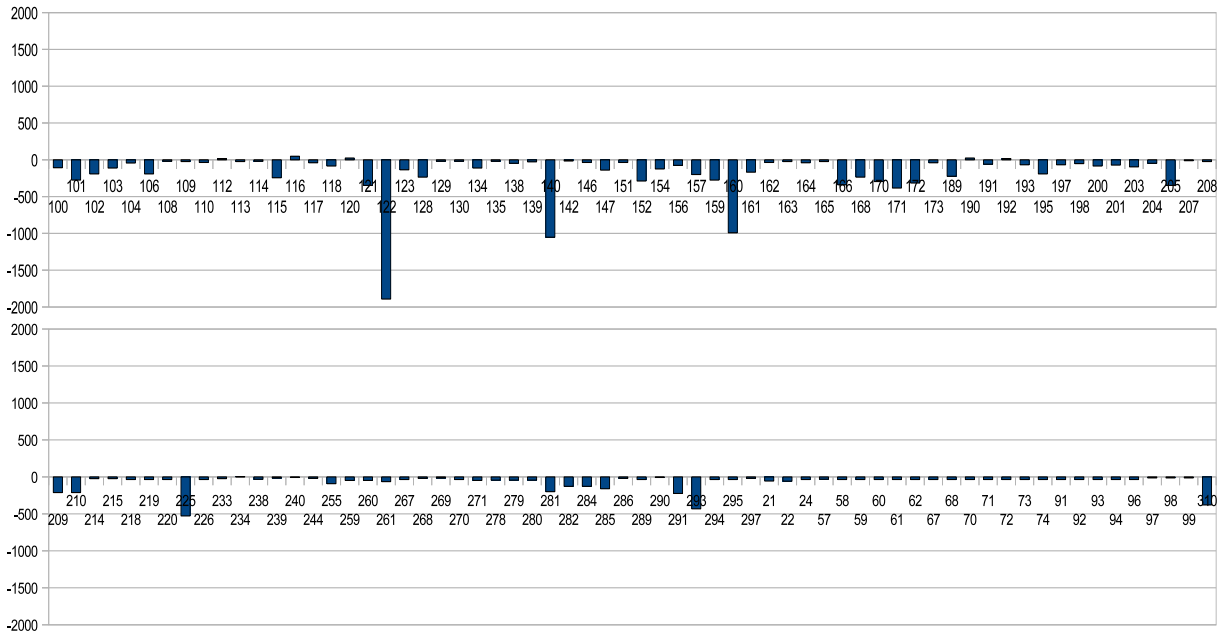


FIGURE 6.10: $\text{rp}(c)$ for BNLC & MPBNGC ($T = |y|_2$)

6.3.3.2 BNLC vs. SolvOpt

In the case $T = 1$ we conclude from Figure 6.11 that BNLC is

- significantly better in 25 examples (59:61, 70, 108, 110, 112, 116, 123, 146, 151, 154, 156, 159, 207, 238, 255, 267:268, 270, 279, 289:290, 293:294)
- better in 13 examples (22, 24, 68, 72:73, 114, 147, 152, 165, 191, 203, 239, 310)
- a bit better in 25 examples (62, 71, 94, 109, 113, 117:118, 120, 138, 140, 142, 162, 168, 193, 195, 200, 233, 259, 261, 269, 271, 278, 280, 291, 295)

in comparison with SolvOpt, while SolvOpt is

- significantly better in 1 example (121)
- better in 3 examples (157, 281, 285)

in comparison with BNLC. In the 61 remaining examples the costs of BNLC and SolvOpt are practically the same.

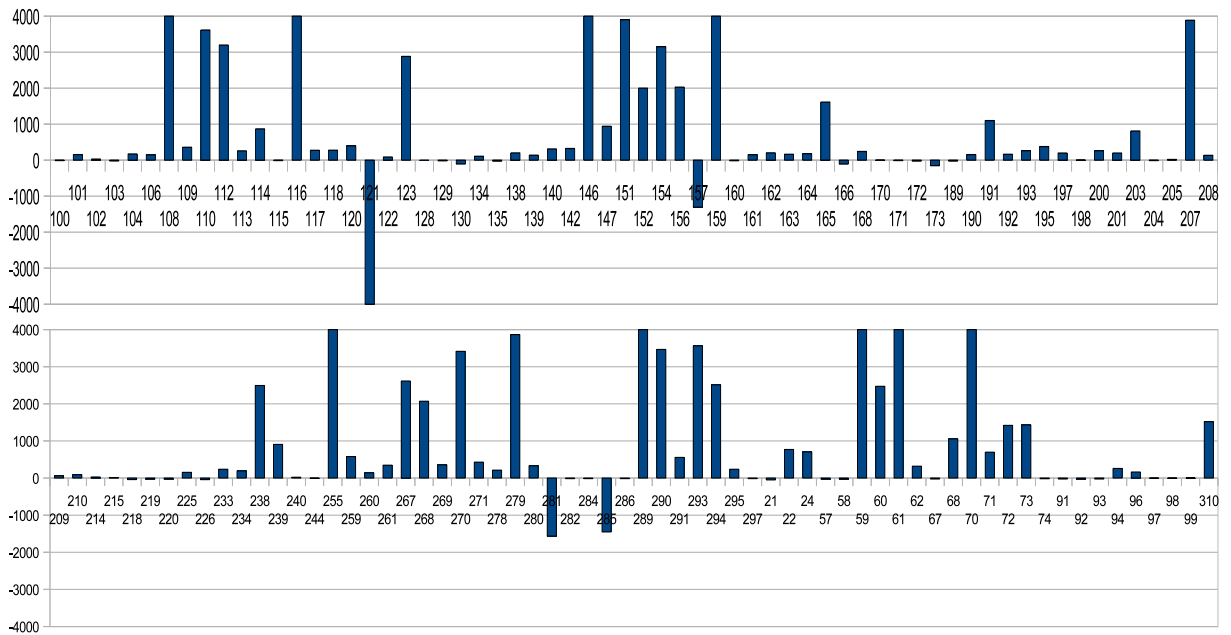


FIGURE 6.11: $rp(c)$ for BNLC & SolvOpt ($T = 1$)

In the case $T = |y|_2$ it follows from Figure 6.12 that BNLC is

- significantly better in 9 examples (24, 59, 62, 68, 70, 96, 134, 208, 310)
- better in 49 examples (22, 60:61, 71, 73, 92, 102, 104, 108:109, 112:114, 117, 120, 138:139, 142, 154, 156, 161:162, 164, 189:193, 197, 200:201, 209, 233:234, 238, 255, 259:261, 267, 269:271, 278:280, 289:290, 295)
- a bit better in 34 examples (57, 72, 100:101, 110, 115:116, 118, 121, 123, 128, 146:147, 151:152, 157, 159, 163, 165:166, 168, 172, 195, 198, 203:204, 207, 210, 239, 268, 281, 285, 291, 294)

in comparison with SolvOpt, while SolvOpt is

- significantly better in 1 example (122)
- better in 2 examples (160, 225)
- a bit better in 1 example (140)

in comparison with BNLC. In the 32 remaining examples the costs of BNLC and SolvOpt are practically the same.

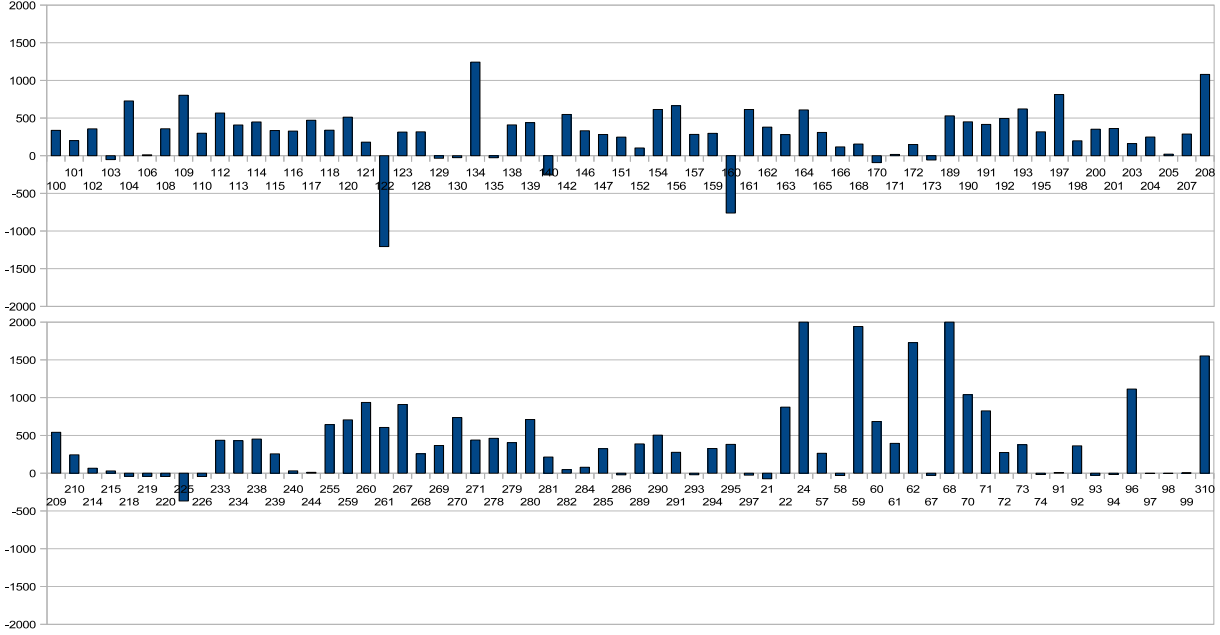


FIGURE 6.12: $rp(c)$ for BNLC & SolvOpt ($T = |y|_2$)

6.3.4 Linearly constrained case (variable box)

We observe that

- for 80 examples (1:20, 23, 25:56, 63:66, 69, 75:90, 95, 97, 119, 167, 169, 206) the starting point is feasible for the CSP and
- for 113 examples (105, 107, 111, 124:127, 131:133, 136:137, 141, 143:145, 148:150, 153, 155, 158, 173:188, 194, 196, 199, 202, 211:217, 220:224, 226:232, 235:237, 241:254, 256:258, 262:266, 271:277, 282:284, 287:288, 292, 296:309) the evaluation of the certificate at the starting point identifies the boxes as infeasible

and hence there remain 117 test problems of the 310 original examples from GloptLab.

6.3.4.1 BNLC vs. MPBNGC

In the case $T = 1$ we conclude from Figure 6.13 that MPBNGC is

- a bit better in 1 example (310)

in comparison with BNLC. In the 116 remaining examples the costs of BNLC and MPBNGC are practically the same.

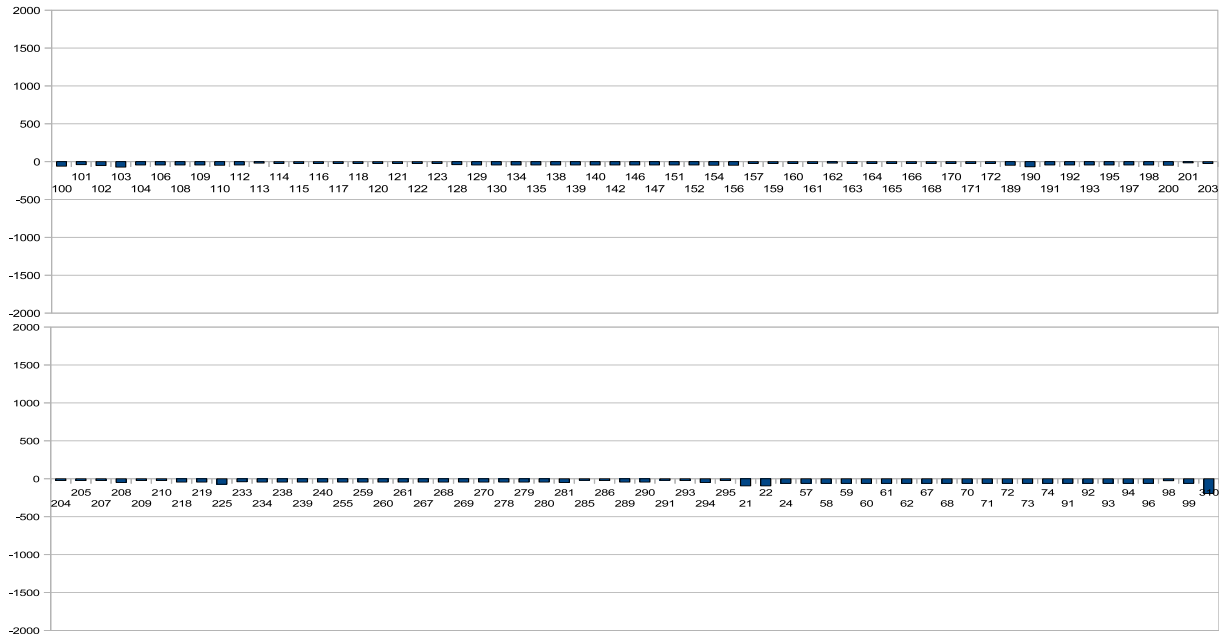


FIGURE 6.13: $rp(c)$ for BNLC & MPBNGC ($T = 1$)

In the case $T = |y|_2$ it follows from Figure 6.14 that MPBNGC is

- a bit better in 5 examples (21, 101:102, 189, 310)

in comparison with BNLC. In the 112 remaining examples the costs of BNLC and MPBNGC are practically the same.

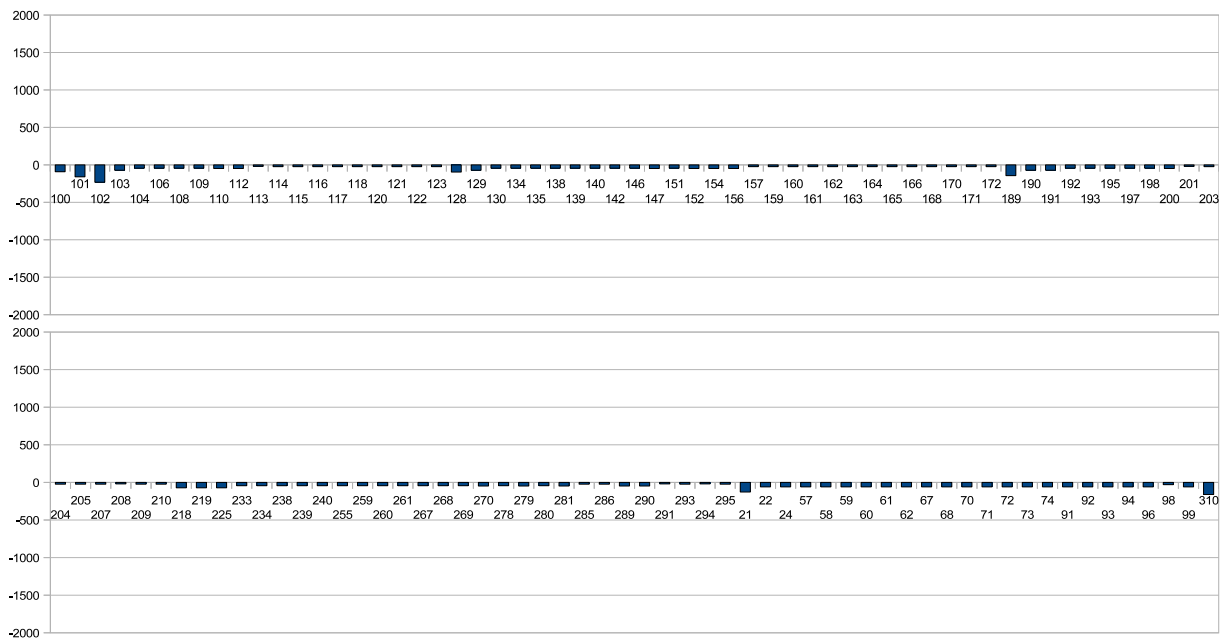


FIGURE 6.14: $rp(c)$ for BNLC & MPBNC ($T = |y|_2$)

6.3.4.2 BNLC vs. SolvOpt

In the case $T = 1$ we conclude from Figure 6.15 that BNLC is

- significantly better in 8 examples (72, 116, 121, 164:165, 203, 205, 207)
- better in 24 examples (24, 59, 61, 68, 70:71, 73, 96, 114, 117:118, 120, 142, 146, 156, 161, 163, 166, 168, 170:171, 192, 295, 310)
- a bit better in 37 examples (58, 60, 62, 99, 106, 109:110, 112:113, 115, 122, 147, 151:152, 154, 157, 159, 162, 197:198, 200:201, 204, 208:210, 261, 268:270, 279:280, 285, 289, 291, 293:294)

in comparison with SolvOpt. In the 48 remaining examples the costs of BNLC and SolvOpt are practically the same.

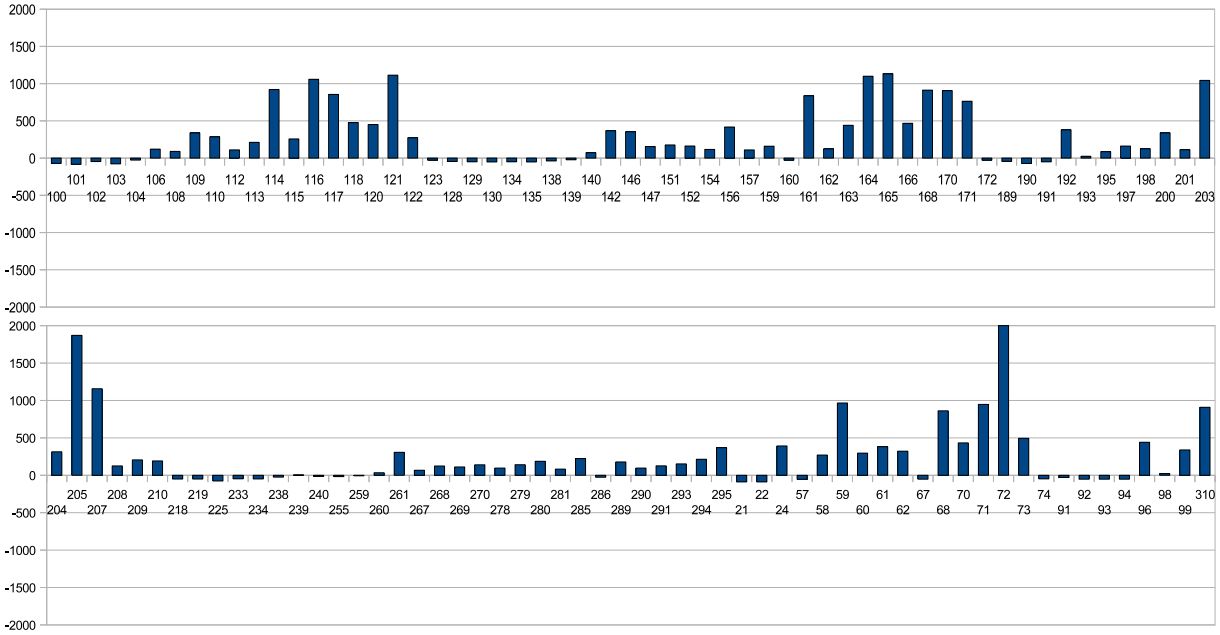


FIGURE 6.15: $rp(c)$ for BNLC & SolvOpt ($T = 1$)

In the case $T = |y|_2$ it follows from Figure 6.16 that BNLC is

- significantly better in 20 examples (57, 61, 70, 73, 115:116, 118, 121:122, 156, 161:163, 165:166, 168, 170:171, 203, 207)
- better in 19 examples (24, 58, 62, 71, 96, 110, 114, 117, 120, 159, 164, 192, 195, 197, 204:205, 239, 295, 310)
- a bit better in 32 examples (60, 72, 99, 109, 113, 123, 142, 151:152, 154, 157, 160, 172, 198, 200:201, 208:210, 268:270, 278:281, 285:286, 290:291, 293:294)

in comparison with SolvOpt, while SolvOpt is

- a bit better in 5 examples (21, 101:102, 128, 189)

in comparison with BNLC. In the 41 remaining examples the costs of BNLC and SolvOpt are practically

the same.

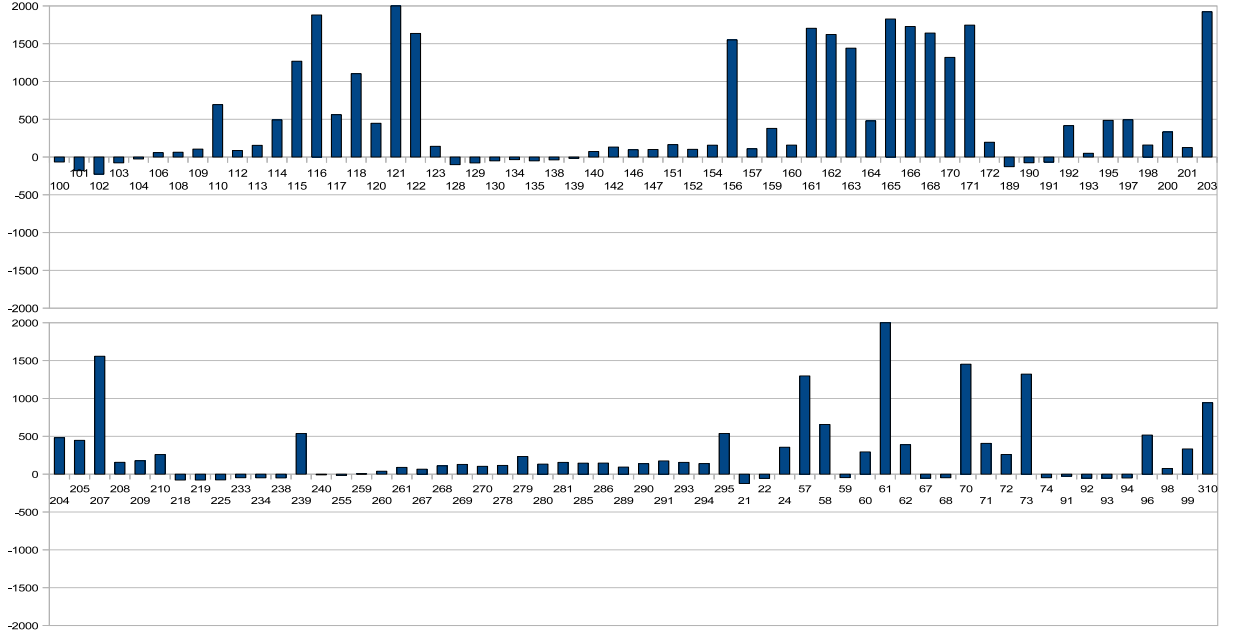


FIGURE 6.16: $\text{rp}(c)$ for BNLC & SolvOpt ($T = |y|_2$)

6.3.5 Nonlinearly constrained case

Since we were not able to find a starting point, i.e. an infeasible sub-box, for 109 examples (1:99, 117:120, 164, 167, 169, 205:206, 310), we exclude them from the following tests for which there remain 201 examples of the 310 original examples from GloptLab.

6.3.5.1 Reduced algorithm vs. MPBNGC

In the case $T = 1$ MPBNGC does not satisfy any of its termination criterions for 32 examples (107, 109, 124:125, 132:133, 145, 147, 155, 171, 183, 208, 214:216, 239, 248, 262:263, 268, 271, 273, 275, 277:278, 283:285, 290, 292, 303:304) within the given number of iterations and function evaluations (also cf. Subsection 6.3.1). For the remaining 169 examples we conclude from Figure 6.17 that the reduced algorithm is

- significantly better in 3 examples (156, 172, 281)
- better in 2 examples (234, 287)
- a bit better in 10 examples (110:111, 185, 201, 240, 289, 291, 294:296)

in comparison with MPBNGC, while MPBNGC is

- significantly better in 6 examples (274, 276, 280, 288, 297, 299)
- better in 28 examples (101, 106, 108, 116, 121:122, 129, 146, 161, 165:166, 168, 170, 192, 195, 200, 226, 232, 250:251, 253:254, 269:270, 272, 298, 300:301)

- a bit better in 89 examples (100, 102:105, 113, 123, 126:127, 136:137, 139:142, 148:149, 152:154, 162:163, 173, 175:180, 186:190, 193, 196:197, 199, 202, 209:211, 213, 217:225, 227:231, 233, 235:238, 241:243, 245:246, 249, 252, 255:261, 264:267, 279, 282, 293, 302, 305:309)

in comparison with the reduced algorithm, and in 31 examples the costs of the reduced algorithm and MPBNGC are practically the same.

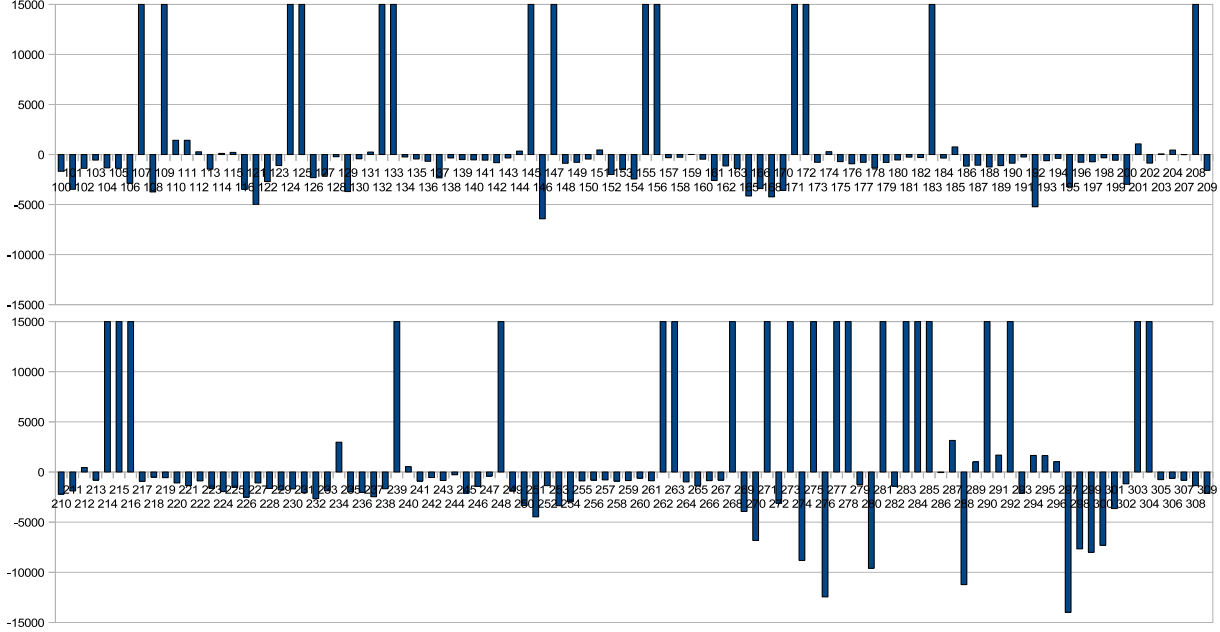


FIGURE 6.17: $rp(c)$ for Red Alg & MPBNGC ($T = 1$)

In the case $T = |y|_2$ MPBNGC does not satisfy any of its termination criterions for 43 examples (106, 109, 122, 125, 129, 131:133, 136, 144, 149:150, 152:153, 155, 161, 182:183, 191, 200:203, 208, 212:214, 217:218, 241, 243:244, 252, 269:270, 274, 276, 281, 283, 291, 301, 305, 308) within the given number of iterations and function evaluations. For the remaining 158 examples it follows from Figure 6.18 that the reduced algorithm is

- significantly better in 8 examples (100, 110, 112, 147, 166, 275, 280, 282)
- better in 14 examples (111, 115, 124, 170, 172:173, 176, 204, 277, 279, 284, 289, 296:297)
- a bit better in 15 examples (113:114, 123, 148, 156, 163, 174:175, 177, 210, 285, 287, 290, 293, 295)

in comparison with MPBNGC, while MPBNGC is

- significantly better in 4 examples (154, 207, 240, 271)
- better in 28 examples (101, 107, 121, 127, 146, 165, 168, 215:216, 221, 224, 226:227, 242, 245, 247, 249:251, 253:254, 272:273, 288, 298:300, 309)
- a bit better in 59 examples (102:103, 108, 116, 126, 128, 130, 134:135, 137:138, 141:143, 145, 157:160, 180, 184, 186:190, 195:196, 209, 211, 219:220, 222:223, 225, 228:233, 235:239, 246, 248,

261:268, 278, 304, 306)

in comparison with the reduced algorithm, and in 30 examples the costs of the reduced algorithm and MPBNGC are practically the same.

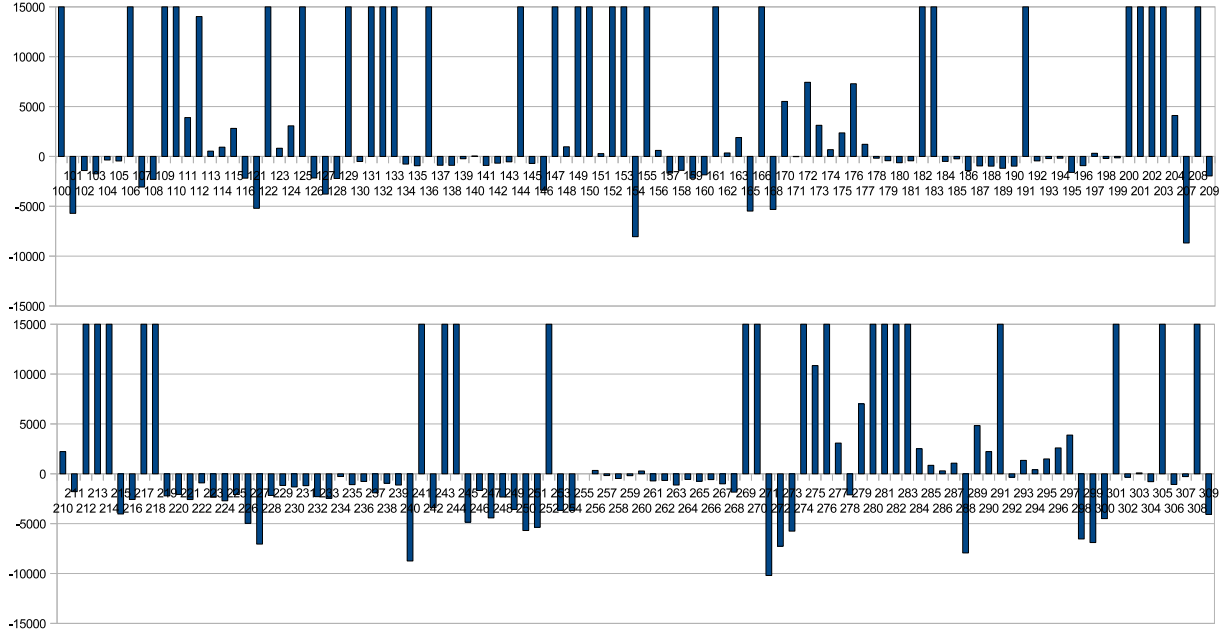


FIGURE 6.18: $rp(c)$ for Red Alg & MPBNGC ($T = |y|_2$)

6.3.5.2 Reduced algorithm vs. SolvOpt

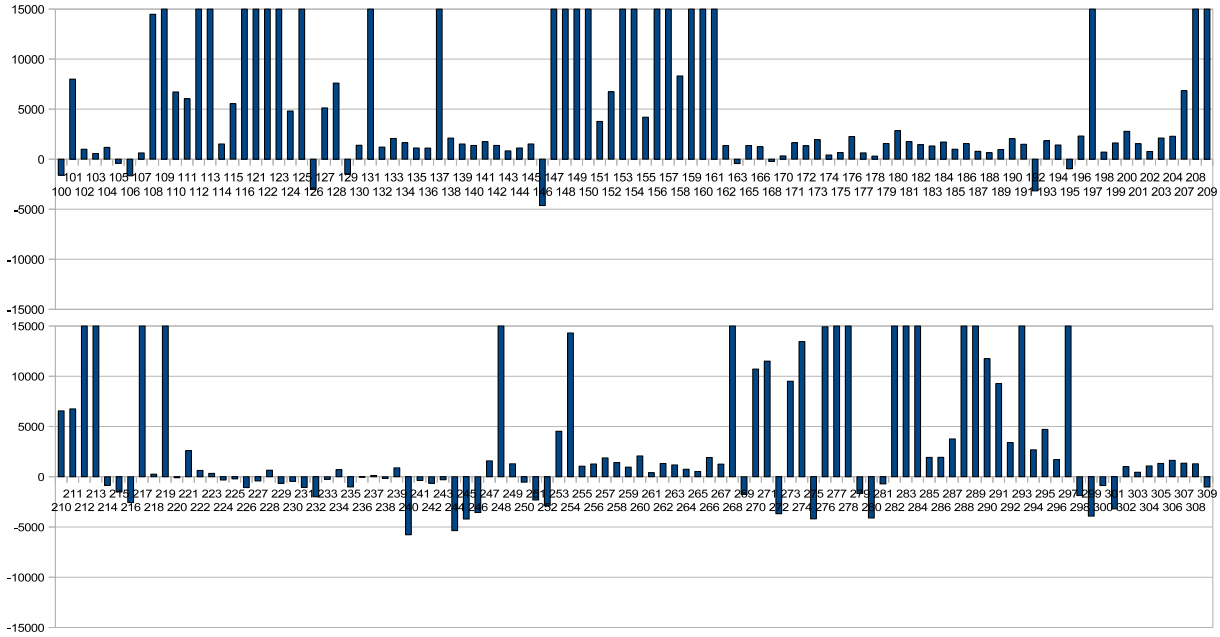
In the case $T = 1$ we conclude from Figure 6.19 that the reduced algorithm is

- significantly better in 50 examples (101, 108:109, 112:113, 116, 121:123, 125, 131, 137, 147:150, 153:154, 156:161, 197, 208:209, 212:213, 217, 219, 248, 254, 268, 270:271, 273:274, 276:278, 282:284, 288:291, 293, 297)
- better in 20 examples (110:111, 115, 124, 127:128, 151:152, 155, 180, 200, 207, 210:211, 221, 253, 287, 292, 294:295)
- a bit better in 76 examples (102:104, 107, 114, 130, 132:136, 138:145, 162, 165:166, 171:173, 175:177, 179, 181:191, 193:194, 196, 198:199, 201:204, 222, 228, 234, 239, 247, 249, 255:260, 262:267, 285:286, 296, 302, 304:308)

in comparison with SolvOpt, while SolvOpt is

- better in 14 examples (126, 146, 192, 216, 240, 244:246, 252, 272, 275, 280, 299, 301)
- a bit better in 20 examples (100, 106, 129, 195, 214:215, 226, 229, 231:232, 235, 242, 250:251, 269, 279, 281, 298, 300, 309)

in comparison with the reduced algorithm. In the 21 remaining examples the costs of the reduced algorithm and SolvOpt are practically the same.

FIGURE 6.19: $rp(c)$ for Red Alg & SolvOpt ($T = 1$)

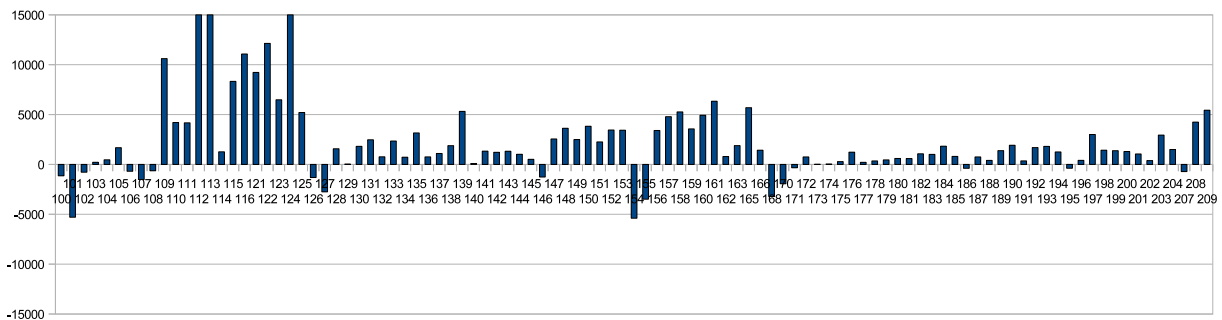
In the case $T = |y|_2$ it follows from Figure 6.20 that the reduced algorithm is

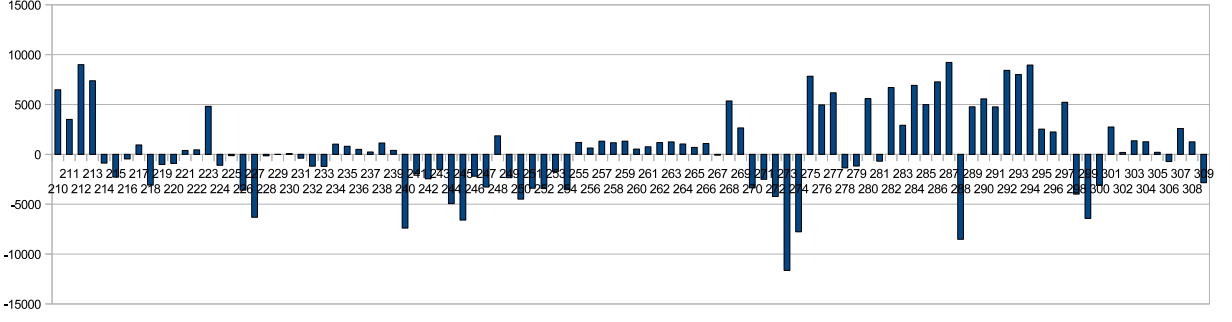
- significantly better in 12 examples (109, 112:113, 115:116, 121:122, 124, 212, 287, 292, 294)
- better in 45 examples (110:111, 123, 125, 135, 139, 147:148, 150, 152:153, 156:161, 165, 197, 203, 208:211, 213, 223, 268:269, 275:277, 280, 282:286, 289:291, 293, 295, 297, 301, 307)
- a bit better in 61 examples (105, 114, 128, 130:134, 136:138, 141:145, 149, 151, 162:163, 166, 172, 176, 180:185, 187, 189:190, 192:194, 198:201, 204, 217, 234:235, 238, 248, 255:266, 296, 303:304, 308)

in comparison with SolvOpt, while SolvOpt is

- significantly better in 2 examples (273, 288)
- better in 24 examples (101, 127, 154:155, 168, 218, 226:227, 240, 244:245, 247, 250:252, 254, 270:272, 274, 298:300, 309)
- a bit better in 26 examples (100, 102, 106:108, 126, 146, 170, 207, 214:215, 219:220, 224, 232:233, 241:243, 246, 249, 253, 278:279, 281, 306)

in comparison with the reduced algorithm. In the 31 remaining examples the costs of the reduced algorithm and SolvOpt are practically the same.



FIGURE 6.20: $\text{rp}(c)$ for Red Alg & SolvOpt ($T = |y|_2$)

6.4 Higher dimensional piecewise quadratic examples

We want to give numerical results for the nonsmooth optimization problem (2.35) with

$$f(x) := \max_{i=1,\dots,m_1} f_i(x)$$

$$F(x) := \max_{j=1,\dots,m_2} F_j(x),$$

where

$$f_i(x) := \alpha_i + a_i^T(x - x_i) + \frac{1}{2}(x - x_i)^T A_i(x - x_i)$$

$$F_j(x) := \beta_j + b_j^T(x - x_j) + \frac{1}{2}(x - x_j)^T B_j(x - x_j)$$

and $\alpha_i, \beta_j \in \mathbb{R}$, $a_i, b_j \in \mathbb{R}^N$, $A_i, B_j \in \mathbb{R}_{\text{sym}}^{N \times N}$, $x_i, x_j \in \mathbb{R}^N$.

The underlying data of the test examples was produced by a random number generator with the following restrictions concerning the data corresponding to F :

- At least one B_j is chosen as a positive definite matrix to guarantee that the feasible set is bounded.
- After choosing b_j , B_j , x_j as well as a starting point $x^0 \in \mathbb{R}^N$, β_j is chosen such that x^0 is strictly feasible.

We made tests for the dimensions $N \in \{20, 40, 60, 80, 100\}$ to investigate the behavior of the following algorithms

- Reduced algorithm (blue)
- MPBNGC (red)
- SolvOpt (yellow),

where we use the colors to distinguish the results of the different solvers, with respect to

- solving time t_1 (cf. (6.2))
- successful termination

focusing on the larger values of N (due to the magnitude of N , we did not test the full version of Algorithm 3.20). Moreover, we chose $m_1 := \frac{N}{10}$ and $m_2 \in \{\frac{N}{2}, N\}$, so that the emphasis of the examples lies on the handling of the constraint.

Furthermore, due to the magnitude of the test examples, we weakened the optimality tolerance of the reduced algorithm to $\varepsilon := 10^{-3}$. Since the reduced algorithm terminated for all examples of this class of test functions with satisfying its termination criterion (which guarantees the stationarity of the computed point due to Theorem 3.36 and Theorem 3.55), we denote the minimizer (of the corresponding example) that was computed by the reduced algorithm by $\hat{x}$.

Before the actual tests, we performed a few runs of the whole test set, where we started with very weak termination criteria for MPBNGC and SolvOpt and then sharpened them, with the goal to make the results between the different solvers comparable in the following way: If the computed minimizer is close to $\hat{x}$, then approximately the same F_j should be active. Based on these empirical observations, we made the final choices for the termination criteria of MPBNGC and SolvOpt, where we were quite successful to achieve this goal for MPBNGC, while we were not able to achieve it for SolvOpt in many cases (although putting a lot of effort into it).

For every pair (N, m_2) we tested 20 different examples for two levels of difficulty that is classified by the average number of $j \in \{1, \dots, m_2\}$ with

$$|F_j(\hat{x}) - F(\hat{x})| \leq 10^{-3},$$

which yields the following overview of our overall 400 different examples

Level	m_2	N				
		20	40	60	80	100
Easy	$\frac{N}{2}$	4	4	6	6	7
	N	5	6	8	9	10
Difficult	$\frac{N}{2}$	4	8	12	15	19
	N	7	14	19	26	31

i.e. for given N and m_2 we regard an example as more difficult, the more impact the constraint has at $\hat{x}$ (in the case of the successful termination of one of the solvers, there was always at least one F_j active). Moreover, for a given level of difficulty, N , and m_2 , the corresponding examples are sorted by the numbers $N - 20 + 1, \dots, N$.

Before making detailed comparisons of the obtained results (s. Tables A.8–A.11 in Appendix A) in Subsections 6.4.1–6.4.4, we give a short overview of them as a reason of clarity of the presentation: For all $N \in \{20, 40, 60, 80, 100\}$ we summarize the easy examples and the difficult examples, where we use two different line types for a better distinction of the comparisons of m_2

- $m_2 = \frac{N}{2}$ (dashed line)
- $m_2 = N$ (solid line)

in Figures 6.21 and 6.22, which yields the following two summary tables consisting of the number of examples for which the reduced algorithm is better than MPBNGC resp. SolvOpt (and vice versa)

with respect to the solving time t_1

(Color code: Violet)		MPBNGC					Red Alg		
Level	m_2	no termination	significantly better	better	a bit better	nearly equal	a bit better	better	significantly better
Easy	$\frac{N}{2}$	1	18	17	18	27	6	5	8
	N	2	8	26	26	20	8	5	5
Difficult	$\frac{N}{2}$	73	0	4	5	4	3	2	9
	N	78	0	1	1	5	1	4	10

(Color code: Orange)		SolvOpt					Red Alg		
Level	m_2	no termination	significantly better	better	a bit better	nearly equal	a bit better	better	significantly better
Easy	$\frac{N}{2}$	18	14	25	11	15	6	7	4
	N	11	16	21	11	15	10	11	5
Difficult	$\frac{N}{2}$	3	4	16	3	8	15	28	23
	N	0	5	8	11	15	7	34	20

that are visualized in Figure 6.21 and Figure 6.22

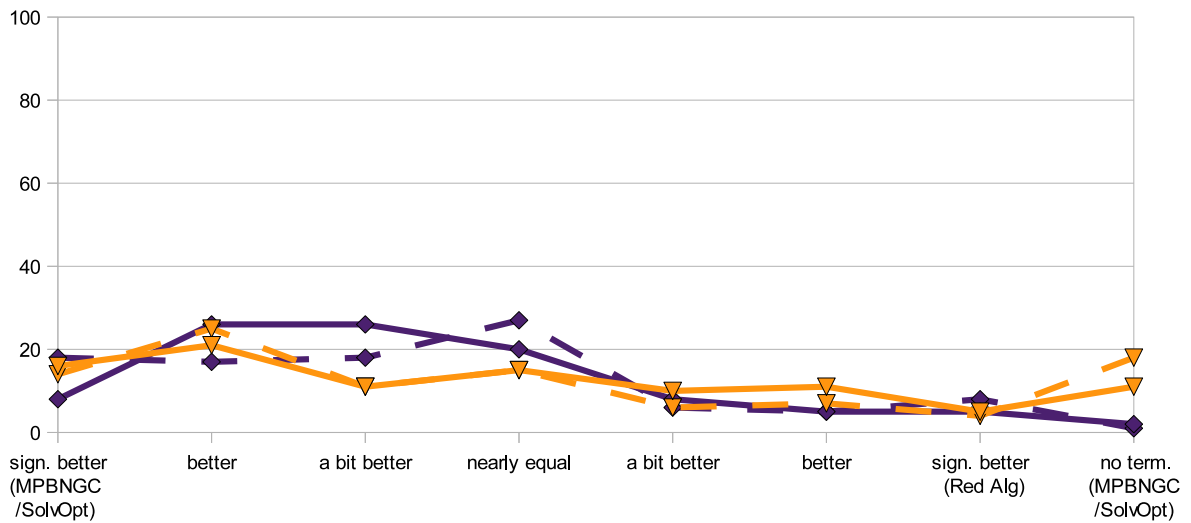


FIGURE 6.21: Easy examples (summary)

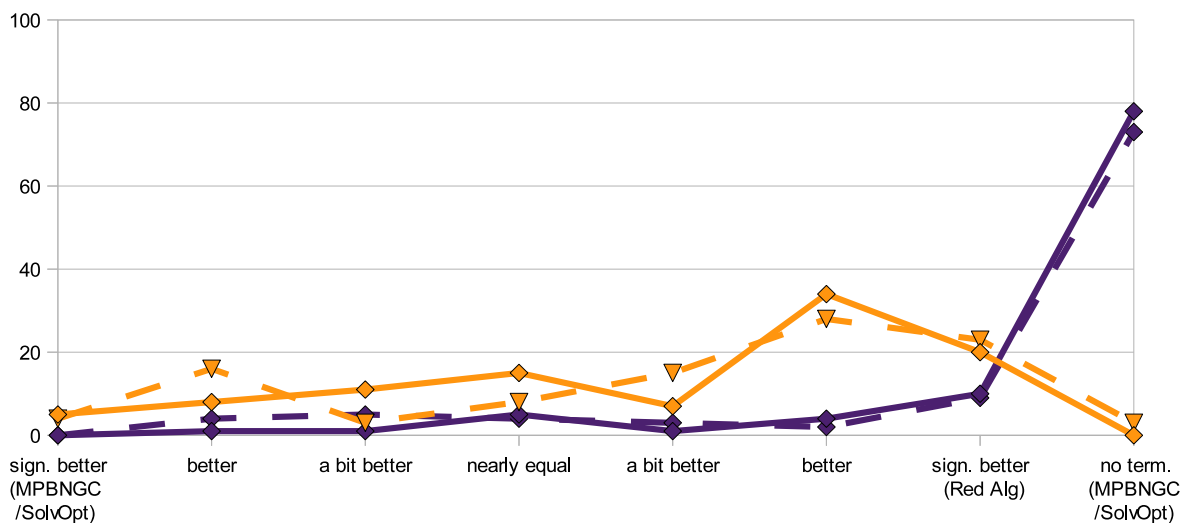


FIGURE 6.22: Difficult examples (summary)

and that let us together with Figure 6.23, in which the solving times t_1 for all examples are plotted

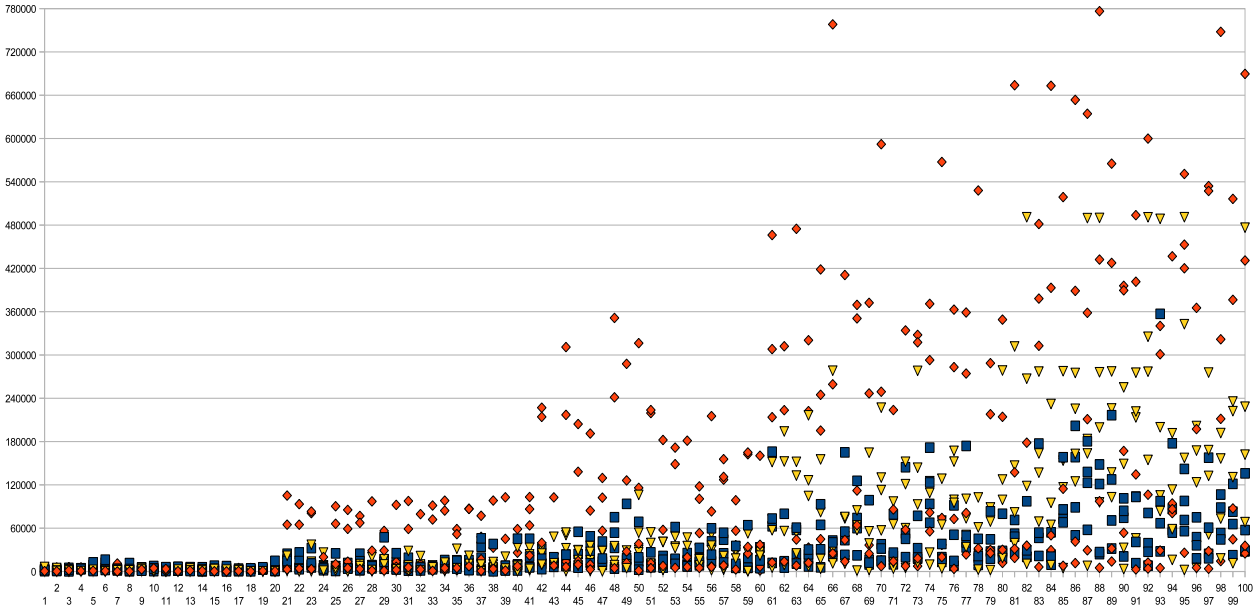


FIGURE 6.23: Solving time t_1 for all higher dimensional piecewise quadratic examples

draw the following conclusions:

- Reduced algorithm & MPBNGC:
 - For the easy examples the reduced algorithm is superior to MPBNGC in thirteen percent of the examples, for thirty percent of the examples one of these two solvers has small advantages over the other (in most cases MPBNGC is the slightly more successful one), the performance differences between the considered two algorithms can be completely neglected for one quarter of the examples, and for one third of the examples MPBNGC beats the reduced algorithm clearly.
 - MPBNGC was not able to terminate successfully for many of the difficult examples in particular for $N \in \{60, 80, 100\}$ despite significantly longer running times as it can be seen in Figure 6.23 (in additional test runs with a softer termination criterion MPBNGC did terminate for approximately half of the difficult examples, but the quality of the obtained minimizers was not comparable with the corresponding $\hat{x}$ produced by the reduced algorithm, while for the comparisons presented here this quality is comparable) and therefore the reduced algorithm is superior to MPBNGC in 88 percent of these examples. Furthermore, for five percent of the examples one of these two solvers has small advantages over the other, the performance differences between the considered two algorithms can be completely neglected for further five percent of the examples, and for the remaining two percent of the examples MPBNGC beats the reduced algorithm clearly.
- Reduced algorithm & SolvOpt:
 - For the easy examples the reduced algorithm is superior to SolvOpt in thirty percent of the examples, for fifteen percent of the examples one of these two solvers has small advantages over the other, the performance differences between the considered two algorithms can be

- completely neglected for further fifteen percent of the examples, and in the remaining forty percent of the examples SolvOpt beats the reduced algorithm clearly.
- For the difficult examples the reduced algorithm is superior to SolvOpt in a bit more than half of the examples (including many examples with $N \in \{80, 100\}$), for twenty percent of the examples one of these two solvers has small advantages over the other, the performance differences between the considered two algorithms can be completely neglected for ten percent of the examples, and in the remaining (a bit less than) twenty percent of the examples SolvOpt beats the reduced algorithm clearly.
 - In particular note that only very few F_j are active at the points which SolvOpt found at termination for the easy examples (in comparison to both the reduced algorithm and MPBNGC), which might indicate that SolvOpt has some problems coming very close to the boundary. Although this behavior improves for the difficult examples, there still remains a clear gap in the number of active F_j between SolvOpt and the other two solvers.
 - We want to emphasize the reduced algorithm was the only solver that terminated for all higher dimensional examples successfully, i.e. with a stationary point that is sufficiently accurate. Moreover, the solving times of the reduced algorithm are quite stable over all dimensions $N \in \{20, 40, 60, 80, 100\}$.

Remark 6.8. Since MOSEK supports multiple CPUs in particular for solving QCQPs (cf. ANDERSEN [2, p.152, 8.1.4 Using multiple CPU's]), we may expect faster solving times for the reduced algorithm on such a system in particular for higher dimensional problems. Nevertheless, we have not been able to test this yet.

6.4.1 Easy examples with $N/2$ constraint components

We summarize the investigations of the results of the easy examples with $m_2 := \frac{N}{2}$, which can be found in Table A.8 in Appendix A and which are visualized in Figure 6.24, in Subsubsections 6.4.1.1 and 6.4.1.2.

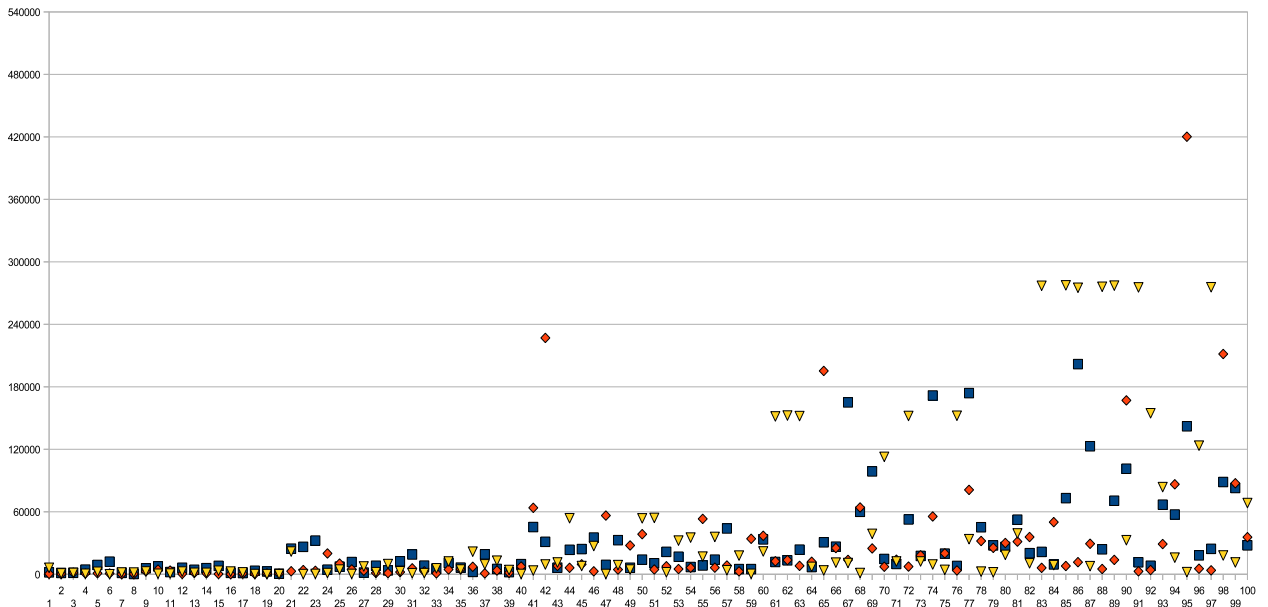


FIGURE 6.24: Easy piecewise quadratic constraint with $m_2 := \frac{N}{2}$

6.4.1.1 Reduced algorithm vs. MPBNGC

MPBNGC does not satisfy its termination criterion for one example (10) within the given number of iterations and function evaluations. For the remaining 99 examples we obtain that the reduced algorithm is

- significantly better in 8 examples (24, 42, 47, 55, 59, 65, 95, 98)
- better in 5 examples (41, 49:50, 84, 90)
- a bit better in 6 examples (11, 25, 36, 60, 82, 94)

in comparison with MPBNGC, while MPBNGC is

- significantly better in 18 examples (5:6, 12, 15, 21:23, 30:31, 37, 46, 48, 57, 67, 69, 74, 77, 86)
- better in 17 examples (4, 9, 13:14, 18, 26, 28, 32, 34, 44:45, 52:53, 72, 85, 87, 89)
- a bit better in 18 examples (1, 7, 16, 19, 29, 33, 40, 51, 56, 58, 63, 78, 81, 83, 88, 93, 96:97)

in comparison with the reduced algorithm, and in 27 examples the solving times of both algorithms do not differ significantly.

6.4.1.2 Reduced algorithm vs. SolvOpt

SolvOpt does not satisfy its termination criterion for 18 examples (1, 21, 36, 44, 50:51, 61:63, 72, 76, 83, 85:86, 88:89, 91, 97) within the given number of iterations and function evaluations. For the remaining 82 examples we obtain that the reduced algorithm is

- significantly better in 4 examples (54, 70, 92, 96)
- better in 7 examples (27, 29, 38, 53, 56, 58, 100)
- a bit better in 6 examples (8, 16, 39, 43, 55, 93)

in comparison with SolvOpt, while SolvOpt is

- significantly better in 14 examples (5:6, 10, 22:23, 26, 31, 41, 57, 67, 74, 77, 87, 95)
- better in 25 examples (4, 9, 12:15, 18:19, 28, 30, 32, 37, 40, 42, 45, 48, 52, 60, 68:69, 78, 90, 94, 98:99)
- a bit better in 11 examples (11, 24:25, 46:47, 59, 65:66, 75, 79, 81)

in comparison with the reduced algorithm, and in 15 examples the solving times of both algorithms do not differ significantly.

6.4.2 Easy examples with N constraint components

We summarize the investigations of the results of the easy examples with $m_2 := N$, which can be

found in Table A.9 in Appendix A and which are visualized in Figure 6.25, in Subsubsections 6.4.2.1 and 6.4.2.2.

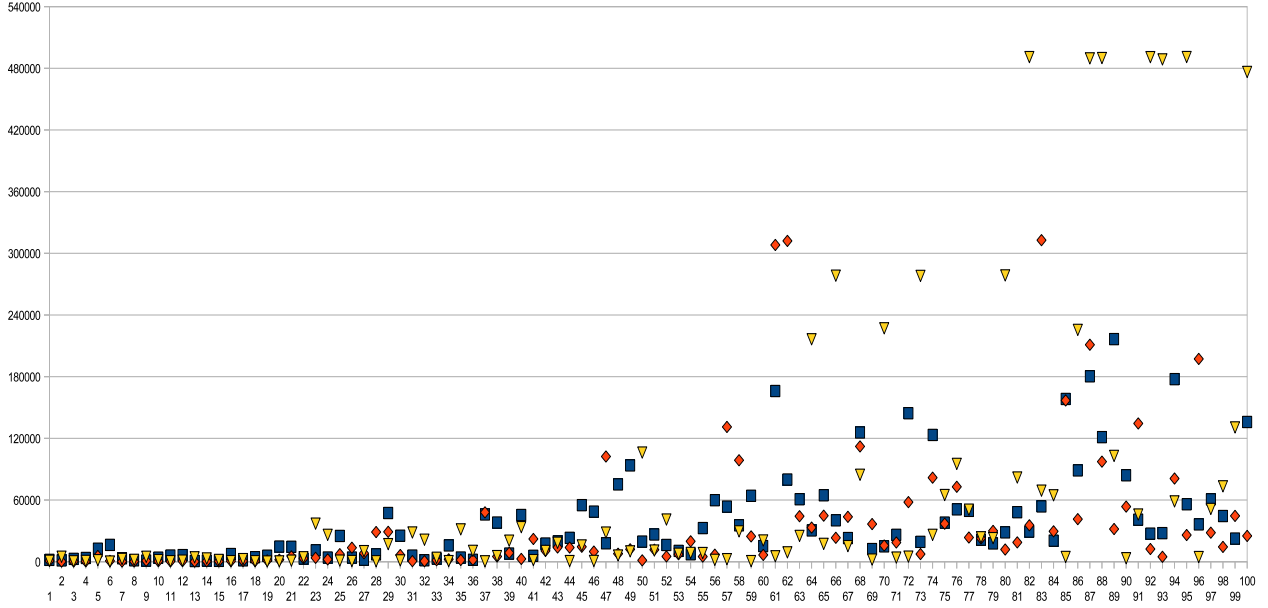


FIGURE 6.25: Easy piecewise quadratic constraint with $m_2 := N$

6.4.2.1 Reduced algorithm vs. MPBNGC

MPBNGC does not satisfy its termination criterion for two examples (61:62) within the given number of iterations and function evaluations. For the remaining 98 examples we obtain that the reduced algorithm is

- significantly better in 5 examples (47, 57:58, 83, 96)
- better in 5 examples (26, 28, 41, 54, 91)
- a bit better in 8 examples (27, 37, 67, 69, 76, 79, 87, 99)

in comparison with MPBNGC, while MPBNGC is

- significantly better in 8 examples (6, 20, 38, 40, 48:49, 56, 89)
- better in 26 examples (4:5, 7, 10:12, 16, 18:19, 21, 25, 29:30, 34, 45:46, 50:52, 55, 59, 72, 74, 86, 94, 100)
- a bit better in 26 examples (2:3, 8, 23, 31, 35, 42:44, 53, 60, 63, 65:66, 68, 73, 77, 80:81, 88, 90, 92:93, 95, 97:98)

in comparison with the reduced algorithm, and in 20 examples the solving times of both algorithms do not differ significantly.

6.4.2.2 Reduced algorithm vs. SolvOpt

SolvOpt does not satisfy its termination criterion for 11 examples (23, 50, 66, 73, 80, 82, 87:88, 92:93,

95) within the given number of iterations and function evaluations. For the remaining 89 examples we obtain that the reduced algorithm is

- significantly better in 5 examples (35, 64, 70, 86, 100)
- better in 11 examples (2, 9, 13, 24, 31:32, 39, 52, 76, 84, 99)
- a bit better in 10 examples (14, 17, 27, 36, 47, 60, 75, 81, 83, 98)

in comparison with SolvOpt, while SolvOpt is

- significantly better in 16 examples (5:6, 20, 29, 37:38, 46, 48:49, 56:57, 59, 61, 72, 74, 85)
- better in 21 examples (11:12, 16, 18:19, 21, 25, 30, 34, 40, 44:45, 51, 55, 62:63, 65, 68, 89:90, 94)
- a bit better in 11 examples (3:4, 10, 26, 28, 41:42, 58, 69, 71, 96)

in comparison with the reduced algorithm, and in 15 examples the solving times of both algorithms do not differ significantly.

6.4.3 Difficult examples with $N/2$ constraint components

We summarize the investigations of the results of the difficult examples with $m_2 := \frac{N}{2}$, which can be found in Table A.10 in Appendix A and which are visualized in Figure 6.26, in Subsubsections 6.4.3.1 and 6.4.3.2.

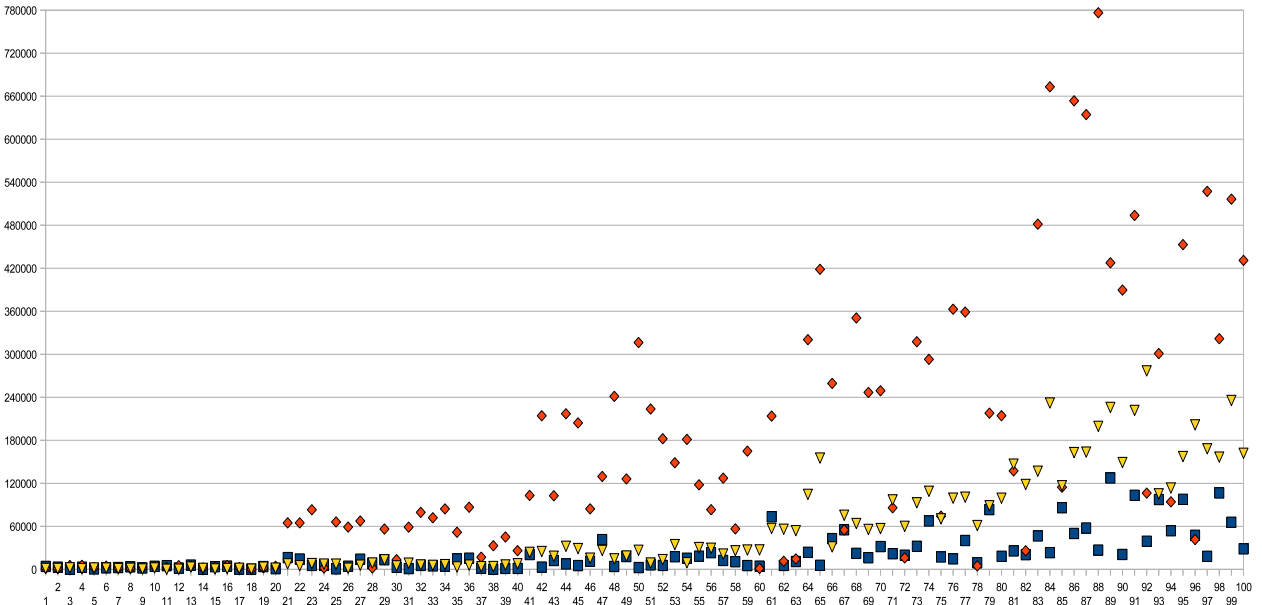


FIGURE 6.26: Difficult piecewise quadratic constraint with $m_2 := \frac{N}{2}$

6.4.3.1 Reduced algorithm vs. MPBNGC

MPBNGC does not satisfy its termination criterion for 73 examples (2:4, 6, 10:17, 20, 22:23, 26, 29,

31:37, 39:40, 42:59, 61, 64:71, 73:74, 76:77, 79:80, 83:84, 87:93, 95, 97:98, 99:100) within the given number of iterations and function evaluations. For the remaining 27 examples we obtain that the reduced algorithm is

- significantly better in 9 examples (21, 25, 27, 30, 38, 41, 75, 81, 86)
- better in 2 examples (5, 94)
- a bit better in 3 examples (62:63, 85)

in comparison with MPBNGC, while MPBNGC is

- better in 4 examples (1, 8, 19, 28)
- a bit better in 5 examples (7, 24, 60, 72, 78)

in comparison with the reduced algorithm, and in 4 examples the solving times of both algorithms do not differ significantly.

6.4.3.2 Reduced algorithm vs. SolvOpt

SolvOpt does not satisfy its termination criterion for 3 examples (28, 65, 92) within the given number of iterations and function evaluations. For the remaining 97 examples we obtain that the reduced algorithm is

- significantly better in 23 examples (62, 64, 71, 73, 75:78, 80:84, 86:91, 96:97, 99:100)
- better in 28 examples (3, 5, 14, 17, 25, 31, 40, 42, 44:45, 48, 50, 53, 55, 58:60, 63, 67:70, 72, 74, 85, 94:95, 98)
- a bit better in 15 examples (20, 23, 30, 34, 37:39, 41, 43, 46, 51:52, 56:57, 79)

in comparison with SolvOpt, while SolvOpt is

- significantly better in 4 examples (11, 15:16, 35)
- better in 16 examples (1, 2, 4, 7:10, 12:13, 21:22, 27, 36, 47, 61, 66)
- a bit better in 3 examples (19, 26, 54)

in comparison with the reduced algorithm, and in 8 examples the solving times of both algorithms do not differ significantly.

6.4.4 Difficult examples with N constraint components

We summarize the investigations of the results of the difficult examples with $m_2 := N$, which can be found in Table A.11 in Appendix A and which are visualized in Figure 6.27, in Subsubsections 6.4.4.1

and 6.4.4.2.

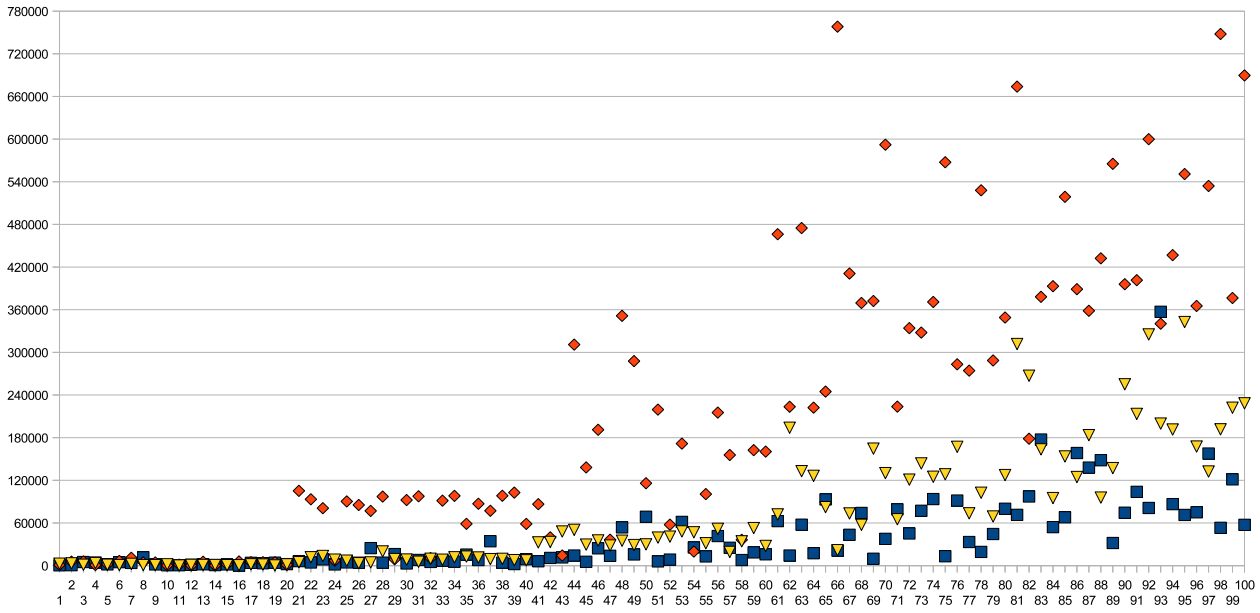


FIGURE 6.27: Difficult piecewise quadratic constraint with $m_2 := N$

6.4.4.1 Reduced algorithm vs. MPBNGC

MPBNGC does not satisfy its termination criterion for 78 examples (2:3, 5:9, 13, 15:19, 21:23, 25:31, 33:39, 44, 46, 48:49, 51:54, 56:57, 59, 61, 63, 65:81, 83:100) within the given number of iterations and function evaluations. For the remaining 22 examples we obtain that the reduced algorithm is

- significantly better in 10 examples (40:42, 45, 50, 55, 58, 60, 62, 64)
- better in 4 examples (24, 32, 47, 82)
- a bit better in 1 example (43)

in comparison with MPBNGC, while MPBNGC is

- better in 1 example (4)
- a bit better in 1 example (20)

in comparison with the reduced algorithm, and in 5 examples the solving times of both algorithms do not differ significantly.

6.4.4.2 Reduced algorithm vs. SolvOpt

For our 100 examples we obtain that the reduced algorithm is

- significantly better in 20 examples (28, 41, 43:44, 51:52, 58:59, 62, 64, 69:70, 75, 81:82, 90, 92, 95, 98, 100)

- better in 34 examples (1:2, 22, 24, 30, 34, 38:39, 42, 45:47, 49, 54:56, 60, 63, 67, 72:74, 76:78, 80, 84:85, 87, 89, 91, 94, 96, 99)
- a bit better in 7 examples (10, 23, 25, 32, 36, 61, 79)

in comparison with SolvOpt, while SolvOpt is

- significantly better in 5 examples (8, 27, 37, 50, 93)
- better in 8 examples (3, 6, 17, 19, 29, 48, 53, 88)
- a bit better in 11 examples (7, 18, 31, 35, 57, 65, 68, 71, 83, 86, 97)

in comparison with the reduced algorithm, and in 15 examples the solving times of both algorithms do not differ significantly.

Chapter 7

Perspectives

In this last chapter we give a brief overview on possible, future development plans, which might be worth considering.

General extensions of Algorithm 3.20 could contain the following points:

- It would (even for the original, unconstrained bundle-Newton method by LUKŠAN & VLČEK [64]) be interesting to find good estimators for the symmetric $n \times n$ -matrices $G_j \in \partial^2 f(y_j)$ and $\hat{G}_j \in \partial^2 F(y_j)$ (which serve as substitutes for the Hessians according to (2.28)) in, e.g., a BFGS-like fashion (Note that we only assumed the boundedness of the damped versions $\{\rho_j G_j\}$ and $\{\hat{\rho}_j \hat{G}_j\}$ of these matrices in the initialization of Algorithm 3.20 for proofing convergence in Section 3.6).
- The use of a different bundle for the objective function and the constraint (cf. Remark 3.8) could make Algorithm 3.20 applicable to problems for which the objective function is not defined outside of the feasible set and for problems with an infeasible starting point.
- A rather tough goal would be to find a way for directly incorporating the handling of a nonsmooth, nonlinear **equality** constraint

$$F_E(x) = 0 \tag{7.1}$$

with a locally Lipschitz continuous function $F_E : \mathbb{R}^n \longrightarrow \mathbb{R}$ in Algorithm 3.20 (Note: Algorithm 3.20 only accepts strictly feasible iteration points, which do not exist for (7.1)) without just demanding the inequality constraint

$$|F_E(x)| - \bar{\varepsilon} \leq 0$$

with a given feasibility tolerance $\bar{\varepsilon} > 0$, which might be numerically unstable.

The convergence result of Algorithm 3.20 might also be improved:

- A further investigation of the proof of Theorem 3.55 might lead to weaker assumptions (e.g., it would be interesting to know, if it is possible to deduce the assumption $\bar{\kappa}^{k+2} - \bar{\kappa}^{k+1} \xrightarrow{K} 0$ in

(3.317) already from other properties — cf., e.g., a similar result for the sequence of iteration points $\{x_k\}$ in (3.266)).

- If we were able to establish an estimate similar to $w_{k+1} < w_k - \bar{c}^2$ from (3.398) in the proof of Theorem 3.55 in the case that a null step with respect to the constraint is taken, it might be possible to get rid of the line search selection parameter i_l (cf. the line search described in Algorithm 3.25). Furthermore, in this case it might even be possible to get rid of the modification of t_0 from (3.113) in the line search and therefore also of the assumption (3.318) in Theorem 3.55. Some results in these directions have already been worked out.

The following suggestions could speed up the current implementation of Algorithm 3.20:

- Since the current implementation of Algorithm 3.20 uses dense linear algebra until now (except for the third party software MOSEK by ANDERSEN et al. [3], which is used for determining the search direction), we might write an additional version of the algorithm which exploits sparsity for increasing the performance.
- Although it seems to be an enormously challenging task, it might be worthwhile to write an own QCQP-solver to get rid of the overhead caused by MOSEK (which is due to the fact that MOSEK, e.g., internally sets up sparse matrix structures) at least for low dimensional problems and which exploits the special structure of the reduced search direction problem (4.12) (as, e.g., PLQDF1 by LUKŠAN [60], which is used for the computation of the search direction in the original, unconstrained bundle-Newton method by LUKŠAN & VLČEK [64], does for QPs).
- There are some additional possibilities for improving the performance of the current implementation of Algorithm 3.20 like, e.g., a still more careful memory management, at the price of significantly reducing the readability of the code, which makes the integration of new ideas into the algorithm much more difficult.

Finally, we list two important points, which are worth to be investigated, concerning our application of nonsmooth optimization to CSPs (cf. Chapter 5):

- It would be interesting to develop strategies in GloptLab by DOMES [18] that significantly benefit from our method for finding and enlarging exclusion boxes using the certificate of infeasibility f from (5.37) in the case of quadratic CSPs (i.e. the matrix valued functions C_k in the CSP (5.9) are constant according to (5.46)).
- A further suggestion is to integrate the certificate in the general case (which requires a very careful implementation) into a solver for global optimization (cf. Remark 5.25), e.g., into the COCONUT environment by SCHICHL [91, 92, 93].

Appendix A

Result tables

TABLE A.1: Different search direction problems

	Median solving time							
	L	D	E	R	L	D	E	R
$n = 50$	$m = 25$				$m = 50$			
MOSEK QCQP	23	102	106	25	32	214	222	39
MOSEK SOCP	28	954	986	52	35	1665	1708	95
SEDUMI	175	480	528	232	226	858	862	254
SDPT3	414	735	804	449	498	1273	1470	624
$n = 100$	$m = 50$				$m = 100$			
MOSEK QCQP	51	793	830	76	94	1761	1627	148
MOSEK SOCP	50	7200	7556	312	95	16631	16694	703
SEDUMI	306	2700	2933	392	398	5975	6330	523
SDPT3	754	6388	7904	863	1143	16910	20441	1385

TABLE A.2: Search direction problems with a QP resp. a reduced QCQP

	Solving time (MOSEK)								
	L	R	L	R	L	R	L	R	
	$n = 50$				$n = 100$				
	$m = 25$		$m = 50$		$m = 50$		$m = 100$		
	minimal	14	13	27	30	40	62	92	135
	median	23	25	32	39	51	76	94	148
	maximal	26	37	38	40	63	84	116	174
	$n = 150$				$n = 200$				
	$m = 75$		$m = 150$		$m = 100$		$m = 200$		
	minimal	117	166	232	340	222	353	557	769
	median	126	186	281	418	249	394	586	870
	maximal	170	216	416	463	316	456	820	973
	$n = 250$				$n = 300$				
	$m = 125$		$m = 250$		$m = 150$		$m = 300$		
	minimal	427	666	942	1403	685	1129	1493	2054
	median	450	700	1188	1592	718	1181	1976	2328
	maximal	651	968	1500	1755	823	1256	2337	2786
	$n = 350$				$n = 400$				
	$m = 175$		$m = 350$		$m = 200$		$m = 400$		
	minimal	932	1520	2226	3003	1382	2093		
	median	1162	1646	2500	3681	1648	2363		
	maximal	1628	2015	3346	3904	2053	2692		

TABLE A.3: Hock-Schittkowski collection

ex	N	Full Alg						Red Alg						MPBNGC						SolvOpt						
		Nit	Na	c	t ₁	t ₂	R	Nit	Na	c	t ₁	t ₂	R	Nit	Nb	c	t ₁	R		Nit	Nc	Ng	N $\hat{g}$	c	t ₁	R
E1	2	6	7	140	0	0		6	7	140	15	0		20	22	176	0			39	141	42	21	471	0	
E2	2	11	12	240	5	0		13	14	280	15	0		57	72	576	0			35	120	37	18	405	0	
E3	2	8	9	180	0	0		8	9	180	0	0		16	23	184	0			45	392	47	22	991	0	-14
10	2	8	9	180	0	0		8	9	180	0	0		9	16	128	0			29	103	31	15	344	0	
11	2	4	5	100	0	0		4	5	100	15	15		43	45	360	0			25	83	27	12	283	0	
12	2	2	3	60	15	0		2	3	60	0	0		19	22	176	0			32	104	34	17	361	0	
15	2	5	6	120	0	0		5	6	120	0	0		3999	4000	32000	62	nt		32	101	34	20	364	0	
16	2	5	6	120	0	0		5	6	120	15	0		261	262	2096	0			69	287	72	22	856	0	
17	2	13	14	280	6	0		13	14	280	15	0		74	80	640	0			32	107	35	18	373	0	
18	2	38	39	780	17	1		38	39	780	31	1		19	42	336	0			148	611	150	77	1903	0	
20	2	4	5	100	0	0		7	8	160	0	0		736	4000	32000	15	nt		36	117	39	21	414	0	
22	2	1	2	40	0	0		1	2	40	0	0		10	11	88	0			35	135	37	18	435	0	
23	2	11	12	240	6	0		13	14	280	31	1		70	190	1520	0			40	138	42	21	465	0	
57	2	1	3	60	0	0		1	3	60	0	0		4	9	72	0			31	117	32	10	360	0	
59	2	11	12	240	8	8		11	12	240	15	15		17	127	1016	15			35	157	36	12	458	0	
215	2	8	9	180	10	0		8	9	180	15	0		16	17	136	0			34	290	36	14	730	0	-14
218	2	11	12	240	3	0		11	12	240	15	15		27	44	352	0			61	462	63	30	1203	0	-14
222	2	5	7	140	0	0		5	8	160	0	0		31	131	1048	0			35	115	37	20	401	0	-13
223	2	5	10	200	8	0		7	15	300	15	0		15	28	224	0			65	308	68	34	922	0	-13
225	2	11	12	240	4	0		13	14	280	15	0		70	190	1520	0			40	138	42	21	465	0	
226	2	6	7	140	0	0		6	7	140	0	0		12	16	128	0			33	117	35	17	390	0	
227	2	5	10	200	5	0	t_0^k	8	16	320	15	0	t_0^k	30	32	256	0			41	137	44	23	475	0	
228	2	2	3	60	0	0		3	5	100	0	0		3	4	32	0			17	78	19	9	240	0	-13
230	2	11	16	320	4	0	t_0^k	13	18	360	15	15	t_0^k	27	128	1024	0			42	150	44	23	501	0	-13
233	2	11	14	280	0	0		11	14	280	0	0		30	34	272	0			16	87	17	0	225	0	
234	2	3	4	80	0	0		1	2	40	0	0		1	2	16	0			44	167	47	25	550	0	
236	2	1	2	40	0	0		1	2	40	0	0		6	7	56	0			38	137	40	22	460	15	
237	2	1	2	40	0	0		1	2	40	0	0		16	116	928	0			49	220	51	25	668	0	
238	2	1	2	40	0	0		1	2	40	0	0		13	114	912	0			46	204	48	22	618	0	
239	2	16	17	340	9	0		16	17	340	15	0		13	32	256	0			40	145	43	23	488	0	
315	2	7	8	160	0	0		7	8	160	0	0		12	15	120	0			40	137	42	21	463	0	-13
323	2	5	6	120	0	0		5	6	120	0	0		37	41	328	0			32	108	34	17	369	0	
324	2	38	39	780	31	11		38	39	780	46	16		19	42	336	0			148	611	150	77	1903	0	
326	2	10	12	240	0	0		10	14	280	0	0		24	28	224	0			35	123	37	18	411	0	
327	2	1	3	60	0	0		1	3	60	0	0		4	9	72	0			31	117	32	10	360	15	
330	2	4	8	160	7	0		5	11	220	15	15	t_0^k	26	34	272	0			37	134	39	16	433	0	
332	2	7	10	200	7	0		6	10	200	15	0		379	380	3040	62			56	386	59	34	1051	31	
29	3	16	17	442	4	0		16	26	676	15	0		27	43	344	0			37	135	39	19	444	0	
30	3	11	12	312	18	8		11	12	312	62	16		40	49	392	0			34	117	37	11	378	0	
31	3	11	12	312	8	0		11	12	312	15	0		110	129	1032	0			37	127	39	14	413	0	
33	3	22	23	598	0	0		4	5	130	0	0		25	33	264	0			81	375	83	45	1134	15	-13
34	3	14	44	1144	6	0	t_0^k	73	224	5824	93	3	t_0^k	24	141	1128	0			67	251	69	38	823	0	-13
65	3	2	3	78	0	0		3	5	130	0	0		15	16	128	0			39	152	41	11	460	0	
66	3	23	48	1248	12	0	t_0^k	31	73	1898	30	0	t_0^k	17	135	1080	0			47	154	50	30	548	0	-13
249	3	11	12	312	4	0		11	12	312	15	0		34	35	280	0			36	122	38	10	388	0	
337	3	11	12	312	0	0		11	12	312	0	0		103	123	984	0			35	116	37	13	382	0	
339	3	14	17	442	12	0		14	17	442	31	1		34	61	488	0			62	231	64	17	705	0	
341	3	16	17	442	4	0		18	34	884	15	0	t_0^k	31	47	376	0			37	135	39	19	444	0	
343	3	9	10	260	16	0		53	56	1456	109	4		322	1812	14496	15			45	190	49	24	599	15	

TABLE A.3: Hock-Schittkowski collection (continued)

ex	N	Full Alg						Red Alg						MPBNGC					SolvOpt						
		Nit	Na	c	t ₁	t ₂	R	Nit	Na	c	t ₁	t ₂	R	Nit	Nb	c	t ₁	R	Nit	Nc	Ng	N $\hat{g}$	c	t ₁	R
43	4	19	50	1600	15	0	t_0^k	20	42	1344	31	1	t_0^k	85	142	1136	0		46	154	48	24	524	0	
264	4	20	81	2592	9	10	t_0^k	21	45	1440	15	15	t_0^k	60	71	568	0		43	142	45	22	485	0	
83	5	17	23	874	26	16	t_0^k	15	27	1026	31	16	t_0^k	1074	4000	32000	46	nt	45	140	47	29	508	15	-13
270	5	8	9	342	10	0		8	9	342	15	0		9	16	128	0		80	1303	81	1	2852	0	-13
100	7	20	80	4000	46	0	t_0^k	21	50	2500	46	1	t_0^k	86	137	1096	0		61	197	63	24	655	0	
113	10	15	22	1496	41	70	t_0^k	22	38	2584	31	16	t_0^k	263	625	5000	31		187	685	189	80	2177	0	
284	15	14	16	1568	92	2		15	29	2842	46	1	t_0^k	124	3990	31920	31		1026	3384	1028	801	12255	46	wm
285	15	67	211	20678	796	62	t_0^k	66	96	9408	140	35	t_0^k	2274	4000	32000	578	nt	5769	18740	5771	4405	68008	265	wm
384	15	66	140	13720	589	179	t_0^k	68	114	11172	125	35	t_0^k	2531	4000	32000	531	nt	635	2117	637	471	7558	15	wm
Total		703	1216	59912	1853	367		824	1323	53046	1198	237		13453	29970	239760	1386		10092	36051	10216	6933	123549	432	

TABLE A.4: Certificate — Linearly constrained case with fixed box

ex	N	$T = 1$										$T = y _2$																					
		BNLC					MPBNGC					SolvOpt					BNLC					MPBNGC					SolvOpt						
		Nit	Na	c	t ₁	t ₂	Nit	Nb	c	t ₁		Nit	Nc	Ng	c	t ₁		Nit	Na	c	t ₁	t ₂		Nit	Nb	c	t ₁		Nit	Nc	Ng	c	t ₁
100	4	17	21	336	31	16	26	28	112	0		38	209	39	326	0		6	8	128	15	0		4	5	20	0		78	228	79	465	0
101	4	11	13	208	15	0	18	20	80	15		53	200	54	362	15		20	21	336	31	1		14	15	60	0		78	284	84	536	15
102	4	11	18	288	15	0	20	23	92	0		49	168	50	318	0		11	14	224	15	0		7	8	32	0		50	422	53	581	0
103	4	15	20	320	31	1	20	21	84	0		41	172	42	298	0		16	19	304	15	0		34	48	192	0		40	129	41	252	15
104	4	8	11	176	0	0	11	14	56	0		34	233	38	347	15		2	4	64	15	0		4	5	20	0		54	626	55	791	15
106	4	3	10	160	15	15	4	6	24	0		36	197	37	308	0		10	15	240	15	0		9	12	48	0		39	129	41	252	0
108	4	1	2	32	0	0	5	7	28	0		220	3798	228	4482	78		1	2	32	0	0		2	3	12	0		34	279	37	390	0
109	4	1	2	32	15	0	3	4	16	0		27	299	30	389	0		1	2	32	15	0		1	2	8	0		39	712	41	835	15
110	4	7	16	256	15	0	2	3	12	0		111	3524	115	3869	78		2	3	48	0	0		2	3	12	0		33	242	35	347	0
112	4	1	2	32	0	0	4	8	32	0		179	2667	187	3228	62		1	2	32	0	0		5	12	48	0		40	476	41	599	15
113	4	1	2	32	0	0	7	21	84	0		29	197	31	290	15		1	2	32	15	0		1	2	8	0		31	338	34	440	0
114	4	3	12	192	15	0	2	3	12	0		52	895	55	1060	15		1	2	32	0	0		2	3	12	0		39	354	42	480	15
115	4	0	1	16	0	0	4	5	20	0		1	2	1	5	0		6	16	256	15	0		2	3	12	0		78	344	82	590	0
116	4	1	2	32	0	0	5	10	40	0		241	3907	250	4657	78		2	3	48	0	0		8	24	96	0		30	279	32	375	15
117	4	1	2	32	15	0	6	21	84	0		32	201	35	306	0		2	3	48	0	0		1	2	8	0		32	414	35	519	0
118	4	0	1	16	0	0	5	11	44	0		23	217	25	292	15		3	6	96	15	0		2	3	12	0		33	324	37	435	15
120	4	3	4	64	15	0	4	8	32	0		51	301	54	463	0		2	3	48	15	0		6	18	72	0		40	434	42	560	0
121	4	99	369	5904	109	19	174	681	2724	15		31	204	33	303	15		8	27	432	15	0		5	21	84	0		38	492	40	612	15
122	4	3	4	64	0	0	4	5	20	0		18	90	20	150	0		59	119	1904	78	3		2	3	12	0		46	556	47	697	15
123	4	1	3	48	0	0	2	3	12	0		102	2620	103	2929	46		6	14	224	15	0		7	22	88	0		34	430	36	538	15
128	4	17	21	336	31	16	27	31	124	0		47	191	48	335	15		13	19	304	15	0		16	17	68	0		101	312	103	621	15
129	4	1	2	32	0	0	5	6	24	0		2	8	2	14	15		2	3	48	0	0		5	7	28	0		2	8	2	14	0
130	4	2	7	112	0	0	2	3	12	0		1	3	1	6	0		1	2	32	15	15		2	3	12	0		1	3	1	6	0
134	4	13	15	240	15	0	14	15	60	0		50	195	51	348	0		7	9	144	15	0		7	8	32	0		66	1186	67	1387	31
135	4	1	2	32	0	0	2	3	12	0		1	2	1	5	0		1	2	32	0	0		2	3	12	0		1	2	1	5	0
138	4	7	9	144	0	0	8	10	40	0		43	210	45	345	0		2	5	80	0	0		7	8	32	15		39	369	40	489	15
139	4	6	14	224	0	0	5	6	24	0		30	266	32	362	15		2	3	48	0	0		4	5	20	0		35	379	36	487	0
140	4	2	4	64	15	0	6	7	28	0		42	240	45	375	0		36	67	1072	62	2		4	5	20	0		42	687	43	816	15
142	4	8	10	160	15	0	7	8	32	0		33	376	36	484	15		4	6	96	0	0		8	20	80	0		37	530	38	644	0
146	4	3	11	176	15	0	2	3	12	0		158	4031	164	4523	93		2	3	48	0	0		2	3	12	0		33	268	37	379	15
147	4	5	9	144	0	0	4	6	24	0		48	934	50	1084	31		5	10	160	15	0		4	5	20	0		34	337	35	442	0

TABLE A.4: Certificate — Linearly constrained case with fixed box (continued)

ex	N	$T = 1$						$T = y _2$																					
		BNLC			MPBNGC			SolvOpt			BNLC			MPBNGC			SolvOpt												
		Nit	Na	c	t ₁	t ₂	Nit	Nb	c	t ₁	Nit	Nc	Ng	c	t ₁	Nit	Na	c	t ₁	t ₂	Nit	Nb	c	t ₁	Nit	Nc	Ng	c	t ₁
151	4	3	4	64	15	0	2	3	12	0	185	3393	191	3966	78	2	3	48	15	0	2	3	12	0	27	205	30	295	0
152	4	5	11	176	0	0	3	4	16	0	93	1887	96	2175	31	10	19	304	15	0	3	4	16	0	31	305	34	407	15
154	4	5	10	160	15	0	8	20	80	0	110	2973	113	3312	62	5	9	144	0	0	4	5	20	0	36	644	38	758	15
156	4	5	15	240	15	0	4	7	28	15	99	1957	103	2266	46	4	9	144	0	0	7	17	68	0	38	688	40	808	15
157	4	37	91	1456	46	16	42	87	348	15	18	93	19	150	0	8	18	288	15	0	8	22	88	0	82	319	84	571	0
159	4	7	9	144	15	0	7	14	56	0	229	3475	237	4186	78	8	24	384	15	0	8	27	108	0	94	393	96	681	15
160	4	0	1	16	0	0	1	1	4	0	0	1	0	1	0	30	68	1088	46	16	6	24	96	0	32	225	34	327	0
161	4	2	3	48	0	0	4	5	20	0	22	120	26	198	0	5	14	224	15	0	4	14	56	0	48	691	49	838	15
162	4	3	10	160	15	15	2	3	12	0	27	270	31	363	15	2	3	48	0	0	2	3	12	0	31	326	34	428	0
163	4	4	6	96	0	0	2	3	12	0	22	179	27	260	0	1	2	32	15	0	1	2	8	0	30	216	32	312	0
164	4	4	6	96	15	0	2	3	12	0	24	187	29	274	0	2	3	48	0	0	1	2	8	0	39	536	40	656	15
165	4	7	22	352	0	0	3	4	16	0	69	1751	71	1964	46	1	2	32	0	0	1	2	8	0	36	230	37	341	15
166	4	9	16	256	15	0	29	58	232	0	17	94	18	148	0	8	26	416	15	0	4	19	76	0	81	284	83	533	0
168	4	0	1	16	0	0	1171	4388	17552	125	28	167	30	257	0	6	19	304	15	0	5	17	68	0	64	258	67	459	15
170	4	0	1	16	15	0	5	6	24	0	1	8	1	11	0	8	27	432	0	0	7	34	136	0	53	178	54	340	15
171	4	0	1	16	0	0	3	4	16	0	1	2	1	5	0	12	29	464	31	1	4	20	80	0	71	256	75	481	0
172	4	1	2	32	0	0	2	4	16	0	1	4	1	7	0	8	23	368	15	0	4	13	52	0	75	280	79	517	15
173	4	3	10	160	15	0	4	6	24	0	1	4	1	7	0	1	4	64	15	0	3	6	24	0	1	4	1	7	0
189	4	13	20	320	15	0	17	18	72	0	44	160	45	295	15	14	17	272	30	0	11	12	48	0	47	651	50	801	15
190	4	14	19	304	15	0	21	22	88	0	60	275	61	458	15	2	3	48	0	0	17	18	72	0	76	261	79	498	15
191	4	12	14	224	15	15	13	16	64	0	55	1150	58	1324	15	5	6	96	15	0	8	9	36	0	40	383	43	512	0
192	4	16	17	272	31	1	6	7	28	0	32	333	34	435	0	1	2	32	0	0	11	12	48	0	32	424	34	526	15
193	4	5	10	160	15	0	7	10	40	0	43	288	45	423	15	2	6	96	0	0	5	7	28	0	40	591	42	717	15
195	4	3	9	144	0	0	5	10	40	0	41	391	43	520	0	8	14	224	0	0	6	8	32	0	41	415	42	541	15
197	4	4	6	96	15	0	2	3	12	0	21	225	23	294	0	3	5	80	15	15	2	3	12	0	51	736	52	892	15
198	4	11	20	320	15	0	4	5	20	0	31	216	34	318	15	3	4	64	0	0	2	3	12	0	23	184	26	262	0
200	4	3	4	64	15	0	2	3	12	0	20	257	23	326	0	3	6	96	0	0	2	3	12	0	33	340	36	448	15
201	4	7	20	320	0	0	13	32	128	15	49	366	50	516	0	5	11	176	15	0	6	26	104	0	29	447	30	537	0
203	4	13	40	640	15	0	10	32	128	0	53	1281	56	1449	31	6	13	208	0	0	8	28	112	0	36	255	38	369	15
204	4	0	1	16	0	0	4	5	20	0	1	2	1	5	0	3	4	64	15	0	3	4	16	0	48	159	51	312	0
205	4	6	21	336	15	0	11	35	140	0	31	253	34	355	15	14	28	448	15	0	7	25	100	0	29	367	34	469	15
207	4	1	2	32	0	0	1	2	8	0	207	3274	214	3916	62	0	1	16	0	0	0	1	4	0	26	208	32	304	0
208	4	1	2	32	0	0	1	2	8	0	17	106	20	166	0	1	2	32	15	0	1	2	8	0	53	941	57	1112	15
209	4	1	2	32	15	0	2	3	12	0	12	50	15	95	0	4	14	224	0	0	2	3	12	0	106	433	111	766	15
210	4	1	2	32	0	0	2	3	12	0	17	63	20	123	0	4	14	224	15	0	2	3	12	0	69	251	72	467	15
214	4	1	2	32	0	0	1	2	8	0	8	32	10	62	0	1	2	32	15	0	1	2	8	0	10	62	12	98	15
215	4	1	2	32	0	0	1	2	8	0	5	25	7	46	0	1	2	32	15	0	1	2	8	0	7	35	9	62	15
218	4	1	3	48	0	0	3	4	16	0	1	4	1	7	0	2	3	48	0	0	2	3	12	0	1	4	1	7	0
219	4	1	2	32	0	0	2	3	12	0	1	3	1	6	0	2	3	48	15	0	2	3	12	0	1	3	1	6	0
220	4	1	2	32	15	0	2	3	12	0	1	2	1	5	0	1	3	48	0	0	2	3	12	0	1	2	1	5	0
225	4	7	9	144	15	0	23	26	104	0	40	176	41	299	0	30	37	592	46	1	12	16	64	0	34	110	38	224	0
226	4	1	3	48	0	0	2	3	12	0	1	2	1	5	0	2	3	48	0	0	2	3	12	0	1	2	1	5	0
233	4	8	10	160	0	0	23	40	160	0	49	244	51	397	0	2	4	64	0	0	8	10	40	0	44	362	46	500	15
234	4	11	13	208	15	0	33	44	176	0	46	251	50	401	0	2	3	48	15	0	12	13	52	0	79	235	81	478	0
238	4	19	52	832	31	1	10	11	44	0	140	2900	142	3326	62	2	3	48	15	0	3	4	16	0	40	377	41	500	15
239	4	2	3	48	0	0	4	5	20	0	46	814	47	955	31	2	3	48	0	0	6	8	32	0	29	211	31	304	0
240	4	1	2	32	0	0	4	5	20	0	4	39	5	54	0	1	2	32	0	0	5	6	24	0	6	42	7	63	0
244	4	1	2	32	0	0	2	3	12	0	3	13	4	25	0	0	2	32	0	0	2	3	12	0	6	23	7	44	0

TABLE A.4: Certificate — Linearly constrained case with fixed box (continued)

ex	N	BNLC					$T = 1$ MPBNGC					SolvOpt					$T = y _2$ BNLC					$T = y _2$ MPBNGC					SolvOpt				
		Nit	Na	c	t_1	t_2	Nit	Nb	c	t_1	Nit	Nc	Ng	c	t_1	Nit	Na	c	t_1	t_2	Nit	Nb	c	t_1	Nit	Nc	Ng	c	t_1		
255	4	3	10	160	15	0	5	6	24	0	183	5288	186	5846	125	5	7	112	15	15	4	5	20	0	55	584	57	755	15		
259	4	3	11	176	15	0	4	5	20	0	65	554	67	755	15	2	4	64	15	0	3	4	16	0	41	641	43	770	15		
260	4	6	9	144	15	0	4	5	20	0	28	196	30	286	0	2	4	64	0	0	3	4	16	0	47	854	49	1001	15		
261	4	2	4	64	0	0	2	3	12	0	33	303	36	411	15	3	5	80	15	0	3	4	16	0	38	563	41	686	15		
267	4	3	4	64	0	0	2	3	12	0	106	2355	108	2679	46	2	3	48	15	0	2	3	12	0	52	798	53	957	15		
268	4	5	8	128	15	0	2	3	12	0	68	1985	71	2198	31	1	2	32	0	0	2	3	12	0	25	205	28	289	0		
269	4	1	2	32	0	0	2	3	12	0	25	301	29	388	0	1	2	32	0	0	3	4	16	0	28	306	31	399	15		
270	4	1	2	32	15	0	2	3	12	0	170	2910	179	3447	62	2	3	48	15	0	2	3	12	0	44	646	46	784	15		
271	4	1	3	48	0	0	2	4	16	0	38	354	41	477	0	1	4	64	0	0	2	4	16	0	42	371	44	503	0		
278	4	3	4	64	15	0	3	5	20	0	21	206	24	278	0	3	6	96	15	15	6	13	52	0	35	444	38	558	15		
279	4	4	5	80	0	0	3	6	24	0	196	3330	205	3945	62	3	6	96	0	0	6	13	52	0	42	365	45	500	0		
280	4	4	6	96	15	0	48	88	352	0	30	325	34	427	0	3	5	80	0	0	3	9	36	0	39	661	43	790	15		
281	4	99	106	1696	124	34	21	41	164	0	13	86	15	131	15	7	16	256	15	0	6	14	56	0	65	266	68	470	15		
282	4	0	1	16	15	0	1	1	4	0	0	1	0	1	0	2	9	144	15	0	2	4	16	0	18	131	20	191	0		
284	4	0	1	16	0	0	1	1	4	0	0	1	0	1	0	2	9	144	0	0	2	4	16	0	21	154	23	223	0		
285	4	90	91	1456	62	32	2	3	12	0	1	2	1	5	0	4	11	176	15	0	3	4	16	0	75	269	77	500	15		
286	4	0	1	16	0	0	1	1	4	0	0	1	0	1	0	0	2	32	0	0	2	3	12	0	2	4	2	10	0		
289	4	5	10	160	0	0	2	3	12	0	185	5885	186	6443	125	2	3	48	15	15	2	3	12	0	42	304	44	436	15		
290	4	6	12	192	15	0	4	8	32	0	132	3256	134	3658	78	2	4	64	15	0	6	14	56	0	36	457	37	568	15		
291	4	74	147	2352	93	3	15	50	200	0	223	2200	236	2908	46	10	18	288	15	0	8	16	64	0	43	430	45	565	0		
293	4	35	96	1536	46	16	43	124	496	15	236	4367	244	5099	93	10	30	480	15	0	4	12	48	0	37	342	39	459	15		
294	4	4	5	80	15	0	2	3	12	0	108	2255	112	2591	46	2	3	48	0	0	2	3	12	0	29	278	32	374	0		
295	4	4	5	80	15	0	2	3	12	0	26	233	28	317	0	2	3	48	15	0	2	3	12	0	35	310	40	430	15		
297	4	0	1	16	0	0	1	1	4	0	0	1	0	1	0	1	2	32	0	0	2	3	12	0	1	3	1	6	0		
21	6	2	3	66	15	0	6	7	28	0	0	10	2	16	0	3	4	88	0	0	7	8	32	0	0	8	2	14	0		
22	6	2	3	66	0	0	1	2	8	0	41	710	42	836	15	2	3	66	0	0	1	2	8	0	46	800	47	941	31		
24	6	4	5	110	15	0	1	2	8	0	39	691	41	814	15	1	2	44	15	0	1	2	8	0	76	2050	79	2287	62		
57	6	1	2	44	0	0	2	3	12	0	0	10	2	16	0	1	2	44	30	0	2	3	12	0	29	215	31	308	0		
58	6	1	2	44	31	0	1	2	8	0	0	9	2	15	0	1	2	44	15	0	1	2	8	0	0	7	2	13	0		
59	6	1	2	44	0	0	1	2	8	0	213	4085	219	4742	141	1	2	44	0	0	1	2	8	0	70	1769	72	1985	62		
60	6	1	2	44	0	0	1	2	8	0	112	2174	114	2516	62	1	2	44	0	0	1	2	8	0	32	630	33	729	15		
61	6	1	2	44	15	0	1	2	8	0	182	4416	187	4977	109	1	2	44	0	0	1	2	8	0	33	331	36	439	15		
62	6	1	2	44	0	0	1	2	8	0	26	279	28	363	15	1	2	44	0	0	1	2	8	0	60	1586	62	1772	46		
67	6	1	2	44	15	0	1	2	8	0	1	10	3	19	0	1	2	44	15	15	1	2	8	0	1	7	3	16	16		
68	6	1	2	44	0	0	1	2	8	0	44	963	46	1101	31	1	2	44	0	0	1	2	8	0	103	3295	104	3607	93		
70	6	1	2	44	15	15	1	2	8	0	191	4167	196	4755	125	1	2	44	15	0	1	2	8	0	45	945	46	1083	31		
71	6	1	2	44	0	0	1	2	8	0	45	604	46	742	15	1	2	44	0	0	1	2	8	0	40	743	42	869	31		
72	6	1	2	44	0	0	1	2	8	0	57	1287	59	1464	31	1	2	44	15	15	1	2	8	0	29	225	31	318	0		
73	6	1	2	44	0	0	1	2	8	0	55	1304	57	1475	31	1	2	44	0	0	1	2	8	0	30	322	33	421	0		
74	6	1	2	44	0	0	1	2	8	0	2	14	4	26	0	1	2	44	15	0	1	2	8	0	2	11	5	26	16		
91	6	1	2	44	0	0	1	2	8	0	1	10	3	19	16	1	2	44	0	0	1	2	8	0	5	33	7	54	0		
92	6	1	2	44	0	0	1	2	8	0	0	9	2	15	0	1	2	44	15	0	1	2	8	0	32	304	34	406	0		
93	6	1	2	44	15	0	1	2	8	0	1	10	3	19	16	1	2	44	0	0	1	2	8	0	1	7	3	16	16		
94	6	1	2	44	0	0	1	2	8	0	25	222	27	303	0	1	2	44	15	0	1	2	8	0	2	11	5	26	16		
96	6	1	2	44	15	0	1	2	8	0	23	135	24	207	0	1	2	44	0	0	1	2	8	0	51	996	54	1158	31		
97	6	0	1	22	0	0	30	77	308	16	2	10	3	19	16	0	1	22	31	0	1	2	8	0	2	10	3	19	16		
98	6	0	1	22	16	0	1	2	8	0	2	10	3	19	16	0	1	22	16	0	1	2	8	0	2	10	3	19	16		
99	6	0	1	22	0	0	1	2	8	0	2	15	3	24	0	0	1	22	0	0	1	2	8	0	4	18	3	27	16		

TABLE A.4: Certificate — Linearly constrained case with fixed box (continued)

ex	N	$T = 1$					$T = y _2$					$T = y _2$					$T = y _2$				
		BNLC					MPBNGC					SolvOpt					BNLC				
		Nit	Na	c	t ₁	t ₂	Nit	Nb	c	t ₁		Nit	Nc	Ng	c	t ₁	Nit	Na	c	t ₁	t ₂
310	11	6	8	296	15	0	7	8	32	0		101	1505	104	1817	93	10	11	407	15	0

TABLE A.5: Certificate — Linearly constrained case with variable box

ex	N	$T = 1$					$T = y _2$					$T = y _2$					$T = y _2$				
		BNLC					MPBNGC					BNLC					MPBNGC				
		Nit	Na	c	t ₁	t ₂	Nit	Nb	c	t ₁		Nit	Nc	Ng	c	t ₁	Nit	Na	c	t ₁	t ₂
100	8	2	3	84	0	0	5	6	24	0		2	7	2	13	0	2	4	112	15	0
101	8	2	4	112	15	0	17	18	72	15		6	7	4	29	0	8	9	196	15	0
102	8	2	3	84	15	0	7	8	32	0		8	9	6	38	0	4	5	252	15	15
103	8	2	3	84	15	0	2	3	12	0		2	3	84	15	0	2	3	84	15	0
104	8	1	2	56	15	15	2	3	12	0		1	2	5	30	0	2	3	12	0	0
106	8	1	2	56	15	15	2	3	12	0		25	95	27	176	15	2	3	12	0	0
108	8	1	2	56	0	0	2	3	12	0		22	74	24	146	0	2	3	12	0	0
109	8	1	2	56	15	0	2	3	12	0		45	254	47	395	15	2	3	12	0	0
110	8	0	2	56	0	0	1	2	8	0		49	189	51	342	0	0	2	56	0	0
112	8	0	2	56	15	0	2	3	12	0		25	88	26	166	0	0	2	56	15	0
113	8	0	1	28	0	0	1	2	8	0		37	125	38	239	15	0	1	28	0	0
114	8	0	1	28	0	0	0	1	4	0		134	535	138	949	31	0	1	28	0	0
115	8	0	1	28	0	0	0	1	4	0		44	150	45	285	0	0	1	28	15	0
116	8	0	1	28	15	0	0	1	4	0		154	609	159	1086	31	0	1	28	0	0
117	8	0	1	28	0	0	0	1	4	0		59	703	60	883	31	0	1	28	0	0
118	8	0	1	28	0	0	0	1	4	0		54	340	55	505	15	0	1	28	0	0
120	8	0	1	28	15	0	0	1	4	0		53	317	54	479	15	0	1	28	0	0
121	8	0	1	28	0	0	0	1	4	0		163	634	169	1141	31	0	1	28	0	0
122	8	0	1	28	0	0	0	1	4	0		48	154	49	301	15	0	1	28	0	0
123	8	0	1	28	0	0	1	1	4	0		0	1	0	1	0	0	1	28	15	0
128	8	1	2	56	15	0	3	4	16	0		3	8	1	11	0	3	4	112	0	0
129	8	1	2	56	0	0	2	3	12	0		1	4	1	7	0	2	3	84	15	0
130	8	1	2	56	15	0	2	3	12	0		1	2	1	5	0	1	2	56	0	0
134	8	1	2	56	15	0	2	3	12	0		1	4	1	7	0	1	2	56	0	0
135	8	1	2	56	0	0	2	3	12	0		1	2	1	5	0	1	2	56	15	15
138	8	1	2	56	15	15	2	3	12	0		2	9	3	18	15	1	2	56	15	0
139	8	1	2	56	0	0	2	3	12	0		5	16	6	34	15	2	3	12	15	0
140	8	1	2	56	15	0	2	3	12	0		18	72	19	129	15	1	2	56	0	0
142	8	1	2	56	0	0	2	3	12	0		44	286	46	424	0	1	2	56	0	0
146	8	1	2	56	0	0	2	3	12	0		43	276	45	411	15	0	2	56	0	0
147	8	1	2	56	15	0	2	3	12	0		32	112	33	211	15	0	2	56	15	0
151	8	0	2	56	0	0	2	3	12	0		35	124	36	232	15	0	2	56	15	0
152	8	1	2	56	0	0	2	3	12	0		34	113	35	218	0	0	2	56	0	0
154	8	0	2	56	0	0	1	2	8	0		25	91	27	172	0	0	2	56	15	0
156	8	1	2	56	15	0	1	2	8	0		40	347	42	473	15	1	2	56	0	0
157	8	0	1	28	0	0	0	1	4	0		22	68	23	137	0	0	1	28	15	0
159	8	0	1	28	15	0	0	1	4	15		29	97	30	187	15	0	1	28	0	0
160	8	0	1	28	15	0	1	1	4	0		0	1	0	1	0	0	1	28	0	0
161	8	0	1	28	0	0	0	1	4	0		123	484	127	865	15	0	1	28	15	0

TABLE A.5: Certificate — Linearly constrained case with variable box (continued)

ex	N	$T = 1$															$T = y _2$														
		BNLC					MPBNGC					SolvOpt					BNLC					MPBNGC					SolvOpt				
		Nit	Na	c	t ₁	t ₂	Nit	Nb	c	t ₁	Nit	Nc	Ng	c	t ₁	Nit	Na	c	t ₁	t ₂	Nit	Nb	c	t ₁	Nit	Nc	Ng	c	t ₁		
162	8	0	1	28	0	0	1	2	8	0	24	78	25	153	15	0	1	28	0	0	0	1	4	0	272	828	274	1650	46		
163	8	0	1	28	0	0	0	1	4	0	63	276	64	468	15	0	1	28	0	0	0	1	4	0	248	722	249	1469	46		
164	8	0	1	28	15	0	0	1	4	0	123	749	126	1127	31	0	1	28	0	0	0	1	4	0	47	359	49	506	31		
165	8	0	1	28	0	0	0	1	4	0	160	665	165	1160	31	0	1	28	15	0	0	1	4	0	297	958	299	1855	46		
166	8	0	1	28	0	0	0	1	4	0	72	276	73	495	0	0	1	28	0	0	0	1	4	0	281	906	283	1755	46		
168	8	0	1	28	15	0	0	1	4	0	138	505	145	940	31	0	1	28	0	0	0	1	4	0	261	880	263	1669	46		
170	8	0	1	28	0	0	0	1	4	0	138	505	143	934	31	0	1	28	0	0	0	1	4	0	225	669	226	1347	31		
171	8	0	1	28	0	0	0	1	4	0	115	431	120	791	15	0	1	28	0	0	0	1	4	0	298	877	299	1774	46		
172	8	0	1	28	15	0	1	1	4	0	0	1	0	1	0	0	1	28	0	0	0	1	4	0	32	125	33	224	15		
189	8	2	3	84	0	0	8	9	36	0	5	23	6	41	0	5	6	168	15	0	5	6	24	0	5	22	6	40	0		
190	8	2	3	84	15	0	3	4	16	0	1	9	1	12	0	2	3	84	15	0	2	3	12	0	1	5	1	8	0		
191	8	1	2	56	0	0	2	3	12	0	1	3	1	6	0	2	3	84	0	0	2	3	12	0	2	8	3	17	0		
192	8	1	2	56	15	0	2	3	12	0	39	312	41	435	15	1	2	56	15	0	2	3	12	0	40	345	42	471	15		
193	8	1	2	56	15	0	2	3	12	0	11	41	13	80	0	1	2	56	0	0	2	3	12	0	14	55	17	106	15		
195	8	1	2	56	15	0	2	3	12	0	20	77	22	143	0	1	2	56	15	0	2	3	12	0	49	386	52	542	31		
197	8	1	2	56	0	0	2	3	12	0	33	115	34	217	15	1	2	56	0	0	2	3	12	0	59	369	60	549	15		
198	8	1	2	56	15	0	2	3	12	0	26	101	27	182	0	1	2	56	15	0	2	3	12	0	35	107	36	215	0		
200	8	1	2	56	0	0	1	2	8	0	44	257	46	395	15	1	2	56	0	0	1	2	8	0	44	250	46	388	15		
201	8	0	1	28	0	0	2	3	12	0	23	69	24	141	0	0	1	28	15	0	1	2	8	0	23	78	25	153	0		
203	8	0	1	28	0	0	0	1	4	0	161	579	164	1071	31	0	1	28	0	0	0	1	4	0	331	954	332	1950	62		
204	8	0	1	28	0	0	0	1	4	0	49	189	51	342	15	0	1	28	0	0	0	1	4	0	76	278	77	509	15		
205	8	0	1	28	0	0	0	1	4	0	250	1121	259	1898	46	0	1	28	15	0	0	1	4	0	48	328	49	475	15		
207	8	0	1	28	0	0	0	1	4	0	168	669	172	1185	46	0	1	28	0	0	0	1	4	0	260	803	261	1586	46		
208	8	0	2	56	0	0	1	2	8	0	29	91	30	181	0	0	1	28	0	0	1	2	8	0	30	92	31	185	15		
209	8	0	1	28	15	15	0	1	4	0	35	126	36	234	0	0	1	28	15	0	0	1	4	0	31	109	32	205	15		
210	8	0	1	28	15	0	0	1	4	0	34	115	35	220	15	0	1	28	0	0	0	1	4	0	44	151	45	286	0		
218	8	1	2	56	15	0	2	3	12	0	1	3	1	6	0	1	3	84	0	0	2	3	12	0	1	3	1	6	0		
219	8	1	2	56	0	0	2	3	12	0	1	3	1	6	0	2	3	84	0	0	2	3	12	0	1	3	1	6	0		
225	8	2	3	84	15	0	2	3	12	0	1	5	1	8	0	2	3	84	0	0	2	3	12	0	1	5	1	8	0		
233	8	1	2	56	0	0	3	4	16	0	1	6	1	9	0	1	2	56	0	0	2	3	12	0	1	5	1	8	0		
234	8	1	2	56	15	0	2	3	12	0	1	4	1	7	0	1	2	56	15	0	2	3	12	0	1	4	1	7	0		
238	8	1	2	56	15	0	2	3	12	0	4	16	5	31	15	1	2	56	0	0	2	3	12	0	1	3	1	6	0		
239	8	1	2	56	0	0	2	3	12	0	7	37	8	61	15	1	2	56	0	0	2	3	12	0	46	448	48	592	15		
240	8	1	2	56	15	0	2	3	12	0	5	23	6	41	0	1	2	56	15	0	2	3	12	0	7	22	8	46	15		
255	8	1	2	56	0	0	2	3	12	0	5	20	6	38	0	1	2	56	15	15	2	3	12	0	5	22	6	40	0		
259	8	1	2	56	0	0	2	3	12	0	7	25	8	49	0	1	2	56	15	0	2	3	12	0	9	30	10	60	0		
260	8	1	2	56	15	0	2	3	12	0	14	45	15	90	0	1	2	56	0	0	2	3	12	0	15	47	16	95	0		
261	8	1	2	56	0	0	2	3	12	0	41	228	45	363	15	1	2	56	15	0	2	3	12	0	21	74	24	146	0		
267	8	1	2	56	15	0	2	3	12	0	17	69	18	123	15	1	2	56	0	0	2	3	12	0	18	65	19	122	0		
268	8	1	2	56	0	0	2	3	12	0	28	93	29	180	0	1	2	56	15	0	2	3	12	0	26	85	27	166	0		
269	8	1	2	56	15	0	2	3	12	0	26	86	27	167	0	1	2	56	0	0	2	3	12	0	28	98	29	185	15		
270	8	0	2	56	0	0	2	3	12	0	27	112	28	196	15	0	2	56	0	0	1	2	8	0	24	85	25	160	0		
278	8	1	2	56	0	0	2	3	12	0	21	83	23	152	0	0	2	56	15	0	2	3	12	0	26	91	27	172	15		
279	8	0	2	56	0	0	2	3	12	0	30	104	31	197	15	0	2	56	15	0	1	2	8	0	41	160	43	289	0		
280	8	1	2	56	15	0	2	3	12	0	36	129	38	243	0	0	2	56	0	0	2	3	12	0	28	102	29	189	15		
281	8	0	2	56	0	0	1	2	8	0	21	69	23	138	0	0	2	56	15	0	1	2	8	0	28	122	30	212	0		
285	8	0	1	28	0	0	1	2	8	0	38	132	40	252	0	0	1	28	15	0	0	1	4	0	27	87	29	174	0		
286	8	0	1	28	0	0	1	1	4	0	0	1	0	1	0	0	1	28	0	0	0	1	4	0	28	88	29	175	0		

TABLE A.5: Certificate — Linearly constrained case with variable box (continued)

		$T = 1$															$T = y _2$														
		BNLC					MPBNGC					SolvOpt					BNLC					MPBNGC					SolvOpt				
ex	N	Nit	Na	c	t ₁	t ₂	Nit	Nb	c	t ₁	Nit	Nc	Ng	c	t ₁	Nit	Na	c	t ₁	t ₂	Nit	Nb	c	t ₁	Nit	Nc	Ng	c	t ₁		
289	8	1	2	56	0	0	2	3	12	0	36	123	37	234	15	0	2	56	15	0	1	2	8	0	24	75	25	150	0		
290	8	0	2	56	15	0	2	3	12	0	24	77	25	152	0	0	2	56	0	0	1	2	8	0	30	101	31	194	15		
291	8	0	1	28	0	0	1	2	8	0	25	76	26	154	0	0	1	28	0	0	1	2	8	0	32	101	34	203	0		
293	8	0	1	28	0	0	0	1	4	0	28	93	29	180	0	0	1	28	0	0	0	1	4	0	30	91	31	184	0		
294	8	0	2	56	15	0	1	2	8	0	39	147	41	270	15	0	1	28	0	0	1	2	8	0	28	80	29	167	15		
295	8	0	1	28	0	0	0	1	4	0	53	235	54	397	15	0	1	28	0	0	0	1	4	0	51	409	52	565	31		
21	10	2	3	102	30	0	1	2	8	0	0	7	2	13	16	3	4	136	15	0	1	2	8	0	0	7	2	13	0		
22	10	2	3	102	0	0	1	2	8	0	0	7	2	13	16	1	2	68	15	0	1	2	8	0	0	7	2	13	0		
24	10	1	2	68	0	0	1	2	8	0	37	340	40	460	15	1	2	68	15	0	1	2	8	0	31	322	34	424	31		
57	10	1	2	68	0	0	1	2	8	0	0	8	2	14	0	1	2	68	31	0	1	2	8	0	58	1186	60	1366	93		
58	10	1	2	68	0	0	1	2	8	0	28	248	30	338	46	1	2	68	15	0	1	2	8	0	52	562	54	724	31		
59	10	1	2	68	0	0	1	2	8	0	47	887	49	1034	46	1	2	68	15	0	1	2	8	0	2	10	4	22	16		
60	10	1	2	68	15	0	1	2	8	0	36	250	38	364	15	1	2	68	0	0	1	2	8	0	35	251	37	362	15		
61	10	1	2	68	15	0	1	2	8	0	37	333	39	450	15	1	2	68	0	0	1	2	8	0	82	2101	86	2359	109		
62	10	1	2	68	0	0	1	2	8	0	44	249	47	390	15	1	2	68	15	0	1	2	8	0	46	310	49	457	15		
67	10	1	2	68	0	0	1	2	8	0	1	6	3	15	16	1	2	68	0	0	1	2	8	0	1	6	3	15	16		
68	10	1	2	68	31	0	1	2	8	0	52	762	56	930	46	1	2	68	0	0	1	2	8	0	2	8	4	20	16		
70	10	1	2	68	0	0	1	2	8	0	32	399	34	501	15	1	2	68	15	0	1	2	8	0	59	1338	61	1521	78		
71	10	1	2	68	0	0	1	2	8	0	54	847	56	1015	31	1	2	68	0	0	1	2	8	0	35	364	37	475	31		
72	10	1	2	68	15	0	1	2	8	0	125	3466	127	3847	171	1	2	68	0	0	1	2	8	0	32	226	34	328	15		
73	10	1	2	68	0	0	1	2	8	0	43	424	46	562	15	1	2	68	15	0	1	2	8	0	59	1207	61	1390	62		
74	10	1	2	68	15	0	1	2	8	0	2	8	4	20	16	1	2	68	0	0	1	2	8	0	2	8	4	20	16		
91	10	1	2	68	15	0	1	2	8	0	4	21	6	39	0	1	2	68	0	0	1	2	8	0	6	17	8	41	47		
92	10	1	2	68	0	0	1	2	8	0	1	6	3	15	0	1	2	68	0	0	1	2	8	0	1	6	3	15	0		
93	10	1	2	68	0	0	1	2	8	0	1	6	3	15	0	1	2	68	16	0	1	2	8	0	1	6	3	15	0		
94	10	1	2	68	15	0	1	2	8	0	1	6	3	15	0	1	2	68	15	0	1	2	8	0	1	8	3	17	0		
96	10	1	2	68	0	0	1	2	8	0	47	359	50	509	15	1	2	68	15	0	1	2	8	0	50	423	54	585	31		
98	10	0	1	34	0	0	1	2	8	0	7	26	10	56	0	0	1	34	0	0	0	1	4	0	12	61	16	109	16		
99	10	1	2	68	31	0	1	2	8	0	38	284	41	407	15	1	2	68	30	0	1	2	8	0	36	284	39	401	15		
310	21	2	3	201	15	0	1	2	8	0	162	615	165	1110	171	2	3	201	15	0	6	10	40	15	101	826	106	1144	218		

TABLE A.6: Certificate — Nonlinearly constrained case for $T = 1$

ex	N	Red Alg					MPBNGC				SolvOpt					
		Nit	Na	c	t ₁	t ₂	Nit	Nb	c	t ₁	Nit	Nc	N _g	N _g	c	t ₁
100	8	35	39	2184	77	17	25	63	504	0	40	175	41	37	584	0
101	8	62	71	3976	171	36	34	62	496	0	838	3896	839	556	11977	234
102	8	26	30	1680	78	18	21	37	296	0	217	825	220	122	2676	62
103	8	9	15	840	31	1	21	35	280	15	88	481	91	56	1403	31
104	8	27	28	1568	77	32	18	29	232	0	227	838	230	128	2750	46
105	8	28	29	1624	78	33	22	31	248	0	94	380	96	53	1207	15
106	8	31	58	3248	93	18	23	46	368	0	136	454	138	86	1580	31
107	8	57	58	3248	171	21	4001	8000	64000	656	299	1211	301	183	3874	78
108	8	66	99	5544	187	22	62	225	1800	15	1491	5750	1493	1349	20026	374
109	8	435	437	24472	1187	272	3995	8000	64000	734	3706	16320	3709	3669	54774	1062
110	8	2	4	224	15	0	45	206	1648	15	528	1897	530	520	6944	140
111	8	2	3	168	0	0	46	199	1592	0	496	1646	498	470	6196	109
112	8	15	16	896	31	16	35	146	1168	0	5792	25061	5794	5787	84865	1671
113	8	30	31	1736	93	18	13	33	264	0	4125	17746	4129	4077	60110	1155
114	8	2	5	280	15	0	18	49	392	0	155	506	157	106	1801	46
115	8	2	5	280	0	0	21	62	496	0	424	1831	426	296	5828	109
116	8	67	68	3808	187	52	18	40	320	0	3318	15068	3320	3302	50002	968
121	8	96	97	5432	358	58	17	56	448	0	3095	13358	3097	3096	45295	890
122	8	49	50	2800	140	20	10	13	104	0	1650	7542	1653	1622	24909	484
123	8	27	28	1568	93	3	17	58	464	15	1475	7032	1477	1461	22878	437
124	8	15	16	896	46	1	123	8000	64000	343	349	1854	352	315	5709	109
125	8	140	146	8176	390	75	3992	8000	64000	640	2668	10452	2670	2662	36900	703
126	8	61	86	4816	171	21	93	313	2504	15	140	543	142	96	1800	31
127	8	32	44	2464	93	33	26	48	384	0	590	2359	592	364	7586	156
128	8	16	19	1064	46	1	42	105	840	0	682	2682	684	418	8670	156
129	8	60	75	4200	156	66	33	62	496	0	223	828	225	120	2691	62
130	8	13	20	1120	46	16	37	86	688	0	203	770	207	118	2515	46
131	8	13	15	840	46	1	41	136	1088	0	2078	9097	2081	2052	30593	578
132	8	15	17	952	31	1	7926	8000	64000	1078	173	639	176	117	2157	31
133	8	12	18	1008	31	1	7986	8000	64000	952	263	877	265	177	3080	46
134	8	8	9	504	15	0	18	32	256	15	177	645	180	107	2151	46
135	8	17	18	1008	46	16	32	70	560	0	183	624	186	106	2124	31
136	8	20	23	1288	46	1	31	76	608	15	200	727	203	112	2399	46
137	8	17	47	2632	61	31	25	37	296	0	2763	11155	2765	2735	38810	749
138	8	8	9	504	15	0	13	20	160	0	211	794	213	128	2611	46
139	8	13	14	784	46	16	20	34	272	0	187	687	189	120	2301	46
140	8	13	14	784	31	1	24	31	248	15	180	645	182	107	2157	46
141	8	13	14	784	46	16	21	27	216	0	207	763	209	126	2531	46
142	8	15	22	1232	46	1	26	52	416	0	226	762	228	133	2607	62
143	8	15	21	1176	31	16	31	104	832	15	173	584	175	105	2008	31
144	8	14	15	840	30	0	39	148	1184	15	167	570	169	103	1956	31
145	8	13	14	784	46	1	1152	8000	64000	421	200	667	202	121	2303	46
146	8	116	119	6664	311	116	13	30	240	0	167	591	169	111	2022	46
147	8	25	26	1456	62	2	7983	8000	64000	1046	1278	5617	1279	1274	18893	359
148	8	21	22	1232	46	1	24	44	352	0	2019	8851	2021	2017	29816	578
149	8	21	22	1232	62	17	25	56	448	15	1314	5998	1316	1312	19880	375
150	8	21	22	1232	62	17	35	97	776	0	1866	8102	1868	1864	27400	530
151	8	2	3	168	15	0	27	78	624	0	311	1156	313	230	3941	78
152	8	47	48	2688	140	50	28	86	688	0	587	2953	589	582	9419	171
153	8	47	48	2688	124	34	40	147	1176	15	1323	5592	1324	1318	19110	375
154	8	97	98	5488	280	55	41	380	3040	31	1447	6288	1448	1445	21255	406
155	8	96	97	5432	280	55	162	8000	64000	343	643	2883	644	643	9627	187
156	8	33	34	1904	93	18	436	3897	31176	218	1587	6817	1589	1583	23150	453
157	8	33	34	1904	93	3	36	198	1584	15	2129	9251	2130	2129	31279	609
158	8	27	28	1568	93	3	37	161	1288	0	630	3050	632	630	9886	203
159	8	41	42	2352	125	20	51	295	2360	31	1438	6360	1440	1439	21357	406
160	8	35	36	2016	109	4	34	193	1544	15	1536	6573	1538	1537	22371	437
161	8	52	53	2968	155	20	17	46	368	0	1999	8472	2001	2000	28947	578
162	8	14	25	1400	31	1	15	32	256	0	218	817	220	157	2765	46
163	8	19	34	1904	62	17	16	63	504	0	123	406	125	91	1460	31
165	8	74	75	4200	359	74	5	7	56	0	246	2053	248	237	5561	109
166	8	103	104	5824	374	74	39	302	2416	15	445	2192	447	446	7063	125
168	8	42	81	4536	124	4	15	38	304	0	337	1278	339	245	4308	78
170	8	37	66	3696	109	19	10	12	96	0	307	1217	309	220	4021	78
171	8	19	34	1904	61	16	645	8000	64000	406	282	1045	284	201	3545	109
172	8	14	25	1400	46	1	450	3718	29744	187	228	789	230	156	2736	46
173	8	17	24	1344	46	16	24	70	560	15	251	992	253	183	3292	62
174	8	19	28	1568	46	16	48	232	1856	15	165	563	167	120	1987	46
175	8	20	30	1680	62	32	33	121	968	0	189	687	191	133	2346	46
176	8	17	24	1344	46	16	18	52	416	0	297	1049	299	203	3604	62
177	8	15	22	1232	46	1	23	56	448	0	150	545	152	106	1864	31
178	8	21	28	1568	62	17	17	28	224	0	154	538	156	110	1874	46
179	8	14	18	1008	46	1	18	27	216	0	219	735	221	145	2568	46
180	8	11	16	896	31	1	23	44	352	15	314	1074	316	215	3741	78
181	8	10	12	672	31	1	27	51	408	0	208	721	210	117	2423	31
182	8	11	13	728	31	1	22	53	424	0	184	659	186	102	2182	46

TABLE A.6: Certificate — Nonlinearly constrained case for $T = 1$ (continued)

ex	N	Red Alg					MPBNGC				SolvOpt					
		Nit	Na	c	t ₁	t ₂	Nit	Nb	c	t ₁	Nit	Nc	Ng	N $\bar{g}$	c	t ₁
183	8	17	20	1120	46	16	4003	8000	64000	656	206	726	208	122	2442	46
184	8	10	12	672	31	1	22	39	312	15	203	710	205	117	2386	46
185	8	11	14	784	31	1	46	193	1544	0	151	533	153	85	1780	46
186	8	21	25	1400	46	16	21	30	240	0	242	885	247	147	2952	46
187	8	21	25	1400	62	2	23	41	328	0	179	671	181	101	2188	46
188	8	23	27	1512	62	2	21	33	264	0	178	663	180	98	2160	46
189	8	19	26	1456	46	16	23	43	344	15	191	751	194	111	2417	46
190	8	24	27	1512	62	2	40	80	640	0	284	1090	287	173	3560	78
191	8	23	24	1344	62	17	36	137	1096	15	244	826	247	147	2834	62
192	8	13	98	5488	46	16	20	33	264	0	197	678	199	121	2316	46
193	8	14	16	896	31	1	18	34	272	0	215	835	217	138	2735	46
194	8	11	14	784	46	16	28	50	400	0	184	667	186	104	2204	46
195	8	59	64	3584	156	21	21	44	352	0	211	780	213	140	2619	46
196	8	17	21	1176	46	1	24	50	400	15	291	1004	293	201	3490	78
197	8	23	24	1344	61	1	30	77	616	0	1961	8258	1962	1957	28273	546
198	8	13	18	1008	31	31	29	85	680	15	145	487	147	100	1715	31
199	8	14	19	1064	46	16	23	61	488	0	216	799	218	144	2684	62
200	8	73	74	4144	202	52	44	145	1160	15	460	2092	461	455	6932	125
201	8	2	3	168	15	0	37	154	1232	0	137	495	140	101	1713	31
202	8	25	36	2016	61	1	39	144	1152	15	221	816	223	160	2781	62
203	8	2	4	224	0	0	14	36	288	0	185	686	188	133	2335	46
204	8	2	4	224	15	0	25	84	672	15	207	723	210	146	2514	46
207	8	0	1	56	15	0	3	5	40	0	298	2598	301	266	6897	140
208	8	17	18	1008	46	16	1009	8000	64000	421	1224	5117	1227	1226	17593	328
209	8	38	39	2184	109	4	20	75	600	0	2510	10815	2512	2509	36693	718
210	8	44	45	2520	125	35	16	40	320	0	631	2646	633	629	9078	171
211	8	35	36	2016	93	18	11	16	128	0	483	3003	485	437	8772	171
212	8	19	20	1120	46	31	49	196	1568	15	3096	13418	3099	3097	45424	874
213	8	23	24	1344	61	1	35	67	536	0	1492	6114	1493	1490	21177	406
214	8	80	82	4592	218	23	2004	8000	64000	499	304	1128	307	185	3732	62
215	8	81	82	4592	218	23	285	8000	64000	359	259	910	261	157	3074	62
216	8	80	85	4760	218	68	7977	8000	64000	1125	168	680	171	109	2200	46
217	8	22	23	1288	62	17	26	45	360	15	2341	10104	2342	2275	34059	656
218	8	38	52	2912	108	18	126	301	2408	15	254	986	257	143	3172	62
219	8	19	21	1176	61	1	38	76	608	0	3658	14968	3661	3620	51779	999
220	8	19	22	1232	46	16	10	17	136	0	71	392	75	40	1129	15
221	8	31	53	2968	93	33	64	202	1616	15	410	1775	414	257	5563	93
222	8	20	21	1176	61	1	21	36	288	15	112	636	114	65	1809	31
223	8	28	37	2072	78	18	31	57	456	0	147	854	149	86	2413	46
224	8	35	38	2128	93	18	13	21	168	0	107	634	110	67	1799	31
225	8	28	32	1792	93	33	23	33	264	0	86	596	87	41	1576	31
226	8	46	52	2912	109	4	33	51	408	15	148	574	150	78	1832	31
227	8	19	21	1176	62	17	7	10	80	0	48	262	49	28	755	15
228	8	33	34	1904	77	17	20	32	256	0	211	791	214	112	2560	46
229	8	30	35	1960	93	3	18	21	168	0	84	450	87	48	1305	31
230	8	28	34	1904	61	16	17	27	216	15	88	517	90	48	1448	31
231	8	37	38	2128	93	3	6	8	64	0	66	379	67	36	1067	15
232	8	47	49	2744	124	34	8	11	88	0	47	263	49	25	748	15
233	8	32	36	2016	78	18	15	17	136	0	104	639	105	56	1761	31
234	8	14	18	1008	31	1	125	498	3984	31	146	508	149	82	1709	31
235	8	45	46	2576	124	4	33	74	592	15	127	484	129	68	1559	31
236	8	46	47	2632	125	35	32	75	600	0	207	803	210	116	2584	46
237	8	51	52	2912	140	20	25	55	440	0	232	969	236	132	3042	62
238	8	36	37	2072	108	3	25	50	400	0	159	579	162	88	1908	46
239	8	23	32	1792	62	17	7975	8000	64000	1233	230	796	232	130	2678	46
240	8	68	158	8848	202	37	335	1172	9376	78	252	932	254	149	3073	62
241	8	31	55	3080	78	3	85	272	2176	15	222	827	225	126	2707	46
242	8	33	54	3024	93	18	97	309	2472	15	203	704	205	115	2368	46
243	8	32	55	3080	78	18	72	279	2232	15	233	824	235	141	2776	46
244	8	57	133	7448	171	6	262	896	7168	62	161	639	163	105	2082	31
245	8	59	135	7560	203	53	149	681	5448	46	279	1002	281	167	3348	62
246	8	46	93	5208	125	20	109	472	3776	31	133	490	135	87	1646	31
247	8	43	92	5152	124	19	175	594	4752	46	483	2142	485	332	6735	140
248	8	53	110	6160	156	6	4002	8000	64000	625	1758	7421	1761	1677	25156	468
249	8	44	88	4928	125	50	108	384	3072	31	439	1855	441	395	6218	125
250	8	58	100	5600	171	36	74	282	2256	15	427	1479	430	269	5055	93
251	8	59	109	6104	156	36	53	204	1632	15	304	1144	307	192	3785	62
252	8	59	87	4872	171	36	135	438	3504	31	150	608	152	104	1984	31
253	8	50	69	3864	140	5	33	62	496	0	590	2462	593	563	8392	156
254	8	65	86	4816	186	36	72	230	1840	15	1334	5565	1337	1328	19125	359
255	8	18	19	1064	61	16	15	22	176	15	175	636	177	105	2118	46
256	8	17	18	1008	46	1	16	25	200	0	191	696	193	102	2277	46
257	8	17	18	1008	46	16	18	27	216	0	231	896	233	132	2887	62
258	8	19	20	1120	62	2	17	27	216	0	214	754	216	124	2528	62
259	8	17	18	1008	46	31	14	19	152	0	167	580	169	100	1967	31
260	8	14	15	840	46	16	16	29	232	15	241	889	243	135	2912	46

TABLE A.6: Certificate — Nonlinearly constrained case for $T = 1$ (continued)

ex	N	Red Alg					MPBNGC				SolvOpt					
		Nit	Na	c	t ₁	t ₂	Nit	Nb	c	t ₁	Nit	Nc	N _g	N $\hat{g}$	c	t ₁
261	8	19	20	1120	46	1	21	31	248	0	122	461	124	75	1519	31
262	8	23	24	1344	62	17	4001	8000	64000	655	211	827	213	123	2662	46
263	8	22	23	1288	61	1	4000	8000	64000	609	206	738	208	121	2463	46
264	8	23	25	1400	61	1	26	51	408	15	178	648	180	104	2148	31
265	8	24	32	1792	78	33	21	50	400	0	189	709	191	110	2321	46
266	8	18	19	1064	46	1	15	25	200	0	257	880	259	147	2978	62
267	8	17	19	1064	46	1	17	32	256	0	193	704	195	110	2323	46
268	8	54	61	3416	155	50	7968	8000	64000	1077	2963	12509	2965	2953	42772	828
269	8	73	77	4312	203	68	25	46	368	0	207	710	209	165	2542	46
270	8	151	156	8736	405	135	64	239	1912	15	1361	5645	1363	1358	19453	374
271	8	156	157	8792	437	92	4457	8000	64000	781	1351	6097	1353	1348	20297	390
272	8	108	110	6160	312	12	85	382	3056	15	177	758	180	139	2473	46
273	8	238	240	13440	656	146	175	8000	64000	359	1575	6743	1577	1576	22945	437
274	8	173	174	9744	483	33	54	116	928	15	1644	6666	1646	1639	23187	437
275	8	296	297	16632	827	122	2004	8000	64000	468	874	3593	876	874	12436	234
276	8	255	257	14392	702	162	49	242	1936	15	2099	8366	2101	2094	29317	562
277	8	17	18	1008	46	16	132	8000	64000	343	1738	6998	1740	1738	24430	468
278	8	108	109	6104	312	42	3991	8000	64000	625	2467	10466	2469	2457	35710	686
279	8	61	71	3976	171	36	75	341	2728	31	189	676	191	130	2315	46
280	8	260	261	14616	780	120	109	628	5024	31	733	3081	735	714	10509	203
281	8	25	49	2744	77	17	582	5970	47760	296	171	570	173	123	2028	46
282	8	39	40	2240	124	19	41	103	824	0	2351	10238	2353	2352	34591	671
283	8	36	37	2072	108	3	1996	8000	64000	468	1203	5232	1205	1200	17679	343
284	8	43	44	2464	140	5	169	8000	64000	343	2768	11037	2770	2768	38688	733
285	8	2	4	224	15	0	683	8000	64000	390	177	621	179	125	2154	46
286	8	2	5	280	15	15	13	28	224	15	172	644	174	136	2218	46
287	8	2	4	224	0	0	53	422	3376	31	323	1171	325	224	3989	78
288	8	218	219	12264	655	130	39	129	1032	0	2778	11695	2780	2762	40016	781
289	8	17	18	1008	46	1	51	251	2008	15	3401	15010	3403	3385	50384	984
290	8	17	18	1008	46	16	123	8000	64000	343	860	3792	862	860	12750	234
291	8	28	29	1624	93	33	62	414	3312	15	750	3192	752	751	10893	203
292	8	25	26	1456	78	18	2655	8000	64000	562	290	1564	292	283	4853	93
293	8	38	39	2184	125	50	8	10	80	0	2111	9033	2113	2112	30741	593
294	8	2	4	224	0	0	57	235	1880	15	245	842	247	161	2908	62
295	8	2	5	280	15	0	37	239	1912	0	401	1448	403	297	4996	93
296	8	2	4	224	15	0	39	156	1248	0	165	565	167	101	1934	31
297	8	264	265	14840	749	104	43	106	848	15	2806	12075	2808	2782	40920	796
298	8	154	155	8680	437	92	46	127	1016	15	520	1848	522	513	6801	125
299	8	166	167	9352	468	168	50	168	1344	15	428	1442	430	419	5431	93
300	8	149	150	8400	421	46	53	137	1096	0	584	2017	586	576	7520	140
301	8	86	90	5040	249	54	41	174	1392	0	150	502	152	126	1838	31
302	8	26	27	1512	78	3	21	41	328	15	203	773	205	118	2515	46
303	8	28	29	1624	77	2	4000	8000	64000	640	169	627	171	102	2073	31
304	8	28	29	1624	78	18	4001	8000	64000	593	220	821	222	131	2701	46
305	8	17	19	1064	46	1	20	37	296	15	194	722	196	121	2395	46
306	8	14	15	840	46	31	16	27	216	0	207	747	209	117	2472	46
307	8	19	20	1120	62	2	21	38	304	0	206	742	208	120	2468	46
308	8	30	31	1736	78	18	21	46	368	15	238	953	240	133	3025	62
309	8	40	41	2296	109	19	14	20	160	0	80	447	82	43	1269	15

TABLE A.7: Certificate — Nonlinearly constrained case for $T = |y|_2$

ex	N	Red Alg					MPBNGC				SolvOpt					
		Nit	Na	c	t ₁	t ₂	Nit	Nb	c	t ₁	Nit	Nc	N _g	N $\hat{g}$	c	t ₁
100	8	37	42	2352	93	3	1212	6000	48000	359	105	351	107	62	1209	31
101	8	105	108	6048	281	86	26	42	336	0	69	218	70	35	751	15
102	8	22	33	1848	62	17	26	54	432	0	89	319	92	51	1067	15
103	8	33	37	2072	93	3	24	40	320	0	193	722	195	85	2284	46
104	8	11	12	672	31	1	26	39	312	15	94	339	96	56	1134	15
105	8	13	14	784	31	16	23	41	328	0	216	757	218	98	2462	62
106	8	16	39	2184	46	1	1205	6000	48000	421	128	428	130	83	1495	31
107	8	67	68	3808	187	52	33	92	736	15	197	663	199	123	2292	46
108	8	53	54	3024	156	51	38	90	720	0	204	686	206	132	2386	62
109	8	11	12	672	31	16	5836	6000	48000	875	890	2972	891	888	11281	218
110	8	2	5	280	0	0	533	5506	44048	296	378	1124	379	366	4483	78
111	8	2	5	280	0	0	120	522	4176	31	366	1136	367	359	4450	78
112	8	13	14	784	46	16	289	1851	14808	125	1905	6323	1907	1904	24079	484
113	8	28	29	1624	77	2	52	269	2152	15	1581	5367	1583	1582	20229	453
114	8	2	6	336	0	0	68	158	1264	15	141	450	143	90	1599	31
115	8	2	6	336	15	0	189	393	3144	46	721	2319	723	619	8664	171
116	8	38	39	2184	124	4	0	0	0	0	1043	3512	1045	1033	13258	265
121	8	92	93	5208	436	46	0	0	0	63	1120	3852	1122	1121	14433	281
122	8	43	44	2464	124	34	83	6000	48000	390	1156	3835	1158	1152	14600	296
123	8	23	24	1344	61	1	71	270	2160	15	605	2095	607	605	7826	156
124	8	13	14	784	46	16	66	481	3848	31	2056	6577	2058	2055	25493	515

TABLE A.7: Certificate — Nonlinearly constrained case for $T = |y|_2$ (continued)

ex	N	Red Alg					MPBNGC				SolvOpt					
		Nit	Na	c	t ₁	t ₂	Nit	Nb	c	t ₁	Nit	Nc	Ng	N $\hat{g}$	c	t ₁
125	8	187	192	10752	530	155	2986	6000	48000	578	1214	4335	1216	1213	15957	312
126	8	53	54	3024	140	50	42	108	864	15	142	496	144	96	1712	31
127	8	50	73	4088	125	20	23	39	312	0	108	417	110	57	1335	15
128	8	45	53	2968	124	19	36	94	752	15	369	1388	372	215	4537	109
129	8	27	33	1848	78	18	5991	6000	48000	750	150	583	152	87	1883	46
130	8	11	14	784	31	1	18	34	272	0	230	785	232	110	2596	46
131	8	13	14	784	31	1	5951	6000	48000	921	269	964	272	170	3254	62
132	8	13	14	784	30	0	1239	6000	48000	390	124	457	127	84	1547	31
133	8	11	12	672	31	1	5973	6000	48000	874	253	866	255	174	3019	46
134	8	25	26	1456	62	32	33	86	688	15	187	643	191	106	2177	46
135	8	16	24	1344	46	1	30	51	408	0	348	1246	350	320	4502	93
136	8	18	20	1120	46	1	1993	6000	48000	484	154	560	157	94	1873	46
137	8	14	24	1344	31	1	26	57	456	0	210	731	212	114	2440	46
138	8	23	24	1344	62	2	28	56	448	0	277	982	279	141	3224	62
139	8	17	18	1008	46	1	34	97	776	15	568	1793	569	345	6328	140
140	8	13	28	1568	31	1	36	201	1608	15	140	489	142	86	1662	46
141	8	13	20	1120	46	1	22	27	216	0	216	721	218	119	2453	46
142	8	16	17	952	46	16	22	34	272	15	188	611	190	125	2167	46
143	8	14	15	840	31	16	20	37	296	0	187	620	189	118	2161	46
144	8	16	26	1456	46	16	5877	6000	48000	1156	220	723	222	120	2472	62
145	8	15	16	896	31	1	18	24	192	0	117	406	119	80	1409	31
146	8	68	71	3976	202	52	31	73	584	0	236	782	239	143	2710	62
147	8	32	33	1848	93	33	88	2643	21144	125	337	1190	338	335	4399	93
148	8	28	29	1624	77	17	59	324	2592	31	411	1405	412	401	5249	109
149	8	27	28	1568	78	3	1507	6000	48000	406	320	1078	321	315	4064	78
150	8	29	30	1680	78	33	96	6000	48000	281	441	1433	443	438	5509	109
151	8	2	3	168	15	0	23	56	448	15	206	705	208	130	2424	46
152	8	72	73	4088	203	8	872	6000	48000	359	601	1961	603	602	7537	156
153	8	73	74	4144	202	52	1186	6000	48000	375	575	2062	577	573	7574	140
154	8	164	165	9240	468	78	36	149	1192	0	306	1003	307	306	3845	78
155	8	164	165	9240	452	92	151	6000	48000	281	454	1508	456	455	5749	109
156	8	30	31	1736	78	18	73	292	2336	31	403	1361	405	401	5140	93
157	8	29	30	1680	77	2	0	0	0	16	517	1679	519	518	6469	125
158	8	24	25	1400	77	2	0	0	0	0	531	1734	533	532	6663	125
159	8	38	39	2184	109	19	0	0	0	47	454	1506	456	455	5745	109
160	8	32	33	1848	93	18	0	0	0	0	532	1795	534	529	6779	125
161	8	48	49	2744	124	19	112	6000	48000	265	708	2412	710	709	9081	187
162	8	14	25	1400	31	1	67	218	1744	15	198	623	200	115	2191	46
163	8	2	3	168	15	0	147	258	2064	15	169	588	172	120	2052	46
165	8	104	105	5880	437	76	15	50	400	0	920	3020	922	921	11569	234
166	8	105	106	5936	483	33	137	3763	30104	187	582	1930	584	583	7361	156
168	8	51	95	5320	155	35	0	0	0	16	185	596	187	116	2101	46
170	8	39	70	3920	109	49	811	1179	9432	125	173	559	175	113	1982	31
171	8	20	36	2016	62	2	131	249	1992	31	144	472	146	100	1682	46
172	8	15	26	1456	62	17	132	1111	8888	62	197	619	199	124	2207	46
173	8	15	23	1288	61	1	95	551	4408	31	111	357	113	83	1302	31
174	8	21	32	1792	62	17	46	309	2472	15	160	516	162	104	1830	31
175	8	13	20	1120	31	1	59	435	3480	15	121	386	123	87	1402	31
176	8	15	22	1232	46	16	124	1065	8520	46	222	704	224	127	2461	62
177	8	12	19	1064	30	0	55	285	2280	15	111	353	113	76	1273	31
178	8	12	16	896	31	16	35	89	712	15	107	361	109	65	1244	15
179	8	12	16	896	31	1	24	58	464	0	118	380	120	77	1351	31
180	8	10	14	784	31	1	13	19	152	0	115	392	117	82	1381	31
181	8	10	12	672	31	1	17	29	232	0	105	371	107	65	1258	31
182	8	8	14	784	15	0	3001	6000	48000	531	159	540	161	95	1848	46
183	8	13	16	896	31	1	1506	6000	48000	406	152	568	154	101	1901	31
184	8	9	14	784	31	1	22	35	280	0	234	771	236	121	2613	46
185	8	7	9	504	15	15	20	33	264	0	108	390	110	66	1308	15
186	8	21	29	1624	46	16	17	26	208	15	100	371	105	59	1234	31
187	8	20	21	1176	62	2	21	28	224	0	160	600	162	80	1926	46
188	8	19	22	1232	46	1	21	32	256	0	133	502	135	75	1634	31
189	8	19	26	1456	62	2	23	33	264	0	246	845	250	132	2836	46
190	8	22	25	1400	62	17	32	53	424	15	281	985	283	168	3323	62
191	8	15	16	896	46	16	90	6000	48000	265	107	369	109	62	1251	31
192	8	11	12	672	31	16	16	28	224	0	208	702	210	108	2358	46
193	8	11	12	672	31	1	27	57	456	15	193	771	195	118	2481	46
194	8	7	9	504	15	0	24	40	320	0	153	512	155	88	1753	46
195	8	36	39	2184	108	48	24	72	576	0	152	516	154	101	1797	31
196	8	17	21	1176	46	1	20	31	248	15	133	456	135	89	1584	31
197	8	31	33	1848	78	3	73	270	2160	31	373	1308	374	370	4848	93
198	8	11	17	952	31	16	31	93	744	0	217	668	219	129	2380	62
199	8	15	21	1176	31	1	39	129	1032	15	228	722	230	140	2554	46
200	8	123	124	6944	343	58	861	6000	48000	343	639	2198	641	640	8239	156
201	8	2	3	168	15	15	474	6000	48000	312	101	342	103	74	1215	31
202	8	25	36	2016	62	17	467	6000	48000	312	209	682	211	134	2399	46
203	8	2	3	168	0	0	163	6000	48000	281	283	814	286	207	3107	62

TABLE A.7: Certificate — Nonlinearly constrained case for $T = |y|_2$ (continued)

ex	N	Red Alg					MPBNGC				SolvOpt					
		Nit	Na	c	t ₁	t ₂	Nit	Nb	c	t ₁	Nit	Nc	N _g	N _g	c	t ₁
204	8	2	4	224	15	0	60	541	4328	31	139	500	142	97	1717	31
207	8	152	155	8680	640	115	0	0	0	31	628	2092	630	629	7961	171
208	8	15	16	896	46	1	103	6000	48000	265	405	1368	407	394	5139	109
209	8	34	35	1960	93	3	0	0	0	0	590	1939	592	581	7397	140
210	8	39	40	2240	109	4	140	558	4464	46	680	2318	682	679	8719	171
211	8	31	32	1792	93	3	0	0	0	0	431	1351	433	432	5297	109
212	8	17	18	1008	46	1	759	6000	48000	343	802	2615	804	788	10006	203
213	8	30	31	1736	78	3	1999	6000	48000	453	721	2397	722	719	9117	187
214	8	69	70	3920	187	7	5973	6000	48000	843	246	885	248	175	3039	62
215	8	77	78	4368	203	23	27	42	336	0	167	622	169	110	2081	46
216	8	77	78	4368	218	23	49	224	1792	15	308	1040	309	299	3904	78
217	8	20	24	1344	46	16	1988	6000	48000	421	180	714	182	104	2286	46
218	8	53	85	4760	140	50	3003	6000	48000	499	132	514	134	80	1670	31
219	8	49	54	3024	140	20	55	101	808	0	160	612	163	98	2007	46
220	8	34	39	2184	77	62	12	15	120	0	80	443	84	46	1276	31
221	8	29	52	2912	77	2	24	41	328	0	226	1110	228	133	3303	78
222	8	16	19	1064	46	31	12	19	152	0	99	519	102	58	1518	31
223	8	36	45	2520	109	4	15	22	176	0	566	2234	570	386	7336	140
224	8	49	53	2968	140	35	22	32	256	15	123	637	126	72	1868	46
225	8	41	46	2576	125	20	36	62	496	0	202	759	205	103	2442	46
226	8	62	98	5488	156	21	31	65	520	15	161	561	164	93	1893	31
227	8	125	127	7112	311	56	6	8	64	0	53	272	54	31	799	15
228	8	40	43	2408	108	48	18	30	240	0	197	700	200	85	2255	46
229	8	22	24	1344	61	16	15	22	176	0	83	451	87	56	1331	15
230	8	24	26	1456	62	17	13	17	136	0	95	534	98	56	1530	31
231	8	21	23	1288	62	2	10	12	96	0	58	315	59	30	897	15
232	8	38	42	2352	93	18	5	7	56	0	73	398	76	45	1159	31
233	8	35	47	2632	109	34	20	22	176	15	83	515	84	44	1414	31
234	8	11	14	784	31	16	32	64	512	0	150	559	152	79	1811	46
235	8	24	25	1400	62	2	28	39	312	0	185	679	187	96	2207	46
236	8	21	22	1232	62	2	35	58	464	0	139	531	142	81	1731	31
237	8	37	40	2240	108	18	26	41	328	0	194	782	199	105	2476	46
238	8	29	30	1680	77	2	33	91	728	15	253	838	255	126	2819	62
239	8	17	28	1568	61	1	27	55	440	0	163	588	165	99	1968	46
240	8	66	176	9856	186	6	46	139	1112	15	205	734	207	123	2458	62
241	8	50	73	4088	140	5	1208	6000	48000	359	175	631	177	112	2129	46
242	8	51	72	4032	155	5	28	79	632	15	128	483	130	74	1578	31
243	8	54	78	4368	140	20	5953	6000	48000	702	241	850	244	132	2828	62
244	8	53	127	7112	140	35	5957	6000	48000	859	177	645	179	114	2169	46
245	8	65	156	8736	186	66	86	485	3880	31	178	638	181	107	2140	31
246	8	47	83	4648	125	35	92	372	2976	15	202	721	204	132	2450	46
247	8	44	92	5152	125	5	31	91	728	15	149	565	152	95	1871	46
248	8	51	79	4424	140	35	75	262	2096	15	490	1732	493	447	6284	125
249	8	44	77	4312	124	4	35	92	736	15	160	584	162	106	1972	31
250	8	70	111	6216	202	52	32	66	528	0	142	497	144	99	1723	31
251	8	64	102	5712	171	51	23	43	344	0	194	673	196	126	2312	46
252	8	84	115	6440	233	38	2970	6000	48000	484	258	873	260	159	3003	62
253	8	59	69	3864	156	21	19	25	200	0	174	617	176	119	2119	46
254	8	79	104	5824	234	24	67	269	2152	15	194	664	196	122	2282	46
255	8	10	11	616	31	1	30	77	616	15	144	549	146	89	1803	31
256	8	14	15	840	31	16	45	147	1176	15	123	440	125	72	1471	31
257	8	11	13	728	31	1	26	69	552	0	181	603	183	99	2052	46
258	8	12	13	728	31	1	25	32	256	0	158	572	160	90	1894	31
259	8	12	13	728	31	1	26	67	536	15	173	621	175	94	2049	31
260	8	15	16	896	46	1	39	147	1176	0	118	424	120	73	1427	15
261	8	17	18	1008	46	31	24	37	296	0	148	530	150	86	1768	31
262	8	22	23	1288	62	2	35	78	624	0	219	729	221	116	2469	46
263	8	25	26	1456	78	3	25	41	328	15	243	803	245	122	2707	46
264	8	24	25	1400	62	17	40	105	840	0	210	710	212	127	2437	46
265	8	23	24	1344	61	1	30	74	592	15	168	607	170	105	2039	46
266	8	16	17	952	31	1	24	46	368	0	178	604	180	98	2042	46
267	8	19	25	1400	46	1	24	49	392	15	108	395	110	64	1312	31
268	8	57	58	3248	156	51	48	176	1408	15	683	2277	684	666	8604	171
269	8	79	80	4480	218	83	5842	6000	48000	828	564	1879	565	560	7133	140
270	8	204	205	11480	577	97	1987	6000	48000	436	651	2101	653	651	8114	156
271	8	205	206	11536	577	52	48	167	1336	0	719	2351	722	720	9028	187
272	8	148	157	8792	421	76	51	191	1528	15	356	1220	358	350	4564	93
273	8	327	328	18368	937	172	242	1577	12616	93	520	1796	522	521	6721	140
274	8	244	245	13720	702	117	1216	6000	48000	421	439	1662	440	435	5949	109
275	8	17	18	1008	46	16	231	1482	11856	93	713	2278	715	714	8843	171
276	8	16	17	952	62	32	2885	6000	48000	531	472	1547	474	466	5914	125
277	8	18	19	1064	46	1	133	518	4144	31	558	1944	561	556	7239	140
278	8	127	128	7168	359	44	147	632	5056	31	483	1528	484	434	5810	109
279	8	43	58	3248	124	34	175	1283	10264	62	170	609	172	114	2076	46
280	8	2	3	168	15	0	580	5626	45008	312	459	1602	461	392	5763	109
281	8	29	56	3136	77	17	204	6000	48000	296	212	692	214	139	2443	46

TABLE A.7: Certificate — Nonlinearly constrained case for $T = |y|_2$ (continued)

ex	N	Red Alg					MPBNGC				SolvOpt					
		Nit	Na	c	t ₁	t ₂	Nit	Nb	c	t ₁	Nit	Nc	Ng	N $\hat{g}$	c	t ₁
282	8	42	43	2408	140	35	124	3493	27944	171	729	2362	731	729	9104	187
283	8	39	40	2240	109	19	290	6000	48000	296	387	1421	389	383	5158	109
284	8	47	48	2688	140	35	89	650	5200	46	733	2602	735	732	9605	203
285	8	2	6	336	15	0	43	148	1184	15	430	1400	432	408	5320	93
286	8	2	6	336	0	0	33	79	632	0	602	2047	604	566	7604	156
287	8	2	4	224	0	0	34	161	1288	15	757	2568	759	676	9441	187
288	8	258	260	14560	733	118	147	828	6624	46	484	1605	485	457	6036	125
289	8	15	16	896	31	16	96	716	5728	31	425	1562	426	421	5665	125
290	8	16	17	952	46	1	53	397	3176	15	513	1716	515	514	6519	140
291	8	30	31	1736	93	18	104	6000	48000	280	500	1743	502	501	6495	140
292	8	27	28	1568	93	18	48	153	1224	0	751	2747	753	746	9991	187
293	8	41	42	2352	124	19	101	462	3696	31	798	2782	800	796	10352	203
294	8	2	4	224	15	0	21	79	632	0	709	2549	711	649	9178	187
295	8	2	6	336	15	0	102	228	1824	15	258	834	260	140	2868	62
296	8	2	4	224	0	0	52	352	2816	15	216	718	218	127	2471	46
297	8	2	4	224	15	0	124	513	4104	31	449	1381	450	446	5450	109
298	8	166	167	9352	452	77	103	352	2816	15	450	1343	451	441	5362	109
299	8	176	177	9912	483	78	94	377	3016	31	290	883	291	282	3485	62
300	8	155	156	8736	421	106	124	530	4240	46	462	1435	463	452	5615	125
301	8	93	94	5264	249	39	5955	6000	48000	828	623	2146	625	613	8006	171
302	8	31	32	1792	93	18	42	181	1448	15	161	587	163	107	1984	46
303	8	31	32	1792	93	18	51	235	1880	15	279	929	281	151	3154	62
304	8	32	33	1848	77	2	39	132	1056	15	272	932	274	142	3112	62
305	8	19	25	1400	46	1	3002	6000	48000	499	131	479	133	82	1603	46
306	8	12	38	2128	31	1	35	132	1056	15	117	420	119	69	1404	31
307	8	12	13	728	31	1	29	56	448	0	297	997	299	144	3323	62
308	8	20	21	1176	46	16	5997	6000	48000	765	201	754	203	105	2432	46
309	8	69	75	4200	187	22	10	16	128	0	87	473	88	48	1354	31

TABLE A.8: Higher dimensional examples — Easy piecewise quadratic constraint with $m_2 := \frac{N}{2}$

ex	N	Red Alg						MPBNGC					SolvOpt						
		Nit	Na	c	t ₁	t ₂	R	Nit	Nb	c	t ₁	R	Nit	Nc	Ng	N $\hat{g}$	c	t ₁	R
1	20	571	1139	692512	1843	433	1	449	727	5816	453	1	35000	116336	35003	27633	420580	6234	nt
2	20	383	699	424992	1296	261	4	287	1418	11344	359	4	4809	15475	4814	3811	56825	843	1
3	20	416	796	483968	1405	325	4	872	1271	10168	1546	4	7772	25036	7779	6113	91748	1328	3
4	20	1336	2523	1533984	4484	1274	4	412	890	7120	750	4	3064	9890	3068	2375	36109	531	1
5	20	2677	5296	3219968	8921	2186	3	536	707	5656	984	3	13972	45567	13976	11155	166527	2436	3
6	20	3595	7011	4262688	12109	3003	4	549	1400	11200	750	4	339	1133	341	245	4024	62	1
7	20	398	761	462688	1327	262	3	155	341	2728	78	3	9696	31317	9700	7773	115053	1686	2
8	20	83	163	99104	265	55	2	275	408	3264	359	2	9726	31579	9732	7644	115286	1671	2
9	20	1816	2476	1505408	5733	1383	4	1566	2927	23416	2859	4	15623	50717	15625	12248	185053	2717	1
10	20	2269	4421	2687968	7592	1712	4	1153	50000	400000	4500	nt	2966	9652	2969	2280	35051	515	1
11	20	704	1251	760608	2124	579	3	1854	3066	24528	3546	3	4134	13374	4138	3241	48885	702	1
12	20	1817	3431	2086048	6250	1510	5	241	1669	13352	375	5	15017	48759	15023	11731	177780	2624	1
13	20	1201	2268	1378944	4328	953	5	498	940	7520	1328	5	8304	26874	8307	6446	98007	1437	1
14	20	2024	4022	2445376	5874	1734	2	578	705	5640	875	2	8165	26315	8169	6409	96364	1406	1
15	20	2080	3933	2391264	8077	1836	6	336	501	4008	281	5	19810	63987	19814	15736	234624	3452	5
16	20	459	894	543552	1374	354	2	184	288	2304	125	2	15067	48820	15069	12069	179054	2624	2
17	20	319	504	306432	1202	332	7	487	820	6560	609	7	10616	33857	10622	8373	124699	1842	5
18	20	1024	1917	1165536	3374	734	5	215	1074	8592	281	5	1431	4835	1436	1095	17263	250	1
19	20	879	1684	1023872	2687	707	3	423	778	6224	781	3	976	3203	979	732	11539	171	2
20	20	175	371	225568	562	157	2	407	829	6632	609	2	370	1284	374	245	4425	93	1
21	40	2702	5257	42056	24609	7314	4	637	1568	12544	2890	4	35000	110655	35002	29415	414561	21812	nt
22	40	2922	5533	44264	26342	7531	6	699	3979	31832	4046	6	341	1127	345	209	3916	202	1
23	40	3515	6782	54256	32358	9333	6	819	4684	37472	3343	6	539	1776	541	430	6465	343	1
24	40	459	896	7168	4421	1136	6	1913	13071	104568	20000	6	1926	6209	1928	1564	22894	1203	1
25	40	830	1638	13104	7375	2110	3	1360	2233	17864	10281	3	7213	22870	7218	6037	85505	4499	1
26	40	1355	2669	21352	11734	3529	2	1351	5599	44792	4703	2	567	1865	569	453	6796	343	1
27	40	182	344	2752	1656	546	6	545	1154	9232	3421	6	12137	38361	12139	10162	143625	7515	1
28	40	901	1741	13928	8186	2471	5	288	1652	13216	1281	5	3275	10447	3277	2739	38942	2031	1
29	40	422	862	6896	3718	1168	2	466	1787	14296	1140	2	15721	49696	15726	13225	186245	9796	1
30	40	1194	2139	17112	12453	3047	4	523	1411	11288	2125	4	4658	14792	4663	3851	55126	2905	1
31	40	2248	4497	35976	19000	5890	2	1382	2141	17128	5703	2	1660	5291	1662	1381	19711	1031	1
32	40	891	1610	12880	8328	2343	8	514	4639	37112	2812	8	1806	5877	1808	1468	21582	1125	1
33	40	563	1091	8728	5406	1386	5	319	455	3640	906	5	9283	29341	9287	7870	110153	5749	1
34	40	1167	2312	18496	10952	2986	2	925	1353	10824	4625	2	19736	62461	19738	16714	234278	12280	2
35	40	666	1290	10320	6421	1816	5	714	1581	12648	4828	5	7433	23454	7437	6116	87567	4593	1
36	40	254	417	3336	2265	645	3	1143	2027	16216	7218	3	35000	110692	35002	29400	414590	21733	nt
37	40	1960	3833	30664	19171	5266	5	233	1221	9768	859	5	15230	48501	15232	12785	181053	9546	1
38	40	507	973	7784	5124	1268	6	609	1537	12296	3593	6	21025	66442	21028	17589	248735	13062	1
39	40	232	458	3664	2078	653	4	321	1557	12456	1406	4	6618	21035	6620	5503	78439	4124	1
40	40	1126	2126	17008	9749	3463	5	1058	4194	33552	7468	5	512	1666	516	389	6047	312	1
41	60	2122	4150	33200	45359	14856	6	3800	5541	44328	63827	6	2363	7423	2367	1955	27812	3609	2
42	60	1203	2597	20776	31187	8468	9	5465	12290	98320	226984	9	6140	19364	6144	5216	72808	9390	1
43	60	314	596	4768	6374	1975	7	492	1967	15736	8078	7	7443	23348	7448	6396	88228	11358	1
44	60	1076	2064	16512	23499	7522	8	586	1939	15512	6234	8	35000	109867	35004	30616	416594	53828	nt
45	60	1197	2363	18904	24187	8331	3	891	4423	35384	9562	3	5167	16263	5172	4362	61128	7921	1
46	60	1710	3062	24496	35280	11814	4	695	1016	8128	2828	4	17557	55471	17562	15188	209192	27140	3
47	60	426	797	6376	9030	2878	6	2014	3577	28616	56390	6	195	658	197	121	2270	296	1
48	60	1321	2349	18792	32734	8995	10	477	1161	9288	4515	10	5708	17955	5710	4947	67881	8811	1
49	60	298	491	3928	6280	2001	5	1383	2987	23896	27671	5	2923	9374	2925	2515	35068	4546	1

TABLE A.8: Higher dimensional examples — Easy piecewise quadratic constraint with $m_2 := \frac{N}{2}$ (continued)

ex	N	Red Alg						MPBNGC					SolvOpt						
		Nit	Na	c	t ₁	t ₂	R	Nit	Nb	c	t ₁	R	Nit	Nc	Ng	N $\hat{g}$	c	t ₁	R
50	60	678	1436	11488	14061	4655	4	1529	2577	20616	38484	4	35000	110240	35003	30166	415987	53718	nt
51	60	499	961	7688	10593	3382	8	692	919	7352	4609	8	35000	110011	35002	30211	415661	54124	nt
52	60	923	1394	11152	21546	6498	10	1082	3339	26712	7686	10	1266	4157	1268	1026	15196	1968	1
53	60	739	1389	11112	16859	5213	8	733	1448	11584	5031	6	20856	66732	20861	17890	249717	32342	1
54	60	332	624	4992	6952	2092	6	778	1250	10000	6312	6	23000	72236	23007	19781	272836	35233	4
55	60	373	582	4656	8453	2431	6	2037	7382	59056	53186	6	10999	35048	11001	9545	131734	16984	1
56	60	619	1203	9624	14030	4322	9	696	5142	41136	6358	9	23412	73394	23418	20103	277351	35843	1
57	60	1759	3504	28032	43999	12774	7	814	4375	35000	8437	8	2512	8075	2514	2126	30070	3890	1
58	60	278	516	4128	4999	1742	2	468	1489	11912	2796	2	11663	36640	11667	10056	138449	17936	1
59	60	222	424	3392	4983	1609	6	2363	3706	29648	34109	5	289	1013	293	162	3391	453	1
60	60	1591	3258	26064	33484	11226	5	1873	5611	44888	36999	8	14227	44725	14235	12119	168512	21936	1
61	80	291	701	342088	11828	4036	2	1207	1778	14224	12484	2	35000	109744	35004	30666	416498	151593	nt
62	80	363	731	356728	13389	4929	6	749	2420	19360	13656	6	35000	109628	35006	30866	416872	152405	nt
63	80	569	1094	533872	23640	7928	9	516	1621	12968	8093	9	35000	109651	35006	30583	416069	151906	nt
64	80	192	400	195200	6921	2433	4	663	3480	27840	12031	8	1596	5230	1598	1341	19277	7109	1
65	80	780	848	413824	30687	10258	9	3216	13022	104176	195264	9	834	2725	836	693	10037	3703	1
66	80	699	1415	690520	26514	9907	4	1230	7722	61776	25186	9	2533	8123	2535	2225	30526	11202	1
67	80	4000	8206	4004528	165046	57878	7	771	4345	34760	13780	9	2492	7877	2494	2168	29740	10828	1
68	80	1389	2727	1330776	59999	19752	5	1991	8319	66552	64171	5	288	929	290	197	3319	1234	1
69	80	2577	5048	2463424	98780	36507	6	1355	2921	23368	24827	6	8962	28100	8964	7790	106462	38796	1
70	80	390	777	379176	14718	5276	3	515	990	7920	7265	1	25989	81444	25994	22884	309522	112749	1
71	80	264	603	294264	9905	3616	6	775	2756	22048	13999	6	2817	8853	2819	2464	33555	12281	1
72	80	1363	2578	1258064	52671	19783	4	388	642	5136	7406	4	35000	109586	35002	30827	416659	152014	nt
73	80	447	717	349896	17640	6067	8	649	2080	16640	18218	8	2814	8932	2817	2417	33566	12281	1
74	80	4535	9018	4400784	171656	65681	5	2398	6711	53688	55515	5	2111	6798	2113	1828	25419	9405	1
75	80	488	1004	489952	19858	6865	3	1682	3630	29040	20093	3	963	3100	965	829	11582	4250	1
76	80	208	403	196664	8031	2689	4	426	584	4672	3671	4	35000	109784	35006	30816	417034	152249	nt
77	80	4410	8618	4205584	173952	64164	8	2494	3774	30192	81031	3	7699	24395	7701	6759	92170	33702	1
78	80	1160	2318	1131184	45171	16398	7	1806	13868	110944	31952	7	575	1906	577	469	6950	2578	1
79	80	682	1332	650016	27702	9818	10	1244	5465	43720	25390	10	435	1436	438	296	5074	1906	1
80	80	693	1558	760304	26280	9957	3	1235	5648	45184	30062	3	4243	13552	4245	3676	50867	18640	1
81	100	848	1877	1141216	52421	21303	2	2156	5260	42080	31374	3	4953	15538	4956	4363	59033	39202	1
82	100	292	540	328320	20249	6722	9	1593	5638	45104	35796	9	1321	4211	1323	1159	15868	10546	1
83	100	288	448	272384	21514	6215	14	392	1937	15496	6203	14	35000	109652	35004	31337	418327	277030	nt
84	100	151	278	169024	9437	2798	6	1822	5781	46248	50077	6	1108	3516	1110	946	13200	8796	1
85	100	987	1630	991040	73124	24263	8	596	1761	14088	7936	9	35000	109636	35003	31226	417959	277468	nt
86	100	2663	3725	2264800	201796	68214	8	602	1864	14912	11608	8	35000	109620	35002	31068	417450	275124	nt
87	100	1692	3323	2020384	122921	43675	8	1218	5114	40912	29311	8	955	3081	958	798	11430	7796	1
88	100	358	632	384256	23999	8306	6	562	1117	8936	5078	7	35000	109511	35009	30907	416770	276077	nt
89	100	835	1528	929024	70640	20525	6	971	1692	13536	13843	6	35000	109762	35002	31235	418235	277156	nt
90	100	1302	2441	1484128	101343	33441	9	1787	3162	25296	166983	9	4152	13038	4157	3585	49302	32702	1
91	100	166	284	172672	11484	3427	5	316	500	4000	2828	5	35000	109591	35005	30983	417146	275484	nt
92	100	143	248	150784	8202	2674	4	567	1001	8008	4077	4	19544	61152	19549	17221	232614	154593	1
93	100	1097	2046	1243968	66828	27023	7	1221	5236	41888	29031	7	10540	33451	10542	9350	126578	83718	1
94	100	864	1813	1102304	57421	21690	5	2463	9330	74640	86405	6	1996	6459	1998	1714	24054	15983	1
95	100	2085	3980	2419840	142124	53269	9	9357	21128	169024	420155	7	241	895	243	159	2996	2046	1
96	100	299	561	341088	18155	6810	7	453	1329	10632	5359	7	15672	49214	15675	13954	187315	123530	1
97	100	347	587	356896	24421	7842	7	378	780	6240	3781	7	35000	109611	35006	31122	417606	275639	nt
98	100	872	1715	1042720	88608	22396	9	3315	8325	66600	211514	9	2285	7249	2288	1939	27179	18061	1

TABLE A.8: Higher dimensional examples — Easy piecewise quadratic constraint with $m_2 := \frac{N}{2}$ (continued)

		Red Alg						MPBNGC					SolvOpt							
ex	N	Nit	Na	c	t ₁	t ₂	R	Nit	Nb	c	t ₁	R	Nit	Nc	Ng	Nĝ	c	t ₁	R	
99	100	1236	2358	1433664	82874	30884	13	1371	8753	70024	87406	13	1423	4656	1428	1133	16995	11374	1	
100	100	440	834	507072	27796	10683	6	1584	4383	35064	35593	5	8667	27168	8672	7607	103173	68406	1	

TABLE A.9: Higher dimensional examples — Easy piecewise quadratic constraint with $m_2 := N$

		Red Alg						MPBNGC					SolvOpt						
ex	N	Nit	Na	c	t ₁	t ₂	R	Nit	Nb	c	t ₁	R	Nit	Nc	Ng	Nġ	c	t ₁	R
1	20	502	904	549632	1765	295	7	449	907	7256	937	7	3327	10744	3331	2555	39146	921	1
2	20	364	717	435936	1234	274	3	244	960	7680	125	3	17361	55905	17363	13723	205068	4796	2
3	20	777	1515	921120	2734	619	3	191	680	5440	265	3	1436	4732	1440	1092	17060	390	2
4	20	1134	2196	1335168	3624	1164	3	319	1546	12368	375	3	2541	8222	2543	1961	29956	703	1
5	20	2758	5329	3240032	12609	2619	4	2171	19147	153176	5546	4	4980	16120	4982	3887	58847	1406	1
6	20	4395	8422	5120576	16359	3592	6	396	1992	15936	750	6	488	1757	492	336	5998	140	1
7	20	919	1795	1091360	3296	866	3	206	1093	8744	218	3	9508	31116	9513	7503	113280	2656	1
8	20	371	644	391552	1327	307	7	294	538	4304	250	7	6347	21452	6351	4971	76870	1796	1
9	20	298	544	330752	1108	253	6	379	816	6528	890	6	14530	59781	14536	11561	197853	4749	1
10	20	922	1754	1066432	3937	847	3	300	711	5688	515	3	5456	17582	5460	4286	64402	1500	2
11	20	1687	3230	1963840	5983	1678	4	248	655	5240	312	4	1642	5380	1645	1259	19472	453	2
12	20	1833	3487	2120096	6389	1679	4	274	1954	15632	546	4	3742	12128	3747	2855	44062	1031	1
13	20	185	239	145312	624	99	5	260	347	2776	109	5	15296	49097	15298	12131	180481	4233	4
14	20	317	548	333184	1109	269	4	299	439	3512	406	4	11062	37134	11066	8767	133767	3140	2
15	20	244	370	224960	828	198	5	374	528	4224	375	5	6208	19954	6210	4868	73142	1687	4
16	20	1945	3582	2177856	7437	1737	6	314	490	3920	531	6	1773	5753	1775	1391	21004	484	1
17	20	384	621	377568	1281	426	6	508	766	6128	734	5	9126	29414	9129	7100	107515	2500	4
18	20	1234	2396	1456768	4358	1223	5	341	1958	15664	640	5	1690	5490	1693	1293	19938	468	1
19	20	1678	3156	1918848	5812	1432	5	510	659	5272	1234	5	720	2415	722	535	8601	203	3
20	20	4241	8216	4995328	14734	3634	5	691	1311	10488	1968	5	1058	3483	1060	804	12558	296	1
21	40	1327	2483	19864	14499	4014	8	661	2844	22752	5140	9	914	2952	919	698	10755	968	1
22	40	322	695	5560	3046	1051	3	495	1703	13624	3281	13	4049	12829	4053	3340	47837	4265	1
23	40	1040	1898	15184	11061	3049	7	690	4184	33472	3733	7	35000	110474	35002	29437	414265	37000	nt
24	40	442	879	7032	4062	1362	2	582	1066	8528	2203	2	24297	78649	24299	20462	291581	26124	1
25	40	2605	5374	42992	25030	8185	3	928	3422	27376	7312	6	960	3088	964	754	11330	1015	1
26	40	416	849	6792	4046	1256	5	978	1651	13208	13609	8	606	1988	610	448	7150	656	1
27	40	187	357	2856	1937	497	2	1419	1996	15968	8234	2	9672	30759	9677	8107	114870	10264	1
28	40	637	994	7952	7092	1737	9	1852	5772	46176	28843	9	361	1289	365	246	4411	406	1
29	40	4658	9036	72288	47358	14853	5	2231	3611	28888	29046	5	16218	51253	16223	13763	192464	17077	2
30	40	2170	4031	32248	25281	6830	8	727	5421	43368	6468	9	1146	3770	1148	917	13735	1250	1
31	40	593	989	7912	5968	1918	5	164	663	5304	531	6	26756	84341	26759	22603	316768	28406	4
32	40	141	273	2184	1390	340	4	216	285	2280	484	4	19172	65505	19177	16161	237024	21343	1
33	40	261	505	4040	3015	675	8	263	1060	8480	1421	8	3613	11385	3617	3019	42678	3796	5
34	40	1575	2947	23576	15843	4862	8	388	1524	12192	2312	8	525	1655	531	340	5923	531	1
35	40	411	613	4904	4046	1256	5	410	564	4512	1687	6	29720	93753	29726	25169	352191	31436	3
36	40	188	367	2936	1858	643	4	348	902	7216	1843	4	9921	31408	9924	8347	117629	10499	1
37	40	4175	7733	61864	45952	12815	9	3230	27921	223368	47999	9	370	1261	372	275	4463	390	1
38	40	4125	8173	65384	38093	12368	3	1370	8180	65440	5093	2	5177	16392	5182	4271	61143	5437	1
39	40	733	1420	11360	7921	2311	7	1435	4903	39224	8578	7	19422	61267	19424	16384	229958	20436	3
40	40	4145	7684	61472	45406	12705	9	801	1133	9064	2703	8	31874	100134	31880	26809	376335	33562	6

TABLE A.9: Higher dimensional examples — Easy piecewise quadratic constraint with $m_2 := N$ (continued)

ex	N	Red Alg						MPBNGC					SolvOpt						
		Nit	Na	c	t ₁	t ₂	R	Nit	Nb	c	t ₁	R	Nit	Nc	Ng	$\hat{N}\hat{g}$	c	t ₁	R
41	60	266	417	3336	5749	2219	5	1320	2450	19600	22125	5	493	1629	498	317	5703	1500	1
42	60	703	1051	8408	17671	5560	13	582	2232	17856	10625	13	3384	10752	3389	2830	40161	10421	1
43	60	892	1679	13432	19734	7140	9	747	1204	9632	13796	9	6031	18974	6035	5061	71236	18437	1
44	60	1014	1917	15336	23217	8212	8	1088	6356	50848	13749	8	184	683	186	116	2272	593	1
45	60	2562	4944	39552	55015	21006	7	790	4986	39888	14515	7	5196	16415	5200	4441	61753	15765	1
46	60	2116	3968	31744	48718	17444	12	668	3624	28992	9859	12	224	794	227	123	2638	703	1
47	60	866	1736	13888	17889	6950	3	7254	10822	86576	102312	3	9237	29027	9239	7937	109582	28312	1
48	60	3238	6103	48824	75187	26999	10	525	2494	19952	7749	10	1952	6179	1954	1664	23212	6015	1
49	60	3924	7281	58248	93686	31754	12	765	4927	39416	12046	12	3248	10210	3251	2783	38522	10014	1
50	60	788	1430	11440	19562	6546	5	258	558	4464	1281	6	35000	109953	35002	30446	416250	106140	nt
51	60	1078	1701	13608	26467	8555	10	1163	2651	21208	11671	10	3515	11110	3517	3041	41894	10827	1
52	60	678	1349	10792	16249	5627	9	417	2148	17184	5109	8	13278	41698	13283	11450	157595	40968	1
53	60	474	887	7096	10515	3948	7	492	2365	18920	7218	7	2534	8027	2538	2153	30127	7702	1
54	60	302	627	5016	7217	2246	6	1473	5394	43152	19827	7	2804	8865	2807	2405	33366	8515	1
55	60	1346	1629	13032	32686	10458	11	712	2733	21864	4921	11	2708	8610	2711	2274	32175	8421	1
56	60	2786	5458	43664	59843	23800	6	588	2497	19976	6843	6	566	1880	568	448	6808	1781	1
57	60	2247	4520	36160	53608	18790	8	7213	14105	112840	131109	8	810	2582	814	617	9457	2437	1
58	60	1513	3196	25568	35390	12664	8	4771	8181	65448	98687	8	9578	30116	9583	8261	113764	29500	2
59	60	3239	6350	50800	64108	26076	4	1498	11283	90264	24593	4	182	639	184	106	2148	562	1
60	60	671	1235	9880	15186	5423	10	466	3112	24896	6546	10	6610	20780	6614	5639	78319	20608	1
61	80	3782	7853	3832264	166046	67185	10	10000	20953	167624	308156	nt	692	2170	697	429	7718	5265	1
62	80	1831	3926	1915888	79937	32616	4	10000	16668	133344	312109	nt	1127	3625	1129	961	13520	9062	1
63	80	1447	2740	1337120	60764	24950	9	1781	5465	43720	44233	9	3163	10009	3165	2728	37697	25046	1
64	80	658	1281	625128	30406	10940	14	945	7471	59768	33156	16	27133	84990	27137	24058	323565	216343	1
65	80	1504	2962	1445456	64593	26012	8	1083	4014	32112	44780	9	2178	6866	2180	1904	25984	17327	1
66	80	842	959	467992	40265	13552	14	898	2810	22480	23234	12	35000	110047	35002	30768	417404	278249	nt
67	80	475	979	477752	22983	8128	4	1894	5707	45656	43625	8	1910	6048	1915	1593	22620	15108	1
68	80	2642	5003	2441464	125671	45619	12	3465	8882	71056	112109	7	10646	33493	10650	9319	126893	84561	1
69	80	312	645	314760	12265	5133	3	1705	6280	50240	36577	8	275	878	277	159	3064	2078	1
70	80	328	560	273280	14905	5343	8	1104	2139	17112	15734	6	28589	89590	28594	25314	340904	227031	3
71	80	645	1030	502640	26155	10622	7	694	3234	25872	18671	7	495	1611	497	376	5841	3953	1
72	80	3604	7068	3449184	144453	63202	7	1868	7715	61720	57952	7	625	2046	630	410	7212	4875	1
73	80	447	769	375272	19312	7410	7	378	879	7032	7390	7	35000	109623	35007	30512	415803	277999	nt
74	80	2625	5233	2553704	123312	46106	16	2265	4098	32784	81733	8	3246	10427	3248	2839	39115	26108	1
75	80	868	1776	866688	38062	14845	12	1111	4162	33296	36936	12	8177	25839	8181	7109	97548	64811	1
76	80	1201	2098	1023824	50905	20474	10	2802	6918	55344	72858	10	12004	37701	12006	10503	142929	95124	1
77	80	948	1377	671976	49468	15699	18	654	5886	47088	23531	18	6374	20031	6378	5504	75708	50609	1
78	80	496	966	471408	21405	8429	8	997	3473	27784	22702	8	2994	9449	2996	2546	35524	23749	1
79	80	445	861	420168	17765	7532	3	1042	3632	29056	29921	3	2851	9105	2853	2478	34203	22890	1
80	80	683	1154	563152	28515	11299	7	809	1240	9920	11828	7	35000	109689	35003	30738	416601	278499	nt
81	100	629	1006	611648	48155	17777	12	803	2972	23776	18765	13	5842	18368	5844	5182	69814	81874	1
82	100	339	601	365408	29296	9305	7	1200	2088	16704	35421	7	35000	109517	35007	31210	417685	490874	nt
83	100	686	1054	640832	53780	19680	16	1942	4300	34400	312796	16	4891	15535	4894	4357	58823	69078	1
84	100	290	391	237728	20499	7475	9	961	3396	27168	29515	9	4613	14471	4615	4072	55003	64562	1
85	100	2081	4074	2476992	158296	64761	9	2529	18780	150240	156859	9	307	1116	309	217	3810	4578	1
86	100	1164	2306	1402048	88812	35292	13	856	5046	40368	41390	14	16203	50647	16208	14330	192908	225483	1
87	100	2593	4574	2780992	180342	76699	7	6796	12918	103344	211046	8	35000	109902	35006	30935	417627	489655	nt
88	100	1670	2691	1636128	121171	48711	9	4453	6033	48264	97311	10	35000	109798	35004	30997	417599	489952	nt
89	100	2601	4316	2624128	216562	77779	18	874	1236	9888	31749	14	7354	23062	7357	6475	87620	102906	1

TABLE A.9: Higher dimensional examples — Easy piecewise quadratic constraint with $m_2 := N$ (continued)

ex	N	Red Alg						MPBNGC					SolvOpt						
		Nit	Na	c	t ₁	t ₂	R	Nit	Nb	c	t ₁	R	Nit	Nc	Ng	N $\hat{g}$	c	t ₁	R
90	100	1125	2211	1344288	84077	34558	13	1299	5981	47848	53577	13	233	819	235	141	2766	3359	1
91	100	490	923	561184	40921	14292	7	3231	14762	118096	134452	10	3263	10309	3267	2829	38906	45858	1
92	100	364	639	388512	27280	10099	7	582	728	5824	12328	8	35000	109609	35005	31151	417686	490843	nt
93	100	372	588	357504	27780	10175	7	367	751	6008	4765	12	35000	109521	35002	31222	417714	488655	nt
94	100	2199	4236	2575488	177593	67363	11	1626	13061	104488	80905	11	4179	13169	4183	3659	49864	58640	1
95	100	728	1226	745408	55858	20686	8	906	1247	9976	25936	10	35000	109554	35002	31277	417945	490952	nt
96	100	479	924	561792	36296	13844	11	2444	11142	89136	197280	11	341	1084	345	186	3761	4515	1
97	100	759	1238	752704	60671	21908	13	995	4679	37432	28218	13	3662	11529	3666	3213	43695	51312	1
98	100	582	1014	616512	44406	16957	9	648	1627	13016	14359	10	5242	16468	5245	4611	62504	73327	1
99	100	366	733	445664	22515	10554	4	1319	1856	14848	44561	4	9337	29296	9341	8247	111356	130624	1
100	100	1867	3516	2137728	135937	56685	7	834	2549	20392	25031	7	33942	106311	33948	29886	404124	476452	1

TABLE A.10: Higher dimensional examples — Difficult piecewise quadratic constraint with $m_2 := \frac{N}{2}$

ex	N	Red Alg						MPBNGC					SolvOpt						
		Nit	Na	c	t ₁	t ₂	R	Nit	Nb	c	t ₁	R	Nit	Nc	Ng	Nĝ	c	t ₁	R
1	20	1124	2141	1301728	4124	944	4	1108	1719	13752	1656	4	7285	23557	7287	5907	86696	1311	4
2	20	910	1798	1093184	3061	661	3	10000	10016	80128	1187	nt	7125	22885	7127	5773	84470	1296	3
3	20	44	89	54112	140	35	3	2755	50000	400000	5421	nt	12069	38886	12071	9759	143262	2124	3
4	20	739	1262	767296	3014	659	5	3165	50000	400000	5640	nt	4731	15060	4733	3845	55854	874	5
5	20	146	231	140448	546	141	4	1574	2417	19336	2500	4	10641	34101	10643	8463	125520	1906	5
6	20	519	979	595232	2062	502	8	2489	50000	400000	4499	nt	10429	33297	10431	8349	122934	1843	7
7	20	771	1491	906528	2593	448	3	1012	1421	11368	1625	3	6371	20451	6373	5188	75585	1124	3
8	20	1024	1943	1181344	3499	784	4	945	1534	12272	1375	4	8471	27437	8473	6827	100774	1500	4
9	20	508	964	586112	1952	317	6	1231	11306	90448	1953	6	3147	10014	3149	2494	36957	578	5
10	20	1165	2095	1273760	4312	1072	6	2174	50000	400000	3984	nt	11348	36406	11350	9159	134339	2014	5
11	20	1381	2592	1575936	5156	1106	5	10000	10178	81424	2515	nt	959	3043	961	776	11297	171	3
12	20	591	1104	671232	1859	494	4	2486	50000	400000	5140	nt	1421	4544	1423	1140	16777	249	4
13	20	1553	2908	1768064	6046	1081	7	2606	50000	400000	4703	nt	18354	58849	18356	14750	217016	3265	7
14	20	47	68	41344	156	21	2	10000	10313	82504	1921	nt	7799	25458	7801	6306	93237	1405	2
15	20	1030	1970	1197760	3733	673	4	1903	50000	400000	3671	nt	3479	11094	3481	2808	41055	609	4
16	20	1286	2501	1520608	4609	1129	3	2787	50000	400000	5359	nt	6067	19718	6069	4867	72244	1078	3
17	20	77	146	88768	280	55	4	10000	10395	83160	2640	nt	8677	27864	8679	7003	102774	1546	4
18	20	76	146	88768	265	40	4	128	226	1808	46	4	2046	6548	2050	1637	24157	359	4
19	20	1144	2208	1342464	3889	874	4	1317	4394	35152	2046	4	17537	56087	17539	14238	207505	3171	4
20	20	307	580	352640	921	201	4	2311	50000	400000	4312	nt	9496	30786	9498	7699	113163	1718	4
21	40	1507	2637	21096	16421	4226	11	6915	8597	68776	64859	12	12166	37875	12168	10548	143898	7718	11
22	40	1377	2579	20632	14436	3306	8	10000	12781	102248	64858	nt	8048	25270	8050	7017	95741	5046	6
23	40	538	1007	8056	5530	1735	7	10000	13099	104792	83124	nt	12299	39383	12301	10692	147745	8046	4
24	40	596	1137	9096	5577	1512	6	778	979	7832	1953	6	11618	36291	11620	9935	137247	7358	6
25	40	71	116	928	640	205	7	5843	7946	63568	66125	7	11515	36063	11517	9785	136032	7281	5
26	40	412	762	6096	4405	1075	8	10000	12823	102584	59030	nt	2476	7739	2478	2132	29308	1593	5
27	40	1380	2627	21016	14218	3972	7	8816	11011	88088	67421	7	9095	28524	9097	7913	108078	5859	7
28	40	777	1518	12144	8031	2030	4	825	978	7824	2015	4	13578	42452	13580	11668	160648	8843	nt
29	40	1493	2895	23160	13436	3806	4	10000	13138	105104	56233	nt	20778	65134	20780	18023	246677	13296	4
30	40	295	571	4568	2843	893	5	2493	3284	26272	13453	5	7772	24308	7774	6697	92029	4952	5
31	40	118	157	1256	1093	223	7	10000	12605	100840	58921	nt	13495	43031	13497	11759	161830	8686	5

TABLE A.10: Higher dimensional examples — Difficult piecewise quadratic constraint with $m_2 := \frac{N}{2}$ (continued)

ex	N	Red Alg						MPBNGC					SolvOpt						
		Nit	Na	c	t ₁	t ₂	R	Nit	Nb	c	t ₁	R	Nit	Nc	Ng	$\tilde{N}_{\tilde{g}}$	c	t ₁	R
32	40	508	863	6904	5515	1345	12	10000	12832	102656	79500	nt	9617	29889	9619	8331	113628	6140	11
33	40	453	663	5304	4796	1256	10	10000	12816	102528	71983	nt	8884	27774	8886	7614	105048	5624	8
34	40	365	586	4688	4530	998	9	10000	12777	102216	84343	nt	10396	32491	10398	8971	123089	6624	9
35	40	1374	2582	20656	14624	3524	8	10000	11808	94464	51734	nt	5377	16788	5379	4685	63768	3374	7
36	40	1564	2892	23136	15624	4254	8	10000	13339	106712	86625	nt	9332	29202	9334	8103	110715	5952	6
37	40	169	193	1544	1608	408	7	10000	10738	85904	16999	nt	5977	18681	5979	5165	70794	3749	7
38	40	28	44	352	203	68	5	4637	5746	45968	32968	5	6298	19680	6300	5454	74622	3968	5
39	40	138	163	1304	1343	338	10	10000	11971	95768	45187	nt	9556	29939	9558	8325	113527	6155	7
40	40	133	226	1808	1358	368	9	10000	10895	87160	26046	nt	12350	38751	12352	10773	146877	7858	7
41	60	871	1626	13008	20655	6122	12	6806	8758	70064	103125	12	14973	46498	14975	13189	177488	23655	12
42	60	151	247	1976	3202	972	12	10000	12727	101816	214249	nt	15611	48778	15613	13991	186368	24858	8
43	60	522	600	4800	12374	3588	13	10000	11455	91640	102656	nt	11509	35605	11513	10104	136061	18328	13
44	60	329	598	4784	7561	2100	11	10000	12751	102008	217030	nt	20138	63097	20140	18098	240908	32171	9
45	60	208	292	2336	5281	1102	15	10000	12467	99736	204312	nt	18220	56588	18222	16047	215983	28984	12
46	60	508	928	7424	11374	3271	9	10000	12430	99440	84389	nt	9816	30561	9818	8812	117012	15624	8
47	60	1596	2983	23864	41467	10719	12	10000	12325	98600	129562	nt	16458	51325	16460	14767	196331	26390	8
48	60	159	241	1928	4046	942	14	10000	13038	104304	241327	nt	9321	28862	9326	8201	110305	14655	12
49	60	771	1203	9624	17921	5423	10	10000	12240	97920	126155	nt	11754	36701	11756	10486	140128	18593	6
50	60	114	211	1688	2484	655	11	10000	13418	107344	316390	nt	16686	52196	16688	14775	198781	26625	6
51	60	283	535	4280	6171	1830	8	10000	14103	112824	223687	nt	5586	17460	5588	4923	66453	8858	3
52	60	258	446	3568	5515	1873	10	10000	12837	102696	182234	nt	8423	26275	8425	7575	100550	13577	6
53	60	688	1233	9864	17593	4566	13	10000	12858	102864	148686	nt	21932	68294	21934	19552	261046	34499	13
54	60	685	1272	10176	15546	4724	8	10000	13167	105336	181280	nt	5769	17987	5771	5189	68854	9218	4
55	60	764	1424	11392	18500	5383	9	10000	12765	102120	117968	nt	18838	58941	18840	16865	224997	30436	8
56	60	912	1366	10928	23343	6080	14	10000	12304	98432	83233	nt	18425	57355	18427	16408	219215	29421	9
57	60	503	555	4440	12437	3282	16	10000	12746	101968	127265	nt	13191	41165	13193	11788	157273	21343	11
58	60	469	795	6360	10655	3250	13	10000	17549	140392	56515	nt	16266	50872	16268	14565	194243	25905	9
59	60	216	274	2192	5250	1308	16	10000	15525	124200	164937	nt	17028	53140	17030	15176	202898	26952	12
60	60	196	336	2688	4515	1396	12	447	547	4376	1281	11	17334	53632	17340	15088	204548	27139	12
61	80	1448	2355	1149240	73390	20582	20	10000	14489	115912	213796	nt	13070	40519	13072	11855	155819	57078	12
62	80	150	282	137616	5483	1756	10	1142	1564	12512	11343	10	12723	39574	12725	11406	151541	55952	8
63	80	260	462	225456	10937	3312	13	1445	2051	16408	14577	13	12281	38054	12286	10934	145768	53968	13
64	80	463	707	345016	23780	6555	16	10000	12559	100472	320405	nt	23706	73421	23708	21489	282433	104593	14
65	80	145	243	118584	5687	1744	14	10000	13835	110680	418530	nt	35000	108763	35004	31509	417065	155249	nt
66	80	933	1647	803736	42906	13081	17	10000	13671	109368	259358	nt	7147	22170	7149	6503	85296	31342	9
67	80	1192	2174	1060912	55218	17196	18	10000	14076	112608	54968	nt	17366	53770	17368	15661	206627	75484	13
68	80	485	829	404552	22327	6663	15	10000	15753	126024	350671	nt	14624	45655	14626	13132	174584	63921	10
69	80	392	487	237656	16249	4843	11	10000	12011	96088	246780	nt	12766	39648	12768	11582	152346	55656	9
70	80	633	1051	512888	31796	8704	12	10000	13433	107464	249093	nt	12910	40167	12912	11775	154395	56921	6
71	80	481	639	311832	21828	6382	15	10000	10834	86672	85906	nt	22254	68965	22256	19979	264635	96905	14
72	80	469	831	405528	19796	6119	15	1731	2134	17072	15984	14	13808	42749	13810	12504	164440	59984	14
73	80	726	770	375760	32233	9430	17	10000	15512	124096	317515	nt	20880	65266	20882	18742	249404	92859	8
74	80	1325	2509	1224392	67515	19127	17	10000	13587	108696	292952	nt	24922	77438	24924	22620	297508	108889	15
75	80	391	734	358192	17281	5242	15	3122	3725	29800	74156	15	15890	49356	15892	14225	189063	70468	12
76	80	310	416	203008	14640	4101	19	10000	12402	99216	362889	nt	22851	70638	22853	20657	271806	99327	15
77	80	946	1795	875960	40312	13290	10	10000	14133	113064	358921	nt	22698	71600	22700	20688	273364	100640	7
78	80	215	390	190320	9124	2801	16	530	2427	19416	4875	16	13978	43302	13980	12377	165675	61093	14
79	80	1835	2168	1057984	83343	25893	14	10000	14007	112056	217905	nt	19956	62230	19958	18162	238820	88717	9
80	80	426	550	268400	18265	5476	17	10000	13163	105304	214328	nt	22531	70638	22533	20484	270327	99218	11

TABLE A.10: Higher dimensional examples — Difficult piecewise quadratic constraint with $m_2 := \frac{N}{2}$ (continued)

ex	N	Red Alg						MPBNGC					SolvOpt						
		Nit	Na	c	t ₁	t ₂	R	Nit	Nb	c	t ₁	R	Nit	Nc	Ng	Nĝ	c	t ₁	R
81	100	369	651	395808	25718	8453	17	4002	4750	38000	137359	17	18374	56813	18376	16611	218587	146734	13
82	100	311	536	325888	20483	7057	18	1728	2605	20840	25999	17	14942	46231	14944	13510	177824	118405	15
83	100	558	923	561184	46624	13205	18	10000	12495	99960	481546	nt	16974	52735	16976	15666	203396	136905	7
84	100	332	498	302784	23312	7298	14	10000	13524	108192	673109	nt	28878	89973	28880	26594	346368	232062	9
85	100	1107	2107	1281056	85967	27989	14	3045	4606	36848	114608	15	14790	45820	14792	13550	176666	116624	13
86	100	611	1150	699200	50093	14821	21	8139	10220	81760	653530	21	20458	63451	20460	18686	244340	162906	14
87	100	777	901	547808	57702	18669	23	10000	14089	112712	634327	nt	20467	63727	20469	18583	244610	163530	10
88	100	360	535	325280	26889	8020	22	10000	12797	102376	906421	nt	24698	76759	24700	22492	295094	199468	17
89	100	1550	1950	1185600	127483	37580	23	10000	12737	101896	427624	nt	28421	88298	28423	26191	340438	225889	10
90	100	302	418	254144	20780	6837	16	10000	13498	107984	389593	nt	18509	57422	18511	16983	221326	148984	11
91	100	1177	2244	1364352	103327	30626	17	10000	13226	105808	493765	nt	27791	86552	27793	25642	333409	221655	9
92	100	546	691	420128	39218	13112	23	10000	12243	97944	106358	nt	35000	108534	35004	31765	417375	276874	nt
93	100	994	1629	990432	97421	25397	26	10000	14469	115752	301014	nt	13203	41026	13205	12149	158114	105359	9
94	100	688	1316	800128	53890	17214	13	2978	3691	29528	94203	13	13969	43416	13971	12744	166977	113468	10
95	100	1177	1628	989824	97749	29547	19	10000	12724	101792	452952	nt	19802	61552	19804	18186	237074	157328	12
96	100	621	887	539296	47546	15200	18	921	22084	176672	41592	19	25560	79020	25562	23115	304071	201515	17
97	100	256	332	201856	18218	5595	22	10000	12997	103976	527280	nt	21305	66060	21310	19327	254031	168139	15
98	100	1258	1938	1178304	106624	32042	19	10000	13838	110704	321749	nt	19568	60755	19570	17952	234076	156561	10
99	100	666	1254	762432	65703	17290	16	10000	13600	108800	516390	nt	29332	92084	29334	27129	353557	235530	8
100	100	405	555	337440	28609	9554	20	10000	12357	98856	430984	nt	20202	62826	20204	18390	241434	161905	13

TABLE A.11: Higher dimensional examples — Difficult piecewise quadratic constraint with $m_2 := N$

ex	N	Red Alg						MPBNGC					SolvOpt						
		Nit	Na	c	t ₁	t ₂	R	Nit	Nb	c	t ₁	R	Nit	Nc	Ng	Nĝ	c	t ₁	R
1	20	108	197	119776	390	120	7	875	1042	8336	1093	7	9850	31342	9852	7883	115889	2796	7
2	20	163	302	183616	592	112	7	2436	50000	400000	5609	nt	11311	36389	11313	9097	134008	3202	7
3	20	1378	2602	1582016	5030	1145	5	2596	50000	400000	5968	nt	4366	13939	4368	3548	51626	1218	4
4	20	1014	1880	1143040	4374	834	8	665	847	6776	781	8	12276	39266	12278	9887	145027	3592	8
5	20	398	566	344128	1764	354	11	10000	10059	80472	1890	nt	4610	14656	4614	3581	53897	1281	11
6	20	1116	2009	1221472	4874	969	7	3298	50000	400000	6593	nt	2537	8052	2539	2022	29787	718	7
7	20	879	1554	944832	3796	706	9	6606	50000	400000	11061	nt	8803	27932	8805	7138	103693	2515	8
8	20	2678	4920	2991360	11671	2025	8	1832	50000	400000	4531	nt	2606	8161	2608	2068	30350	718	8
9	20	449	788	479104	1671	396	9	1783	50000	400000	4484	nt	2514	8020	2516	2031	29681	703	4
10	20	121	173	105184	468	93	6	282	391	3128	171	8	6681	21171	6684	5234	78096	1875	7
11	20	108	192	116736	531	81	9	1087	1318	10544	1421	9	851	2659	853	661	9860	234	9
12	20	290	536	325888	1092	207	5	219	450	3600	156	5	4087	12996	4091	3245	48000	1125	5
13	20	423	688	418304	1905	345	10	2548	50000	400000	5468	nt	3403	10772	3405	2741	39982	937	8
14	20	189	319	193952	733	253	8	346	467	3736	218	9	3462	10945	3464	2750	40532	953	8
15	20	407	803	488224	1655	335	6	10000	10263	82104	2031	nt	2465	7836	2467	1981	29016	687	7
16	20	17	30	18240	46	31	3	2665	50000	400000	5905	nt	2243	7140	2245	1814	26457	625	3
17	20	1017	1880	1143040	3999	924	6	1879	50000	400000	4827	nt	6609	21065	6611	5347	78004	1859	6
18	20	864	1628	989824	3343	733	6	1929	50000	400000	4796	nt	5829	18743	5831	4706	69097	1640	6
19	20	1018	1807	1098656	4156	1021	7	1921	50000	400000	4718	nt	998	3151	1000	803	11711	281	6
20	20	517	967	587936	2108	413	9	740	1329	10632	1062	9	6611	20974	6613	5200	77387	1843	8
21	40	496	848	6784	6296	1484	10	10000	12976	103808	105202	nt	4557	14198	4559	3948	53917	4874	7
22	40	375	528	4224	4421	1061	15	10000	13169	105352	93327	nt	11103	34550	11105	9411	130648	11796	11

TABLE A.11: Higher dimensional examples — Difficult piecewise quadratic constraint with $m_2 := N$ (continued)

ex	N	Red Alg						MPBNGC					SolvOpt						
		Nit	Na	c	t ₁	t ₂	R	Nit	Nb	c	t ₁	R	Nit	Nc	Ng	N $\hat{g}$	c	t ₁	R
23	40	735	1172	9376	8828	2303	17	10000	11883	95064	80812	nt	12823	39585	12825	11001	150648	13546	15
24	40	150	233	1864	1686	471	13	1406	2181	17448	7062	13	7999	24755	8001	6845	94048	8436	12
25	40	389	558	4464	4499	1184	16	10000	13253	106024	90359	nt	6653	20678	6655	5694	78403	7046	13
26	40	332	419	3352	4093	1048	19	10000	11700	93600	85327	nt	3182	9918	3184	2733	37587	3390	7
27	40	2068	3770	30160	24640	6610	13	10000	12496	99968	76953	nt	4308	13445	4310	3714	50962	4546	8
28	40	346	604	4832	4218	1068	12	10000	13703	109624	97124	nt	18483	57686	18485	15913	218566	20014	7
29	40	1341	2295	18360	16250	3995	13	10000	10448	83584	8921	nt	8657	27069	8659	7476	102543	9280	11
30	40	255	298	2384	3046	781	16	10000	13478	107824	92265	nt	7896	24651	7898	6861	93579	8593	10
31	40	659	1146	9168	7968	2133	16	10000	15177	121416	97609	nt	5133	15963	5135	4429	60618	5484	9
32	40	461	584	4672	5155	1375	14	2087	7072	56576	10578	14	9022	27866	9024	7806	106222	9671	14
33	40	606	909	7272	6968	1953	16	10000	13264	106112	91421	nt	7832	24369	7834	6722	92406	8343	9
34	40	481	863	6904	5640	1484	12	9955	50000	400000	98171	nt	10870	33915	10873	9203	128058	11703	10
35	40	1403	2639	21112	15530	4429	8	10000	11761	94088	58780	nt	11696	37395	11698	10183	140433	12765	4
36	40	658	1165	9320	7858	1948	13	10000	12744	101952	87015	nt	10770	33786	10772	9349	127935	11452	7
37	40	2511	3835	30680	34327	7851	19	10000	12714	101712	77171	nt	8656	26755	8658	7373	101603	9171	13
38	40	364	590	4720	4217	1157	12	10000	12857	102856	98249	nt	8892	27954	8894	7665	105585	9561	8
39	40	198	345	2760	2171	656	10	10000	12580	100640	102749	nt	7271	22787	7273	6314	86335	7749	7
40	40	763	1398	11184	9030	2338	11	5742	7378	59024	58734	11	7146	22173	7148	6173	84309	7530	10
41	60	225	324	2592	6140	1792	15	4123	5623	44984	86421	15	10720	33133	10723	9353	126494	32749	12
42	60	398	641	5128	10859	3076	18	2689	3737	29896	39749	18	10735	33065	10740	9451	126703	33296	18
43	60	426	784	6272	11515	3477	15	842	19084	152672	14296	15	15535	47978	15537	13675	183592	48031	14
44	60	481	705	5640	14218	3943	18	10000	13024	104192	310952	nt	16241	50656	16243	14280	192881	50343	13
45	60	196	334	2672	5453	1553	17	7377	8768	70144	138156	16	9321	28998	9323	8321	110928	29593	11
46	60	829	1189	9512	24343	6039	20	10000	12306	98448	191140	nt	11508	35736	11510	10248	136746	35468	15
47	60	457	660	5280	13843	3824	17	2977	3766	30128	36500	20	9158	28252	9161	8026	108065	28327	15
48	60	1715	2907	23256	53765	13126	20	10000	13121	104968	351389	nt	11292	34819	11294	9963	133409	35046	17
49	60	546	761	6088	15717	4658	24	10000	12725	101800	287718	nt	9374	29182	9376	8259	111269	28983	9
50	60	2334	3726	29808	68796	18451	28	6564	8328	66624	115890	27	9799	30148	9801	8677	115730	29921	21
51	60	211	299	2392	5921	1757	20	10000	12461	99688	219514	nt	12452	38647	12454	11028	147740	39468	13
52	60	349	539	4312	8468	2544	15	10000	11434	91472	57671	nt	12995	40511	12997	11542	154639	40765	9
53	60	1957	2370	18960	61546	14814	23	10000	13130	105040	171687	nt	15471	48033	15473	13655	183450	48108	12
54	60	890	1513	12104	25515	7102	19	10000	11026	88208	19780	nt	15267	46872	15269	13340	179571	46921	18
55	60	438	587	4696	12953	3420	19	4199	5611	44888	100656	21	10200	31586	10202	8924	120550	31515	17
56	60	1326	2086	16688	41593	10152	20	10000	12894	103152	215312	nt	16846	51857	16848	14872	198874	51796	15
57	60	781	1130	9040	25171	6103	25	10000	14123	112984	155639	nt	6178	19210	6180	5506	73478	19327	8
58	60	321	597	4776	7843	2512	12	2464	3344	26752	36359	12	11091	34358	11093	9787	131356	34187	12
59	60	673	1121	8968	18624	5324	17	10000	13183	105464	162453	nt	16778	52530	16780	14970	200310	52827	11
60	60	581	977	7816	15921	4666	19	8663	10525	84200	160421	19	8815	27339	8817	7879	104766	27671	12
61	80	1128	1410	688080	62577	18479	32	10000	14373	114984	466187	nt	9108	28296	9110	8058	108096	72108	14
62	80	322	465	226920	14186	5155	18	4739	6169	49352	223500	19	24397	75680	24399	21759	289834	193983	16
63	80	1168	1593	777384	57609	18907	27	10000	12668	101344	474859	nt	16784	51859	16786	15155	199541	132890	19
64	80	359	494	241072	17452	5572	23	5262	6552	52416	222186	25	15997	49376	15999	14229	189436	126327	22
65	80	1714	2346	1144848	93109	28473	31	10000	12942	103536	244859	nt	10233	31695	10235	9275	121920	82234	15
66	80	423	621	303048	21124	6749	29	10000	12640	101120	758171	nt	2728	8468	2730	2435	32431	21733	9
67	80	790	1328	648064	43374	13330	23	10000	14628	117024	410874	nt	9165	28443	9167	8332	109383	73421	12
68	80	1310	1859	907192	74124	21309	26	10000	13833	110664	369483	nt	7212	22280	7214	6505	85717	57359	12
69	80	197	256	124928	9671	2843	22	10000	13917	111336	372249	nt	20529	64071	20531	18433	245034	164609	11
70	80	744	1167	569496	37702	11992	27	10000	12587	100696	592093	nt	16440	50801	16442	14869	195535	130124	17
71	80	1470	2441	1191208	79140	25310	26	10000	13356	106848	223702	nt	8195	25375	8197	7425	97616	65296	11

TABLE A.11: Higher dimensional examples — Difficult piecewise quadratic constraint with $m_2 := N$ (continued)

ex	N	Red Alg						MPBNGC					SolvOpt						
		Nit	Na	c	t ₁	t ₂	R	Nit	Nb	c	t ₁	R	Nit	Nc	Ng	Nĝ	c	t ₁	R
72	80	911	1306	637328	45218	15073	27	10000	11793	94344	334077	nt	15167	46849	15169	13694	180287	120999	17
73	80	1493	1756	856928	77140	23834	29	10000	15730	125840	327828	nt	18064	56070	18066	16212	214974	143811	14
74	80	1680	2977	1452776	93640	28452	23	10000	13889	111112	370889	nt	15687	48657	15689	14130	186771	125046	17
75	80	267	358	174704	13217	4026	23	10000	12978	103824	567467	nt	16091	49603	16093	14538	191099	128499	17
76	80	1750	2222	1084336	91546	28255	28	10000	13402	107216	283171	nt	21054	64846	21056	18910	249590	167124	19
77	80	669	812	396256	33234	10620	23	10000	12977	103816	274280	nt	9208	28462	9210	8296	109442	73608	16
78	80	378	508	247904	19280	6029	28	10000	13292	106336	527984	nt	12874	39831	12876	11530	152880	102546	15
79	80	795	1318	643184	44499	13329	19	10000	13277	106216	288609	nt	8558	26423	8560	7716	101674	69124	16
80	80	1563	2041	996008	79999	25366	26	10000	13311	106488	349046	nt	15919	49478	15921	14346	189757	127374	13
81	100	668	1142	694336	71421	19034	22	10000	14060	112480	673828	nt	21948	68236	21950	20133	262721	311749	13
82	100	1129	1625	988000	97249	32463	31	5124	6872	54976	178515	33	18969	58513	18971	17194	225521	267156	26
83	100	2010	2121	1289568	177358	57450	41	10000	14323	114584	378140	nt	11551	35555	11553	10460	137149	163561	22
84	100	673	775	471200	54140	18988	33	10000	13422	107376	393109	nt	6718	20848	6721	6083	80108	95062	11
85	100	847	1018	618944	68030	24555	32	10000	13072	104576	518827	nt	10807	33502	10809	9891	129104	153609	15
86	100	1734	2279	1385632	158406	51587	33	10000	15603	124824	388890	nt	8721	27070	8723	8021	104372	124686	10
87	100	1601	1918	1166144	137890	45702	30	10000	13946	111568	358484	nt	12885	39999	12887	11812	154095	183531	16
88	100	1537	2154	1309632	148280	45319	32	10000	17083	136664	432249	nt	6740	20865	6742	6169	80463	95765	12
89	100	393	501	304608	31812	10747	27	10000	13823	110584	565202	nt	9666	29905	9668	8837	115325	137186	16
90	100	915	1053	640224	74437	25944	25	10000	13777	110216	395890	nt	17791	55267	17793	16399	213110	255108	9
91	100	1224	1466	891328	103842	35258	35	10000	17222	137776	401562	nt	15026	46593	15028	13798	179664	213437	18
92	100	978	1161	705888	81078	27825	29	10000	12328	98624	599906	nt	22926	70888	22928	20844	273092	325249	22
93	100	3405	3638	2211904	357061	100012	33	10000	14681	117448	340296	nt	14085	43695	14087	12887	168312	199859	9
94	100	991	1273	773984	86546	28408	34	10000	13864	110912	436780	nt	13536	41955	13538	12367	161625	191530	15
95	100	900	1070	650560	71327	25599	28	10000	12784	102272	550858	nt	24157	74597	24159	22027	287752	342655	20
96	100	792	1157	703456	74999	22481	30	10000	13445	107560	365296	nt	11765	36522	11767	10771	140658	167499	11
97	100	1855	2130	1295040	157687	52694	34	10000	13375	107000	534078	nt	9235	28612	9237	8455	110300	132437	14
98	100	662	800	486400	53218	18028	31	10000	13756	110048	747843	nt	13589	42069	13591	12313	161850	191780	19
99	100	1355	1523	925984	121389	39159	36	10000	14653	117224	376467	nt	15569	48223	15571	14222	185825	222156	17
100	100	633	932	566656	57374	18413	30	10000	11959	95672	689514	nt	16020	49617	16022	14669	191307	228327	19

Bibliography

- [1] F. Alizadeh and D. Goldfarb. Second-order cone programming. *Mathematical Programming*, 95: 3–51, 2001.
- [2] E.D. Andersen. *The MOSEK C API manual*. MOSEK ApS, Denmark, 1998 – 2010. Version 6.0 (Revision 66). URL <http://www.mosek.com/>.
- [3] E.D. Andersen, C. Roos, and T. Terlaky. On implementing a primal-dual interior-point method for conic quadratic optimization. *Math. Program.*, B(95):249–277, 2003.
- [4] A.M. Bagirov. A method for minimization of quasidifferentiable functions. *Optimization Methods and Software*, 17:31–60, 2002.
- [5] A.M. Bagirov. Continuous subdifferential approximations and their applications. *Journal of Mathematical Sciences*, 115:2567–2609, 2003.
- [6] A.M. Bagirov, G. Beliakov, M. Mammadov, and A.M. Rubinov. *Programming library GANSO (Global And Non-Smooth Optimization)*. Centre for Informatics and Applied Optimization, University of Ballarat, Australia, 2005. Version 5.1.0.0. URL <http://www.ballarat.edu.au/ard/itms/CIAO/ganso/>.
- [7] A.M. Bagirov and A. Ganjehlou. A quasisecant method for minimizing nonsmooth functions. *Optimization Methods and Software*, 25(1):3–18, 2010.
- [8] A.M. Bagirov, B. Karasözen, and M. Sezer. Discrete Gradient Method: Derivative-Free Method for Nonsmooth Optimization. *Journal of Optimization Theory and Applications*, 137(2):317–334, 2008.
- [9] M. Berkelaar, K. Eikland, and P. Notebaert. *lp_solve*. Open source (Mixed-Integer) Linear Programming system (Version 5.1.0.0), May 2004. URL <http://lpsolve.sourceforge.net/>.
- [10] L.S. Blackford, J. Demmel, J. Dongarra, I. Duff, S. Hammarling, G. Henry, M. Heroux, L. Kaufman, A. Lumsdaine, A. Petitet, R. Pozo, K. Remington, and R.C. Whaley. An updated set of basic linear algebra subprograms (BLAS). *ACM Trans. Math. Soft.*, 28(2):135–151, 2002. URL <http://www.netlib.org/>.
- [11] J.F. Bonnans, J.C. Gilbert, C. Lemaréchal, and C.A. Sagastizábal. *Numerical Optimization: Theoretical and Practical Aspects*. Springer-Verlag, Berlin Heidelberg New York, 2nd ed., 2006. ISBN 3-540-35445-X.

- [12] J.M. Borwein and A.S. Lewis. *Convex Analysis and nonlinear Optimization: Theory and Examples*. Springer, 2nd ed., 2006.
- [13] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, 2004.
- [14] J.V. Burke, A.S. Lewis, and M.L. Overton. Approximating subdifferentials by random sampling of gradients. *Math. Oper. Res.*, 27(3):567–584, August 2002.
- [15] J.V. Burke, A.S. Lewis, and M.L. Overton. A robust gradient sampling algorithm for nonsmooth, nonconvex optimization. *SIAM J. Optim.*, 15(3):751–779, 2005.
- [16] R.H. Byrd, J. Nocedal, and R.B. Schnabel. Representations of quasi-newton matrices and their use in limited memory methods. *Mathematical Programming*, 63:129–156, 1994.
- [17] F.E. Curtis and M.L. Overton. A robust sequential quadratic programming algorithm for non-convex, nonsmooth constrained optimization. *SIAM Journal on Optimization*, 2010. Submitted for publication.
- [18] F. Domes. GloptLab – a configurable framework for the rigorous global solution of quadratic constraint satisfaction problems. *Optimization Methods and Software*, 24(4–5):727–747, 2009. URL <http://www.mat.univie.ac.at/~dferi/gloptlab.html>.
- [19] H. Fendl, A. Neumaier, and H. Schichl. Certificates of infeasibility via nonsmooth optimization. In preparation, 2011.
- [20] H. Fendl and H. Schichl. A feasible second order bundle algorithm for nonsmooth, nonconvex optimization problems with inequality constraints. In preparation, 2011.
- [21] H. Fendl and H. Schichl. Numerical results of the feasible second order bundle algorithm for nonsmooth, nonconvex optimization problems with inequality constraints. In preparation, 2011.
- [22] R. Fletcher and S. Leyffer. A bundle filter method for nonsmooth nonlinear optimization. Numerical Analysis Report NA/195, University of Dundee, Department of Mathematics, December 1999.
- [23] C.A. Floudas. *Nonlinear and Mixed-Integer Optimization: Fundamentals and Applications*. Oxford Univ. Press, Oxford, 1995.
- [24] C.A. Floudas. *Deterministic Global Optimization: Theory, Algorithms and Applications*. Kluwer, Dordrecht, 1999.
- [25] P.E. Gill and W. Murray. Newton-type methods for unconstrained and linearly constrained Optimization. *Math. Programming*, 28:1974, 311–350.
- [26] 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.

- [27] G.H. Golub and C.F. Van Loan. *Matrix Computations (Johns Hopkins Studies in Mathematical Sciences)(3rd Edition)*. The Johns Hopkins University Press, 3rd ed., October 1996. ISBN 0801854148.
- [28] K. Goto and R.A. van de Geijn. Anatomy of high-performance matrix multiplication. *ACM Trans. Math. Softw.*, 34(3):12:1–12:25, May 2008. ISSN 0098-3500. URL <http://www.tacc.utexas.edu/tacc-projects/gotoblas2/>.
- [29] M. Grant and S. Boyd. *CVX Users' Guide for CVX version 1.2 (build 711)*, June 2009. URL <http://cvxr.com/cvx/>.
- [30] A. Griewank and G.F. Corliss, editors. *Automatic Differentiation of Algorithms: Theory, Implementation, and Application*. SIAM, Philadelphia, PA, 1991. ISBN 0-89871-284-X.
- [31] M. Haarala. *Large-scale nonsmooth optimization: Variable metric bundle method with limited memoery*. PhD thesis, University of Jyväskylä, Department of Mathematical Information Technology, 2004.
- [32] M. Haarala, M. Miettinen, and M.M. Mäkelä. New limited memory bundle method for large-scale nonsmooth optimization. *Optimization Methods and Software*, 19(6):673–692, 2004.
- [33] N. Haarala, M. Miettinen, and M.M. Mäkelä. Globally convergent limited memory bundle method for large-scale nonsmooth optimization. *Mathematical Programming*, 109(1):181–205, 2007.
- [34] E.R. Hansen. *Global Optimization Using Interval Analysis*. Dekker, New York, 1992.
- [35] J. Herskovits. Feasible direction interior-point technique for nonlinear optimization. *Journal of Optimization Theory and Applications*, 99(1):121–146, 1998.
- [36] J. Herskovits and G. Santos. On the computer implementation of feasible direction interior point algorithms for nonlinear optimization. *Structural Optimization*, 14:165–172, 1997.
- [37] H. Heuser. *Lehrbuch der Analysis Teil 1*. B. G. Teubner, Stuttgart Leipzig, 1998.
- [38] H. Heuser. *Lehrbuch der Analysis Teil 2*. B. G. Teubner, Stuttgart Leipzig Wiesbaden, 2000.
- [39] N.J. Higham. *Functions of Matrices: Theory and Computation*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008. ISBN 978-0-898716-46-7.
- [40] R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, February 1990. ISBN 0521386322.
- [41] F. Kappel and A.V. Kuntsevich. An Implementation of Shor's r-Algorithm. *Comput. Optim. Appl.*, 15(2):193–205, February 2000. ISSN 0926-6003.
- [42] E. Karas, A. Ribeiro, C. Sagastizábal, and M. Solodov. A bundle-filter method for nonsmooth convex constrained optimization. *Math. Program.*, B(116):297–320, 2009.

- [43] N. Kar Mitsa. Decision tree for nonsmooth optimization software. WWW-Page. URL <http://napsu.karmitsa.fi/solveromatic/>.
- [44] N. Kar Mitsa, A.M. Bagirov, and M.M. Mäkelä. Empirical and Theoretical Comparisons of Several Nonsmooth Minimization Methods and Software. TUCS Technical Report 959, Turku Centre for Computer Science, October 2009.
- [45] N. Kar Mitsa and M.M. Mäkelä. Adaptive limited memory bundle method for bound constrained large-scale nonsmooth optimization. *Optimization: A Journal of Mathematical Programming and Operations Research*, 59(6):945–962, 2010.
- [46] N. Kar Mitsa and M.M. Mäkelä. Limited memory bundle method for large bound constrained nonsmooth optimization: Convergence analysis. *Optimization Methods and Software*, 25(6): 895–916, 2010.
- [47] N. Kar Mitsa, M.M. Mäkelä, and M.M. Ali. Limited Memory Bundle Algorithm for Inequality Constrained Nondifferentiable Optimization. Reports of the Department of Mathematical Information Technology Series B Scientific Computing 3, University of Jyväskylä, Department of Mathematical Information Technology, P.O. Box 35 (Agora), FI-40014 University of Jyväskylä, Finland, 2007.
- [48] R.B. Kearfott. *Rigorous Global Search: Continuous Problems*. Kluwer, Dordrecht, 1996.
- [49] K.C. Kiwiel. *Methods of Descent for Nondifferentiable Optimization*. Lecture Notes in Mathematics 1133. Springer-Verlag, Berlin, 1985.
- [50] K.C. Kiwiel. A constraint linearization method for nondifferentiable convex minimization. *Numerische Mathematik*, 51:395–414, 1987.
- [51] K.C. Kiwiel. Restricted step and Levenberg-Marquardt techniques in proximal bundle methods for nonconvex nondifferentiable optimization. *SIAM J. Optimization*, 6(1):227–249, February 1996.
- [52] K.C. Kiwiel and A. Stachurski. Issues of Effectiveness Arising in the Design of a System of Nondifferentiable Optimization algorithms. In A. Lewandowski and A.P. Wierzbicki, editors, *Aspiration Based Decision Support Systems: Theory, Software and Applications*, Lecture Notes in Economics and Mathematical Systems 331, pp. 180–192. Springer-Verlag, Berlin, 1989.
- [53] D.L. Kroshko. *ralg*. Software package. URL <http://openopt.org/ralg/>.
- [54] A.V. Kuntsevich and F. Kappel. *SolvOpt The Solver For Local Nonlinear Optimization Problems*. Karl-Franzens Universität Graz, 1997. URL <http://www.kfunigraz.ac.at/imawww/kuntsevich/solvopt/>.
- [55] C.T. Lawrence and A.L. Tits. A computationally efficient feasible sequential quadratic programming algorithm. *SIAM J. Optim.*, 11(4):1092–1118, 2001.
- [56] A.S. Lewis and M.L. Overton. Nonsmooth optimization via BFGS. *SIAM Journal on Optimization*, 2008. Submitted for publication.

- [57] M.S. Lobo, L. Vandenberghe, and S. Boyd. *socp Software for Second-Order Cone Programming User's Guide*, April 1997. URL http://stanford.edu/~boyd/old_software/SOCP.html.
- [58] M.S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebert. Applications of second-order cone programming. *Linear Algebra and its Applications*, 284:193–228, 1998.
- [59] J. Löfberg. YALMIP: A Toolbox for Modeling and Optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004. URL <http://users.isy.liu.se/johanl/yalmip/>.
- [60] L. Lukšan. Dual method for solving a special problem of quadratic programming as a subproblem at linearly constrained nonlinear minimax approximation. *Kybernetika*, 20:445–457, 1984.
- [61] L. Lukšan. An implementation of recursive quadratic programming variable metric methods for linearly constrained nonlinear minimax approximation. *Kybernetika*, 21(1):22–40, 1985.
- [62] L. Lukšan and J. Vlček. PBUN, PNEW – Bundle-Type Algorithms for Nonsmooth Optimization. Technical report 718, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, Czech Republic, September 1997. URL <http://www.uivt.cas.cz/~luksan/subroutines.html>.
- [63] L. Lukšan and J. Vlček. PMIN – A Recursive Quadratic Programming Variable Metric Algorithm for Minimax Optimization. Technical report 717, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, Czech Republic, September 1997. URL <http://www.uivt.cas.cz/~luksan/subroutines.html>.
- [64] L. Lukšan and J. Vlček. A bundle-Newton method for nonsmooth unconstrained minimization. *Mathematical Programming*, 83:373–391, 1998.
- [65] L. Lukšan and J. Vlček. Globally convergent variable metric method for convex nonsmooth unconstrained minimization. *Journal of Optimization Theory and Applications*, 102(3):593–613, 1999.
- [66] L. Lukšan and J. Vlček. NDA: Algorithms for nondifferentiable optimization. Technical report 797, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, Czech Republic, 2000. URL <http://www.uivt.cas.cz/~luksan/subroutines.html>.
- [67] L. Lukšan and J. Vlček. Test Problems for Nonsmooth Unconstrained and Linearly Constrained Optimization. Technical report 798, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, Czech Republic, January 2000.
- [68] L. Lukšan and J. Vlček. Test problems for unconstrained optimization. Technical report 897, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, Czech Republic, November 2003.
- [69] J.R. Magnus and H. Neudecker. *Matrix Differential Calculus with Applications in Statistics and Econometrics, 2nd Edition*. John Wiley & Sons, 3rd ed., March 1999. ISBN 047198633X.

- [70] M.M. Mäkelä. Multiobjective Proximal Bundle Method for Nonconvex Nonsmooth Optimization: Fortran Subroutine MPBNGC 2.0. Reports of the Department of Mathematical Information Technology, Series B. Scientific computing, B 13/2003 University of Jyväskylä, Jyväskylä, 2003. URL <http://napsu.karmita.fi/proxbundle/>.
- [71] M.M. Mäkelä and P. Neittaanmäki. *Nonsmooth Optimization: Analysis and Algorithms with Applications to Optimal Control*. World Scientific Publishing Co., Singapore, 1992.
- [72] N. Maratos. *Exact penalty function algorithms for finite dimensional and control optimization problems*. PhD thesis, University of London, 1978.
- [73] O.H. Merrill. *Applications and extensions of an algorithm that computes fixed points of certain upper semicontinuous point to set mappings*. PhD thesis, Univ. of Michigan, Ann Arbor, 1972.
- [74] R. Mifflin. An algorithm for constrained optimization with semismooth functions. *Mathematics of Operations Research*, 2(2):191–207, May 1977.
- [75] R. Mifflin. Semismooth and semiconvex functions in constrained Optimization. *SIAM J. Control and Optimization*, 15(6):959–972, November 1977.
- [76] R. Mifflin. A Modification and an extension of Lemarechal’s algorithm for nonsmooth minimization. *Mathematical Programmin Study*, 17:77–90, 1982.
- [77] R. Mifflin, D. Sun, and L. Qi. Quasi-Newton bundle-type methods for nondifferentiable convex optimization. *SIAM J. Optim.*, 8(2):583–603, May 1998.
- [78] H.D. Mittelmann. Benchmarking of Optimization Software. INFORMS Annual Meeting, Pittsburgh, PA, November 2006.
- [79] H.D. Mittelmann. Recent Developments in SDP and SOCP Software. INFORMS Annual Meeting, Pittsburgh, PA, November 2006.
- [80] Y. Nesterov and A. Nemirovsky. *Interior-point polynomial methods in convex programming*, vol. 13 of *Studies in Applied Mathematics*. SIAM, Philadelphia, PA, January 1994.
- [81] Y. Nesterov and J.-P. Vial. Homogeneous analytic center cutting plane methods for convex problems and variational inequalities. *SIAM Journal on Optimization*, 9:707–728, 1999.
- [82] A. Neumaier. *Interval methods for systems of equations*, vol. 37 of *Encyclopedia of Mathematics and its Applications*. Cambridge Univ. Press, Cambridge, 1990.
- [83] A. Neumaier. Complete search in continuous global optimization and constraint satisfaction. *Acta Numerica*, 13:271–369, 2004. ISSN 0962-4929.
- [84] J. Nocedal and S.J. Wright. *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2nd ed., 2006. ISBN 978-0-387-30303-1.
- [85] M. Overton. Hanso: Hybrid Algorithm for Non-Smooth Optimization, September 2010. New York University, Department of Computer Science. URL <http://www.cs.nyu.edu/overton/software/hanso/>.

- [86] I. Pólik. *Addendum to the SeDuMi user guide version 1.1*, June 2005. URL <http://sedumi.ie.lehigh.edu/>.
- [87] S.M. Rump. INTLAB - INTerval LABoratory. In Tibor Csendes, editor, *Developments in Reliable Computing*, pp. 77–104. Kluwer Academic Publishers, Dordrecht, 1999. URL <http://www.ti3.tu-harburg.de/~rump/intlab/>.
- [88] C. Sagastizábal and M. Solodov. An infeasible bundle method for nonsmooth convex constrained optimization without a penalty function or a filter. *SIAM J. Optim.*, 16:146–169, 2006.
- [89] N.V. Sahinidis. BARON. Branch and reduce optimization navigator. User’s manual. WWW-Document. URL <http://archimedes.scs.uiuc.edu/baron/baron.html>.
- [90] N.V. Sahinidis. BARON: A general purpose global optimization software package. *J. Global Optimization*, 8:201–205, 1996.
- [91] H. Schichl. The COCONUT environment. Software package. URL <http://www.mat.univie.ac.at/coconut-environment/>.
- [92] H. Schichl. *Mathematical Modeling and Global Optimization*. Habilitation thesis, Universität Wien, November 2003.
- [93] H. Schichl. Global optimization in the COCONUT project. *Numerical Software with Result Verification*, pp. 277–293, 2004.
- [94] H. Schichl and A. Neumaier. Transposition theorems and qualification-free optimality conditions. *SIAM J. Optimization*, 17:1035–1055, 2006.
- [95] K. Schittkowski. Test Examples for Nonlinear Programming Codes – All Problems from the Hock-Schittkowski-Collection. Department of Computer Science, University of Bayreuth, D - 95440 Bayreuth, February 2009.
- [96] K. Schittkowski. An updated set of 306 test problems for nonlinear programming with validated optimal solutions - user’s guide. Department of Computer Science, University of Bayreuth, D - 95440 Bayreuth, November 2009.
- [97] L. Schrage. *Optimization Modeling with LINGO*. LINDO Systems, Inc., Chicago, IL, 1999.
- [98] H. Schramm. Eine Kombination von Bundle- und Trust-Region-Verfahren zur Lösung nichtdifferenzierbarer Optimierungsprobleme. *Bayreuther Mathematische Schriften*, 30, 1989.
- [99] H. Schramm and J. Zowe. A version of the bundle idea for minimizing a nonsmooth function: Conceptual idea, convergence analysis, numerical results. *SIAM J. Optimization*, 2(1):121–152, February 1992.
- [100] N.Z. Shor. *Minimization Methods for Non-Differentiable Functions*. Springer-Verlag, Berlin Heidelberg New York Tokyo, 1985.
- [101] M.V. Solodov. On the sequential quadratically constrained quadratic programming methods. *Mathematics of Operations Research*, 29(1), February 2004.

- [102] J.F. Sturm. *Using SeDuMi 1.02, A MATLAB Toolbox for optimization over symmetric cones (Updated for Version 1.05)*. Department of Econometrics, Tilburg University, Tilburg, The Netherlands, 1998 – 2001.
- [103] A.L. Tits. Feasible sequential quadratic programming. In C.A. Floudas and P.M. Pardalos, editors, *Encyclopedia of Optimization*, pp. 1001–1005. Springer, 2nd ed., 2009. ISBN 978-0-387-74758-3.
- [104] K.C. Toh, R.H. Tütüncü, and M.J. Todd. *On the implementation and usage of SDPT3 – a MATLAB software package for semidefinite-quadratic-linear programming, version 4.0*, July 2006. Draft. URL <http://www.math.nus.edu.sg/~mattohkc/sdpt3.html>.
- [105] J.-P. Vial and N. Sawhney. *OBOE User Guide Version 1.0*, June 2007. URL <https://projects.coin-or.org/OBOE/>.
- [106] J. Vlček. Bundle algorithms for nonsmooth unconstrained minimization. Research Report 608, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, Czech Republic, 1995.
- [107] J. Vlček and L. Lukšan. Globally convergent variable metric method for nonconvex nondifferentiable unconstrained minimization. *Journal of Optimization Theory and Applications*, 111(2): 407–430, 2001.
- [108] A. Wächter. *Introduction to IPOPT: A tutorial for downloading, installing, and using IPOPT*, October 2009. Revision : 1585. URL <https://projects.coin-or.org/Ipopt/>.
- [109] A. Wächter and L.T. Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Program.*, 106(1):25–57, May 2006. ISSN 0025-5610.
- [110] R.C. Whaley and A. Petitet. Minimizing development and maintenance costs in supporting persistently optimized BLAS. *Software: Practice and Experience*, 35(2):101–121, February 2005. URL <http://math-atlas.sourceforge.net/>.
- [111] J. Zowe. The BT-Algorithm for minimizing a nonsmooth functional subject to linear constraints. In F.H. Clarke, V.F. Demyanov, and F. Giannessi, editors, *Nonsmooth optimization and related topics*, pp. 459–480. Plenum Press, New York, 1989.

Hannes Fendl

Curriculum Vitae

Persönliche Daten

Geburtstag	19.05.1983
Staatsbürger-schaft	Österreich
Eltern	<i>Johann Fendl</i> , geboren am 24.10.1952. Landesbeamter der NÖ Agrarbezirksbehörde. <i>Gottfrieda Fendl (geb. Übelbacher)</i> , geboren am 09.09.1955. Angestellte der NÖ Gebietskrankenkasse.

Ausbildung

09/1989– 06/1993	<i>Volksschule des Instituts der Englischen Fräulein</i> , 3100 St. Pölten.
09/1993– 06/2001	<i>Gymnasium des Instituts der Englischen Fräulein</i> , 3100 St. Pölten.
19.06.2001	<i>Matura</i> . Mit gutem Erfolg bestanden.
10/2001– 06/2004	<i>Universität Wien</i> . Erster Studienabschnitt des Diplomstudiums der Mathematik.
16.06.2004	<i>Abschluss des ersten Studienabschnitts</i> . Mit Auszeichnung bestanden.
10/2004– 10/2006	<i>Universität Wien</i> . Zweiter Studienabschnitt des Studiums der Mathematik.
24.10.2006	<i>Abschluss des Studiums (Mag. rer. nat.)</i> . Mit Auszeichnung bestanden.
01/2007– 09/2007	<i>Arbeiter-Samariterbund NÖ</i> , 3203 Rabenstein. Ableistung des ordentlichen Zivildienstes im Rettungs- und Krankentransportdienst.
10/2007– 06/2011	<i>Universität Wien</i> . Doktoratsstudium der Naturwissenschaften.

Diplomarbeit

Titel	<i>Ein Newton-Côtes artiger Algorithmus zur Berechnung von Coulomb-Integralen</i>
Betreuer	Ao. Prof. Dipl.-Ing. Dr. Hermann Schichl

Publikationen

- [1] A. Neumaier, H. Fendl, H. Schilly, and T. Leitner. VXQR: Derivative-free unconstrained optimization based on QR factorizations. *Soft Computing - A Fusion of Foundations, Methodologies and Applications*, pp. 1–12, 2010. ISSN 1432-7643.