



UNIVERSITÄT ZU LÜBECK  
INSTITUTE OF MATHEMATICS  
AND IMAGE COMPUTING

## **An I-BFGS based Tikhonov-regularised interior-point method in large-scale nonlinear inequality-constrained convex optimisation**

*Ein I-BFGS-basiertes Tikhonov-regularisiertes Innere-Punkte-Verfahren für hochdimensionale nichtlineare ungleichungsbeschränkte konvexe Optimierungsprobleme*

### **Masterarbeit**

verfasst am

**Institute of Mathematics and Image Computing**

im Rahmen des Studiengangs

**Mathematik in Medizin und Lebenswissenschaften**

der Universität zu Lübeck

vorgelegt von

**Ole Gildemeister**

ausgegeben und betreut von

**Prof. Dr. Jan Lellmann**

mit Unterstützung von

**Dr. Florian Mannel**

Lübeck, den 07. März 2024

### **Eidesstattliche Erklärung**

*Ich erkläre hiermit an Eides statt, dass ich diese Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.*

O. Gildemeister

Ole Gildemeister

## Zusammenfassung

In den letzten Jahren wurde der Entwicklung von Innere-Punkte-Verfahren für hochdimensionale konvexe Optimierungsprobleme viel Beachtung geschenkt. Innere-Punkte-Verfahren sind dafür bekannt, in Kombination mit dem Newtonverfahren exzellent zu funktionieren, aber durch dessen Voraussetzung, lineare Gleichungssysteme mit Hessematrizen zu lösen, ist ihre Verwendung häufig nicht sinnvoll oder sogar unmöglich. Quasi-Newton-Verfahren stellen hierzu eine geeignete Alternative dar, indem sie Ableitungen zweiter Ordnung ausschließlich anhand bereits berechneter Gradienteninformationen approximieren. Allerdings sind sie dafür bekannt, bei schlechter Konditioniertheit langsam zu konvergieren, wodurch ihre Verwendung in Innere-Punkte-Verfahren besonders herausfordernd ist. In dieser Arbeit wird daher ein Verfahren zur Lösung nichtlinearer ungleichungsbeschränkter Probleme diskutiert, in dem ein primaler Innere-Punkte-Ansatz mit einer Tikhonov-Regularisierung modifiziert und ein strukturiertes l-BFGS-Verfahren zum Lösen der inneren Probleme verwendet wird, das darauf abzielt, die Approximation an die Hessematrix zu verbessern. Es wird bewiesen, dass das Verfahren mit einer in den äußeren Iterierten linearen Konvergenzrate gegen die Minimum-Norm-Lösung konvergiert, und dass die Anzahl der l-BFGS-Iterationen, die zum Erreichen einer gegebenen Genauigkeit in der Zielfunktion benötigt werden, polynomiell in dieser Genauigkeit ist. Abschließend werden numerische Ergebnisse präsentiert, die auf die Eignung des Verfahrens für hochdimensionale nichtlineare konvexe Probleme hindeuten.

## Abstract

In recent years, the development of interior-point methods for large-scale convex optimisation problems has received much attention. Interior-point methods are well-known to perform excellent when combined with Newton's method, but the latter's requirement of solving linear systems of equations involving Hessian matrices often make their use unviable or even impossible. Quasi-Newton methods present a convenient alternative, as they approximate second-order derivatives using only previously computed gradient information. However, they are known to converge slowly under ill-conditionedness, which makes their use particularly challenging in interior-point methods. In this thesis, a method for solving nonlinear inequality-constrained problems is therefore discussed in which a primal interior-point approach is modified with a Tikhonov regularisation, and in which a structured l-BFGS method is used to solve the inner problems, aiming at improving the Hessian approximation. It is proven that the method converges to the minimal-norm solution at a linear convergence rate in the outer iterates, and that the number of l-BFGS iterations required to reach a given accuracy in the objective function is polynomial in the accuracy. Finally, numerical results of the method are presented which indicate its suitability for dealing with large-scale nonlinear convex problems.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Background and motivation	1
1.2	Contributions	2
1.3	Main results and structure	4
1.4	Related work	5
1.5	Notation	6
<b>2</b>	<b>Preliminaries on convex optimisation and interior-point methods</b>	<b>7</b>
2.1	Convexity and strong convexity	7
2.2	Nonlinear inequality-constrained convex programs	9
2.3	Interior-point methods	11
2.3.1	General concept	12
2.3.2	Existence of barrier minimisers	12
2.3.3	Computation of barrier minimisers	14
2.3.4	Generic interior-point method	15
2.4	Convergence of the barrier minimisers	15
2.4.1	Rate of convergence	16
2.4.2	Convergence to the analytic centre	19
<b>3</b>	<b>A Tikhonov-regularised interior-point method</b>	<b>21</b>
3.1	Tikhonov regularisation	21
3.2	The Tikhonov-regularised interior-point method	23
3.2.1	The regularised barrier problem	23
3.2.2	Existence of regularised barrier minimisers	24
3.2.3	General structure of the method	25
3.3	Convergence of the regularised barrier minimisers	26
3.3.1	Reduction to convergence of regularised minimisers	26
3.3.2	Convergence to the minimal norm solution	27
3.3.3	Rate of convergence	30
3.4	Convergence in case of inexact solutions	31
3.4.1	Convergence to the minimal norm solution	32
3.4.2	Rate of convergence	34
3.4.3	Specification of the outer method	36
3.5	Modifications	38
3.5.1	Shifted regularisation	38
3.5.2	Proximal-point regularisation	38
3.5.3	Linear equality constraints	39
3.5.4	Augmented Lagrangian approach	40

<b>4</b>	<b>A structured l-BFGS method for solving the inner problems</b>	<b>41</b>
4.1	L-BFGS methods	41
4.1.1	Quasi-Newton methods	42
4.1.2	The l-BFGS update	43
4.1.3	Seed matrix and two-loop recursion	44
4.2	The structured l-BFGS method	45
4.2.1	General concept	45
4.2.2	Choice of the scaling factor	46
4.3	Structured l-BFGS as inner solver	48
4.3.1	Choice of the seed matrix	49
4.3.2	Armijo backtracking	50
4.3.3	Specification of the structured l-BFGS method	51
4.4	Convergence of the inner iterates	51
4.4.1	Matrix norm and Lipschitz continuity	52
4.4.2	Convergence rate for the inner iterations	53
4.5	A bound on the number of required inner iterations	55
4.5.1	Bounds on the l-BFGS matrices	56
4.5.2	Assumptions required for the quantitative analysis	57
4.5.3	Derivation of the quantitative estimates	60
4.5.4	A global rate for the number of inner iterations	66
4.6	Polynomial complexity of the Tikhonov-regularised interior-point method	70
<b>5</b>	<b>Numerical results</b>	<b>73</b>
5.1	Implementation details	73
5.2	Test problems	74
5.3	Results for small-scale problems	75
5.4	Results for large-scale problems	81
<b>6</b>	<b>Conclusion</b>	<b>85</b>
	<b>Bibliography</b>	<b>88</b>

# List of symbols

Throughout this thesis, the following symbols are used:

Symbol	Description
$\mathbb{N}$	Set of natural numbers excluding 0
$\mathbb{N}_0$	Set of natural numbers including 0
$\mathbb{R}$	Set of real numbers
$n$	Number of variables
$m$	Number of inequality constraints
$f$	Objective function of the optimisation problem
$g$	Inequality-constraint function of the optimisation problem
$\mathcal{F}$	Set of feasible points, defined in Section 2.2
$\mathcal{F}_+$	Set of strictly feasible points, defined in Section 2.2
$\mathcal{S}$	Solution set, defined in Section 2.2
$\mu$	Barrier parameter
$\varepsilon$	Regularisation parameter
$b$	Barrier term
$\varphi_\mu, \varphi_{\mu_j}$	Barrier function, defined in Section 2.3
$f_\varepsilon$	Tikhonov-regularised function, defined in Section 3.1
$f_{\varepsilon,\mu}, f_{\varepsilon_j,\mu_j}$	Tikhonov-regularised barrier function, defined in Section 3.2
$(\mathcal{P})$	Original optimisation problem, defined in Chapter 1
$(\mathcal{P}_\mu), (\mathcal{P}_{\mu_j})$	Barrier problem, defined in Section 2.3
$(\mathcal{P}_\varepsilon)$	Tikhonov-regularised problem, defined in Section 3.1
$(\mathcal{P}_{\varepsilon,\mu}), (\mathcal{P}_{\varepsilon_j,\mu_j})$	Tikhonov-regularised barrier problem, defined in Section 3.2
$x^*$	Minimiser of $(\mathcal{P})$
$x_M^*$	Minimal-norm solution of $(\mathcal{P})$
$\bar{x}_\mu, \bar{x}_{\mu_j}$	Barrier minimiser, i.e. minimiser of $(\mathcal{P}_\mu) / (\mathcal{P}_{\mu_j})$
$\bar{x}_\varepsilon$	Regularised minimiser, i.e. minimiser of $(\mathcal{P}_\varepsilon)$
$\bar{x}_{\varepsilon,\mu}, \bar{x}_{\varepsilon_j,\mu_j}$	Regularised barrier minimiser, i.e. minimiser of $(\mathcal{P}_{\varepsilon,\mu}) / (\mathcal{P}_{\varepsilon_j,\mu_j})$
$i$	If not stated otherwise, index of inequality constraint
$j$	Outer iteration index
$k$	Inner iteration index
$x^j$	Outer iterate
$x_k^j, x_k$	Inner iterate
$\ell, \ell_{\max}$	Current / maximum memory length for l-BFGS update
$B_k$	l-BFGS matrix
$B_k^{(0)}$	Seed matrix in l-BFGS iteration
$S_k$	Structured part of seed matrix in l-BFGS iteration

# 1

## Introduction

Constrained convex optimisation has been a widely-researched area for decades. It is encountered in numerous areas, such as “automatic control systems, estimation and signal processing, communications and networks, electronic circuit design, data analysis and modelling, statistics, and finance” [8], engineering [4, 40], and image processing [12, 14].

In this thesis, smooth inequality-constrained problems will be considered, in particular, problems for which the objective and constraint functions are convex and twice continuously differentiable. The goal is to minimise the objective function subject to all constraint functions being non-negative, that is, to find a solution to

$$\inf_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad g(x) \leq 0, \quad (\mathcal{P})$$

where  $g(x) \leq 0$  means that  $g_i(x) \leq 0$  for  $i = 1, \dots, m$ . It is assumed that the problem  $(\mathcal{P})$  has an actual minimiser.

### 1.1 Background and motivation

Commonly-used approaches to solve these problems are *interior-point methods* (IPMs). In these, the inequality constraints are associated with a barrier function and included directly into the objective, thus transforming the entire constrained problem into an unconstrained one and allowing for the use of unconstrained optimisation techniques. As mentioned in [24, page 588], “the use of a logarithmic barrier method in the context of optimization had already been proposed in 1955 by Frisch [19] and studied extensively by Fiacco and McCormick [16] in the context of nonlinear optimization.”

Back then referred to as *sequential unconstrained minimization technique* [8, page 569], the approach lost most of its popularity during the 1970s, mainly due to the discovery of the increasing ill-conditionedness in the underlying sub-problems and the development of new methods [18, page 528]. It wasn’t until the groundbreaking paper from Karmarkar [28] in 1984 that IPMs regained interest, as he proved polynomial time complexity for such an algorithm in linear programming, in contrast to the combinatorial worst-case complexity of simplex methods, cf. [18, page 526].

In 1994, the complexity theory of IPMs was seminally extended to nonlinear convex optimisation problems by Nesterov and Nemirovski [35], using the convergence theory of Newton’s method for self-concordant functions, cf. [8, page 621]. Today, these methods are usually referred to as *barrier*, *path-following* or *interior-point methods*. They enjoy

great popularity in linear, quadratic, and nonlinear convex optimisation [8, 24, 37] and are included in widely-used solvers such as MOSEK, IPOPT or CVXOPT [3, 37].

Most often, so-called *infeasible primal-dual algorithms* are used, as they are considered the most efficient IPMs [24, page 588]. In these, typically one or multiple Newton steps are applied to the *perturbed KKT system* of the optimisation problem (see Section 2.3), thereby avoiding the two main drawbacks of primal IPMs: the increasing nonlinearity and the increasingly ill-conditioned Hessian of the barrier function [24, 37]. This proved very efficient especially because of the properties of Newton’s method related to self-concordant functions. For many IPMs, polynomial time complexity results were proven [8, 24, 35, 43].

However, these methods generally involve the solution of large linear systems of equations which might become prohibitively expensive for large-scale problems [14, page 956]. Furthermore, Newton’s method requires the knowledge and computation of second derivatives of the objective and all constraint functions which in practice may be unavailable or unsuitably expensive to compute for nonlinear problems [4, 40].

Therefore, in recent years, much attention has been focused on the development of IPMs for large-scale nonlinear optimisation without the need of Hessian matrices [2, 4, 40]. A convenient and efficient alternative to Newton’s method is available in the form of Quasi-Newton methods which internally build and maintain approximations to the Hessian using only previously-computed gradient information [40]. In recent years, a number of Quasi-Newton based primal-dual IPMs was developed [4, 5, 6, 25, 26, 40].

But even though all these methods do not require Hessian matrices, they still involve the solution of large linear systems of equations, limiting their suitability especially when a large number of nonlinear inequality-constraints is present [6].

An alternative which does not require the solution of linear systems is given when a Quasi-Newton method is applied directly to the primal barrier function. However, this is accompanied by a major downfall: Already primal IPMs which employ Newton’s method fell out of favour due to the nonlinearity of the barrier function close to the boundary [37, page 584], as this makes the underlying quadratic approximations less accurate and the overall method numerically less stable than primal-dual methods [4, page 201]. The use of a Quasi-Newton method further worsens this issue, as the latter is known to struggle with ill-conditioned problems [37, page 180], and the Hessian of the barrier function becomes increasingly ill-conditioned in IPMs [37, 41, 49].

Therefore, a method is discussed in this thesis which eliminates the need for solving linear systems while simultaneously reducing the impact of the increasing nonlinearity and ill-conditionedness. The method aims especially at large-scale inequality-constrained convex optimisation problems.

## 1.2 Contributions

In this thesis, an l-BFGS based primal interior-point method is proposed and analysed. It does not rely on the solution of linear systems and reduces the memory cost compared to a BFGS based method. To deal with the issue of increasing nonlinearity and ill-conditioning,



two modifications are made:

- On the one hand, a Tikhonov regularisation is added to the problem in order to improve its quadratic approximability and to reduce the overall ill-conditionedness of the barrier function. Throughout the iteration process, the weight of both regularisation and barrier term is simultaneously reduced, but the barrier parameter is forced to decrease faster than the regularisation parameter. This ensures that the regularisation effect dominates the barrier term from which the ill-conditionedness arises, ultimately allowing for the convergence results derived in this thesis.
- On the other hand, a structured l-BFGS method is employed in which the *seed matrix* is chosen in each l-BFGS iteration in a way that captures the spectrum of the Hessian more properly than a scaled identity matrix, thereby making the method more viable under ill-conditionedness. The structured l-BFGS method used here was very recently discussed and analysed in [34], with a focus on the solution of inverse problems. In this thesis, it is adapted for the application to primal IPMs. The main concept in this is to include the information about the current Hessian in the l-BFGS seed matrix which is already at hand, can be used unproblematically in computations and likely contributes most to the problem's ill-conditioning. The more complicated and possibly unavailable parts of the Hessian are left to be approximated by a scaled identity in the seed matrix.

Besides improving the problem's conditioning, it is proven in this thesis that the Tikhonov regularisation also brings two more benefits:

- Whereas in IPMs it is typically required that the solution set of the underlying problem is bounded, cf. [18, 46, 48], this assumption is not necessary here to prove convergence. Losing this requirement is substantial, as there are many cases in which boundedness is not given. This includes Linear Programs, as already the simple example

$$\min_{x \in \mathbb{R}^n} c^T x \quad \text{s.t.} \quad x \geq 0 \quad \text{with} \quad c = (0, 1, \dots, 1)^T$$

illustrates.

- Furthermore, due to the regularisation, the iterates ultimately converge to the minimal-norm solution of the problem, thereby uniquely characterising the limit point without the need for further assumptions. This stands in contrast to other IPMs which, under appropriate assumptions, typically converge to the analytic centre of the solution set [4, 18], as is further discussed in Section 2.4.

In this thesis, the impact of Tikhonov regularisation on interior-point methods is analysed. An algorithm which integrates this concept is proposed, and a suitable update and stopping strategy for the method is derived. It is proven that the iterates of the method converge to the minimal-norm solution, that the method reaches any given accuracy in the function value after finitely many iterations, and that the number of required l-BFGS iterations grows at most polynomially in the accuracy.

To the best of our knowledge, this is the first time in which a convergence analysis for Tikhonov-regularised IPMs is conducted in such detail. Furthermore, there appear to be no

complexity results on purely Quasi-Newton based IPMs for nonlinear convex optimisation in the literature, indicating that the result derived in this thesis is the first of its kind.

### 1.3 Main results and structure

This thesis is structured as follows. **Chapter 2** covers preliminaries on convex functions and convex optimisation problems (Sections 2.1 and 2.2) as well as an introduction to interior-point methods (Sections 2.3 and 2.4).

In **Chapter 3**, the Tikhonov-regularised interior-point method is presented and analysed in detail. At first, a brief description of Tikhonov regularisation is given in Section 3.1, followed by the introduction of the Tikhonov-regularised IPM in Section 3.2. Its general structure is outlined in Algorithm 3.1. Furthermore, it is shown that each regularised barrier problem has a unique minimiser. The next part of the chapter covers the convergence analysis of the outer iterations: in Section 3.3 for the case that the minimiser of each sub-problem is computed exactly, and in Section 3.4 for the practically more relevant case that the minimisers are only approximated in each iteration.

It is proven that the iterates converge to the minimal-norm solution, and a convergence rate in the objective is established which is linear in the regularisation parameter. Based on that, suitable choices for the parameter update and the inner and outer stopping criteria are discussed. The detailed structure of the overall method is presented in Algorithm 3.2. It is proven in Theorem 3.12 that the method reaches a given accuracy after finitely many outer iterations and that it supports an  $r$ -linear convergence rate in the objective.

The chapter concludes with a brief discussion on possible modifications to the Tikhonov-regularised IPM in Section 3.5, including proximal-point regularisation and the integration of linear equality constraints.

While Chapter 3 focusses on the outer structure of the method, **Chapter 4** covers the *inner solver*: After the concept behind l-BFGS methods is described in Section 4.1, the structured l-BFGS method analysed in [34] is introduced in Section 4.2. Its implementation to solve the inner problems of the Tikhonov-regularised IPM is addressed in detail in Section 4.3. The precise inner procedure is presented in Algorithm 4.2.

In the rest of Chapter 4, the convergence behaviour of the inner solver is analysed: Based on the results from [34], it is proven in Section 4.4 that each inner loop terminates after finitely many iterations, yielding that also the overall method terminates after finitely many iterations. Furthermore, it is shown that each inner iteration procedure supports an  $r$ -linear convergence rate, with the main result presented in Theorem 4.5.

In Section 4.5, the convergence of the inner iterations is quantified in the context of the overall method, culminating in Section 4.6 where it is proven that, under certain assumptions on the algorithm, the total number of l-BFGS iterations required for the entire Tikhonov-regularised IPM grows at most polynomially in the imposed tolerance. This result is stated in Theorem 4.12.

Finally in **Chapter 5**, numerical results are presented: In Section 5.1, some notes on

the implementation are given. Section 5.2 covers a detailed description of the considered test problems. In Section 5.3, the theoretical convergence results of Chapters 3 and 4 are validated numerically based on low-dimensional problems. Lastly, results for large-scale problems are presented and discussed in Section 5.4.

The Tikhonov-regularised IPM discussed in this thesis was implemented in Python, and the code is available on GitHub [22].

## 1.4 Related work

In 2000, a **primal-dual IPM** for nonlinear inequality-constrained convex problems was proposed and discussed in which the Hessian of the problem’s Lagrangian was substituted by BFGS approximations [4]. The authors proved convergence of the method to the analytic centre of the primal-dual optimal set when the Lagrangian is strongly convex and strict complementarity holds. In 2003, they extended their results to an infeasible method [5] and published a limited-memory variant [6].

The method discussed in this thesis differs from their approach as follows: In contrast to them, a primal IPM is discussed here which does not require the solution of linear systems of equations. Furthermore, the Tikhonov regularisation yields convergence to the minimal-norm solution instead of the analytic centre, and the assumption of a strongly convex Lagrangian [4, Assumption 2.1] is not necessary here. Furthermore, in [4, 5], a BFGS method is employed as inner solver for which q-superlinear convergence is proven, in contrast to the l-BFGS method used here which converges only q-linearly. Lastly, a polynomial complexity result is proven for the overall algorithm here.

The software package **IPOPT** [47] is designed for large-scale nonlinear optimisation based on interior-point methods. Amongst others, it supports Quasi-Newton based Hessian approximations, although there appear to exist no complexity results for it. The application of a Quasi-Newton based IPM can for example be found in [11].

More recently, a primal-dual IPM for convex quadratic problems which **combines Newton and Quasi-Newton steps** was proposed and discussed in [25], and the polynomial iteration complexity of a similar method for linear programming was proven in [26]. Again, these primal-dual approaches involve the solution of linear systems. Moreover, in contrast to the method discussed in this thesis, they are limited to quadratic and linear programming, respectively, and not solely based on Quasi-Newton steps. However, they allow for the integration of linear equality constraints.

The paper [40] which was published in 2019 deals with a BFGS based primal-dual IPM **in generic Hilbert spaces**. Thereby, it differs substantially from the scenario considered in this thesis, and furthermore only covers “pointwise inequality constraints” [40]. Also in 2019, the **parallelisation** of Quasi-Newton based primal-dual IPMs for nonlinear optimisation with a small number of general constraints was discussed in [39]. It focusses on the parallelised implementation, block structure and numerical analysis of the method, and also allows for non-convex problems; in contrast to the theoretical focus and results which are given in this thesis.

Regularisation approaches in interior-point methods are not new. In 1998, convergence for a **combination of proximal and interior-point methods** was proven [27], and recent modifications of this idea can for example be found in [13], with an application to image processing in [12]. As is discussed in Section 3.5, proximal-point regularisation has similarities to the Tikhonov regularisation employed in this thesis, but they differ especially in terms of the convergence analysis and do not yield convergence to the minimal-norm solution.

The integration of Tikhonov regularisation into interior-point frameworks is especially found in the context of **linear ill-posed problems** [10, 44]. This problem class, however, is substantially different to the one considered in this thesis.

Recently, the combination of proximal and interior-point methods was further modified with an **Augmented Lagrangian approach**, and different variations of this so-called *interior-point proximal method of multipliers* were proposed and discussed for linear and quadratic convex programming in [7, 41], for semi-definite programming in [42] and for sparse approximations in [14]. In detail, a primal-dual Augmented Lagrangian IPM for large-scale nonlinear convex problems is proposed and discussed in [31]. The integration of an Augmented Lagrangian approach aims at solving convex optimisation problems in the additional presence of linear equality constraints which are not covered by this thesis. All of these methods are based on Newton steps and thereby rely on linear systems, and, in case of nonlinear optimisation, on the knowledge of the involved Hessian matrices.

## 1.5 Notation

Mainly following [37], the order notations  $O(\cdot)$  and  $o(\cdot)$  for an asymptotic upper estimate, and  $\Omega(\cdot)$  for a corresponding lower estimate are used in this thesis as follows. For non-negative sequences  $(a_j)_{j \in \mathbb{N}}$  and  $(b_j)_{j \in \mathbb{N}}$ , it is

$$\begin{aligned} a_j = O(b_j) & : \Leftrightarrow \exists C > 0 : a_j \leq Cb_j \quad \text{for all sufficiently large } j, \\ a_j = o(b_j) & : \Leftrightarrow \lim_{j \rightarrow \infty} \frac{a_j}{b_j} = 0, \\ a_j = \Omega(b_j) & : \Leftrightarrow \exists C > 0 : a_j \geq Cb_j \quad \text{for all sufficiently large } j. \end{aligned}$$

Sometimes, the order notation is also used as dependency in a continuous variable instead of sequences. Then, it is defined analogously to the discrete case.

Again following [37], a sequence  $(x_k)_{k \in \mathbb{N}_0}$  is said to *converge q-linearly in the objective f* to the infimum  $f^*$  if there is a constant  $r \in (0, 1)$  such that

$$f(x_{k+1}) - f^* \leq r(f(x_k) - f^*) \quad \text{for all sufficiently large } k.$$

Furthermore, a sequence  $(x_k)_{k \in \mathbb{N}_0}$  is said to *converge r-linearly in the objective f* to the infimum  $f^*$  if it is dominated by a q-linear sequence, i.e. if there is a non-negative sequence  $(r_k)_{k \in \mathbb{N}_0}$  which converges q-linearly to zero and for which it holds

$$f(x_k) - f^* \leq r_k \quad \text{for all } k.$$

# 2

## Preliminaries on convex optimisation and interior-point methods

In this chapter, the basis for the later proposed and discussed Tikhonov-regularised interior-point method is laid. Starting with preliminaries on convex and strongly convex functions in Section 2.1, Section 2.2 covers the description and related definitions of the setting considered in this thesis: a general inequality-constrained convex optimisation problem. Furthermore, basic results on the existence and character of solutions to this problem are stated which are later in Chapters 3 and 4 needed for the analysis.

As the method discussed in this thesis is built on interior-point methods, their general concept is described in Section 2.3, followed by an outline and discussion on the convergence behaviour of those methods in Section 2.4. This is later in Chapter 3 used to analyse how the results derived for the Tikhonov-regularised method differ from the behaviour of other IPMs.

### 2.1 Convexity and strong convexity

As this thesis deals with convex optimisation, the basic definitions of convexity shall give the starting point. These follow [20, Chapter 3].

**Definition 2.1.** A set  $X \subseteq \mathbb{R}^n$  is called *convex* if for all  $x, y \in X$ ,  $\lambda \in (0, 1)$  it holds

$$\lambda x + (1 - \lambda)y \in X.$$

Thus, a set is convex if and only if it contains the line segment between any points in it. Similarly, convexity of a function  $f$  is defined by the line segments between  $(x, f(x))$  and  $(y, f(y))$  of any points  $x$  and  $y$  lying above the graph of  $f$ , cf. [8].

**Definition 2.2.** A function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is called *convex* if for all  $x, y \in \mathbb{R}^n$ ,  $\lambda \in (0, 1)$  it holds

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

Whenever  $f$  is continuously differentiable, this convexity can be expressed in terms of its gradient or Hessian matrix:

**Lemma 2.3.** [cf. 20, Satz 3.5, Satz 3.7, Satz 3.8] *Let  $f \in C^1(\mathbb{R}^n, \mathbb{R})$ . Then the following statements are equivalent:*

- a)  $f$  is convex.
- b)  $f(x) \geq f(y) + \nabla f(y)^T(x - y) \quad \forall x, y \in \mathbb{R}^n$ .
- c)  $(\nabla f(x) - \nabla f(y))^T(x - y) \geq 0 \quad \forall x, y \in \mathbb{R}^n$ .
- d) *If  $f$  furthermore is twice continuously differentiable:  $h^T(\nabla^2 f(x))h \geq 0 \quad \forall x, h \in \mathbb{R}^n$ .*

Here,  $C^1(\mathbb{R}^n, \mathbb{R})$  denotes the space of continuously differentiable functions from  $\mathbb{R}^n$  to  $\mathbb{R}$ . By statement b), any linear Taylor approximation of a convex function  $f$  lies entirely below the function  $f$ , thus yielding the non-negative curvature described in statement d).

An important concept for Tikhonov regularisation and the Tikhonov-regularised IPM discussed later in this thesis is the so-called *strong convexity* which essentially guarantees that a function does not only have a non-negative curvature, but maintains a certain positive curvature on its entire domain. This property is defined as follows.

**Definition 2.4.** A function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is called  $\kappa$ -strongly convex for  $\kappa > 0$  if for all  $x, y \in \mathbb{R}^n$ ,  $\lambda \in (0, 1)$  it holds

$$f(\lambda x + (1 - \lambda)y) + \kappa\lambda(1 - \lambda)\|x - y\|_2^2 \leq \lambda f(x) + (1 - \lambda)f(y).$$

A function  $f$  is called *strongly convex* if it is  $\kappa$ -strongly convex for some  $\kappa > 0$ .

Similarly to Lemma 2.3, the property of strong convexity can be expressed in terms of the gradient or Hessian matrix for continuously differentiable functions:

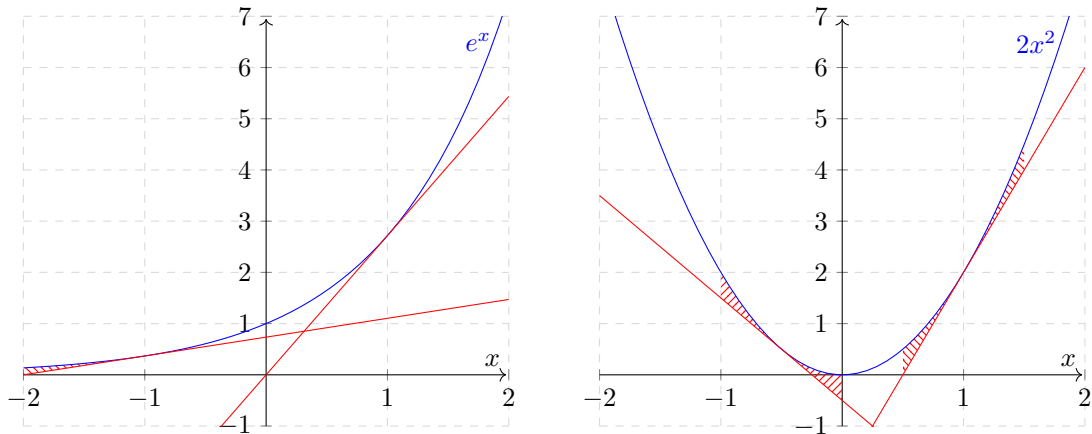
**Lemma 2.5.** [cf. 20, Satz 3.5, Satz 3.7, Satz 3.8] *Let  $f \in C^1(\mathbb{R}^n, \mathbb{R})$ . Then the following statements are equivalent:*

- a)  $f$  is  $\kappa$ -strongly convex for  $\kappa > 0$ .
- b)  $f(x) \geq f(y) + \nabla f(y)^T(x - y) + \kappa\|x - y\|_2^2 \quad \forall x, y \in \mathbb{R}^n$ .
- c)  $(\nabla f(x) - \nabla f(y))^T(x - y) \geq 2\kappa\|x - y\|_2^2 \quad \forall x, y \in \mathbb{R}^n$ .
- d) *If  $f$  furthermore is twice continuously differentiable:*  
 $h^T(\nabla^2 f(x))h \geq 2\kappa\|h\|_2^2 \quad \forall x, h \in \mathbb{R}^n$ .

In contrast to Lemma 2.3, in statement b), the strong convexity now implies that the function  $f$  does not only lie above any linear approximation of itself, but that it further maintains a margin of at least the scaled squared distance from the approximation point. Thus, intuitively speaking, strong convexity guarantees that  $f$  grows at least quadratically. This concept and the difference between a convex and a strongly convex function are visualised in Figure 2.1.

Statement c) of Lemma 2.5 corresponds to the *strong monotonicity* of the gradient of  $f$ , and d) is referred to as *strong positive definiteness* of the Hessian of  $f$ , cf. [20, pages 15 and 17]. The latter ensures that all of its eigenvalues are bounded away from 0; in particular, the minimal eigenvalue is at least  $2\kappa$  [8, Section 9.1.2].

An important property of strongly convex functions is that their level sets are convex and compact.



**Figure 2.1:** Visualisation of a convex and a strongly convex function. Left: The function  $e^x$  is convex and all linear Taylor approximations lie below it, but their distance to  $e^x$  becomes arbitrarily small as the expansion point approaches  $-\infty$ . Right: The function  $2x^2$  is 2-strongly convex and at every point, the distance to its linear Taylor approximation, shaded in red, is bounded below by twice the squared distance to that point.

**Lemma 2.6.** [cf. 20, Lemma 3.9 and page 19] *Let  $f \in C^1(\mathbb{R}^n, \mathbb{R})$  be strongly convex, and let  $x^0 \in \mathbb{R}^n$ . Then the level set*

$$\{x \in \mathbb{R}^n \mid f(x) \leq f(x^0)\}$$

*is compact—that is, closed and bounded—and convex.*

From this property, it follows that strongly convex functions always have a unique minimiser on a closed convex set:

**Theorem 2.7.** [cf. 21, Satz 2.13] *Let  $f \in C^1(\mathbb{R}^n, \mathbb{R})$  be strongly convex, and let  $X \neq \emptyset$  be a closed convex set. Then the optimisation problem*

$$\inf_{x \in X} f(x)$$

*has a unique minimiser.*

This will later be needed in the discussions on regularisation and the regularised IPM.

## 2.2 Nonlinear inequality-constrained convex programs

In the context of this thesis, general inequality-constrained convex optimisation problems are considered where the objective and constraint functions are all convex and twice continuously differentiable. Therefore, let

- $f \in C^2(\mathbb{R}^n, \mathbb{R})$  be convex, and
- $g \in C^2(\mathbb{R}^n, \mathbb{R}^m)$ , where all  $g_1, \dots, g_m$  are convex.

Here,  $C^2(\mathbb{R}^n, \mathbb{R}^m)$  denotes the space of twice continuously differentiable functions from  $\mathbb{R}^n$  to  $\mathbb{R}^m$ .

Throughout the further discussions, the following sets related to the convex optimisation problem  $(\mathcal{P})$ , defined in the introduction, will be needed.

**Definition 2.8.** Let

- $\mathcal{F} := \{x \in \mathbb{R}^n \mid g_i(x) \leq 0, i = 1, \dots, m\}$  denote the *set of feasible points*,
- $\mathcal{F}_+ := \{x \in \mathbb{R}^n \mid g_i(x) < 0, i = 1, \dots, m\}$  denote the *set of strictly feasible points*,  
and
- $\mathcal{S} := \arg \min_{x \in \mathcal{F}} f(x)$  denote the *set of solutions*

of the optimisation problem  $(\mathcal{P})$ .

*Remark 2.9.* The set of feasible and strictly feasible points  $\mathcal{F}$  and  $\mathcal{F}_+$  as well as the solution set  $\mathcal{S}$  are all convex, cf. [21, Lemma 2.14, Satz 2.13] and [18, page 545]. Furthermore,  $\mathcal{F}$  as intersection over the closed level sets  $\{x \in \mathbb{R}^n \mid g_i(x) \leq 0\}$  is closed. Thereby, also  $\mathcal{S}$  is closed as it consists of the intersection of  $\mathcal{F}$  and the closed inverse image of  $f$  under the infimum of  $(\mathcal{P})$ .

The Lagrangian  $L : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$  of problem  $(\mathcal{P})$  is given by [37, page 320]

$$L(x, \lambda) := f(x) + \lambda^T g(x).$$

The further discussions in this thesis are mainly based on two basic assumptions: Firstly, the so-called *Slater Constraint Qualification condition* [21, Definition 2.44] is required to hold which enforces the strictly feasible set  $\mathcal{F}_+$  to be non-empty. While this clearly does not generally hold for any convex optimisation problem, it is essential for using feasible interior-point methods, as the latter operate solely within  $\mathcal{F}_+$ . The second assumption postulates the existence of a minimiser of problem  $(\mathcal{P})$ , meaning that the infimum is actually attained. As the optimisation aims at finding such a minimiser, this represents a reasonable restriction in the context of this thesis.

**Assumption 2.10.**

- a) The Slater Constraint Qualification (Slater CQ) condition holds, i.e.  $\mathcal{F}_+ \neq \emptyset$ . Thus, there exists  $x^{\text{Slat}} \in \mathbb{R}^n$  with

$$g_i(x^{\text{Slat}}) < 0 \quad \forall i = 1, \dots, m.$$

- b) The optimisation problem  $(\mathcal{P})$  has a global minimiser, i.e.  $\mathcal{S} \neq \emptyset$ .

*Remark.* If the Slater condition holds, then the strictly feasible set  $\mathcal{F}_+$  corresponds to the *interior* of  $\mathcal{F}$ , and  $\mathcal{F}$  is precisely the closure of  $\mathcal{F}_+$  [30, Lemma C.4.2].

Although these assumptions are the basis for most results in this thesis, at any time, it will be clearly stated when and which parts are required for the theorem or lemma.

The Slater condition is not only a key requirement for feasible IPMs, but it is also a constraint qualification which ensures that any minimiser of problem  $(\mathcal{P})$  satisfies the first-order necessary optimality conditions.



**Theorem 2.11.** [cf. 21, Satz 2.45] *Let  $x^*$  be a local minimiser of the convex minimisation problem  $(\mathcal{P})$ , and let the Slater CQ condition hold. Then there exists a multiplier  $\lambda^* \in \mathbb{R}^m$  such that the following KKT conditions hold:*

$$\nabla f(x^*) + \sum_{i=1}^m \lambda_i^* \nabla g_i(x^*) = 0, \quad (2.1)$$

$$\lambda_i^* g_i(x^*) = 0 \quad \text{for } i = 1, \dots, m, \quad (2.2)$$

$$g_i(x^*) \leq 0 \quad \text{for } i = 1, \dots, m, \quad (2.3)$$

$$\lambda_i^* \geq 0 \quad \text{for } i = 1, \dots, m. \quad (2.4)$$

Any tuple  $(x^*, \lambda^*) \in \mathbb{R}^n \times \mathbb{R}^m$  satisfying the conditions (2.1)–(2.4) is called KKT point of  $(\mathcal{P})$ .

An essential feature of convex optimisation is that above KKT conditions are not only necessary, but also sufficient for a point  $x^* \in \mathbb{R}^n$  to be a minimiser of problem  $(\mathcal{P})$ , and furthermore, that any local minimiser is also a global minimiser.

**Theorem 2.12.** [cf. 21, Satz 2.46] *Let  $(x^*, \lambda^*) \in \mathbb{R}^n \times \mathbb{R}^m$  be a KKT point of  $(\mathcal{P})$ . Then  $x^*$  is a global minimiser of problem  $(\mathcal{P})$ .*

Consequently, every KKT point makes for a global minimiser, and any minimiser yields a KKT point. This, however, does neither imply the existence nor the uniqueness of such a minimiser.

In case there exists a minimiser, however, an interesting question for methods that solve  $(\mathcal{P})$  is the nature of the point which is obtained with the method. As will be shown later in Section 3.3, the Tikhonov-regularised IPM discussed in this thesis converges to the minimiser with minimal Euclidean norm—which, in fact, is unique:

**Lemma 2.13.** *Let  $\mathcal{S} \neq \emptyset$ . Then problem  $(\mathcal{P})$  has a unique minimiser  $x_M^*$  with minimal Euclidean norm, i.e.*

$$x_M^* = \arg \min_{x^* \in \mathcal{S}} \|x^*\|_2. \quad (2.5)$$

*Proof.* As the solution set  $\mathcal{S}$  is convex and closed by Remark 2.9, there exists a unique orthogonal projection of 0 onto  $\mathcal{S}$ , see e.g. [21, Lemma 2.17]. By definition,  $x_M^*$  is precisely this projection, and the statement follows.  $\square$

## 2.3 Interior-point methods

As the method discussed in this thesis is based on an interior-point method, an introduction to these is given next.

### 2.3.1 General concept

One way to tackle the general inequality-constrained convex problem  $(\mathcal{P})$  is to transform it into an unconstrained problem by adding a term to the objective function that intrinsically enforces the constraints. In so-called *barrier methods*, a barrier term is employed which ensures that the constraints remain fulfilled all the time.

In this thesis, a barrier method with a logarithmic barrier will be considered. For this, the general convex inequality constraints  $g_1(x), \dots, g_m(x) \leq 0$  are associated with the logarithmic barrier term  $b : \mathcal{F}_+ \rightarrow \mathbb{R}$  with  $b(x) := -\sum_{i=1}^m \ln(-g_i(x))$ . Clearly, this term can only be defined on the strictly feasible set  $\mathcal{F}_+$  (or is else set to infinity) since the logarithm function is only defined for positive values.

In general, the barrier  $b$  penalises a point  $x$ , the closer it is to some boundary  $g_i(x) = 0$ , as is depicted in Figure 2.2.

The *barrier function*  $\varphi_\mu : \mathcal{F}_+ \rightarrow \mathbb{R}$  of problem  $(\mathcal{P})$  with *barrier parameter*  $\mu > 0$  is defined as

$$\begin{aligned} \varphi_\mu(x) &:= f(x) + \mu b(x) \\ &= f(x) - \mu \sum_{i=1}^m \ln(-g_i(x)). \end{aligned}$$

The function  $\varphi_\mu$  is convex for  $\mu > 0$ , as  $f$  and all  $g_1, \dots, g_m$  are convex and the function  $t \mapsto -\ln(-t)$  is monotonically increasing and convex on  $(-\infty, 0)$  [48, Theorem 5 i)]. Now, the *barrier problem*

$$\inf_{x \in \mathbb{R}^n} \varphi_\mu(x) \quad \text{s.t.} \quad g(x) < 0 \quad (\mathcal{P}_\mu)$$

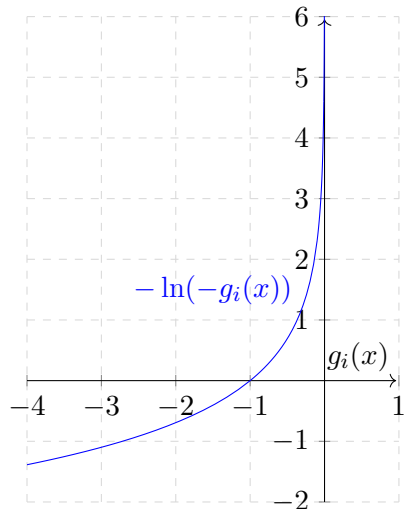
arises as a modification to the original problem  $(\mathcal{P})$  whose solutions in some sense approximate a solution to  $(\mathcal{P})$ , as will be discussed later in Section 2.4. Since the logarithmic barrier implicitly enforces the inequality constraint  $g(x) < 0$ , it may be omitted here.

### 2.3.2 Existence of barrier minimisers

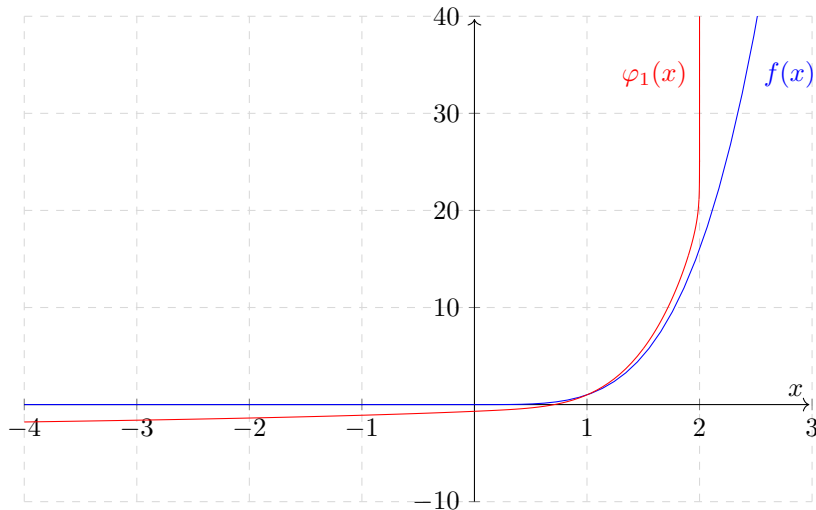
In this context, it becomes clear why requiring both parts from Assumption 2.10 is reasonable, namely that  $\mathcal{F}_+$  as well as the solution set  $\mathcal{S}$  both are non-empty. Since this means that the domain of  $\varphi_\mu$  is non-empty and problem  $(\mathcal{P})$  has a minimiser, one could expect that also the barrier problem  $(\mathcal{P}_\mu)$  has a minimiser for any  $\mu > 0$ . That this is not the case in general, however, illustrates the following example.

**Example 2.14.** Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  and  $g : \mathbb{R} \rightarrow \mathbb{R}$  with

$$f(x) := \begin{cases} x^4 & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases}, \quad g(x) := x - 2.$$



**Figure 2.2:** Visualisation of the logarithmic barrier function for an inequality constraint  $g_i(x) \leq 0$ . The closer  $g_i(x)$  is to 0, the greater is the penalisation.



**Figure 2.3:** Visualisation of the objective function  $f$  (blue) and barrier function  $\varphi_\mu$  (red) for the problem described in Example 2.14 with barrier parameter  $\mu = 1$ . Although the original problem  $(\mathcal{P})$  has a non-empty solution set  $\mathcal{S} = (-\infty, 0]$ , the barrier function has no minimiser for any  $\mu > 0$ .

The function  $f$  is convex and twice continuously differentiable with

$$\nabla f(x) := \begin{cases} 4x^3 & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases}, \quad \nabla^2 f(x) := \begin{cases} 12x^2 & \text{for } x \geq 0 \\ 0 & \text{for } x < 0 \end{cases},$$

and  $g$  as linear function is also convex and twice continuously differentiable.

As  $g(x) \leq 0$  means that  $x \leq 2$ , the optimisation problem

$$\inf_{x \in \mathbb{R}} f(x) \quad \text{s.t.} \quad g(x) \leq 0$$

reaches its minimal value 0 at any  $x \leq 0$ , so its solution set is  $\mathcal{S} = (-\infty, 0]$ .

For any  $\mu > 0$ , the corresponding barrier function is given on the strictly feasible set  $\mathcal{F}_+ = (-\infty, 2)$  by

$$\varphi_\mu(x) = f(x) - \mu \ln(-g(x)) = \begin{cases} x^4 - \mu \ln(2 - x) & \text{for } 0 \leq x < 2 \\ -\mu \ln(2 - x) & \text{for } x < 0 \end{cases},$$

and it holds  $\lim_{x \rightarrow -\infty} \varphi_\mu(x) = \lim_{x \rightarrow -\infty} -\mu \ln(2 - x) = -\infty$ .

Therefore, even though Assumption 2.10 holds, that is, the solution set of the underlying problem  $(\mathcal{P})$  is non-empty and  $\mathcal{F}_+ \neq \emptyset$ , the barrier problem  $\inf_{x \in \mathcal{F}_+} \varphi_\mu(x)$  is unbounded and has no minimiser.

The objective and barrier function are visualised in Figure 2.3.

To overcome the issue of non-existing minimisers of the barrier problem, in theoretical results found in the literature, it is typically assumed that the solution set  $\mathcal{S}$  is not only non-empty, but also bounded [18, 46, 48]. From this, it follows that also the solution set of

problem  $(\mathcal{P}_\mu)$  is non-empty and bounded (see Theorem 2.15 a)). A common way to meet this requirement in convex optimisation is to assume that at least one of the functions  $f, g_1, \dots, g_m$  is strongly convex [4]; if necessary, a fictive constraint could also be added for that purpose. However, an advantage of the Tikhonov-regularised method proposed in this thesis is that such a restriction is not necessary, as is later shown in Theorem 3.4.

### 2.3.3 Computation of barrier minimisers

To compute a minimiser of problem  $(\mathcal{P}_\mu)$ , typically, the first-order optimality conditions are used. For this, gradient and Hessian of  $\varphi_\mu$  are given in  $x \in \mathcal{F}_+$  by

$$\begin{aligned}\nabla\varphi_\mu(x) &= \nabla f(x) + \mu\nabla b(x) \\ &= \nabla f(x) - \mu \sum_{i=1}^m \frac{1}{g_i(x)} \nabla g_i(x)\end{aligned}$$

and

$$\begin{aligned}\nabla^2\varphi_\mu(x) &= \nabla^2 f(x) + \mu\nabla^2 b(x) \\ &= \nabla^2 f(x) + \mu \sum_{i=1}^m \left( \frac{1}{g_i(x)^2} \nabla g_i(x) \nabla g_i(x)^T - \frac{1}{g_i(x)} \nabla^2 g_i(x) \right),\end{aligned}$$

respectively.

The necessary optimality condition [37, Theorem 2.2] for a minimiser  $\bar{x}_\mu$  of the unconstrained problem  $(\mathcal{P}_\mu)$  is then given by

$$\nabla\varphi_\mu(\bar{x}_\mu) = 0. \tag{2.6}$$

By defining  $(\bar{\lambda}_\mu)_i := -\frac{\mu}{g_i(\bar{x}_\mu)} > 0$ , this corresponds to solving the system

$$\begin{aligned}\nabla f(\bar{x}_\mu) + \sum_{i=1}^m (\bar{\lambda}_\mu)_i \nabla g_i(\bar{x}_\mu) &= 0, \\ (\bar{\lambda}_\mu)_i &= -\frac{\mu}{g_i(\bar{x}_\mu)} \quad \text{for } i = 1, \dots, m,\end{aligned}$$

which, on the other hand, is equivalent to solving

$$\nabla f(\bar{x}_\mu) + \sum_{i=1}^m (\bar{\lambda}_\mu)_i \nabla g_i(\bar{x}_\mu) = 0, \tag{2.7}$$

$$(\bar{\lambda}_\mu)_i g_i(\bar{x}_\mu) = -\mu \quad \text{for } i = 1, \dots, m. \tag{2.8}$$

The equations (2.7)–(2.8) are precisely the KKT conditions of  $(\mathcal{P})$  stated in Theorem 2.11, where the complementarity conditions  $(\bar{\lambda}_\mu)_i g_i(\bar{x}_\mu) = 0$  are perturbed by  $-\mu$ , and the, now implicitly enforced, conditions (2.3) and (2.4) are omitted.

Therefore, computing a minimiser  $\bar{x}_\mu$  of problem  $(\mathcal{P}_\mu)$  by solving (2.7)–(2.8) can be interpreted as approximating a solution  $x^*$  of the original problem  $(\mathcal{P})$  while relaxing the complementarity condition [25, page 94].

### 2.3.4 Generic interior-point method

A common strategy to solve problem  $(\mathcal{P})$  is to alternately approximate a solution  $\bar{x}_{\mu_j}$  to the corresponding barrier problem  $(\mathcal{P}_{\mu_j})$  and decrease the barrier parameter  $\mu_{j+1} < \mu_j$ , starting with an initial barrier parameter  $\mu_1 > 0$ . Since this strategy implies starting with a strictly feasible guess  $x^0 \in \mathcal{F}_+$  and approaching the solution of  $(\mathcal{P})$  “from the inside”, these barrier methods are also referred to as *interior-point methods* [4, 8, 37].

This concept is illustrated in Figure 2.5 where the resultant barrier functions and their minimisers are shown for a simple exemplary problem. Following [48, page 354], the basic structure of a typical IPM is given as follows Algorithm 2.1.

---

#### Algorithm 2.1 Generic interior-point method

---

**Given:** Optimisation problem  $(\mathcal{P})$

**Choose:**  $x^0 \in \mathcal{F}_+$ ,  $\mu_1 > 0$

- 1: **for**  $j = 1, 2, 3, \dots$  **do**
- 2:     **if**  $x^{j-1}$  satisfies stopping criteria **then** let  $x^* \leftarrow x^{j-1}$  and **break**
- 3:     Compute  $x^j$  as approximation to  $\bar{x}_{\mu_j}$  ▷ Typically initialised with  $x^{j-1}$
- 4:     Update  $0 < \mu_{j+1} < \mu_j$
- 5: **end for**

**Return:**  $x^*$

---

This general scheme raises a number of questions:

- It requires a feasible starting point  $x^0 \in \mathcal{F}_+$  which, depending on the nature of the constraints, might be hard to find. To this extent, so-called *Phase I methods* can be employed in which a strictly feasible point is computed, see e.g. [8, Section 11.4]. However, this topic will not be covered in this thesis.
- How should the barrier parameter  $\mu$  be initialised, i.e.  $\mu_1$  be chosen, and how should it be updated during the iteration (line 4)?
- How is the approximation to  $\bar{x}_{\mu_j}$  (line 3) computed?

The computation in line 3 is usually referred to as *inner step*, and the method employed for this is called *inner method*, *inner solver* or, when an iterative method is used, *inner iteration*. Typical IPMs employ one or multiple Newton steps for that, often as so-called primal-dual methods on the perturbed KKT system (2.7)–(2.8), starting at the latest outer iterate  $x^{j-1}$  [8, Section 11.7].

Since this might become unsuitable for large-scale problems, however, in this thesis, an l-BFGS method is used to solve the inner problems. This inner solver is described and analysed in detail in Chapter 4.

## 2.4 Convergence of the barrier minimisers

If the barrier problems  $(\mathcal{P}_\mu)$ , defined in the last section, are solved exactly in each iteration (line 3 in Algorithm 2.1), then the barrier approach supports many results on the convergence of their solutions  $\bar{x}_\mu$  as well as on the so-called *barrier trajectory* as  $\mu \rightarrow 0$ .

In this section, some basic theoretical results on that will be presented that are, with some variations, often found in the literature.

A key requirement for typical interior-point methods is the actual existence of a minimiser  $\bar{x}_\mu$  of each barrier problem  $(\mathcal{P}_\mu)$ . As already mentioned in Section 2.3, this is typically ensured by assuming that the solution set  $\mathcal{S}$  of the original problem  $(\mathcal{P})$  is bounded [18, 46, 48]. For the Tikhonov-regularised approach later discussed in Chapter 3, this requirement will not be necessary; but in the discussion on the convergence of general IPMs carried out in this section, the boundedness of  $\mathcal{S}$  must be assumed.

The following well-known theorem (see e.g. [46, 48]) shows the existence of *barrier minimisers*  $\bar{x}_\mu$  as well as their limit behaviour as  $\mu \rightarrow 0$ , and thereby forms the theoretical foundation for the functionality of general IPMs designed as in Algorithm 2.1.

**Theorem 2.15.** [cf. 48, Theorem 5] *Let Assumption 2.10 hold and let  $\mathcal{S}$  be bounded. Let  $(\mu_j)_{j \in \mathbb{N}}$  be a monotonically decreasing sequence of barrier parameters with  $\mu_j \rightarrow 0^+$ . Then it holds:*

- a) *The set of minimisers of  $(\mathcal{P}_{\mu_j})$  is non-empty, convex and compact for any  $j \in \mathbb{N}$ .*
- b) *Any sequence  $\{\bar{x}_{\mu_j}\}$  of minimisers of  $(\mathcal{P}_{\mu_j})$  has at least one convergent subsequence.*
- c) *For such a subsequence,  $\lim_{j \rightarrow \infty} \varphi_{\mu_j}(\bar{x}_{\mu_j}) = \lim_{j \rightarrow \infty} f(\bar{x}_{\mu_j}) = \min_{x \in \mathcal{F}} f(x)$ .*

It should be noted that this theorem does not state that every sequence of unconstrained minimisers necessarily converges, and in fact, this is not the case in general [48, page 365]. However, as  $\mu \rightarrow 0$ , the function values of  $f$  evaluated at any unconstrained minimiser  $\bar{x}_\mu$  converge to the minimal function value of problem  $(\mathcal{P})$ .

### 2.4.1 Rate of convergence

In fact, this convergence in  $f$  can be characterised by a rate in  $O(\mu)$ , and the actual difference is even bounded by  $m\mu$  [48, page 366]. As parts of the convergence analysis in Chapter 3 are conducted similarly, a proof for the result is presented next.

To show this, a rather remarkable property of the barrier term's gradient is needed which will also be used for discussions later in this thesis. It ensures that, on the entire feasible set, any directional derivative of  $b$ , that is, the inner product between  $\nabla b(x)$  and the vector pointing from  $x$  to anywhere else in the feasible set, is bounded above by the number of constraints  $m$ .

In general, this property can be deduced from the concept of *self-concordance*, see e.g. [30, Lemma 2.3.16], which are of special interest in the context of IPMs combined with Newton's method, but will not be further discussed in this thesis. For the logarithmic barrier term, however, it is straightforward to prove the statement without self-concordance, as is presented next.

**Lemma 2.16.** [cf. 30, Lemma 2.3.16] *Let  $g_1, \dots, g_m$  be convex functions. For the logarithmic barrier term  $b(x) := -\sum_{i=1}^m \ln(-g_i(x))$ , it holds:*

$$\nabla b(x)^T(y - x) \leq m \quad \forall x \in \mathcal{F}_+ \quad \forall y \in \mathcal{F}.$$

*Proof.* Let  $x \in \mathcal{F}_+$  and  $y \in \mathcal{F}$ . The gradient of the logarithmic barrier term  $b$  in  $x$  is given by

$$\nabla b(x) = - \sum_{i=1}^m \frac{1}{g_i(x)} \nabla g_i(x).$$

For each  $i = 1, \dots, m$  it holds by definition of  $\mathcal{F}$  and the convexity of  $g_i$  that

$$0 \stackrel{y \in \mathcal{F}}{\geq} g_i(y) \stackrel{\text{Lemma 2.3 b)}}{\geq} g_i(x) + \nabla g_i(x)^T (y - x).$$

Thus,

$$\nabla g_i(x)^T (y - x) \leq -g_i(x),$$

and since  $x \in \mathcal{F}_+$  and thereby  $-g_i(x) > 0$  one obtains

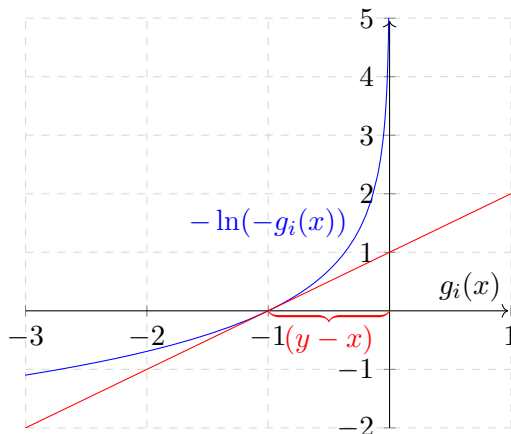
$$-\frac{1}{g_i(x)} \nabla g_i(x)^T (y - x) \leq 1.$$

Taking the sum over  $i = 1, \dots, m$  finally gives

$$\nabla b(x)^T (y - x) = \left( - \sum_{i=1}^m \frac{1}{g_i(x)} \nabla g_i(x) \right)^T (y - x) \leq m. \quad \square$$

At first glance, this property might seem surprising, especially since  $b$  as well as its gradient go to infinity when approaching the boundary of a constraint. But note that by Lemma 2.16, only the “non-normalised” directional derivatives are bounded above. The greater the increase in the barrier in one direction is, the closer it must be to a corresponding boundary in this direction. This limits the maximal difference  $(y - x)$  in that direction as  $y$  must still lie within the feasible set  $\mathcal{F}$ . Conversely, whenever a greater distance between two points  $x$  and  $y$  is given, then at point  $x$ , the barrier term is either slowly or not at all increasing in the corresponding direction  $y - x$ . This idea is also illustrated in Figure 2.4

Utilising Lemma 2.16, the aforementioned upper bound on the deviation in the objective function value  $f$  at a barrier minimiser  $\bar{x}_\mu$  from its minimal value can now be derived.



**Figure 2.4:** Visualisation of Lemma 2.16: The directional derivative of the barrier term at a strictly feasible point  $x$  towards a feasible point  $y$  is bounded by 1. Here, this corresponds to the product of the slope at  $x$  and the distance  $y - x$ . When the slope is large, then  $x$  is already close to the boundary, limiting the maximally possible distance  $y - x$ . Conversely, when the distance is large and  $x$  thereby lies further in the interior, then the slope of the barrier function is small.

**Lemma 2.17.** *Let the solution set of  $(\mathcal{P})$  be non-empty, i.e.  $\mathcal{S} \neq \emptyset$ . Let  $x^* \in \mathcal{S}$  and  $\mu > 0$ . For any minimiser  $\bar{x}_\mu$  of the barrier problem  $(\mathcal{P}_\mu)$ , it holds:*

$$0 \leq f(\bar{x}_\mu) - f(x^*) \leq m\mu.$$

*Proof.* The first inequality follows directly from  $x^*$  being a minimiser of the original problem  $(\mathcal{P})$  and  $\bar{x}_\mu$  also being a feasible point. For the second inequality, first note that from the optimality of  $\bar{x}_\mu$  it follows that  $0 = \nabla\varphi_\mu(\bar{x}_\mu)$  (see (2.6)) and thus

$$\nabla f(\bar{x}_\mu) = -\mu\nabla b(x).$$

By the convexity of  $f$  and Lemma 2.16, it now holds:

$$f(\bar{x}_\mu) - f(x^*) \stackrel{\text{Lemma 2.3 b)}}{\leq} -\nabla f(\bar{x}_\mu)^T(x^* - \bar{x}_\mu) = \mu\nabla b(x)^T(x^* - \bar{x}_\mu) \stackrel{\text{Lemma 2.16}}{\leq} m\mu. \quad \square$$

This result is remarkably strong, and it is well-known for interior-point methods [8, 48], as it does not only give a bound on the deviation of  $f(\bar{x}_\mu)$  from its constrained minimum, but also restricts the worst rate at which the function values of  $f$  converge to its minimum along the barrier minimisers to  $O(\mu)$ . In fact, the upper bound of Lemma 2.17 is sharp, as the following example illustrates.

**Example 2.18.** Consider the linear problem

$$\inf_{x \in \mathbb{R}} f(x) \quad \text{s.t.} \quad x \geq 0$$

with  $f(x) := ax$  for  $a > 0$ .

Clearly, its minimiser is  $x^* = 0$ . The corresponding barrier function

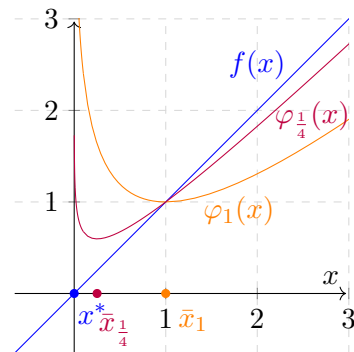
$$\varphi_\mu(x) = ax - \mu \ln x$$

becomes minimal at  $\bar{x}_\mu = \frac{\mu}{a}$ . Thus,

$$f(\bar{x}_\mu) - f(x^*) = f\left(\frac{\mu}{a}\right) - f(0) = \mu,$$

so equality holds in the upper estimate of Lemma 2.17.

The objective and barrier function as well as the minimisers  $x^*$  and  $\bar{x}_\mu$  are shown in Figure 2.5 for  $a = 1$  and  $\mu \in \{1, \frac{1}{4}\}$ .



**Figure 2.5:** Illustration of Example 2.18 for  $\mu = 1$  (orange) and  $\mu = \frac{1}{4}$  (purple) with  $a = 1$ . The difference in the objective  $f(x^*) - f(\bar{x}_\mu)$  is precisely given by  $\mu$ .

For rather “well-behaving” functions that satisfy further assumptions on the character of the optimisation problem  $(\mathcal{P})$ , even stronger results can be obtained: If the *Linear Independent constraint qualification* [37, Definition 12.4] or the *Mangasarian-Fromovitz constraint qualification* [37, Definition 12.6] are satisfied, the Lagrangian  $L$  is strongly convex with respect to  $x$ , and a KKT point  $(x^*, \lambda^*)$  is given, then  $x^*$  is the unique minimiser of problem  $(\mathcal{P})$ , at least a subsequence of barrier minimisers  $(\bar{x}_{\mu_j})$  converges to  $x^*$ , and simultaneously, the parameters  $(\bar{\lambda}_{\mu_j})_i := \frac{-\mu_j}{g_i(\bar{x}_{\mu_j})}$  converge to the Lagrange multiplier  $\lambda_i^*$  [18, Theorem 3.12].



If furthermore strict complementarity holds, that is, for every  $i = 1, \dots, m$  it is either  $\lambda_i^* > 0$  or  $g_i(x^*) < 0$ , then the rate of convergence in the minimisers  $\|\bar{x}_{\mu_j} - x^*\|_2$  is of order  $O(\mu_j)$ . Moreover, then a unique, continuously differentiable function  $\bar{x}(\mu)$  of unconstrained minimisers of  $\varphi_\mu$ , the so-called *barrier trajectory*, exists for sufficiently small  $\mu > 0$ , it converges to  $x^*$  as  $\mu \rightarrow 0$ , and it approaches  $x^*$  non-tangentially, cf. [18, Theorem 3.12, Lemma 3.13, page 558]. On the contrary, if strict complementarity does not hold, then the barrier trajectory might become tangential to the strongly active constraints, and a convergence rate of  $\|\bar{x}_{\mu_j} - x^*\|_2 = O(\sqrt{\mu_j})$  might be obtained [18, page 561].

### 2.4.2 Convergence to the analytic centre

In case the above assumptions are not met and there is no unique minimiser, the question arises whether and to which point the interior-point method ultimately converges. The discussion in this section will therefore be ended with a result that characterises the limit point of the barrier trajectory in case of a strongly convex Lagrangian and a strict complementarity condition.

The following theorem is a modification of [4, Theorem 5.3], and it shows that this limit point is precisely the *analytic centre* of the solution set  $\mathcal{S}$ . This is especially interesting in the context of this thesis as, in contrast, the later considered Tikhonov-regularised IPM converges to the minimiser with minimal Euclidean norm  $x_M^*$  (see Section 3.3). An illustration of these different convergence behaviours is later given in Example 3.7.

**Theorem 2.19.** *Let Assumption 2.10 hold. Further assume that at least one of the functions  $f, g_1, \dots, g_m$  is strongly convex, and that for every  $i = 1, \dots, m$  there is a KKT point  $(x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^m$  of  $(\mathcal{P})$  such that either  $g_i(x) < 0$  or  $\lambda_i > 0$  holds.*

*Then the tuple  $(\bar{x}_\mu, \bar{\lambda}_\mu)$  with  $\bar{x}_\mu$  minimising  $(\mathcal{P}_\mu)$  and  $(\bar{\lambda}_\mu)_i := -\frac{\mu}{g_i(\bar{x}_\mu)}$  converges to the unique analytic centre  $(x^*, \lambda^*)$  of the solution set  $\mathcal{S}$  as  $\mu \rightarrow 0^+$ , that is,  $x^*$  is either the only solution of  $(\mathcal{P})$ , or*

$$x^* = \arg \max_{\substack{x \in \mathcal{S} \\ g_B(x) < 0}} \sum_{i \in B} \ln(-g_i(x)), \quad (2.9)$$

where  $B := \{i \in \{1, \dots, m\} \mid \exists x \in \mathcal{S} \text{ such that } g_i(x) < 0\}$ ,

$\lambda^*$  analogously.

*Proof.* The proof mainly follows the idea of [4, pages 216–218].

First note that from Assumption 2.10 and the strong convexity assumption it follows that  $\mathcal{S}$  is bounded [4, Lemma 2.2], so Theorem 2.15 applies and all barrier minimisers  $\bar{x}_\mu$  exist. Furthermore, the assumptions imply that  $\varphi_\mu$  is strictly convex and thereby  $\bar{x}_\mu$  is uniquely given for all  $\mu > 0$  [4, Lemma 2.4].

Now, if  $\mathcal{S}$  contains only one element, then this element must be  $x^*$  since  $\bar{x}_\mu$  converges to a solution of  $(\mathcal{P})$ . Otherwise,  $f$  cannot be strongly convex, so at least one of the constraint functions, say  $g_{i_0}$ , is strongly convex and  $g_{i_0}(\hat{x}) < 0$  for some  $\hat{x} \in \mathcal{S}$ . Thus,  $B \neq \emptyset$ , and the objective of (2.9) is strongly concave and has a unique solution (for details, see [4, Lemma 5.2]).

Next, let  $\mathcal{M}$  denote the set of KKT points of  $(\mathcal{P})$ , verify that

$$B = \{i \in \{1, \dots, m\} \mid \exists (x, \lambda) \in \mathcal{M} \text{ such that } g_i(x) < 0\},$$

and correspondingly define the set of indices for which a strict dual point exists as

$$N := \{i \in \{1, \dots, m\} \mid \exists(x, \lambda) \in \mathcal{M} \text{ such that } \lambda_i > 0\}.$$

Let  $(\hat{x}, \hat{\lambda}) \in \mathcal{M}$  be an arbitrary KKT point of  $(\mathcal{P})$ . Since  $\hat{x}$  minimises  $L(\cdot, \hat{\lambda})$ , it follows:

$$f(\bar{x}_\mu) + \hat{\lambda}^T g(\bar{x}_\mu) = L(\bar{x}_\mu, \hat{\lambda}) \geq L(\hat{x}, \hat{\lambda}) \stackrel{(2.2)}{=} f(\hat{x}).$$

Furthermore, the convexity of  $L$  in the variable  $x$  and the fact that  $\nabla_x L(\bar{x}_\mu, \bar{\lambda}_\mu) = 0$  by (2.6) yield:

$$\begin{aligned} f(\hat{x}) + \bar{\lambda}_\mu^T g(\hat{x}) &= L(\hat{x}, \bar{\lambda}_\mu) \stackrel{\text{Lemma 2.3 b)}}{\geq} L(\bar{x}_\mu, \bar{\lambda}_\mu) + \underbrace{\nabla_x L(\bar{x}_\mu, \bar{\lambda}_\mu)^T}_{=0} (\hat{x} - \bar{x}_\mu) \\ &= f(\bar{x}_\mu) + \underbrace{\bar{\lambda}_\mu^T g(\bar{x}_\mu)}_{=-m\mu} = f(\bar{x}_\mu) - m\mu. \end{aligned}$$

By adding both equations, subtracting  $f(\bar{x}_\mu) + f(\hat{x})$  on both sides, and using that  $\hat{\lambda}_i = 0$  for  $i \notin N$  and  $g_i(\hat{x}) = 0$  for  $i \notin B$ , one obtains

$$\hat{\lambda}_N^T g_N(\bar{x}_\mu) + (\bar{\lambda}_\mu)_B^T g_B(\hat{x}) = \hat{\lambda}^T g(\bar{x}_\mu) + \bar{\lambda}_\mu^T g(\hat{x}) \geq -m\mu.$$

Using  $g_i(\bar{x}_\mu) = -\frac{\mu}{(\bar{\lambda}_\mu)_i}$  for  $i \in N$ ,  $(\bar{\lambda}_\mu)_i = -\frac{\mu}{g_i(\bar{x}_\mu)}$  for  $i \in B$  and dividing by  $-\mu$  yields

$$\sum_{i \in N} \frac{\hat{\lambda}_i}{(\bar{\lambda}_\mu)_i} + \sum_{i \in B} \frac{g_i(\hat{x})}{g_i(\bar{x}_\mu)} \leq m.$$

Now suppose that  $(x^*, \lambda^*)$  is a limit point of  $((\bar{x}_\mu, \bar{\lambda}_\mu))_{\mu > 0}$  for  $\mu \rightarrow 0^+$ . Taking the limit in above estimate gives

$$\sum_{i \in N} \frac{\hat{\lambda}_i}{\lambda_i^*} + \sum_{i \in B} \frac{g_i(\hat{x})}{g_i(x^*)} \leq m,$$

which directly implies that  $g_B(x^*) < 0$  and  $\lambda_N^* > 0$  (otherwise, above limit would go to infinity). Furthermore, note that by the strict complementarity assumption, there are exactly  $m$  terms being added on the left-hand side in above estimate. Hence, by the arithmetic-geometric mean inequality it follows that

$$\left( \prod_{i \in N} \frac{\hat{\lambda}_i}{\lambda_i^*} \right) \left( \prod_{i \in B} \frac{g_i(\hat{x})}{g_i(x^*)} \right) \leq \frac{1}{m} \left( \sum_{i \in N} \frac{\hat{\lambda}_i}{\lambda_i^*} + \sum_{i \in B} \frac{g_i(\hat{x})}{g_i(x^*)} \right) \leq 1$$

and thus

$$\left( \prod_{i \in N} \hat{\lambda}_i \right) \left( \prod_{i \in B} g_i(\hat{x}) \right) \leq \left( \prod_{i \in N} \lambda_i^* \right) \left( \prod_{i \in B} g_i(x^*) \right).$$

By taking  $\hat{\lambda}_N = \lambda_N^*$  or  $g_B(\hat{x}) = g_B(x^*)$ , one finally obtains:

$$\prod_{i \in B} g_i(\hat{x}) \leq \prod_{i \in B} g_i(x^*) \quad \text{and} \quad \prod_{i \in N} \hat{\lambda}_i \leq \prod_{i \in N} \lambda_i^*.$$

Therefore,  $x^*$  is a solution of (2.9), and since (2.9) has a unique solution, the entire sequence  $(\bar{x}_\mu)_{\mu > 0}$  must converge to  $x^*$ . Analogously,  $(\bar{\lambda}_\mu)_{\mu > 0}$  converges to the unique analytic centre  $\lambda^*$ .  $\square$

# 3

## A Tikhonov-regularised interior-point method

In this chapter, the Tikhonov-regularised interior-point method is introduced and analysed in detail. While the focus lies on the outer structure of the method here, Chapter 4 afterwards covers how its inner problems are tackled.

At first, in Section 3.1, the general concept of Tikhonov regularisation is briefly described, and properties of Tikhonov-regularised methods are presented, including especially their convergence to the minimal-norm solution. Next in Section 3.2, the actual Tikhonov-regularised IPM is introduced, its general structure is presented in Algorithm 3.1, and the existence of the unique minimisers of the corresponding regularised barrier problem is shown. Based on that, Section 3.3 covers the convergence analysis for these *regularised barrier minimisers*. It is proven that they converge to the minimal-norm solution of problem  $(\mathcal{P})$ , and a convergence rate in the objective is established which is linear in the regularisation parameter.

In Section 3.4, the consequences are discussed which emerge from solving the inner problems only inexactly—a relevant aspect for using the method in practice. It is shown in Theorem 3.9 how convergence also of those inexact solutions to the minimal-norm solution can be ensured, followed by the derivation of a convergence rate in Theorem 3.11 which again is linear in the regularisation parameter. Furthermore, suitable choices for the parameter update as well as for inner and outer stopping criteria are discussed, and the detailed structure of the proposed method is presented in Algorithm 3.2. Afterwards, it is proven in Theorem 3.12 that the method reaches a given accuracy after finitely many outer iterations, and that it supports an  $r$ -linear convergence rate in the objective.

The chapter concludes with a brief discussion on possible modifications to the Tikhonov-regularised IPM in Section 3.5, including proximal-point regularisation and the integration of linear equality constraints.

### 3.1 Tikhonov regularisation

In this section, an introduction to the concept of Tikhonov regularisation is given. It mainly follows [21, Section 6.4.3], simplified for the continuously differentiable case considered in this thesis.

Generally, to improve the conditioning and obtain strong convexity as well as a unique minimiser, it is a common strategy to modify the optimisation problem  $(\mathcal{P})$ , defined in the introduction, with a *Tikhonov regularisation*. In this, the squared Euclidean norm of  $x$

is added to the objective function, scaled with a *regularisation parameter*  $\varepsilon > 0$ . The *Tikhonov-regularised function*  $f_\varepsilon : \mathbb{R}^n \rightarrow \mathbb{R}$  is thus defined by

$$f_\varepsilon(x) := f(x) + \frac{\varepsilon}{2} \|x\|_2^2,$$

yielding the *Tikhonov-regularised problem*

$$\inf_{x \in \mathbb{R}^n} f_\varepsilon(x) \quad \text{s.t.} \quad g(x) \leq 0. \quad (\mathcal{P}_\varepsilon)$$

Since  $f$  is assumed to be convex and  $\|\cdot\|_2^2$  is 1-strongly convex, the function  $f_\varepsilon$  is  $\frac{\varepsilon}{2}$ -strongly convex. By Theorem 2.7, any strongly convex function has exactly one minimiser on a closed convex set; hence, for fixed  $\varepsilon > 0$ , the regularised problem  $(\mathcal{P}_\varepsilon)$  has a unique solution which will from now on be denoted by  $\bar{x}_\varepsilon$ . By definition,  $\bar{x}_\varepsilon$  also lies in the feasible set  $\mathcal{F}$  as the constraints of  $(\mathcal{P}_\varepsilon)$  and  $(\mathcal{P})$  are the same, and analogously to Theorem 2.11, optimality conditions for  $\bar{x}_\varepsilon$  can be established:

**Theorem 3.1.** *Let the set of feasible points be non-empty, i.e.  $\mathcal{F} \neq \emptyset$ . Then the Tikhonov-regularised convex optimisation problem  $(\mathcal{P}_\varepsilon)$  has a unique solution  $\bar{x}_\varepsilon$ . If, in addition, the Slater CQ condition  $\mathcal{F}_+ \neq \emptyset$  holds, then there exists a multiplier  $\bar{\lambda}_\varepsilon \in \mathbb{R}^m$  such that the following KKT conditions hold:*

$$\nabla f_\varepsilon(\bar{x}_\varepsilon) + \sum_{i=1}^m (\bar{\lambda}_\varepsilon)_i \nabla g_i(\bar{x}_\varepsilon) = 0, \quad (3.1)$$

$$(\bar{\lambda}_\varepsilon)_i g_i(\bar{x}_\varepsilon) = 0 \quad \text{for } i = 1, \dots, m, \quad (3.2)$$

$$g_i(\bar{x}_\varepsilon) \leq 0 \quad \text{for } i = 1, \dots, m, \quad (3.3)$$

$$(\bar{\lambda}_\varepsilon)_i \geq 0 \quad \text{for } i = 1, \dots, m. \quad (3.4)$$

*Proof.* Existence and uniqueness of  $\bar{x}_\varepsilon$  follow directly from Theorem 2.7 since the set of feasible points  $\mathcal{F}$  is non-empty by assumption and closed and convex by Remark 2.9. Existence of  $\bar{\lambda}_\varepsilon$  and the KKT conditions then follow from Theorem 2.11.  $\square$

In Tikhonov-regularised methods, the idea is to decrease the regularisation parameter  $\varepsilon$  step by step to 0. Then, the corresponding solutions  $\bar{x}_\varepsilon$  of the regularised problem  $(\mathcal{P}_\varepsilon)$  remain bounded, approach the minimal function value of the unregularised problem  $(\mathcal{P})$  in  $O(\varepsilon)$ , and ultimately converge to its unique minimal-norm solution.

**Lemma 3.2.** [cf. 21, Lemma 6.46, Satz 6.48] *Let the solution set of  $(\mathcal{P})$  be non-empty, i.e.  $\mathcal{S} \neq \emptyset$ , and let  $x_M^*$  denote the minimal-norm solution. Furthermore, let  $\varepsilon > 0$ , and let  $\bar{x}_\varepsilon$  denote the minimiser of the regularised problem  $(\mathcal{P}_\varepsilon)$ . Then it holds:*

- a) *The deviation of the objective function  $f$  at  $\bar{x}_\varepsilon$  from the solution of  $(\mathcal{P})$  is bounded by*

$$0 \leq f(\bar{x}_\varepsilon) - \min_{x \in \mathcal{F}} f(x) \leq \frac{\varepsilon}{2} \|x_M^*\|_2^2.$$

*Thus, the minimisers  $\bar{x}_\varepsilon$  of the respective regularised problems  $(\mathcal{P}_\varepsilon)$  approach the minimal function value of  $f$  as  $\varepsilon \rightarrow 0^+$ , that is,*

$$\lim_{\varepsilon \rightarrow 0} f(\bar{x}_\varepsilon) = \min_{x \in \mathcal{F}} f(x).$$

*The rate of this convergence lies in  $O(\varepsilon)$ .*

b) All minimisers  $\bar{x}_\varepsilon$  are bounded by the minimal-norm solution of  $(\mathcal{P})$ , i.e.

$$\|\bar{x}_\varepsilon\|_2 \leq \|x_M^*\|_2,$$

and in particular, they converge to  $x_M^*$  as  $\varepsilon \rightarrow 0^+$ , that is,

$$\lim_{\varepsilon \rightarrow 0} \bar{x}_\varepsilon = x_M^*.$$

*Proof.* Firstly, since  $\mathcal{S} \neq \emptyset$ , Lemma 2.13 states that the minimal-norm solution  $x_M^*$  exists. Next note that the set of feasible points  $\mathcal{F}$  is non-empty as the solution set  $\mathcal{S}$  consists only of feasible points, so by Theorem 3.1, also the unique minimiser  $\bar{x}_\varepsilon$  of  $(\mathcal{P}_\varepsilon)$  exists for any  $\varepsilon > 0$ . Since both  $x^*$  and  $\bar{x}_\varepsilon$  are feasible and solve problems  $(\mathcal{P})$  and  $(\mathcal{P}_\varepsilon)$ , respectively, it holds:

$$f(x_M^*) \leq f(\bar{x}_\varepsilon) \leq f_\varepsilon(\bar{x}_\varepsilon) \leq f_\varepsilon(x_M^*) = f(x_M^*) + \frac{\varepsilon}{2} \|x_M^*\|_2^2.$$

From this, the rest of statement a) follows.

A proof for part b) can for example be found in [21, Satz 6.48].  $\square$

*Remark.* Instead of the regularisation with  $\|x\|_2^2$ , also the “shifted regularisation”  $\|x - x^{(0)}\|_2^2$  could be used for any  $x^{(0)} \in \mathbb{R}^n$ . Then,  $x_M^*$  would be given as the unique projection of  $x^{(0)}$  onto  $\mathcal{S}$ . This substitution is also compatible with all theoretical results presented in this chapter, with the only difference that  $\|x\|_2$  needs to be replaced by  $\|x - x^{(0)}\|_2$  in all occurrences except for when a difference of the kind  $\|x - y\|_2$  is given. An outlook on this idea is discussed in Section 3.5.

This lemma will later be used in Theorem 3.6 to show that also the Tikhonov-regularised IPM described in the next section ultimately converges to the minimal-norm solution. A similar result to part a) for interior-point methods was proven in Lemma 2.17 which yielded an upper bound on the deviation in the function value by  $m\mu$ . Both estimates will be combined in Lemma 3.8 to an analogue bound for the minimisers of the Tikhonov-regularised IPM.

## 3.2 The Tikhonov-regularised interior-point method

In this section, the central topic of this thesis is presented: The combination of a Tikhonov regularisation and an interior-point approach for tackling the general inequality-constrained convex optimisation problem  $(\mathcal{P})$  which was defined in the introduction.

### 3.2.1 The regularised barrier problem

As in Chapter 2, let  $b : \mathcal{F}_+ \rightarrow \mathbb{R}$  with  $b(x) := -\sum_{i=1}^m \ln(-g_i(x))$  denote the logarithmic barrier term associated with the constraint functions  $g_1, \dots, g_m$ . We define the *Tikhonov-regularised barrier function*  $f_{\varepsilon, \mu} : \mathcal{F}_+ \rightarrow \mathbb{R}$  with regularisation parameter  $\varepsilon > 0$  and barrier parameter  $\mu > 0$  as

$$\begin{aligned} f_{\varepsilon, \mu}(x) &:= f_\varepsilon(x) + \mu b(x) \\ &= f(x) + \frac{\varepsilon}{2} \|x\|_2^2 - \mu \sum_{i=1}^m \ln(-g_i(x)). \end{aligned} \tag{3.5}$$

Gradient and Hessian of  $f_{\varepsilon,\mu}$  are given by

$$\begin{aligned}\nabla f_{\varepsilon,\mu}(x) &= \nabla f_{\varepsilon}(x) + \mu \nabla b(x) \\ &= \nabla f(x) + \varepsilon x - \mu \sum_{i=1}^m \frac{1}{g_i(x)} \nabla g_i(x)\end{aligned}\tag{3.6}$$

and

$$\begin{aligned}\nabla^2 f_{\varepsilon,\mu}(x) &= \nabla^2 f_{\varepsilon}(x) + \mu \nabla^2 b(x) \\ &= \nabla^2 f(x) + \varepsilon I + \mu \sum_{i=1}^m \left( \frac{1}{g_i(x)^2} \nabla g_i(x) \nabla g_i(x)^T - \frac{1}{g_i(x)} \nabla^2 g_i(x) \right),\end{aligned}\tag{3.7}$$

respectively.

Since the functions  $f$  and  $b$  are convex,  $f_{\varepsilon,\mu}$  is  $\frac{\varepsilon}{2}$ -strongly convex and the corresponding *Tikhonov-regularised barrier problem*

$$\inf_{x \in \mathbb{R}^n} f_{\varepsilon,\mu}(x) \quad \text{s.t.} \quad g(x) < 0 \tag{\mathcal{P}_{\varepsilon,\mu}}$$

has a unique global minimiser, from now on denoted by  $\bar{x}_{\varepsilon,\mu}$ . This is proven next.

### 3.2.2 Existence of regularised barrier minimisers

In contrast to the solely Tikhonov-regularised case described before in Theorem 3.1, the strong convexity and Theorem 2.7 cannot directly be used to show the existence of the unique minimiser, as the domain  $\mathcal{F}_+$  which underlies  $f_{\varepsilon,\mu}$  is not closed in general. Instead, the compactness of level sets of  $f_{\varepsilon,\mu}$  will be needed. These follow from the well-known property of barrier functions that their level sets are compact when their solution sets are non-empty and bounded [48, Theorem 4].

**Lemma 3.3.** *For any parameters  $\varepsilon, \mu > 0$  and any constant  $\alpha \in \mathbb{R}$ , the level set*

$$\{x \in \mathcal{F}_+ \mid f_{\varepsilon,\mu}(x) \leq \alpha\}$$

*is compact, i.e. closed and bounded.*

*Proof.* The function  $f_{\varepsilon,\mu}$  corresponds to the barrier function of the Tikhonov-regularised problem  $(\mathcal{P}_{\varepsilon})$ . By Theorem 3.1, the latter has a unique minimiser  $\bar{x}_{\varepsilon}$ , so in particular, it has a non-empty and bounded solution set (only consisting of  $\bar{x}_{\varepsilon}$ ). Thus, the requirements for [48, Theorem 4] are met and thereby all level sets of  $f_{\varepsilon,\mu}$  compact.  $\square$

In general, this property is not too surprising, since at each boundary point  $x$  of the strictly feasible set, it holds  $g_i(x) = 0$  for at least one  $i \in \{1, \dots, m\}$ . Hence, the continuity of all  $g_i$  implies that  $b(x^k) \rightarrow \infty$  and thus  $f_{\varepsilon,\mu}(x^k) \rightarrow \infty$  for any sequence  $(x^k)_{k \in \mathbb{N}} \subseteq \mathcal{F}_+$  approaching the boundary. Therefore, the fact that  $f_{\varepsilon,\mu}$  is continuous and strongly convex on  $\mathcal{F}_+$  yields that any level set of it must be closed.

This property can now be used to show the existence of the unique minimiser  $\bar{x}_{\varepsilon,\mu}$  of problem  $(\mathcal{P}_{\varepsilon,\mu})$ .

**Theorem 3.4.** *Let the Slater CQ condition hold, i.e.  $\mathcal{F}_+ \neq \emptyset$ . Then the Tikhonov-regularised barrier problem  $(\mathcal{P}_{\varepsilon,\mu})$  has a unique solution  $\bar{x}_{\varepsilon,\mu} \in \mathcal{F}_+$  and it holds*

$$\nabla f_{\varepsilon,\mu}(\bar{x}_{\varepsilon,\mu}) = 0. \quad (3.8)$$

*Proof.* By Lemma 3.3, the level set  $\{x \in \mathcal{F}_+ \mid f_{\varepsilon,\mu}(x) \leq f_{\varepsilon,\mu}(x^0)\}$  is closed and convex for any  $x^0 \in \mathcal{F}_+$ , and the optimisation of  $(\mathcal{P}_{\varepsilon,\mu})$  can be limited to any such (non-empty) level set. Existence and uniqueness of  $\bar{x}_{\varepsilon,\mu}$  now follow with Theorem 2.7.

Finally, Equation (3.8) corresponds to the first-order necessary optimality condition [37, Theorem 2.2] of problem  $(\mathcal{P}_{\varepsilon,\mu})$  and therefore must hold for  $\bar{x}_{\varepsilon,\mu}$ .  $\square$

Due to the convexity of  $f_{\varepsilon,\mu}$ , this  $\bar{x}_{\varepsilon,\mu}$  is also the only point  $x \in \mathcal{F}_+$  for which  $\nabla f_{\varepsilon,\mu}(x) = 0$  holds [21, page 27], so a reasonable strategy for computing the minimiser  $\bar{x}_{\varepsilon,\mu}$  is to find a solution to (3.8). This process will be further discussed in Chapter 4.

### 3.2.3 General structure of the method

In Section 2.3, it was described how, in order to find a solution to problem  $(\mathcal{P})$ , in interior-point methods, the barrier parameter  $\mu$  is iteratively decreased, and in each iteration, an approximation to a barrier minimiser  $\bar{x}_\mu$  is computed. This idea is now adopted for the Tikhonov-regularised IPM of this thesis:

Starting with a strictly feasible point  $x^0 \in \mathcal{F}_+$  and initial regularisation and barrier parameters  $\varepsilon_1 > 0$  and  $\mu_1 > 0$ , both parameters  $\varepsilon$  and  $\mu$  are iteratively decreased, and in each iteration, an approximation to the regularised barrier minimiser  $\bar{x}_{\varepsilon,\mu}$  is computed. These iterations will be referred to as *outer iterations* and denoted with iteration index  $j$ , in contrast to the *inner iterations* which describe the process of approximating  $\bar{x}_{\varepsilon,\mu}$  for fixed  $\varepsilon$  and  $\mu$ . The latter is further discussed in Chapter 4, with  $k$  used as inner iteration index.

The general structure of the Tikhonov-regularised IPM is presented in Algorithm 3.1. A detailed description of the proposed method in which the results and discussions from the next section are included, is given in Algorithm 3.2.

---

**Algorithm 3.1** General scheme of the Tikhonov-regularised interior-point method

---

**Given:** Optimisation problem  $(\mathcal{P})$

**Choose:**  $x^0 \in \mathcal{F}_+$ ,  $\varepsilon_1, \mu_1 > 0$

1: **for**  $j = 1, 2, 3, \dots$  **do**

2:     **if**  $x^{j-1}$  satisfies stopping criteria **then** let  $x^* \leftarrow x^{j-1}$  and **break**

3:     Compute  $x^j$  as approximation to  $\bar{x}_{\varepsilon_j, \mu_j}$   $\triangleright$  Typically initialised with  $x^{j-1}$

4:     Update  $0 < \varepsilon_{j+1} < \varepsilon_j$ ;  $0 < \mu_{j+1} < \mu_j$

5: **end for**

**Return:**  $x^*$

---

In the next section, the convergence of this algorithm to the minimal-norm solution  $x_M^*$  of problem  $(\mathcal{P})$  will be shown under the assumption that the minimiser  $\bar{x}_{\varepsilon_j, \mu_j}$  is computed exactly in each iteration. After that, Section 3.4 covers the convergence analysis for the case where  $\bar{x}_{\varepsilon_j, \mu_j}$  are only approximated by the iterates  $x^j$ . Resulting from this, the question will be addressed how to update the parameters  $\varepsilon_j$  and  $\mu_j$  and what qualifies the iterates  $x^j$  to be a sufficiently accurate approximation to  $\bar{x}_{\varepsilon_j, \mu_j}$ .

### 3.3 Convergence of the regularised barrier minimisers

In the last section, the concept of the Tikhonov-regularised IPM was introduced, and its general structure was described in Algorithm 3.1. It was proven that, under the Slater CQ, each optimisation problem  $(\mathcal{P}_{\varepsilon,\mu})$  has a unique minimiser  $\bar{x}_{\varepsilon,\mu}$  for any regularisation and barrier parameter  $\varepsilon > 0$  and  $\mu > 0$ , respectively. In this section, it will be analysed how these regularised barrier minimisers converge as both parameters approach 0. The main results are presented in Theorem 3.6 and Lemma 3.8 and state that  $\bar{x}_{\varepsilon,\mu}$  converges to the minimal-norm solution  $x_M^*$  of problem  $(\mathcal{P})$  when  $\mu$  is reduced faster than  $\varepsilon$ , and that this convergence supports a rate of  $O(\varepsilon)$  in the objective  $f$ .

To prove these results, the convergence of the regularised barrier minimisers  $\bar{x}_{\varepsilon,\mu}$  will be reduced to the convergence of the minimisers  $\bar{x}_\varepsilon$  of the corresponding solely Tikhonov-regularised problem  $(\mathcal{P}_\varepsilon)$  which was described in Section 3.1. As was shown in Theorem 3.1, the Slater CQ is sufficient for the existence of the unique minimiser  $\bar{x}_\varepsilon$  for any regularisation parameter  $\varepsilon > 0$ , and by Lemma 3.2,  $\bar{x}_\varepsilon$  converges to  $x_M^*$  in  $O(\varepsilon)$  as  $\varepsilon \rightarrow 0$ .

#### 3.3.1 Reduction to convergence of regularised minimisers

To obtain the convergence result of Theorem 3.6, the relation between the trajectories of the regularised barrier minimisers  $\bar{x}_{\varepsilon,\mu}$  and the regularised minimisers  $\bar{x}_\varepsilon$  must be analysed. As the following lemma shows, for given parameters  $\varepsilon$  and  $\mu$ , their difference is bounded by  $\sqrt{m\frac{\mu}{\varepsilon}}$ .

**Lemma 3.5.** *Let  $\varepsilon, \mu > 0$ , and let  $\bar{x}_\varepsilon, \bar{x}_{\varepsilon,\mu}$  denote the minimisers of the problems  $(\mathcal{P}_\varepsilon)$  and  $(\mathcal{P}_{\varepsilon,\mu})$ . Then it holds:*

$$\|\bar{x}_{\varepsilon,\mu} - \bar{x}_\varepsilon\|_2^2 \leq m\frac{\mu}{\varepsilon}. \quad (3.9)$$

*Proof.* Subtracting the necessary optimality conditions (3.1) and (3.8) of the problems  $(\mathcal{P}_\varepsilon)$  and  $(\mathcal{P}_{\varepsilon,\mu})$ , and multiplying both sides with  $(\bar{x}_\varepsilon - \bar{x}_{\varepsilon,\mu})$  yields

$$\begin{aligned} 0 &= (\nabla f_\varepsilon(\bar{x}_\varepsilon) - \nabla f_\varepsilon(\bar{x}_{\varepsilon,\mu}))^T (\bar{x}_\varepsilon - \bar{x}_{\varepsilon,\mu}) \\ &\quad + \sum_{i=1}^m (\bar{\lambda}_\varepsilon)_i \nabla g_i(\bar{x}_\varepsilon)^T (\bar{x}_\varepsilon - \bar{x}_{\varepsilon,\mu}) - \mu \nabla b(\bar{x}_{\varepsilon,\mu})^T (\bar{x}_\varepsilon - \bar{x}_{\varepsilon,\mu}). \end{aligned}$$

Due to  $f_\varepsilon$  being  $\frac{\varepsilon}{2}$ -strongly convex, it holds by Lemma 2.5 c) that  $(\nabla f_\varepsilon(\bar{x}_\varepsilon) - \nabla f_\varepsilon(\bar{x}_{\varepsilon,\mu}))^T (\bar{x}_\varepsilon - \bar{x}_{\varepsilon,\mu}) \geq \varepsilon \|\bar{x}_\varepsilon - \bar{x}_{\varepsilon,\mu}\|_2^2$ , so

$$\varepsilon \|\bar{x}_\varepsilon - \bar{x}_{\varepsilon,\mu}\|_2^2 \leq \sum_{i=1}^m (\bar{\lambda}_\varepsilon)_i \nabla g_i(\bar{x}_\varepsilon)^T (\bar{x}_{\varepsilon,\mu} - \bar{x}_\varepsilon) + \mu \nabla b(\bar{x}_{\varepsilon,\mu})^T (\bar{x}_\varepsilon - \bar{x}_{\varepsilon,\mu}).$$

Furthermore, since all  $g_i$  are convex and therefore, by Lemma 2.3 b), it is  $g_i(\bar{x}_\varepsilon) + \nabla g_i(\bar{x}_\varepsilon)^T (\bar{x}_{\varepsilon,\mu} - \bar{x}_\varepsilon) \leq g_i(\bar{x}_{\varepsilon,\mu})$ , one obtains

$$\varepsilon \|\bar{x}_\varepsilon - \bar{x}_{\varepsilon,\mu}\|_2^2 \leq \sum_{i=1}^m (\bar{\lambda}_\varepsilon)_i (g_i(\bar{x}_{\varepsilon,\mu}) - g_i(\bar{x}_\varepsilon)) + \mu \nabla b(\bar{x}_{\varepsilon,\mu})^T (\bar{x}_\varepsilon - \bar{x}_{\varepsilon,\mu}).$$



By optimality of  $\bar{x}_\varepsilon$ , it now holds for  $i = 1, \dots, m$  that  $(\bar{\lambda}_\varepsilon)_i g_i(\bar{x}_\varepsilon) = 0$  and  $(\bar{\lambda}_\varepsilon)_i \geq 0$  (see (3.2) and (3.4)), which, together with  $g_i(\bar{x}_{\varepsilon,\mu}) < 0$ , yields

$$\varepsilon \|\bar{x}_\varepsilon - \bar{x}_{\varepsilon,\mu}\|_2^2 \leq \mu \nabla b(\bar{x}_{\varepsilon,\mu})^T (\bar{x}_\varepsilon - \bar{x}_{\varepsilon,\mu}) \stackrel{\text{Lemma 2.16}}{\leq} m\mu.$$

Hence, the bound (3.9) is proven.  $\square$

Lemma 3.5 implies that the distance between the corresponding minimisers  $\bar{x}_\varepsilon$  and  $\bar{x}_{\varepsilon,\mu}$  lies in  $O(\sqrt{\frac{\mu}{\varepsilon}})$ , meaning that the greater the barrier parameter  $\mu$  is relative to  $\varepsilon$ , the more both points might deviate from each other. Conversely, the smaller the impact of  $\mu$ , the closer the minimisers lie together. This behaviour is also expected as the underlying optimisation problems only differ in the additional barrier term  $\mu b(x)$ .

### 3.3.2 Convergence to the minimal norm solution

Lemma 3.5 can now be utilised to analyse the convergence behaviour of the Tikhonov-regularised IPM described in Algorithm 3.1. As was shown in Lemma 3.2 b), the regularised minimisers  $\bar{x}_\varepsilon$  converge to the minimal-norm solution  $x_M^*$ . Thus, when  $\mu$  is reduced “faster” than  $\varepsilon$  during the outer iterations, then (3.9) yields that the regularised barrier minimisers  $\bar{x}_{\varepsilon,\mu}$  will get closer and closer to the corresponding  $\bar{x}_\varepsilon$  and ultimately converge to the same limit point: the minimal-norm solution  $x_M^*$ .

**Theorem 3.6.** *Let Assumption 2.10 hold, and for any  $\varepsilon, \mu > 0$ , let  $\bar{x}_{\varepsilon,\mu}$  denote the minimiser of  $(\mathcal{P}_{\varepsilon,\mu})$ . It holds:*

- a) *The family of minimisers  $(\bar{x}_{\varepsilon,\mu})_{\substack{\varepsilon \in (0, \varepsilon_1] \\ \mu \in (0, \mu_1]}}$  remains bounded for  $\mu = O(\varepsilon)$  as  $\varepsilon \rightarrow 0^+$ , and*
- b)  *$\lim_{\varepsilon \rightarrow 0^+} \bar{x}_{\varepsilon,\mu} = x_M^*$  for  $\mu = o(\varepsilon)$  as  $\varepsilon \rightarrow 0^+$ .*

*Proof.* First recall that by Lemma 2.13 and Assumption 2.10, the minimal-norm solution  $x_M^*$  exists, and that by Lemma 3.5 it holds  $\|\bar{x}_{\varepsilon,\mu} - \bar{x}_\varepsilon\|_2^2 \leq m\frac{\mu}{\varepsilon}$  for all  $\varepsilon, \mu > 0$ . Thus:

- a) Lemma 3.2 b) implies that  $\|\bar{x}_\varepsilon\|_2 \leq \|x_M^*\|_2$ , so

$$\|\bar{x}_{\varepsilon,\mu}\|_2^2 \leq \|\bar{x}_{\varepsilon,\mu} - \bar{x}_\varepsilon\|_2^2 + \|\bar{x}_\varepsilon\|_2^2 \leq m\frac{\mu}{\varepsilon} + \|x_M^*\|_2^2 = O(1)$$

for  $\mu = O(\varepsilon)$ . Hence, the sequence of minimisers remains bounded.

- b) If furthermore  $\mu = o(\varepsilon)$ , then

$$\|\bar{x}_{\varepsilon,\mu} - \bar{x}_\varepsilon\|_2^2 \leq m\frac{\mu}{\varepsilon} \rightarrow 0$$

as  $\varepsilon \rightarrow 0^+$ , so since  $\lim_{\varepsilon \rightarrow 0^+} \bar{x}_\varepsilon = x_M^*$  by Lemma 3.2 b), also the minimisers  $\bar{x}_{\varepsilon,\mu}$  converge to  $x_M^*$ .  $\square$

Theorem 3.6 is a remarkably strong result: It ensures that if Assumption 2.10 holds, then the solutions of the unconstrained Tikhonov-regularised barrier problems converge to

a solution of  $(\mathcal{P})$  as long as  $\mu = O(\varepsilon)$ . Furthermore, this solution is characterised as the minimal-norm solution  $x_M^*$ .

This stands in contrast to typical results for IPMs as presented in Section 2.4 which additionally require at least the boundedness of the solution set  $\mathcal{S}$  to obtain a subsequence that converges to a minimiser, and usually need even further assumptions on the problem or the limit point in order to classify the solution, such as it is the case in Theorem 2.19.

But Theorem 3.6 also allows for another observation: Whereas IPMs as described in Algorithm 2.1 typically converge to the analytic centre of the solution set  $\mathcal{S}$ , the Tikhonov-regularised IPM yields the minimal-norm solution when  $\mu = o(\varepsilon)$ , analogously to the solely Tikhonov-regularised approach. Therefore, the regularisation enforces the optimisation of the solution set for a minimal Euclidean norm instead of the geometrically weighted greatest distance to all constraints.

This difference in the trajectories and limit points of the minimisers of  $(\mathcal{P}_\mu)$ ,  $(\mathcal{P}_\varepsilon)$  and  $(\mathcal{P}_{\varepsilon,\mu})$  is illustrated in Figure 3.2, depicting the following example:

**Example 3.7.** Consider the two-dimensional problem

$$\inf_{x \in \mathbb{R}^2} f(x) \quad \text{s.t.} \quad x_2 \geq 0,$$

where  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  is given by

$$f(x) := \begin{cases} (\|x - (2, 0)^T\|_2 - 1)^4 & \text{for } \|x - (2, 0)^T\|_2 \geq 1 \\ 0 & \text{otherwise} \end{cases}.$$

Intuitively, the function  $f$  is obtained when inserting a plateau at the interval  $[-1, 1]$  into the one-dimensional function  $x_1^4$ , then extending it into the second dimension by circularly rotating it along the vertical axis, and finally shifting the result along the  $x_1$ -axis to the point  $(2, 0)^T$ . Thus,  $f$  consists of a disc with radius 1, centred in  $(2, 0)^T$ , from whose boundary it increases smoothly like  $z^4$  in every direction.

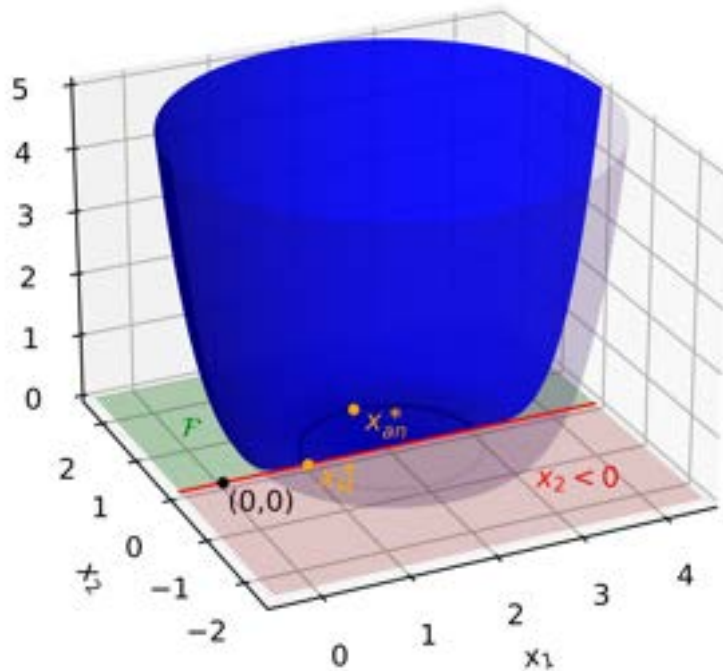
It is straightforward to show that  $f$  is convex and twice continuously differentiable, and its gradient is given by

$$\nabla f(x) := \begin{cases} \frac{4(\|x - (2, 0)^T\|_2 - 1)^3}{\|x - (2, 0)^T\|_2} (x - (2, 0)^T) & \text{for } \|x - (2, 0)^T\|_2 \geq 1 \\ (0, 0)^T & \text{otherwise} \end{cases}.$$

Thus,  $\nabla f(x) = 0 \Leftrightarrow \|x - (2, 0)^T\|_2 \leq 1$ , so the minimisers of  $f$  disregarding the constraint are precisely the points in the unit disc centred in  $(2, 0)^T$ .

The constraint  $x_2 \geq 0$  yields the feasible set  $\mathcal{F} = \{x \in \mathbb{R}^2 \mid x_2 \geq 0\}$  and can be represented by  $g(x) \leq 0$  for example with the linear function  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $g(x) := -x_2$ . It follows that the solutions of the optimisation problem are the points in the unit disc with  $x_2 \geq 0$ , so the solution set is given by the half-disc  $\mathcal{S} = \{x \in \mathbb{R}^2 \mid \|x - (2, 0)^T\|_2 \leq 1, x_2 \geq 0\}$ .

The objective function  $f$  and its feasible set are plotted in Figure 3.1, along with the minimal-norm solution  $x_M^* = (1, 0)^T$  and the *analytic centre*  $x_{an}^* = (2, 1)^T$  of the solution set  $\mathcal{S}$ . The latter is the point in  $\mathcal{S}$  with the greatest distance to the constraint boundary  $x_2 = 0$ ; for details, see Theorem 2.19.

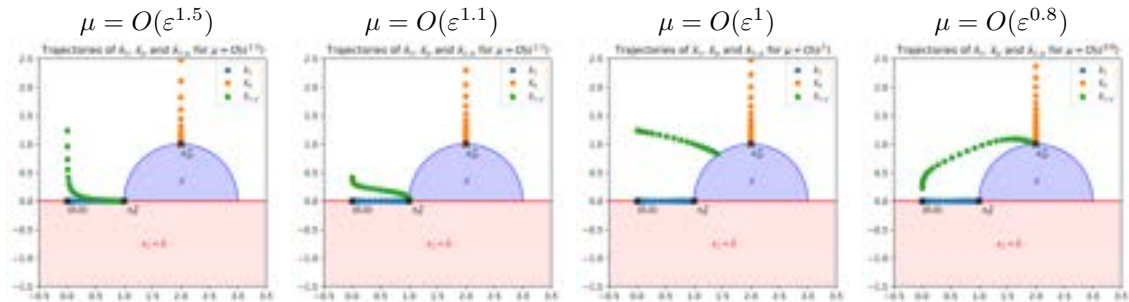


**Figure 3.1:** The objective function from Example 3.7 with the minimal-norm solution  $x_M^*$  and the analytic centre  $x_{an}^*$ . The objective is drawn blue on the feasible set  $\mathcal{F}$  (highlighted green) and sketched transparently over the infeasible area (highlighted red). The solution set is given by the half-disc  $\mathcal{S} = \{x \in \mathbb{R}^2 \mid \|x - (2,0)^T\|_2 \leq 1, x_2 \geq 0\}$ . The minimal-norm solution  $x_M^*$  is the closest point in  $\mathcal{S}$  to the origin, whereas  $x_{an}^*$  maximises the distance to the boundary.

By construction, the objective  $f$  is circular around  $(2,0)$  and thus symmetrical in every direction from there. Hence, when looking at the Tikhonov-regularised function  $f_\varepsilon(x) = f(x) + \frac{\varepsilon}{2} \|x\|_2^2$ , it becomes clear that its minimiser  $\bar{x}_\varepsilon$  lies directly between the origin  $(0,0)$  and the minimal-norm solution  $x_M^* = (1,0)$ , as any deviation from that line on the same level of  $f$  would increase the distance to the origin. For large values of  $\varepsilon$ ,  $\bar{x}_\varepsilon$  lies close to the origin, and as  $\varepsilon$  is reduced to 0, it approaches  $x_M^*$  in accordance with Lemma 3.2 b).

Similarly, the barrier function  $\varphi_\mu(x) = f(x) - \mu \ln(-g(x)) = f(x) - \mu \ln(x_2)$  is minimised along the axis  $x_1 = 2$ , since the barrier term purely strives for maximising the  $x_2$ -component. The greater the barrier parameter  $\mu$ , the farther away from the centre is the minimiser  $\bar{x}_\mu$ , and when  $\mu \rightarrow 0$  like in interior-point methods,  $\bar{x}_\mu$  converges to the analytic centre  $x_{an}^*$ .

Now, when both approaches are combined in the Tikhonov-regularised IPM, the minimiser  $\bar{x}_{\varepsilon,\mu}$  of the corresponding function  $f_{\varepsilon,\mu}(x) = f(x) + \frac{\varepsilon}{2} \|x\|_2^2 - \mu \ln(x_2)$  lies somewhere between those two scenarios, with its position depending on the value of  $\varepsilon$  and  $\mu$ . When  $\varepsilon$  and  $\mu$  are simultaneously reduced as described in Algorithm 3.1, it can be observed how  $\bar{x}_{\varepsilon,\mu}$  approaches the solution set: When  $\mu$  is reduced faster than  $\varepsilon$ , the regularisation term



**Figure 3.2:** Trajectories of  $\bar{x}_\varepsilon$  (blue),  $\bar{x}_\mu$  (orange) and  $\bar{x}_{\varepsilon,\mu}$  (green) for the problem of Example 3.7 and different choices of the reduction speed of  $\varepsilon$  and  $\mu$ . The feasible set consists of the half-plane above the line  $x_2 = 0$ , the infeasible area is highlighted in red, and the solution set  $\mathcal{S}$  is the blue-shaded half-disc. In all scenarios,  $\bar{x}_\varepsilon$  converges to the minimal-norm solution  $x_M^*$ , and  $\bar{x}_\mu$  approaches the analytic centre  $x_{an}^*$ . When  $\mu = o(\varepsilon)$ , then  $\bar{x}_{\varepsilon,\mu}$  converges to  $x_M^*$  as stated in Theorem 3.6. However, when  $\mu$  is reduced slower than  $\varepsilon$ , then  $\bar{x}_{\varepsilon,\mu}$  converges to  $x_{an}^*$ . When both parameters are reduced at the same speed, then  $\bar{x}_{\varepsilon,\mu}$  approaches the boundary of  $\mathcal{S}$  at a point between  $x_{an}^*$  and  $x_M^*$ .

becomes dominant and convergence to  $x_M^*$  is obtained in accordance with Theorem 3.6. Conversely, when  $\varepsilon$  is reduced faster than  $\mu$ , then the barrier term dominates, and  $\bar{x}_{\varepsilon,\mu}$  approaches  $x_{an}^*$ . When, however, both parameters are reduced at the same speed, the effect of both terms decreases equally, and a point between  $x_M^*$  and  $x_{an}^*$  is approached.

The trajectories of the respective minimisers  $\bar{x}_\varepsilon$ ,  $\bar{x}_\mu$  and  $\bar{x}_{\varepsilon,\mu}$  for all three scenarios are shown in Figure 3.2.

### 3.3.3 Rate of convergence

After general convergence of the Tikhonov-regularised IPM for  $\mu = o(\varepsilon)$  has been proven in Theorem 3.6, the last part of this section is dedicated to the question how this convergence can be quantified.

As presented in Section 2.4, interior-point methods do not only provide a convergence in the function value of the objective  $f$  along the barrier trajectory in  $O(\mu)$ , but in particular, the deviation between  $f(\bar{x}_\mu)$  and  $\min_{x \in \mathcal{F}} f(x)$  is bounded by  $m\mu$  for each barrier parameter  $\mu > 0$  (see Lemma 2.17). For the solely Tikhonov-regularised problem  $(\mathcal{P}_\varepsilon)$ , on the other hand, Lemma 3.2 a) similarly bounds the deviation in the objective function value by  $\frac{\varepsilon}{2} \|x_M^*\|_2^2$ .

Now, it seems natural that an analogue result can also be obtained for the Tikhonov-regularised barrier problem  $(\mathcal{P}_{\varepsilon,\mu})$  in which both approaches are combined. And in fact, the next lemma shows that such a bound is given by a term which almost matches the sum of both aforementioned terms, except for a factor 2 on the norm of  $x_M^*$ .

**Lemma 3.8.** *Let Assumption 2.10 hold, and let  $x^* \in \mathcal{S}$ . For any parameters  $\varepsilon, \mu > 0$ , the minimiser  $\bar{x}_{\varepsilon,\mu}$  of the regularised barrier problem  $(\mathcal{P}_{\varepsilon,\mu})$  satisfies:*

$$0 \leq f(\bar{x}_{\varepsilon,\mu}) - f(x^*) \leq m\mu + \varepsilon \|x_M^*\|_2^2, \quad (3.10)$$

which lies in  $O(\varepsilon)$  for  $\mu = O(\varepsilon)$ .

*Proof.* The first inequality follows directly from  $\bar{x}_{\varepsilon,\mu}$  being feasible and  $x^*$  being a minimiser of  $f$  on the feasible set. The second part will be proven by separating the Tikhonov regularisation from the barrier approach. For this, the minimiser  $\bar{x}_\varepsilon$  of problem  $(\mathcal{P}_\varepsilon)$  is used (its existence follows directly from  $\mathcal{F} \neq \emptyset$ , see Theorem 3.1). The function value difference can be split into

$$f(\bar{x}_{\varepsilon,\mu}) - f(x^*) = (f(\bar{x}_{\varepsilon,\mu}) - f(\bar{x}_\varepsilon)) + (f(\bar{x}_\varepsilon) - f(x^*)).$$

In Lemma 3.2 a), it was derived that the second term is bounded by

$$f(\bar{x}_\varepsilon) - f(x^*) \leq \frac{\varepsilon}{2} \|x_M^*\|_2^2. \quad (3.11)$$

For the first difference, on the other hand, the result from Lemma 2.17 cannot directly be applied, as the function which underlies the barrier method is here given by  $f_\varepsilon$  and not by  $f$ . But since  $\bar{x}_\varepsilon$  is the solution of problem  $(\mathcal{P}_\varepsilon)$ , and  $\bar{x}_{\varepsilon,\mu}$  is the solution of the to  $(\mathcal{P}_\varepsilon)$  corresponding barrier problem  $(\mathcal{P}_{\varepsilon,\mu})$ , Lemma 2.17 yields

$$f_\varepsilon(\bar{x}_{\varepsilon,\mu}) - f_\varepsilon(\bar{x}_\varepsilon) \leq m\mu. \quad (3.12)$$

As an estimate on the difference in the function value  $f$  instead of  $f_\varepsilon$  is sought, (3.12) can now be used by adding and subtracting the regularisation terms to the function values:

$$\begin{aligned} f(\bar{x}_{\varepsilon,\mu}) - f(\bar{x}_\varepsilon) &= f_\varepsilon(\bar{x}_{\varepsilon,\mu}) - f_\varepsilon(\bar{x}_\varepsilon) - \frac{\varepsilon}{2} \left( \|\bar{x}_{\varepsilon,\mu}\|_2^2 - \|\bar{x}_\varepsilon\|_2^2 \right) \\ &\stackrel{(3.12)}{\leq} m\mu - \frac{\varepsilon}{2} \left( \|\bar{x}_{\varepsilon,\mu}\|_2^2 - \|\bar{x}_\varepsilon\|_2^2 \right). \end{aligned}$$

Dropping the non-positive summand, one finally obtains

$$f(\bar{x}_{\varepsilon,\mu}) - f(\bar{x}_\varepsilon) \leq m\mu + \frac{\varepsilon}{2} \|\bar{x}_\varepsilon\|_2^2 \stackrel{\text{Lemma 3.2 b)}}{\leq} m\mu + \frac{\varepsilon}{2} \|x_M^*\|_2^2.$$

Together with (3.11), the upper bound of (3.10) follows.  $\square$

*Remark.* The lemma states that, when solved exactly in every iteration, the Tikhonov-regularised IPM converges to the minimal function value as long as both the regularisation parameter  $\varepsilon_j$  and the barrier parameter  $\mu_j$  are reduced to 0 as  $j \rightarrow \infty$ . When  $\mu_j = O(\varepsilon_j)$ , a convergence rate of  $O(\varepsilon_j)$  is obtained in the objective  $f$ . The stronger restriction of  $\mu_j = o(\varepsilon_j)$  which underlies Theorem 3.6 b) is only required to obtain convergence of the minimisers  $\bar{x}_{\varepsilon_j,\mu_j}$  to the minimal-norm solution  $x_M^*$ . However, this faster reduction does not improve the convergence rate of  $O(\varepsilon_j)$  in Lemma 3.8.

Together, Theorem 3.6 and Lemma 3.8 build the foundation for the convergence analysis of the Tikhonov-regularised IPM of Algorithm 3.1. However, they only apply when the regularised barrier minimiser  $\bar{x}_{\varepsilon,\mu}$  is computed exactly in every iteration. The next section therefore discusses the case when the minimisers are only approximated in every iteration.

### 3.4 Convergence in case of inexact solutions

In the last section, it was analysed how the minimisers  $\bar{x}_{\varepsilon,\mu}$  of the regularised barrier function  $f_{\varepsilon,\mu}$  converge to the minimal-norm solution  $x_M^*$  of the underlying problem  $(\mathcal{P})$  when

the barrier parameter  $\mu$  is reduced faster than the regularisation parameter  $\varepsilon$ . Furthermore, it was shown that the rate of convergence in the objective function then lies in  $O(\varepsilon)$ .

In practice, however, it is usually not possible or reasonable to compute the regularised barrier minimisers  $\bar{x}_{\varepsilon,\mu}$  exactly, especially since the result might only be used as starting point for the next iteration. Hence, the question arises whether the Tikhonov-regularised IPM of Algorithm 3.1 also converges if the inner problems (cf. line 3) are only solved approximately, that is, that in each iteration with fixed parameters  $\varepsilon$  and  $\mu$ , only an approximation to  $\bar{x}_{\varepsilon,\mu}$  is computed.

### 3.4.1 Convergence to the minimal norm solution

As is proven next in Theorem 3.9, convergence is in fact obtained as long as the accuracy of the approximations  $x^j$  to the respective  $\bar{x}_{\varepsilon_j,\mu_j}$  increases sufficiently much with every iteration. Specifically, the gradient norm of  $f_{\varepsilon_j,\mu_j}$  evaluated at the iterate  $x^j$  is required to decrease faster than  $\varepsilon_j$ , that is,  $\|\nabla f_{\varepsilon_j,\mu_j}(x^j)\|_2 = o(\varepsilon_j)$ . This is a useful requirement since computing the minimiser  $\bar{x}_{\varepsilon_j,\mu_j}$  is equivalent to solving the equation  $\nabla f_{\varepsilon_j,\mu_j}(x) = 0$  (cf. Theorem 3.4), so an inner solver would typically decrease the gradient's norm; more on that in Chapter 4.

The idea behind proving the convergence of the iterates  $x^j$  is to reduce it to the convergence of the respective exact minimisers  $\bar{x}_{\varepsilon_j,\mu_j}$ . To do so, it is necessary that the iterates converge to the respective  $\bar{x}_{\varepsilon_j,\mu_j}$  as  $j \rightarrow \infty$ . For this, the  $\frac{\varepsilon_j}{2}$ -strong convexity of  $f_{\varepsilon,\mu}$  comes into play: As is shown in the proof of Theorem 3.9, it provides that the distance between  $x^j$  and  $\bar{x}_{\varepsilon_j,\mu_j}$  can be limited to

$$\|x^j - \bar{x}_{\varepsilon_j,\mu_j}\|_2 \leq \frac{1}{\varepsilon_j} \|\nabla f_{\varepsilon_j,\mu_j}(x^j)\|_2.$$

Therefore, when the gradient norm is reduced faster than  $\varepsilon_j$ , the right-hand side goes to 0, so the convergence of  $\bar{x}_{\varepsilon_j,\mu_j}$  to  $x_M^*$ , which was proven in Theorem 3.6 for  $\mu_j = o(\varepsilon_j)$ , ultimately yields that also the iterates  $x^j$  converge to  $x_M^*$ .

**Theorem 3.9.** *Let Assumption 2.10 hold. Consider the Tikhonov-regularised interior-point method of Algorithm 3.1 with parameter sequences  $(\varepsilon_j)_{j \in \mathbb{N}}$  and  $(\mu_j)_{j \in \mathbb{N}}$  where  $\varepsilon_j \rightarrow 0^+$  and  $\mu_j \rightarrow 0^+$  as  $j \rightarrow \infty$ . If  $\mu_j = o(\varepsilon_j)$  and the gradient norm in the iterates  $x^j$  produced by the algorithm satisfy  $\|\nabla f_{\varepsilon_j,\mu_j}(x^j)\|_2 = o(\varepsilon_j)$  for  $j \rightarrow \infty$ , then the iterates  $x^j$  converge to  $x_M^*$ .*

*Proof.* First observe that Assumption 2.10 ensures the existence of the minimal-norm solution  $x_M^*$  as well as of the regularised barrier minimisers  $\bar{x}_{\varepsilon_j,\mu_j}$  for any  $\varepsilon_j, \mu_j > 0$  (see Lemma 2.13 and Theorem 3.4).

Next, a bound on the distance between any point  $x \in \mathcal{F}_+$  and  $\bar{x}_{\varepsilon_j,\mu_j}$  will be derived: Using the  $\frac{\varepsilon_j}{2}$ -strong convexity of  $f_{\varepsilon_j,\mu_j}$ , the fact that  $\nabla f_{\varepsilon_j,\mu_j}(\bar{x}_{\varepsilon_j,\mu_j}) = 0$  (due to the optimality of  $\bar{x}_{\varepsilon_j,\mu_j}$ , see Theorem 3.4), and the Cauchy-Schwarz inequality [37, page 600], it

follows that

$$\begin{aligned} \|x - \bar{x}_{\varepsilon_j, \mu_j}\|_2 &\stackrel{\text{Lemma 2.5 c)}}{\leq} \frac{1}{\varepsilon_j} \left( \nabla f_{\varepsilon_j, \mu_j}(x) - \underbrace{\nabla f_{\varepsilon_j, \mu_j}(\bar{x}_{\varepsilon_j, \mu_j})}_{=0} \right)^T (x - \bar{x}_{\varepsilon_j, \mu_j}) \\ &\leq \frac{1}{\varepsilon_j} \|\nabla f_{\varepsilon_j, \mu_j}(x)\|_2 \|x - \bar{x}_{\varepsilon_j, \mu_j}\|_2, \end{aligned}$$

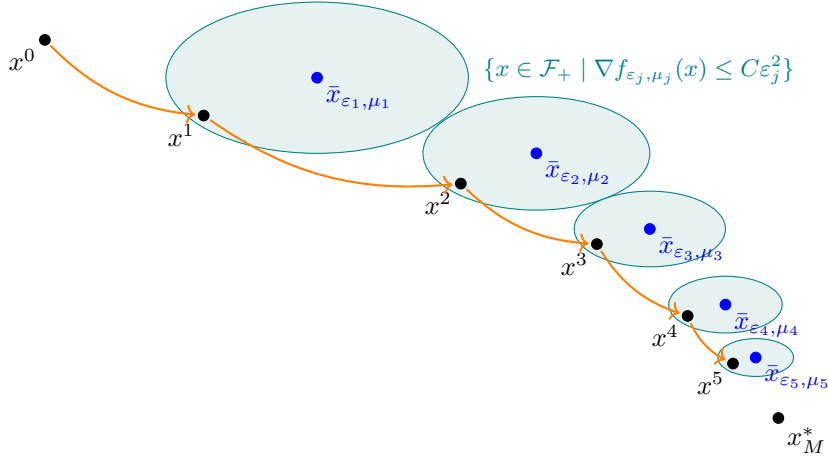
and thus

$$\|x - \bar{x}_{\varepsilon_j, \mu_j}\|_2 \leq \frac{1}{\varepsilon_j} \|\nabla f_{\varepsilon_j, \mu_j}(x)\|_2. \quad (3.13)$$

Hence,  $\|x^j - \bar{x}_{\varepsilon_j, \mu_j}\|_2 \leq \frac{1}{\varepsilon_j} \|\nabla f_{\varepsilon_j, \mu_j}(x^j)\|_2$ , so when  $\|\nabla f_{\varepsilon_j, \mu_j}(x^j)\|_2 = o(\varepsilon_j)$  as  $j \rightarrow \infty$ , the difference goes to 0.

Finally, with  $\mu_j = o(\varepsilon_j)$ , Theorem 3.6 b) provides that  $\lim_{j \rightarrow \infty} \bar{x}_{\varepsilon_j, \mu_j} = x_M^*$ , and the convergence of the iterates  $x^j$  to  $x_M^*$  follows.  $\square$

Hence, instead of solving the optimisation problem  $(\mathcal{P}_{\varepsilon_j, \mu_j})$  exactly in every iteration, it suffices that the iterates  $x^j$  only lie within a certain neighbourhood of their respective regularised barrier minimiser  $\bar{x}_{\varepsilon_j, \mu_j}$ . This neighbourhood can be characterised by the gradient norm  $\|\nabla f_{\varepsilon_j, \mu_j}(\cdot)\|_2$  lying below a certain threshold, and this threshold must be reduced faster than the regularisation parameter  $\varepsilon_j$  during the iteration. An illustration of this concept is given in Figure 3.3.



**Figure 3.3:** Illustration of the iteration progress in the Tikhonov-regularised IPM. For each iteration  $j$ , the regularised barrier minimiser  $\bar{x}_{\varepsilon_j, \mu_j}$  (blue) is depicted with its corresponding neighbourhood (teal) at which the norm of the gradient of  $f_{\varepsilon_j, \mu_j}$  lies below a certain threshold. This threshold is defined by  $C\varepsilon_j^2$  for some  $C > 0$ , hence, the size of the neighbourhood shrinks with progressing iteration. In each step, an approximation  $x^j$  to  $\bar{x}_{\varepsilon_j, \mu_j}$  is computed such that it lies in the corresponding neighbourhood, resulting in a progression as depicted in orange. Ultimately, both  $\bar{x}_{\varepsilon_j, \mu_j}$  and  $x^j$  converge to the minimal-norm solution  $x_M^*$ , as stated in Theorem 3.9.

For this convergence result, it is essential that both the barrier parameter  $\mu$  as well as the allowed approximation error are ultimately dominated by the regularisation parameter  $\varepsilon$ ,



making the latter the central parameter in the method. The reason for this necessity on the one hand lies in the goal to obtain convergence to  $x_M^*$ , and on the other hand is given by the fact that the regularisation ensures strong convexity, and that this strong convexity is central for the neighbourhood around  $\bar{x}_{\varepsilon,\mu}$  to suffice for the convergence. Therefore, the regularisation parameter must not decrease too fast.

### 3.4.2 Rate of convergence

Theorem 3.9 guarantees general convergence of the iterates  $x^j$ , but it neither gives a statement on the accuracy of the iterates or the rate of the convergence, nor does it yield a suggestion for how fast the barrier parameter and residual gradient norm should be reduced with respect to  $\varepsilon$ , other than that it has to be faster than  $\varepsilon$ . And lastly, the question remains how well an iterate  $x^j$  approximates  $x_M^*$ .

In the last sections, estimates on the convergence rate in the objective  $f$  and on the actual deviation from the minimum were developed for the different components of the regularised barrier method of Algorithm 3.1:

- Lemma 2.17 bounds the difference for barrier methods by  $f(\bar{x}_\mu) - f(x^*) \leq m\mu$ ,
- Lemma 3.2 a) yields an estimate for Tikhonov-regularised methods of  $f(\bar{x}_\varepsilon) - f(x^*) \leq \frac{\varepsilon}{2} \|x_M^*\|_2^2$ , and
- in Lemma 3.8 it was proven that the regularised barrier method, when solved exactly in every iteration, satisfies  $f(\bar{x}_{\varepsilon,\mu}) - f(x^*) \leq m\mu + \varepsilon \|x_M^*\|_2^2$ .

Based on that, an estimate on the deviation in the objective function for the inexactly computed iterates  $x^j$  of the regularised barrier method of Algorithm 3.1 is derived next. The estimate then yields a convergence rate of  $O(\varepsilon_j)$  as well as a possible stopping criterion, both presented in Theorem 3.11.

Since the iterates  $x^j$  are approximations to  $\bar{x}_{\varepsilon_j,\mu_j}$ , and their approximation accuracy in terms of  $\|\nabla f_{\varepsilon_j,\mu_j}(x^j)\|_2$  can be controlled in the algorithm, the already given knowledge on the deviation of  $\bar{x}_{\varepsilon_j,\mu_j}$  is used for the estimate. This yields the following upper bound:

**Lemma 3.10.** *Let Assumption 2.10 hold, and let  $x^* \in \mathcal{S}$ . For any  $\varepsilon, \mu > 0$  and any  $x \in \mathcal{F}_+$ , it holds:*

$$0 \leq f(x) - f(x^*) \leq \frac{1}{\varepsilon} \|\nabla f(x)\|_2 \|\nabla f_{\varepsilon,\mu}(x)\|_2 + m\mu + \varepsilon \|x_M^*\|_2^2.$$

*Proof.* The first inequality follows directly from  $x$  being feasible and  $x^*$  being a minimiser of  $f$  on the feasible set. The upper bound will be proven by reducing it to the already known result on the function value deviation in the minimiser  $\bar{x}_{\varepsilon,\mu}$  of the corresponding problem  $(\mathcal{P}_{\varepsilon,\mu})$ . Its existence follows directly from  $\mathcal{F}_+ \neq \emptyset$ , see Theorem 3.4.

The difference can be split into

$$f(x) - f(x^*) = (f(x) - f(\bar{x}_{\varepsilon,\mu})) + (f(\bar{x}_{\varepsilon,\mu}) - f(x^*)).$$

For the second part, it was already shown in Lemma 3.8 that

$$f(\bar{x}_{\varepsilon,\mu}) - f(x^*) \leq m\mu + \varepsilon \|x_M^*\|_2^2.$$



For the first difference, on the other hand, convexity of  $f$ , the Cauchy-Schwarz inequality [37, page 600], and the estimate (3.13) from the proof of Theorem 3.9 yield

$$\begin{aligned} f(x) - f(\bar{x}_{\varepsilon,\mu}) &\stackrel{\text{Lemma 2.3 b)}}{\leq} \nabla f(x)^T (x - \bar{x}_{\varepsilon,\mu}) \\ &\leq \|\nabla f(x)\|_2 \|x - \bar{x}_{\varepsilon,\mu}\|_2 \\ &\stackrel{(3.13)}{\leq} \frac{1}{\varepsilon} \|\nabla f(x)\|_2 \|\nabla f_{\varepsilon,\mu}(x)\|_2. \end{aligned}$$

Taking both estimates together, one finally obtains

$$\begin{aligned} f(x) - f(x^*) &= (f(x) - f(\bar{x}_{\varepsilon,\mu})) + (f(\bar{x}_{\varepsilon,\mu}) - f(x^*)) \\ &\leq \frac{1}{\varepsilon} \|\nabla f(x)\|_2 \|\nabla f_{\varepsilon,\mu}(x)\|_2 + m\mu + \varepsilon \|x_M^*\|_2^2. \quad \square \end{aligned}$$

This estimate now allows for three things: Firstly, it can be used to prove that for convergence of the Tikhonov-regularised IPM to the minimal function value, it suffices that the barrier parameter  $\mu_j$  lies in  $O(\varepsilon_j)$  and it holds  $\|\nabla f_{\varepsilon_j,\mu_j}(x^j)\|_2 = o(\varepsilon_j)$ ; the stronger assumption  $\mu_j = o(\varepsilon_j)$  as in Theorem 3.9 is not required for that. Secondly, if the accuracy of the approximations  $x^j$  is increased more rapidly at a rate of  $\|\nabla f_{\varepsilon_j,\mu_j}(x^j)\|_2 = O(\varepsilon_j^2)$ , then the convergence of the method in the objective  $f$  is characterised by a rate in  $O(\varepsilon_j)$ .

And finally, the estimate in Lemma 3.10 gives an answer to the question how well the current iterate  $x^j$  at least approximates a minimiser of the original problem ( $\mathcal{P}$ ). Except for the typically unknown term  $\|x_M^*\|_2^2$ , the bound can easily be computed, and since the iterates  $x^j$  converge to  $x_M^*$  (see Theorem 3.9), the latter may be bounded by twice the size of  $x^j$  for sufficiently large  $j$ . This yields an estimate for the accuracy of  $x^j$  which makes for a potential stopping criterion in the overall algorithm.

These results are summarised in the next theorem.

**Theorem 3.11.** *Let Assumption 2.10 hold, and let  $x^* \in \mathcal{S}$ . Consider the Tikhonov-regularised interior-point method of Algorithm 3.1 with parameter sequences  $(\varepsilon_j)_{j \in \mathbb{N}}$  and  $(\mu_j)_{j \in \mathbb{N}}$  where  $\varepsilon_j \rightarrow 0^+$  and  $\mu_j \rightarrow 0^+$  as  $j \rightarrow \infty$ .*

a) *If  $\mu_j = O(\varepsilon_j)$  and  $\|\nabla f_{\varepsilon_j,\mu_j}(x^j)\|_2 = o(\varepsilon_j)$  for  $j \rightarrow \infty$ , then*

$$\lim_{j \rightarrow \infty} f(x^j) = f(x^*),$$

*so the method converges to the optimal value of problem ( $\mathcal{P}$ ).*

b) *If  $\mu_j = O(\varepsilon_j)$  and furthermore  $\|\nabla f_{\varepsilon_j,\mu_j}(x^j)\|_2 = O(\varepsilon_j^2)$  for  $j \rightarrow \infty$ , then*

$$f(x^j) - f(x^*) = O(\varepsilon_j),$$

*so the rate of convergence in the objective  $f$  lies in  $O(\varepsilon_j)$ .*

c) *If  $\mu_j = o(\varepsilon_j)$  and  $\|\nabla f_{\varepsilon_j,\mu_j}(x^j)\|_2 = o(\varepsilon_j)$  for  $j \rightarrow \infty$ , then it holds for sufficiently large  $j$  that*

$$f(x^j) - f(x^*) \leq \frac{1}{\varepsilon_j} \|\nabla f(x^j)\|_2 \|\nabla f_{\varepsilon_j,\mu_j}(x^j)\|_2 + m\mu_j + 2\varepsilon_j \|x^j\|_2^2. \quad (3.14)$$

*Proof.*

- a) First note that by Theorem 3.6 a), the assumption  $\mu_j = O(\varepsilon_j)$  implies that the sequence  $(\bar{x}_{\varepsilon_j, \mu_j})_{j \in \mathbb{N}}$  remains bounded. Hence, with  $\|\nabla f_{\varepsilon_j, \mu_j}(x^j)\|_2 = o(\varepsilon_j)$  and (3.13), the iterates  $x^j$  also remain bounded and thus all lie in some compact set. Since  $f$  is continuously differentiable, its gradient is bounded on any compact set [29, page 31] and therefore, the term  $\|\nabla f(x^j)\|_2$  is bounded for all  $j$  by some constant. It follows that all summands in the upper bound of the estimate

$$0 \leq f(x^j) - f(x^*) \leq \frac{1}{\varepsilon_j} \|\nabla f(x^j)\|_2 \|\nabla f_{\varepsilon_j, \mu_j}(x^j)\|_2 + m\mu_j + \varepsilon_j \|x_M^*\|_2^2$$

of Lemma 3.10 approach 0 as  $j \rightarrow \infty$ , so  $f(x^j)$  converges to  $f(x^*)$ .

- b) This also proves the second statement, as the assumption  $\|\nabla f_{\varepsilon_j, \mu_j}(x^j)\|_2 = O(\varepsilon_j^2)$  now additionally yields that in above upper bound, all three summands are characterised by a convergence rate of  $O(\varepsilon_j)$ .
- c) Finally, with  $\mu_j = o(\varepsilon_j)$  and  $\|\nabla f_{\varepsilon_j, \mu_j}(x^j)\|_2 = o(\varepsilon_j)$  for  $j \rightarrow \infty$ , the requirements of Theorem 3.9 are satisfied. Thus, the iterates  $x^j$  converge to  $x_M^*$ , so for sufficiently large  $j$  it holds  $\|x_M^*\|_2^2 \leq 2\|x^j\|_2^2$ . The estimate (3.14) then follows directly from Lemma 3.10.  $\square$

*Remark.*

- Under the assumptions of parts a) and b), convergence is only obtained in the function value; the iterates do not necessarily converge. The requirements for Theorem 3.9 include that  $\mu_j$  must be reduced strictly faster than  $\varepsilon_j$ , so only in part c), convergence of the iterates to  $x_M^*$  is guaranteed.
- In the estimate (3.14), all terms are already known or can easily be computed in every iteration. Therefore, it can be used as stopping criterion for Algorithm 3.1. However, one should keep in mind that the estimate might not hold for the first (unknown many) iterations.

### 3.4.3 Specification of the outer method

Based on these observations, a reasonable rule for the approximation accuracy and following parameter update in each iteration of Algorithm 3.1 is given by

$$\|\nabla f_{\varepsilon_j, \mu_j}(x^j)\|_2 \stackrel{!}{\leq} C\varepsilon_j^2, \quad \varepsilon_{j+1} \leftarrow \beta\varepsilon_j, \quad \mu_{j+1} \leftarrow \beta^\gamma\mu_j \quad (3.15)$$

for some update parameters  $\beta \in (0, 1)$ ,  $\gamma > 1$  and constant  $C > 0$ .

With these choices, it holds  $\|\nabla f_{\varepsilon_j, \mu_j}(x^j)\|_2 = O(\varepsilon_j^2)$  and  $\mu_j = o(\varepsilon_j)$ , so all requirements of Theorems 3.9 and 3.11 are satisfied and convergence of the iterates to  $x_M^*$  as well as a linear convergence rate in the objective is obtained for Algorithm 3.1. Furthermore, the update choices for  $\varepsilon_j$  and  $\mu_j$  imply that  $\mu_j = \beta^{j(\gamma-1)} \frac{\mu_0}{\varepsilon_0} \varepsilon_j < \frac{\mu_0}{\varepsilon_0} \varepsilon_j$ , so (3.14), holding for

sufficiently large  $j \in \mathbb{N}$ , is available as possible stopping criterion and takes the form

$$f(x^j) - f(x^*) \leq \varepsilon_j \left( C \|\nabla f(x^j)\|_2 + 2 \|x^j\|_2^2 \right) + m\mu_j \quad (3.16)$$

$$\begin{aligned} &= \varepsilon_j \left( C \|\nabla f(x^j)\|_2 + 2 \|x^j\|_2^2 + \beta^{j(\gamma-1)} m \frac{\mu_0}{\varepsilon_0} \right) \\ &< \beta^j \left( C\varepsilon_0 \|\nabla f(x^j)\|_2 + 2\varepsilon_0 \|x^j\|_2^2 + m\mu_0 \right). \end{aligned} \quad (3.17)$$

The last inequality thereby highlights the linear convergence rate in the objective  $f$  with respect to the outer iterations.

As conclusion to this chapter, the entire procedure of the Tikhonov-regularised IPM with the above choices for the update and stopping criteria is presented in Algorithm 3.2. It describes the outer structure of the method proposed in this thesis in detail.

---

**Algorithm 3.2** Outer loop of the Tikhonov-regularised interior-point method

---

**Given:** Optimisation problem  $(\mathcal{P})$ , tolerance  $\tau > 0$

**Choose:** Starting point  $x^0 \in \mathcal{F}_+$ , initial parameters  $\varepsilon_1, \mu_1 > 0$ , update parameters  $\beta \in (0, 1)$ ,  $\gamma > 1$ , constant  $C > 0$

1: **for**  $j = 1, 2, 3, \dots$  **do**

2:   Compute  $x^j$  satisfying  $\|\nabla f_{\varepsilon_j, \mu_j}(x^j)\|_2 \leq C\varepsilon_j^2$  using the inner solver described in Algorithm 4.2, starting at  $x^{j-1}$

3:   **if**  $\varepsilon_j \left( C \|\nabla f(x^j)\|_2 + 2 \|x^j\|_2^2 \right) + m\mu_j \leq \tau$  **then** let  $x^* \leftarrow x^{j-1}$  and **break**

4:   Update  $\varepsilon_{j+1} \leftarrow \beta\varepsilon_j$ ;  $\mu_{j+1} \leftarrow \beta^\gamma\mu_j$

5: **end for**

**Return:**  $x^*$

---

With the results derived in this section, it follows directly that the algorithm is well-defined and terminates at an approximation to a solution of  $(\mathcal{P})$  after finitely many outer iterations. Furthermore, the iterates  $x^j$  converge to the minimal norm solution  $x_M^*$ , and this convergence is  $r$ -linear in the objective  $f$  with respect to the number of outer iterations and quantified by the convergence rate  $\beta$ .

**Theorem 3.12.** *Let Assumption 2.10 hold. Then the Tikhonov-regularised interior-point method of Algorithm 3.2 is well-defined and terminates after finitely many outer iterations with an approximation  $x^* \in \mathcal{F}_+$  to a solution of  $(\mathcal{P})$ . If the tolerance  $\tau$  is chosen sufficiently small, then  $x^*$  satisfies*

$$f(x^*) - \min_{x \in \mathcal{F}} f(x) \leq \tau.$$

*Furthermore, if the algorithm is applied with  $\tau = 0$ , then it either terminates after finitely many outer iterations with the minimal-norm solution  $x^* = x_M^*$ , or the outer iterates  $x^j$  converge to  $x_M^*$  with an  $r$ -linear convergence rate  $\beta \in (0, 1)$  in the objective  $f$ .*

*Proof.* Assumption 2.10 ensures that a feasible starting point  $x^0 \in \mathcal{F}_+$  exists. With Theorem 3.4, the regularised barrier minimiser  $\bar{x}_{\varepsilon_j, \mu_j}$  further exists for all parameters  $\varepsilon_j, \mu_j > 0$ . Thus, for each  $j \in \mathbb{N}$ , there exists also an outer iterate  $x^j \in \mathcal{F}_+$  which satisfies  $\|\nabla f_{\varepsilon_j, \mu_j}(x^j)\|_2 \leq C\varepsilon_j^2$ . Hence, Algorithm 3.2 is well-defined.

The required “inner approximation accuracy” (line 2) and the parameter update (line 4) of Algorithm 3.2 satisfy  $\|\nabla f_{\varepsilon_j, \mu_j}(x^j)\|_2 = O(\varepsilon_j^2)$  and  $\mu_j = o(\varepsilon_j)$ . Therefore, Theorem 3.9 implies that  $\lim_{j \rightarrow \infty} x^j = x_M^*$ , so in particular, the outer iterates  $(x^j)_{j \in \mathbb{N}_0}$  and thereby also the gradient of  $f$  evaluated at those iterates are bounded.

With  $\varepsilon_j \rightarrow 0^+$  and  $\mu_j \rightarrow 0^+$  as  $j \rightarrow \infty$ , it follows that the left-hand side in the stopping criterion in line 3 will at some point fall below the fixed tolerance  $\tau$ , so Algorithm 3.2 terminates after finitely many iterations for any  $\tau > 0$ . For the resulting point  $x^*$ , (3.16) then implies that  $f(x^j) - \min_{x \in \mathcal{F}} f(x) \leq \tau$  for sufficiently large  $j \in \mathbb{N}$ . As the left-hand side in the stopping criterion is at least  $m\mu_j$ , this iteration number  $j$  is guaranteed to be reached in Algorithm 3.2 when the tolerance  $\tau$  is sufficiently small.

In case that  $\tau = 0$ , on the other hand, the r-linear convergence rate in the objective follows directly from (3.17) and the boundedness of  $x^j$  and  $\|\nabla f(x^j)\|_2$ .  $\square$

Algorithm 3.2 only covers the *outer iteration loop* of the overall method. For the actual computation of the corresponding *outer iterates*  $x^j$  in line 2, a structured l-BFGS method is employed. This *inner solver* will be introduced, described and analysed in detail in Chapter 4.

## 3.5 Modifications

In the following, possible modifications to the Tikhonov-regularised IPM of Algorithm 3.2 are discussed, regarding different regularisation strategies and the integration of linear equality constraints.

### 3.5.1 Shifted regularisation

The Tikhonov regularisation ensures that the iterates produced by Algorithm 3.2 converge to the minimal-norm solution of  $x_M^*$ , meaning that the solution set  $\mathcal{S}$  is optimised for the minimal distance to the origin. If, on the other hand, a priori information on the problem are given which suggest that one desires the solution to be as close as possible to some other point  $x^{(0)} \in \mathbb{R}^n$ , or which indicate that an improved convergence behaviour could be obtained when approaching the solution set from a direction closer to  $x^{(0)}$  instead of the origin, then the regularisation term can be changed as follows.

Instead of including  $\frac{\varepsilon}{2} \|x\|_2^2$  in the function  $f_{\varepsilon, \mu}$ , the “shifted regularisation” term  $\frac{\varepsilon}{2} \|x - x^{(0)}\|_2^2$  can be used. With this modification, the limit point  $x_M^*$  is then given as the unique projection of  $x^{(0)}$  onto  $\mathcal{S}$ . The modification is also compatible with all theoretical results presented in Chapter 3, with the only difference that, wherever the norm of a point  $x \in \mathbb{R}^n$  occurs, it needs to be replaced by  $\|x - x^{(0)}\|_2$ . This adjustment in the norm does not apply to those norms in which a difference of the kind  $\|x - y\|_2$  is considered, or in which the norm of a gradient or matrix is computed.

### 3.5.2 Proximal-point regularisation

A more substantial modification is given by the replacement of the Tikhonov regularisation with a proximal-point regularisation. In this, the centre of the regularisation changes

with every outer iteration  $j$  to the respective latest outer iterate  $x^{j-1}$ , meaning that the term  $\frac{\varepsilon_j}{2} \|x\|_2^2$  is replaced in the function  $f_{\varepsilon_j, \mu_j}$  by the proximal term  $\frac{\varepsilon_j}{2} \|x - x^{j-1}\|_2^2$ , cf. [21, Section 6.4.2].

A proximal-point regularisation changes the behaviour of IPMs less severely than the Tikhonov regularisation, as it does not extrinsically impose a preference towards the origin, but rather damps the step undertaken in each outer iteration. Due to that, it is also not compatible with most of the convergence analysis for the outer iterates which was conducted in this section. Although the regularised barrier function  $f_{\varepsilon_j, \mu_j}$  remains strongly convex and still has a unique minimiser in every iteration, the limit point of the outer iterates can neither be classified nor can even its existence be proven in the same way as in Section 3.3.

Convergence of proximal IPMs under certain conditions was first proven in [27], including an r-linear convergence rate in the objective. Since then, they are commonly used in different variants and applications, see e.g. [12, 13, 41]. A possible benefit of proximal-point methods is that they do not require the regularisation parameter  $\varepsilon_j$  to approach 0 in order to converge to a minimiser of  $(\mathcal{P})$ , but instead allow for the parameter to be kept above a sufficiently small positive value, cf. [21, page 357]. Thus, they can maintain a certain strong convexity throughout the entire method.

Even though the convergence results of Chapter 3 do not apply for proximal-point methods, the latter can still be integrated in Algorithm 3.2, and as long as the solution set  $\mathcal{S}$  is bounded or convergence of the outer iterates guaranteed otherwise, the left-hand side in the outer stopping criterion (line 3 of Algorithm 3.2) approaches 0 as  $j \rightarrow \infty$ . However, it can not be guaranteed that it satisfies the estimate (3.16).

### 3.5.3 Linear equality constraints

Often, interior-point methods not only support the solution of inequality-constrained problems, but can also deal with linear equality constraints [8, Chapter 11]. If such a convex optimisation problem

$$\inf_{x \in \mathbb{R}^n} f(x) \quad \text{s.t.} \quad Ax = b, \quad g(x) \leq 0 \quad (\mathcal{P}^{A,b})$$

is given where  $A \in \mathbb{R}^{p \times n}$  with  $\text{rank } A = p < n$  and  $b \in \mathbb{R}^p$ , then the barrier approach can be used to deal with the inequality constraints, while the equality constraints are explicitly enforced.

For the regularised barrier problem  $(\mathcal{P}_{\varepsilon, \mu})$ , defined in Section 3.2, this leads to the constrained modification

$$\inf_{x \in \mathbb{R}^n} f_{\varepsilon, \mu}(x) \quad \text{s.t.} \quad Ax = b, \quad g(x) < 0.$$

With the correspondingly modified sets  $\mathcal{F}, \mathcal{F}_+$  and  $\mathcal{S}$ , and the adapted KKT conditions (see e.g. [8, equations (11.2) and (11.4)]), one can conduct the convergence theory for the minimisers  $\bar{x}_{\varepsilon, \mu}$  of the modified Tikhonov-regularised IPM similarly to Section 3.3, and analogously obtains convergence to the minimal-norm solution.

However, in order to adopt the convergence analysis for the inexactly solved outer iterates  $x^j$ , the equality constraints  $Ax^j = b$  must be exactly satisfied in every outer iteration. For this, the convex problem  $(\mathcal{P}^{A,b})$  can be reduced to a purely inequality constraint one by

limiting the optimisation to an affine subspace instead of operating on the entire  $\mathbb{R}^n$ ; specifically, to the space  $\{Fz + x_0 \mid z \in \mathbb{R}^{n-p}\}$  where  $x_0 \in \mathbb{R}^n$  is a particular solution of  $Ax = b$ , and  $F \in \mathbb{R}^{n \times (n-p)}$  is a matrix whose range is the nullspace of  $A$ , cf. [8, Section 4.2.4]. For the thereby obtained reduced problem

$$\inf_{z \in \mathbb{R}^{n-p}} f(Fz + x_0) \quad \text{s.t.} \quad g(Fz + x_0) \leq 0,$$

the Tikhonov-regularised IPM discussed in this thesis can be employed equivalently to  $(\mathcal{P})$ . This concept was followed in [4], for example, but as stated in [8, page 143], in many cases, it is better to retain the equality constraints than to reduce the problem in this manner.

In contrast to the reduction, typical Newton based IPMs include the equality constraints directly into the primal or primal-dual linear system of equations [8, 24, 37]. As soon as a full Newton step is accepted once in the process, the equality constraints are and remain satisfied, and in practice, it is common that this is the case before the optimality criteria are met [24, page 593]. However, this would imply to solve linear systems of equations, which is avoided in the method discussed in this thesis.

### 3.5.4 Augmented Lagrangian approach

An alternative approach is the so-called *method of multipliers* or *Augmented Lagrangian method* [37, Section 17.3]. It consists of two parts: As in penalty methods [37, Section 17.1], the term  $\frac{\rho_j}{2} \|Ax - b\|_2^2$  is added to the objective with an increasing penalty parameter  $\rho_j > 0$ , yielding that the constraint will be satisfied asymptotically. Furthermore, in every step  $j$ , a multiplier  $\lambda^j \in \mathbb{R}^p$  is computed and included in the optimisation which estimates the Lagrange multiplier which corresponds to the equality constraints [37, Section 17.3].

In the context of this thesis, the integration of an Augmented Lagrangian approach into the Tikhonov-regularised IPM yields the function

$$f_{\varepsilon_j, \rho_j, \mu_j; \lambda^j}(x) := f(x) + \frac{\varepsilon_j}{2} \|x\|_2^2 - (\lambda^j)^T (Ax - b) + \frac{\rho_j}{2} \|Ax - b\|_2^2 - \mu_j \sum_{i=1}^m \ln(-g_i(x)),$$

where  $\lambda^j \in \mathbb{R}^p$  is the current approximation to the Lagrange multiplier, and  $\varepsilon_j, \rho_j, \mu_j > 0$  are the regularisation, penalty and barrier parameter, respectively. Following [37, Section 17.3], the multiplier is updated in every outer iteration by

$$\lambda^{j+1} := \lambda^j - \rho_j (Ax - b).$$

The penalty parameter  $\rho_j$ , on the other hand, is increased in every iteration, e.g. by setting  $\rho_{j+1} := \frac{1}{\beta^\gamma} \rho_j$ , yielding that it increases at the same speed as  $\mu_j$  is being reduced. To avoid further ill-conditioning, the penalty parameter is only updated when the current residual  $\|Ax^j - b\|_2^2$  does not lie below an imposed tolerance.

When the Tikhonov regularisation is replaced by a proximal-point regularisation as discussed above, then this combination of Augmented Lagrangian, proximal-point and interior-point method leads to the so-called *interior-point proximal method of multipliers (IP-PM)*. These methods have recently been discussed in different contexts [7, 31, 41, 42].

# 4

## A structured l-BFGS method for solving the inner problems

In Chapter 3, the Tikhonov-regularised interior-point method for solving the general inequality-constrained convex problem ( $\mathcal{P}$ ) was introduced. It was shown how, for certain choices of the parameters  $\varepsilon$  and  $\mu$ , the outer iterates  $x^j$  converge to the minimal-norm solution  $x_M^*$ , and that a convergence rate of  $O(\varepsilon_j)$  is obtained in the objective  $f$ . Necessary for this convergence rate is that in each outer iteration  $j$ , an iterate  $x^j$  is computed whose accuracy increases at a rate of  $\|\nabla f_{\varepsilon_j, \mu_j}(x^j)\|_2 = O(\varepsilon_j^2)$ .

This chapter therefore covers the second main topic of this thesis: the question how a structured l-BFGS method can be used to solve the inner problems of the Tikhonov-regularised IPM, and how convergence of the inner iterates can be obtained. As a starting point, in Section 4.1, the concept behind Quasi-Newton methods is briefly described, and the BFGS update and its limited-memory variant are introduced. Based on that, Section 4.2 covers the structured l-BFGS method analysed in [34]. In Section 4.3, it is described in detail how this method is used as inner solver for the proposed Tikhonov-regularised IPM, and the precise procedure is presented in Algorithm 4.2.

The rest of the chapter deals with the convergence analysis of the inner iterations: In Section 4.4, it is proven that each inner loop terminates after finitely many iterations at a sought point  $x^j \in \mathcal{F}_+$ , yielding that also the overall method terminates after finitely many iterations, and that each inner method features an r-linear convergence rate. The main result is presented in Theorem 4.5.

In Section 4.5, the convergence of the inner iterations is quantified in the context of the overall method, focussing on the question how the number of inner iterations required at each outer step evolves throughout the method. The discussion ultimately culminates in Section 4.6 with the result of Theorem 4.12 in which it is proven that the total number of l-BFGS iterations required for the entire Tikhonov-regularised IPM grows at most polynomially in the imposed tolerance  $\tau$ .

### 4.1 L-BFGS methods

In this section, the concept of Quasi-Newton methods is briefly introduced, followed by a description of the well-known BFGS-update and the limited-memory variant arising from this. The descriptions and statements in this section mainly follow [37, Chapters 6 and 7].

### 4.1.1 Quasi-Newton methods

In this thesis, Quasi-Newton methods are used to find a minimiser of an unconstrained and twice differentiable function  $\varphi \in C^2(\mathbb{R}^n, \mathbb{R})$ . Starting point for these methods is the idea to iteratively approximate the objective  $\varphi$  locally around the current iterate  $x_k \in \mathbb{R}^n$  by a strongly convex quadratic function  $Q_k$ , defined by

$$\varphi(x_k + p) \approx Q_k(p) := \varphi(x_k) + \nabla\varphi(x_k)^T p + \frac{1}{2} p^T B_k p, \quad (4.1)$$

where  $B_k \in \mathbb{R}^{n \times n}$  is a symmetric positive definite (spd) matrix.

This quadratic function becomes minimal for

$$0 = \nabla Q_k(p) = \nabla\varphi(x_k) + B_k p,$$

yielding that its unique minimiser ( $Q_k$  is strongly convex by Lemma 2.5 d)) is given by [37, page 136]

$$p_k = -B_k^{-1} \nabla\varphi(x_k).$$

In Quasi-Newton methods, the minimiser  $p_k$  is then used as search direction for the iteration, and the new iterate is obtained by

$$x_{k+1} = x_k + \alpha_k p_k$$

with a step size  $\alpha_k > 0$ . Thereby,  $p_k$  is guaranteed to be a descent direction in  $\varphi$ , since  $B_k$  and hence  $B_k^{-1}$  are positive definite and therefore  $p_k^T \nabla\varphi(x_k) < 0$ .

When  $B_k = \nabla^2\varphi(x_k)$ , that is,  $B_k$  is the Hessian matrix of  $\varphi$  in  $x_k$ , then (4.1) corresponds to the second-order Taylor approximation [37, Theorem 2.1] of  $\varphi$  in  $x_k$ . In this case, the quadratic approximation  $Q_k$  is as accurate as possible among all quadratic functions in the proximity of  $x_k$ , and  $p_k$  is the so-called *Newton direction* which is typically used in Newton's method [37, page 22]. However, a drawback of Newton's method is that, in each iteration, it requires the computation of the Hessian  $\nabla^2\varphi(x_k)$  and the solution of a linear system of equations with it, in order to obtain  $p_k$ . For some problems, this might be impossible as the Hessian is unknown, or, especially in higher dimensions, comes at a prohibitive computational cost [4, 40].

The idea behind Quasi-Newton methods is to use a matrix  $B_k$  as approximation to the Hessian  $\nabla^2\varphi(x_k)$  which is viable and easier than the Hessian to compute, and from which the search direction  $p_k$  can be obtained with acceptable cost. Starting with an initial guess  $B_0$  this approximation to the Hessian is then in each iteration updated to a new matrix  $B_{k+1}$  such that it describes the Hessian at the new iterate  $x_{k+1}$  in a useful way.

Hence, the question arises how the information gained in the iteration could be used for the update. A basic condition for the updated matrix  $B_{k+1}$  is that the next quadratic approximation  $Q_{k+1}$ , which arises from it, should not only have the same gradient as  $\varphi$  in  $x_{k+1}$ , but also in the previous iterate  $x_k$  [37, page 137]. This yields

$$\nabla Q_{k+1}(-\alpha_k p_k) = \nabla\varphi(x_{k+1}) - \alpha_k B_{k+1} p_k \stackrel{!}{=} \nabla\varphi(x_k). \quad (4.2)$$



By defining the latest update in the iterate and the difference of the gradients by

$$s_k := x_{k+1} - x_k = \alpha_k p_k, \quad y_k := \nabla\varphi(x_{k+1}) - \nabla\varphi(x_k),$$

the condition (4.2) is equivalent to the so-called *secant equation* [37, page 137]

$$B_{k+1}s_k = y_k. \tag{4.3}$$

Evidently, if  $B_{k+1}$  is positive definite and satisfies the secant equation, then  $s_k$  and  $y_k$  must satisfy the *curvature condition*  $s_k^T y_k > 0$ . This is intrinsically given when  $\varphi$  is strongly convex, as it follows directly from Lemma 2.5 c), and hence must not be taken into consideration in this thesis. Otherwise, the curvature condition could for example be ensured by a line search which is conform to the Wolfe conditions [37, page 138]. It should be noted that (4.3) is not generally satisfied by the Hessian  $\nabla^2\varphi$  at  $x_{k+1}$ , but only by its average between  $x_k$  and  $x_{k+1}$ ; for details, see [37, page 138]. Nonetheless, the secant equation proved to be a reasonable imposition on  $B_{k+1}$ .

In general, an optimisation method is called *Quasi-Newton method* if it follows the concept of iteratively approximating the objective by a quadratic function, and if the quadratic matrix used for this satisfies the secant equation (4.3), cf. [46, page 65]. However, by only requiring the secant equation, still many degrees of freedom are left for possible choices of  $B_{k+1}$  [37, page 138].

#### 4.1.2 The l-BFGS update

There are several well-known Quasi-Newton update formulas. Among these, the *BFGS update* is the one that is most widely used [37, page 139]. It is named after its inventors Broyden, Fletcher, Goldfarb and Shanno who publicised the concept all independently of each other in 1970 [9, 17, 23, 45].

The BFGS update for the matrix  $B_k$  is defined as [37, page 140]

$$B_{k+1} = B_k - \frac{(B_k s_k)(B_k s_k)^T}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k}, \tag{4.4}$$

and equivalently for the inverse matrix  $H_k := B_k^{-1}$  by

$$H_{k+1} = \left( I - \frac{s_k y_k^T}{s_k^T y_k} \right) H_k \left( I - \frac{y_k s_k^T}{s_k^T y_k} \right) + \frac{s_k s_k^T}{s_k^T y_k}. \tag{4.5}$$

It can be shown that  $H_{k+1}$  is, in some sense, precisely the symmetric matrix which is most similar to its predecessor  $H_k$  while satisfying the secant equation, for details see [37, page 138].

In general, the BFGS update formula is considered most effective for line-search based Quasi-Newton methods and has proven to yield robust results [37, pages 139–140]. It is a rank-two update which preserves symmetry and positive definiteness [37, page 141] and “has very effective self-correcting properties” [37, page 142].

In Quasi-Newton methods, it is usually not necessary to know and save the, generally dense, matrix  $B_k$  or, analogously, its inverse  $H_k$  explicitly. Instead, only the search direction  $p_k$  must be obtained in every iteration by computing

$$p_k = -H_k \nabla\varphi(x_k).$$

When  $H_k$  is defined by the inverse BFGS update (4.5), this computation can be reduced to a recursive approach which only involves vector operations and one matrix-vector-multiplication with the initial matrix  $H_0$  [37, Algorithm 7.4]. Thus, instead of computing and saving each matrix  $H_k$ , it suffices to save all update vector pairs  $(s_i, y_i)_{i=0, \dots, k}$  as well as  $H_0$ , and to compute the search direction  $p_k$  recursively from them.

However, the memory and computation requirement for this approach grows linearly with the number of iterations  $k$ , while especially the first information might only contribute very little to the approximation quality of  $H_k$ . This yields a limited-memory modification of it: the so-called *l-BFGS methods*, first proposed by Nocedal in 1980 [36]. In these, only the last  $\ell$  vector pairs  $(s_i, y_i)_{i=k-\ell, \dots, k-1}$  are stored and used for the computation of  $p_k$ , where the memory length is given by  $\ell := \min\{k, \ell_{\max}\}$  for some  $\ell_{\max} \in \mathbb{N}_0$ . Thus, in the first  $\ell_{\max}$  iterations, the l-BFGS update coincides with the standard BFGS update, and once  $k > \ell_{\max}$ , the oldest information are disregarded.

As described in [37], the main idea behind this is to “use curvature information from only the most recent iterations to construct the Hessian approximation”, whereas the curvature information from earlier iterations “is less likely to be relevant to the actual behaviour of the Hessian at the current iteration” [37, page 176] and can therefore be discarded in order to limit the memory and computation costs.

### 4.1.3 Seed matrix and two-loop recursion

Notably, this limited-memory approach of the l-BFGS update allows for varying the initial matrix  $H_0$  with every iteration. Therefore, in every iteration  $k$ , a new symmetric positive definite matrix  $H_k^{(0)}$  can be chosen as *seed matrix* for the recursion.

As described above, the quadratic matrix  $B_k = H_k^{-1}$  in Quasi-Newton methods is supposed to approximate the Hessian matrix  $\nabla^2 \varphi(x_k)$ , so that the corresponding quadratic approximation of  $\varphi$  near  $x_k$  and thereby the search direction  $p_k = -H_k \nabla \varphi(x_k)$  is as accurate as possible. Since  $H_k$  is computed from the curvature information of the  $\ell$  latest steps and the seed matrix  $H_k^{(0)}$ , it is desirable that already this seed matrix represents an approximation to the inverse Hessian [34].

In practice, the seed matrix is often chosen as  $H_k^{(0)} = \hat{\tau}_k I$  with a scaling factor  $\hat{\tau}_k > 0$ , as this is a simple, effective and easily computable choice which needs no direct knowledge on the Hessian of  $\varphi$  [37]. Most commonly, one of the *Barzilai-Borwein factors*

$$\hat{\tau}_k^y := \frac{s_{k-1}^T y_{k-1}}{y_{k-1}^T y_{k-1}} \quad \text{or} \quad \hat{\tau}_k^s := \frac{s_{k-1}^T s_{k-1}}{s_{k-1}^T y_{k-1}} \quad (4.6)$$

is used as scaling factor [34, Section 2.2].

These factors are obtained by minimising the deviation in the secant equation (4.3) for  $H_k^{(0)}$  in the direction of  $s_k$  resp.  $y_k$ ; more on that in Section 4.2. It is well-known that both factors lie between the smallest and largest eigenvalue of the average inverse Hessian between  $x_{k-1}$  and  $x_k$  [34, Lemma 2.2]. Therefore, they represent an approximation to the spectrum of the inverse Hessian matrix. In practice, the factor  $\hat{\tau}_k^y$  proved to be preferable for inverse l-BFGS methods [34, Section 2.2].

The recursive algorithm by which the search direction  $p_k = -H_k \nabla \varphi(x_k)$  is computed based on the  $\ell$  latest vector pairs  $(s_i, y_i)_{i=k-\ell, \dots, k-1}$  and the initial matrix  $H_k^{(0)}$ , is described

in Algorithm 4.1. It is also used as part of the inner solver specified in Section 4.3.

---

**Algorithm 4.1** l-BFGS two-loop recursion, cf. [37, Algorithm 7.4]

---

**Given:**  $(s_{k-\ell}, y_{k-\ell}), \dots, (s_{k-1}, y_{k-1}) \in \mathbb{R}^n \times \mathbb{R}^n$ ;  $q = -\nabla\varphi(x_k) \in \mathbb{R}^n$ ;  $H_k^{(0)} \in \mathbb{R}^{n \times n}$  spd

```

1: for  $i = k - 1, k - 2, \dots, k - \ell$  do
2:    $\rho_i \leftarrow \frac{1}{s_i^T y_i}$ 
3:    $\alpha_i \leftarrow \rho_i s_i^T q$ 
4:    $q \leftarrow q - \alpha_i y_i$ 
5: end for
6:  $p \leftarrow H_k^{(0)} q$ 
7: for  $i = k - \ell, k - \ell + 1, \dots, k - 1$  do
8:    $\beta \leftarrow \rho_i y_i^T p$ 
9:    $p \leftarrow p + (\alpha_i - \beta) s_i$ 
10: end for

```

**Return:**  $p = -H_k \nabla\varphi(x_k)$

---

As long as the multiplication  $H_k^{(0)} p$  can be computed in  $O(n)$ , for example by choosing  $H_k^{(0)}$  as diagonal matrix, the computation and memory cost for this two-loop recursion lie in  $O((\ell + 1)n)$ . Typical choices for  $\ell_{\max}$  lie between 3 and 20 [37, page 177], and the method is nowadays prevalently used [37, page 179].

In 1989, Liu and Nocedal proved that for strongly convex functions, l-BFGS methods converge r-linearly in the function value to the minimum, as long as the step sizes  $\alpha_k$  satisfy the Wolfe-conditions and the eigenvalues of the seed matrices  $B_k^{(0)}$  remain bounded away from 0 and infinity, cf. [32, Theorem 7.1].

However, as stated in [37], the main weakness of l-BFGS methods is its slow convergence on “ill-conditioned problems; specifically, on problems where the Hessian matrix contains a wide distribution of eigenvalues” [37, page 180]. This issue is especially present with the seed-matrix choice  $H_k^{(0)} = \hat{\tau}_k I$  [34]. Therefore, an alternative choice for  $H_k^{(0)}$  is discussed in Section 4.2. This *structured l-BFGS* approach was recently analysed in [34], and it will be discussed in Section 4.3 how it is used to tackle the problem of ill-conditioning in the context of this thesis.

## 4.2 The structured l-BFGS method

In the last section, the concept of l-BFGS methods was motivated and introduced, and it was described that these methods allow for a new choice of the seed matrix  $H_k^{(0)}$  in every iteration  $k$ . Most commonly,  $H_k^{(0)}$  is chosen as an adaptively scaled identity matrix, but as mentioned above, this can lead to slow convergence, especially for ill-conditioned problems.

### 4.2.1 General concept

To tackle this issue, another possible choice for the seed matrix was recently analysed in [34] as so-called *structured l-BFGS method*. It is based on the idea that the objective  $\varphi$  is given by the sum of two functions where the Hessian matrix of one part is known and supports a

viable method to solve a linear system of equations with it, while the Hessian of the other part is approximated by a scaled identity matrix. Therefore, consider the minimisation problem

$$\inf_{x \in \mathbb{R}^n} \varphi(x),$$

where

$$\varphi : \mathbb{R}^n \rightarrow \mathbb{R}, \quad \varphi(x) = D(x) + S(x)$$

for twice continuously differentiable functions  $D, S \in C^2(\mathbb{R}^n, \mathbb{R})$ .

As described in Section 4.1, in Quasi-Newton methods, the quadratic matrix  $B_k = H_k^{-1}$  is supposed to approximate the Hessian matrix  $\nabla^2 \varphi(x_k)$ . To obtain a more accurate approximation, it is useful when already the seed matrix  $B_k^{(0)} = (H_k^{(0)})^{-1}$  approximates the Hessian in some way. To improve this approximation, the structured l-BFGS approach aims at utilising the available information about the Hessian of  $S$  and approximates only the possibly problematic part of  $D$  by a scaled identity matrix, instead of setting the entire base matrix to such a scaled identity.

Therefore, the authors of [34] suggest to use the non-inverse seed matrix

$$B_k^{(0)} := \tau_k I + S_k, \tag{4.7}$$

where  $\tau_k \geq 0$  and  $S_k$  is a symmetric positive definite matrix, reasonably  $S_k \approx \nabla^2 S(x_k)$ . This choice for  $B_k^{(0)}$  is designed to improve the approximation accuracy of it to the Hessian  $\nabla^2 \varphi(x_k) = \nabla^2 D(x_k) + \nabla^2 S(x_k)$ , particularly for ill-conditioned problems in which a scaled identity is not capable of properly capturing the spectrum of the Hessian.

One main difference and possible computational disadvantage of the structured approach is that, in contrast to the approach described in Section 4.1, the seed matrix  $B_k^{(0)}$  approximates the Hessian of  $\varphi$ , whereas the seed matrix  $H_k^{(0)}$  approximates the inverse of the Hessian. Therefore, the computation of  $p$  in line 6 of the two-loop recursion requires to solve the linear system  $B_k^{(0)} p = q$  instead of simply obtaining  $p$  by the multiplication  $p = H_k^{(0)} q$ .

Therefore, it is important that  $S_k$  is chosen in a way that linear systems involving  $B_k^{(0)}$  can be solved efficiently [34]. In practice, the linear system might also be solved only inexactly [34, page 26], as the result is only used for the seed matrix of the l-BFGS recursion. Hence, when  $S_k$  supports the efficient use of inexact solvers such as MINRES [38], it might already prove useful.

## 4.2.2 Choice of the scaling factor

The question remains how the scaling factor  $\tau_k$  should be chosen in (4.7). Similarly to the Barzilai-Borwein factors  $\hat{\tau}_k^y$  and  $\hat{\tau}_k^s$  defined in (4.6), the scaling factor  $\tau_k$  can be computed adaptively in each iteration based on the latest step. Again, the idea is to choose  $\tau_k$  in a way that already the base matrix  $B_k^{(0)}$  is “similar” to the Hessian of  $\varphi$  in the sense that it explains how the latest step occurred. Hence,  $\tau_{k+1} \geq 0$  is sought such that  $B_{k+1}^{(0)}$  satisfies the secant equation (4.3) as well as possible. It follows [34, Section 3.1]

$$y_k \stackrel{!}{\approx} B_{k+1}^{(0)} s_k = \tau_{k+1} s_k + S_{k+1} s_k \quad \Leftrightarrow \quad \tau_{k+1} s_k \stackrel{!}{\approx} y_k - S_{k+1} s_k =: z_k. \tag{4.8}$$

Based on that idea, the authors of [34] propose four different possible choices for  $\tau_{k+1}$  which, in some sense, optimise the objective  $\tau_{k+1} s_k \stackrel{!}{\approx} z_k$ .

**Lemma 4.1.** [cf. 34, Definition 3.1] *Let  $s_k, z_k \in \mathbb{R}^n$  with  $s_k \neq 0$ . Then the factor*

a)

$$\tau_{k+1}^s := \max \left\{ \frac{s_k^T z_k}{s_k^T s_k}, 0 \right\} = \arg \min_{\tau \geq 0} \|\tau s_k - z_k\|_2$$

*minimises the distance between  $s_k$  and  $z_k$  in the direction of  $s_k$ ,*

b)

$$\tau_{k+1}^g := \sqrt{\frac{z_k^T z_k}{s_k^T s_k}} = \frac{\|z_k\|_2}{\|s_k\|_2} = \arg \min_{\tau \geq 0} \|\sqrt{\tau} s_k - z_k / \sqrt{\tau}\|_2$$

*minimises the distance between  $s_k$  and  $z_k$  geometrically weighted in both directions  $s_k$  and  $z_k$ .*

c) *If furthermore  $s_k^T z_k \neq 0$ , then*

$$\tau_{k+1}^z := \max \left\{ \frac{z_k^T z_k}{s_k^T z_k}, 0 \right\} = \arg \min_{\tau \geq 0} \|s_k - z_k / \tau\|_2$$

*minimises the distance between  $s_k$  and  $z_k$  in the direction of  $z_k$ .*

d) *Again if  $s_k^T z_k \neq 0$ , then*

$$\tau_{k+1}^u := \max \left\{ \frac{z_k^T z_k - s_k^T s_k + \sqrt{(z_k^T z_k - s_k^T s_k)^2 + 4(s_k^T z_k)^2}}{2s_k^T z_k}, 0 \right\}$$

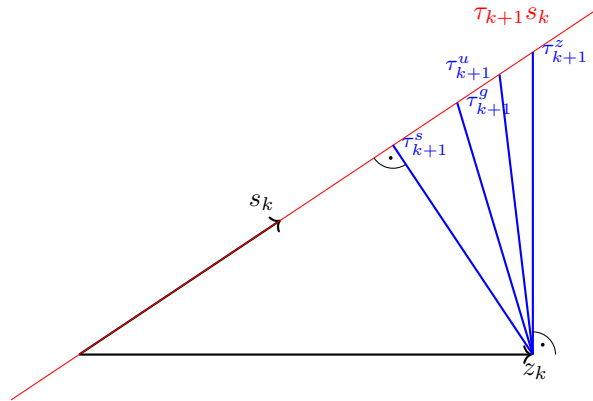
*minimises the distance between  $s_k$  and  $z_k$  unbiased; for details, see [1, Lemma 2].*

*Remark.*

- When  $S_k = 0$ , the factors  $\tau_{k+1}^z$  and  $\tau_{k+1}^s$  correspond to the (inverse) Barzilai-Borwein factors  $\hat{\tau}_k^y$  and  $\hat{\tau}_k^s$  defined in (4.6).
- The statement  $s_k^T z_k > 0$  does not generally hold when  $S_k \neq 0$ , not even for strongly convex functions  $\varphi$ . Therefore, it is necessary to explicitly ensure that  $\tau_{k+1} \geq 0$  for  $\tau_{k+1}^s$ ,  $\tau_{k+1}^z$  and  $\tau_{k+1}^u$ . On the other hand, it always holds  $\tau_{k+1}^g \geq 0$ .
- As long as  $s_k^T z_k > 0$ , the factor  $\tau_{k+1}^g$  corresponds to the geometric mean of  $\tau_{k+1}^s$  and  $\tau_{k+1}^z$  [34].
- It holds  $0 \leq \tau_{k+1}^s \leq \tau_{k+1}^g \leq \tau_{k+1}^z$  and  $0 \leq \tau_{k+1}^s \leq \tau_{k+1}^u \leq \tau_{k+1}^z$  (the latter being only defined for  $s_k^T z_k \neq 0$ ) [34].
- In [34], the authors consider more general cases, even of non-convex functions  $\varphi$ , and therefore confine the factors  $\tau_k$  optionally to an interval  $[\tau_k^{\min}, \tau_k^{\max}] \subseteq [0, \infty]$  which converges to  $[0, \infty]$  as  $k \rightarrow \infty$ . In the scenario considered in this thesis, however, these restrictions are not necessary, and the constraint  $\tau_k \geq 0$  suffices.

The geometrical meaning of all four factors for the objective  $\tau_{k+1} s_k \overset{!}{\approx} z_k$  is illustrated in Figure 4.1.

Now, the question arises which factor should be chosen in order to observe the fastest convergence, but unfortunately, there is no simple answer to that: In [1], it is observed how



**Figure 4.1:** [cf. 34] Geometrical view on the scaling factors from Lemma 4.1. The upper end of the blue lines marks the elongation of  $s_k$  with the factor  $\tau_{k+1}$ . The factors  $\tau_{k+1}^s$  and  $\tau_{k+1}^z$  correspond to the orthogonal projection of  $z_k$  onto  $\mathbb{R}s_k$  and of  $s_k$  onto  $\mathbb{R}z_k$ , respectively. In the method presented in Algorithm 4.2, the factor  $\tau_{k+1}^g$  is always used. As their geometric mean, it lies between the other two factors as long as the angle between  $s_k$  and  $z_k$  is below  $90^\circ$ .

a smaller factor  $\tau_k$  tends to lead to a longer search direction. This is not surprising since a smaller  $\tau_k$  yields greater eigenvalues of  $H_k^{(0)}$ , which ultimately also influences the length of  $p_k = -H_k \nabla \varphi(x_k)$ . Hence, when using  $\tau_k^s$ , more line-searches might be required, while the use of  $\tau_k^z$  tends to result in the necessity for more iterations as it yields shorter steps. In [1], the authors conclude that  $\tau_k^s$  and  $\tau_k^g$  perform best in their numerical experiments, and that  $\tau_k^g$  proves more robust.

This discussion is extended in [34] where, in contrast to constantly choosing the same factor, an algorithm for an adaptive factor choice is presented. For reasons of simplicity, however, this thesis is limited to the usage of  $\tau_k^g$  which has not only shown to be robust, but is also in line with the recommendation in [34] for the case  $s_k^T z_k < 0$ .

In [34], q-linear convergence in the objective function value  $\varphi(x_k)$  and r-linear convergence in the iterates  $x_k$  and in the gradient  $\nabla \varphi(x_k)$  is proven for the structured l-BFGS approach, requiring certain assumptions. These convergence results will later be invoked for the convergence analysis of the inner iterations, as the latter are based on the structured l-BFGS method.

The design of these inner iterations is described next in Section 4.3, and the analysis of their convergence is carried out in Sections 4.4 and 4.5.

### 4.3 Structured l-BFGS as inner solver

In the last section, the concept of the structured l-BFGS method was described which had been analysed in [34]. This section deals with the application of this method for solving the inner problems of the Tikhonov-regularised interior-point method which was introduced and analysed extensively in Chapter 3, and whose structure is specified in Algorithm 3.2.

In Theorems 3.11 and 3.12, it was shown that the outer iterates  $x^j$  produced by Algo-

gorithm 3.2 converge to the minimal-norm solution  $x_M^*$ , and that this convergence is quantified in the objective function  $f$  by a rate of  $O(\varepsilon_j)$ . In line 2, the method requires that in each outer iteration  $j$ , an approximation  $x^j$  to the current regularised barrier minimiser  $\bar{x}_{\varepsilon_j, \mu_j}$  is computed at which the residual gradient norm  $\|\nabla f_{\varepsilon_j, \mu_j}(x^j)\|_2$  is smaller than  $C\varepsilon_j^2$  for some fixed constant  $C$ .

In IPMs, typically one or multiple Newton steps are employed to the respective barrier problem ( $\mathcal{P}_{\mu_j}$ ) in order to obtain an iterate  $x^j$  which is sufficiently close to the barrier trajectory or lies within some trust region [18, 37]. But as Newton's method might become prohibitive for general large-scale problems [14, page 956], in this thesis, instead, an l-BFGS method is used which is applied directly to the primal problem, as this allows to completely proceed without the need of solving linear systems of equations.

In general, this comes with a great problem: As mentioned in Section 4.1, Quasi-Newton methods are known to struggle with and only converge slowly for ill-conditioned problems, and the Hessian of the barrier function becomes increasingly ill-conditioned as the barrier parameter  $\mu$  approaches zero [37, 41, 49]. Partly to deal with this problem, a Tikhonov regularisation was included in the interior-point method, see Chapter 3, but this still does not resolve the issue which active constraints impose on the problem condition.

Therefore, the idea in this thesis is to employ the structured l-BFGS method of Section 4.2 by including the possibly most problematic part of the Hessian of  $f_{\varepsilon, \mu}$  in the structured component  $S_k$ , while leaving the more complicated and possibly unavailable parts of the Hessian to be approximated in the seed matrix  $B_k^{(0)}$  by a scaled identity matrix.

### 4.3.1 Choice of the seed matrix

The Hessian of the Tikhonov-regularised barrier function  $f_{\varepsilon, \mu}$  is given in  $x \in \mathcal{F}_+$  by

$$\nabla^2 f_{\varepsilon, \mu}(x) = \nabla^2 f(x) + \varepsilon I + \underbrace{\mu \sum_{i=1}^m \left( \frac{1}{g_i(x)^2} \nabla g_i(x) \nabla g_i(x)^T - \frac{1}{g_i(x)} \nabla^2 g_i(x) \right)}_{\rightsquigarrow S_k},$$

cf. (3.7). In this, the summands  $\frac{\mu}{g_i(x)^2} \nabla g_i(x) \nabla g_i(x)^T$  can be expected to cause the greatest issues for the conditioning, as  $g_i(\bar{x}_{\varepsilon, \mu}) \rightarrow 0$  in any active constraint. Since furthermore the gradients of all  $g_i$  are available, needed during the l-BFGS iteration anyway, allow for an efficient multiplication with a vector, and their outer product is positive semi-definite, it is reasonable to include this part into the structure matrix  $S_k$ . Similarly, the regularisation part  $\varepsilon I$  should be integrated into  $S_k$ , as it represents the strong convexity of  $f_{\varepsilon, \mu}$  and thereby ensures positive definiteness in the l-BFGS system.

On the other hand, the matrices  $\nabla^2 f$  and  $\nabla^2 g_i$  might be unknown in practice, too complicated or too expensive to compute, or dense and therefore not suitable to be saved or used efficiently in matrix-vector multiplications. For a fixed  $j$ -th outer iteration, the proposal in this thesis hence is to use

$$S_k := \varepsilon_j I + \mu_j \sum_{i=1}^m \frac{1}{g_i(x_k^j)^2} \nabla g_i(x_k^j) \nabla g_i(x_k^j)^T \quad (4.9)$$

as structured Hessian part at the current inner iterate  $x_k^j$ . Note that  $S_k$  is symmetric and positive definite with all eigenvalues being at least  $\varepsilon$ . The entire seed matrix for the



structured l-BFGS method is then given by

$$B_k^{(0)} := \tau_k^g I + S_k \quad (4.10)$$

with  $\tau_k^g$  as defined in Lemma 4.1 c). Thus,  $\tau_k^g I$  approximates the in  $S_k$  missing part

$$\nabla^2 f(x_k^j) - \mu_j \sum_{i=1}^m \frac{1}{g_i(x_k^j)} \nabla^2 g_i(x_k^j).$$

This approach allows to capture the possibly large and for  $j \rightarrow \infty$  increasingly wide spectrum of  $\nabla^2 f_{\varepsilon_j, \mu_j}$  more properly in the seed matrix of the l-BFGS method, and thereby might considerably reduce the amount of inner iterations needed in the Tikhonov-regularised IPM. In case the method is used with the Augmented Lagrangian modification as described in Section 3.5, the term  $\rho_j A^T A$ , added to the Hessian matrix of  $f_{\varepsilon_j, \rho_j, \mu_j; \lambda^j}$  by the penalty term, can also be included in the structured part of the seed matrix  $S_k$ , as this might further improve the convergence speed of the inner iterates.

It should be noted that, as mentioned in Section 4.2, the use of the structured l-BFGS method requires, in each inner iteration  $k$ , the computation of the vector  $p$  as solution to the linear system  $B_k^{(0)} p = q$ , see line 6 of Algorithm 4.1. In practice, it makes sense to let  $p$  only approximate a solution to the linear system, as it is only used as initialisation for the l-BFGS recursion, and the fast matrix-vector multiplications supported by  $S_k$  as defined above allow for an efficient use of inexact solvers such as MINRES [38]. Clearly, this approach marks a less severe trade-off than solving an entire Newton system inexactly, as the curvature information of the last steps remains fully included in the search direction  $p_k$ . Furthermore, the proposed choice for  $S_k$  allows for a matrix-free implementation, so the actual computation and storing of the matrix is not necessary.

### 4.3.2 Armijo backtracking

To obtain convergence to a minimiser, it is well-known for Quasi-Newton methods [32] that the next iterate  $x_{k+1} := x_k + \alpha_k p_k$  should satisfy the *Armijo condition* [37, page 33]

$$f_{\varepsilon_j, \mu_j}(x_{k+1}^j) \leq f_{\varepsilon_j, \mu_j}(x_k^j) + c_{\text{ls}} \alpha_k \nabla f_{\varepsilon_j, \mu_j}(x_k^j)^T p_k \quad (4.11)$$

for an *Armijo* or *line-search constant*  $c_{\text{ls}} \in (0, 1)$ .

This condition enforces a “sufficient decrease” in the objective  $f_{\varepsilon_j, \mu_j}$  and is guaranteed to be satisfied for sufficiently small  $\alpha_k > 0$ , as long as  $p_k$  is a descent direction [37, Lemma 3.1]. Conveniently, the search direction  $p_k = -H_k \nabla f_{\varepsilon_j, \mu_j}(x_k^j)$  in l-BFGS methods is always a descent direction, as  $H_k$  is positive definite.

Hence, a sufficiently small step size  $\alpha_k$  can be computed by an iterative reduction. This procedure is referred to as (*Armijo*) *backtracking line-search* [37, Algorithm 3.1] and here performed by testing  $\alpha_k = 1, \rho_{\text{ls}}, \rho_{\text{ls}}^2, \rho_{\text{ls}}^3, \dots$  with a line-search constant  $\rho_{\text{ls}} \in (0, 1)$ , until (4.11) holds. For convergence results of l-BFGS methods, a curvature condition is often additionally required, leading to the *Wolfe conditions* [37, page 34]. However, since this thesis only deals with the application of an l-BFGS method to the strongly convex functions  $f_{\varepsilon_j, \mu_j}$ , it is not necessary here as will later be seen in Section 4.4.



### 4.3.3 Specification of the structured l-BFGS method

The procedure for the inner iterations of the Tikhonov-regularised IPM (line 2 in Algorithm 3.2) now is as follows. Starting with the iterate of the latest outer step  $x^{j-1}$ , the structured l-BFGS method is applied, producing a sequence of inner iterates  $x_k^j$ , until the stopping criterion  $\|\nabla f_{\varepsilon_j, \mu_j}(x_k^j)\|_2 \leq C\varepsilon_j^2$  is satisfied. The resultant final inner iterate  $x_k^j$  is then returned as new outer iterate  $x^j$ .

The scheme for the inner iterations using the structured l-BFGS method is described in Algorithm 4.2. It is inspired by [34, Algorithm TULIP].

---

**Algorithm 4.2** Inner solver of the Tikhonov-regularised interior-point method

---

**Given:** Optimisation problem  $(\mathcal{P})$ , current parameters  $\varepsilon_j, \mu_j > 0$ , latest outer iterate  $x^{j-1} \in \mathcal{F}_+$ , line-search parameters  $c_{\text{ls}}, \rho_{\text{ls}} \in (0, 1)$ , memory length  $\ell_{\text{max}} \in \mathbb{N}$ , constant  $C > 0$

**Choose:**  $\tau_0 > 0$

- 1: Initialise  $x_0^j \leftarrow x^{j-1}$ , compute  $S_0$  as defined in (4.9)
- 2: **for**  $k = 0, 1, 2, \dots$  **do**
- 3:   **if**  $\|\nabla f_{\varepsilon_j, \mu_j}(x_k^j)\|_2 \leq C\varepsilon_j^2$  **then** let  $x^j \leftarrow x_k^j$  and **break**
- 4:   Let  $\ell := \min\{k, \ell_{\text{max}}\}$
- 5:   Compute  $p_k := -H_k \nabla f_{\varepsilon_j, \mu_j}(x_k^j)$  using the two-loop recursion of Algorithm 4.1 with  $B_k^{(0)} := \tau_k I + S_k$ ,  $q := -\nabla f_{\varepsilon_j, \mu_j}(x_k^j)$ , and the stored pairs  $\{(s_i, y_i)\}_{i=k-\ell, \dots, k-1}$
- 6:   Compute step size  $\alpha_k$  using Armijo backtracking line-search with parameters  $c_{\text{ls}}, \rho_{\text{ls}}$
- 7:   Let  $s_k := \alpha_k p_k$ ,  $x_{k+1}^j := x_k^j + s_k$ ,  $y_k := \nabla f_{\varepsilon_j, \mu_j}(x_{k+1}^j) - \nabla f_{\varepsilon_j, \mu_j}(x_k^j)$
- 8:   Append  $(s_k, y_k)$  to storage
- 9:   **if**  $k \geq \ell_{\text{max}}$  **then** remove  $(s_{k-\ell}, y_{k-\ell})$  from storage
- 10:   Update  $S_{k+1}$  as defined in (4.9), let  $z_k := y_k - S_{k+1} s_k$
- 11:   Update  $\tau_{k+1} \leftarrow \sqrt{\frac{z_k^T z_k}{s_k^T s_k}}$  ▷ Corresponds to  $\tau_{k+1}^g$  in Lemma 4.1 c)
- 12: **end for**

**Return:**  $x^j$

---

Next in Section 4.4, it will be shown that the inner iterates  $x_k^j$  generated by Algorithm 4.2 converge r-linearly to a point  $x^j$  with sufficiently small residual gradient norm. Afterwards in Section 4.5, a bound on the number of necessary inner iterations will be derived in the context of the overall method, and in Section 4.6, it is proven that their total number grows at most polynomially in the imposed tolerance.

## 4.4 Convergence of the inner iterates

In the last section, it was described how the structured l-BFGS method can be used to solve the inner problems of the Tikhonov-regularised IPM, and the exact procedure was defined in Algorithm 4.2. This section deals with the convergence of the inner iterates produced by that algorithm.

In a first step, the analysis is carried out for one isolated outer step  $j$ , that is, for arbitrary parameters  $\varepsilon_j, \mu_j > 0$ . For ease of notation, the outer iteration index  $j$  will be omitted throughout that discussion, so parameters  $\varepsilon, \mu > 0$  and iterates  $x_k$  will be

considered instead of  $\varepsilon_j$ ,  $\mu_j$  and  $x_k^j$ . In a second step, in Section 4.5, the results from this isolated analysis are quantified in the context of the overall algorithm, ultimately proving a polynomial growth of the number of required inner iterations.

The convergence analysis is mainly performed by applying the convergence results stated for the structured l-BFGS method in [34] to the inner solver. For this, it must first be verified that for any parameters  $\varepsilon, \mu > 0$ , the procedure described in Algorithm 4.2 actually is a structured l-BFGS method in terms of the method discussed in [34]. In fact, Algorithm 4.2 corresponds to *Algorithm TULIP* from [34] with  $J := f_{\varepsilon, \mu}$ ,  $S_k$  as in (4.9), the step size consistently chosen as  $\tau_k := \tau_k^g$ , and the respective parameter choices  $\epsilon := C\varepsilon^2$ ,  $\ell := \ell_{\max}$ ,  $c_0 := 0$  and  $C_0 := \infty$ . The parameter  $c_s$  is not needed (or could be chosen arbitrarily from the interval  $(0, \varepsilon)$ ) as it always holds  $y_k^T x_k \geq \varepsilon \|s_k\|_2^2$  by the  $\frac{\varepsilon}{2}$ -strong convexity of  $f_{\varepsilon, \mu}$ , see Lemma 2.5 c). The parameters  $c_1$  and  $c_2$  are irrelevant due to the choices of  $c_0$  and  $C_0$ .

Therefore, the results from [34] can be used for the analysis of Algorithm 4.2. As a starting point, the algorithm for the inner iterations is well-defined and converges in the objective  $f_{\varepsilon, \mu}$ .

**Theorem 4.2.** [cf. 34, Lemma 4.3] *Let the Slater CQ hold, i.e.  $\mathcal{F}_+ \neq \emptyset$ . For any  $\varepsilon, \mu > 0$ , Algorithm 4.2 is well-defined, and the sequence  $(f_{\varepsilon, \mu}(x_k))_{k \in \mathbb{N}_0}$  is strictly monotonically decreasing and convergent.*

*Proof.* The function  $f_{\varepsilon, \mu}$  is continuously differentiable on  $\mathcal{F}_+$ , and due to its strong convexity, it is also bounded below. Since  $S_k$  is spd for all  $k \in \mathbb{N}_0$ , the assumptions of [34, Assumption 4.2] are satisfied and hence, [34, Lemma 4.3] can be applied for Algorithm 4.2, yielding the statement of this theorem.  $\square$

#### 4.4.1 Matrix norm and Lipschitz continuity

The further convergence results of [34] yield convergence rates in the objective, in the iterates, and in the gradient. To apply these to the inner iterations, it must be shown that a number of requirements is satisfied. They involve, amongst other things, the spectral norm for matrices which, for completeness, will briefly be recalled here: For a symmetric positive definite (spd) matrix  $A$ , the spectral norm is given by (cf. [20, page 326])

$$\|A\|_2 := \max_{\substack{x \in \mathbb{R}^n \\ \|x\|_2=1}} \|Ax\|_2 = \lambda_{\max}(A),$$

and conversely for its inverse by

$$\|A^{-1}\|_2 = \lambda_{\max}(A^{-1}) = \frac{1}{\lambda_{\min}(A)},$$

where  $\lambda_{\max}(A)$  and  $\lambda_{\min}(A)$  denote the maximal and minimal eigenvalue of  $A$ , respectively. Due to  $A$  being spd, these eigenvalues are always real and strictly positive [20, Satz A.5].

A major requirement in the following convergence analysis is that the gradient of the inner objective function  $f_{\varepsilon, \mu}$  must be Lipschitz continuous. To this purpose, the following standard result on Lipschitz continuous gradients and its relation to a bounded Hessian for twice continuously differentiable functions  $\varphi$  will be needed.

**Lemma 4.3.** [cf. 29, page 103] *Let  $\varphi \in C^2(\Omega, \mathbb{R})$  be a twice continuously differentiable function on an open set  $\Omega \subseteq \mathbb{R}^n$ , and let  $K \subset \Omega$  be a compact convex subset. Then the gradient of  $\varphi$  is Lipschitz continuous on  $K$  with Lipschitz constant  $L := \sup_{x \in K} \|\nabla^2 \varphi(x)\|_2 < \infty$ , that is, for all  $x, y \in K$  it holds*

$$\|\nabla \varphi(x) - \nabla \varphi(y)\|_2 \leq L \|x - y\|_2.$$

Hence, a bounded Hessian matrix implies that the gradient is Lipschitz continuous. Due to the barrier term in  $f_{\varepsilon, \mu}$ , however, Lipschitz continuity of its gradient clearly does not hold on the entire domain  $\mathcal{F}_+$ , not even if  $f$  and all  $g_i$  had a Lipschitz continuous gradient: For example, the one-dimensional constraint  $g(x) := x \leq 0$  already yields  $\frac{d}{dx}(-\ln(-x)) = -\frac{1}{x}$ , which is neither bounded nor Lipschitz continuous on  $(-\infty, 0)$ .

Therefore, it is necessary to limit the discussion to a compact subset of  $\mathcal{F}_+$ ; specifically, this will be done with the level set of  $f_{\varepsilon, \mu}$ , based on the function value at the first iteration  $x_0$ : Hence, for parameters  $\varepsilon, \mu > 0$  and starting point  $x_0 \in \mathcal{F}_+$ , consider the level set

$$\Omega_{\varepsilon, \mu} := \{x \in \mathcal{F}_+ \mid f_{\varepsilon, \mu}(x) \leq f_{\varepsilon, \mu}(x_0)\}. \quad (4.12)$$

By Lemma 3.3, the set  $\Omega_{\varepsilon, \mu}$  is compact, and since  $f_{\varepsilon, \mu}(x_k)$  decreases strictly in  $k$  as stated in Theorem 4.2, all inner iterates  $x_k$  lie in that set. It should be noted that  $\Omega_{\varepsilon, \mu}$  corresponds to the set  $\Omega$  in [34].

Now, the compactness of  $\Omega_{\varepsilon, \mu}$  and the fact that  $f_{\varepsilon, \mu}$  is twice continuously differentiable on  $\mathcal{F}_+$  and therefore especially on  $\Omega_{\varepsilon, \mu}$ , yield that its Hessian is bounded on the latter, leading to  $f_{\varepsilon, \mu}$  having a Lipschitz continuous gradient on  $\Omega_{\varepsilon, \mu}$ :

**Lemma 4.4.** *For any parameters  $\varepsilon, \mu > 0$ , the gradient of  $f_{\varepsilon, \mu}$  is Lipschitz continuous on the set  $\Omega_{\varepsilon, \mu}$ , that is, there exists a constant  $L_{\varepsilon, \mu} > 0$  such that for all  $x, y \in \Omega_{\varepsilon, \mu}$  it holds*

$$\|\nabla f_{\varepsilon, \mu}(x) - \nabla f_{\varepsilon, \mu}(y)\|_2 \leq L_{\varepsilon, \mu} \|x - y\|_2.$$

*Proof.* Lemma 3.3 implies that  $\Omega_{\varepsilon, \mu}$  is compact, i.e. closed and bounded. Since  $f_{\varepsilon, \mu}$  is twice continuously differentiable on the open set  $\mathcal{F}_+ \supset \Omega_{\varepsilon, \mu}$ , Lemma 4.3 implies that  $\nabla f_{\varepsilon, \mu}$  is Lipschitz continuous on  $\Omega_{\varepsilon, \mu}$  with some Lipschitz constant  $L_{\varepsilon, \mu} > 0$ .  $\square$

#### 4.4.2 Convergence rate for the inner iterations

With a similar argument, the extreme value theorem of continuous functions on compact sets [29, page 31] yields an upper bound for the structured Hessian part  $S_k$ , as all terms in its definition (4.9) are continuous on  $\Omega_{\varepsilon, \mu}$ , and all iterates  $x_k$  remain in this set. Thus, the sequence  $(\|S_k\|_2)_{k \in \mathbb{N}_0}$  is bounded.

Now, all needed prerequisites are prepared, and it follows that Algorithm 4.2 meets all requirements for the aforementioned convergence results, that is namely (cf. [34, Assumption 4.4] and [34, Assumption 4.10]):

- $f_{\varepsilon, \mu}$  is continuously differentiable,
- the gradient of  $f_{\varepsilon, \mu}$  is Lipschitz continuous on  $\Omega_{\varepsilon, \mu}$ ,
- the sequence  $(\|S_k\|_2)_{k \in \mathbb{N}_0}$  is bounded,

- it holds  $\sup_{k \in \mathbb{N}_0} \left\| (B_k^{(0)})^{-1} \right\|_2 < \infty$  (this is the case since  $S_k$  is strongly positive definite with all eigenvalues being greater than  $\varepsilon > 0$ , so  $\left\| (B_k^{(0)})^{-1} \right\|_2 \leq \frac{1}{\varepsilon}$  for all  $k$ ),
- the step size  $\tau_k$  lies in  $[\tau_k^s, \tau_k^g]$  for all  $k$ , and
- the sequences  $(\|B_k\|_2)_{k \in \mathbb{N}_0}$  and  $(\|B_k^{-1}\|_2)_{k \in \mathbb{N}_0}$  are bounded (this holds by [34, Remark 4.11] with the given choice of  $\tau_k$  and the strong positive definiteness of  $S_k$ ; for details, see Lemma 4.6).

It should be noted that the requirements [34, Assumption 4.4 - 4)] and [34, Assumption 4.10 - 4)] involving the uniform continuity of  $f_{\varepsilon, \mu}$  in a neighbourhood of  $\Omega_{\varepsilon, \mu}$  were dropped here as they are not necessary in the scenario considered in this thesis; for details, see [34, Remark 4.9] and [34, Theorem 4.15].

As all requirements are met, the results from [34] can be applied to each inner iteration procedure defined by Algorithm 4.2.

**Theorem 4.5.** [cf. 34, Theorem 4.8 and Theorem 4.12] *Let the Slater CQ hold, i.e.  $\mathcal{F}_+ \neq \emptyset$ , and let  $\varepsilon, \mu, C > 0$  and  $x_0 \in \mathcal{F}_+$ . Then, Algorithm 4.2 terminates after finitely many iterations with an iterate  $x_k$  that satisfies  $\|\nabla f_{\varepsilon, \mu}(x_k)\|_2 \leq C\varepsilon^2$ . Furthermore, if the algorithm is applied until  $\|\nabla f_{\varepsilon, \mu}(x_k)\|_2 = 0$ , then it either terminates after finitely many iterations  $k$ , or it generates an infinite sequence  $(x_k)_{k \in \mathbb{N}_0}$  which satisfies:*

- $\lim_{k \rightarrow \infty} \|\nabla f_{\varepsilon, \mu}(x_k)\|_2 = 0$ .
- The iterates  $(x_k)_{k \in \mathbb{N}_0}$  converge  $r$ -linearly to  $\bar{x}_{\varepsilon, \mu}$ .
- The gradients  $(\nabla f_{\varepsilon, \mu}(x_k))_{k \in \mathbb{N}_0}$  converge  $r$ -linearly to 0.
- The function values  $(f_{\varepsilon, \mu}(x_k))_{k \in \mathbb{N}_0}$  converge  $q$ -linearly to  $f_{\varepsilon, \mu}(\bar{x}_{\varepsilon, \mu})$ .

In particular, these convergence results are quantified for  $k \in \mathbb{N}_0$  by

$$\begin{aligned} \|x_k - \bar{x}_{\varepsilon, \mu}\|_2^2 &\leq \nu^k \left( \frac{\hat{\beta}}{1 - \sqrt{\nu}} \right)^2 \frac{L_{\varepsilon, \mu}}{2} \|x_0 - \bar{x}_{\varepsilon, \mu}\|_2^2, \\ \|\nabla f_{\varepsilon, \mu}(x_k)\|_2^2 &\leq \nu^k \left( \frac{\hat{\beta}}{1 - \sqrt{\nu}} \right)^2 \frac{L_{\varepsilon, \mu}^2}{\varepsilon} \|\nabla f_{\varepsilon, \mu}(x_0)\|_2^2, \\ f_{\varepsilon, \mu}(x_{k+1}) - f_{\varepsilon, \mu}(\bar{x}_{\varepsilon, \mu}) &\leq \left( 1 - \frac{c_{\text{ls}} \alpha_k \varepsilon}{\|B_k\|_2} \right) (f_{\varepsilon, \mu}(x_k) - f_{\varepsilon, \mu}(\bar{x}_{\varepsilon, \mu})) \\ &\leq \nu (f_{\varepsilon, \mu}(x_k) - f_{\varepsilon, \mu}(\bar{x}_{\varepsilon, \mu})), \end{aligned}$$

where

$$\hat{\beta} := \sup_{k \in \mathbb{N}_0} \sqrt{\frac{\alpha_k \|B_k^{-1}\|_2}{c_{\text{ls}}}}, \quad \nu := \sup_{k \in \mathbb{N}_0} \left( 1 - \frac{c_{\text{ls}} \alpha_k \varepsilon}{\|B_k\|_2} \right)$$

with  $c_{\text{ls}} \in (0, 1)$  being the Armijo constant from (4.11).

*Proof.* As discussed above, the procedure of Algorithm 4.2 meets the requirements for the referenced theorems of [34]. The termination after finitely many steps for an inexact solution then follows from [34, Theorem 4.8], as well as the convergence of the gradient of  $f_{\varepsilon, \mu}$  to 0.

As stated in [34, Remark 4.13 and Theorem 4.15],  $f_{\varepsilon,\mu}$  satisfies the required *Kurdyka-Łojasiewicz inequality* for all  $k \in \mathbb{N}_0$  with  $\mu := \varepsilon$  as it is  $\frac{\varepsilon}{2}$ -strongly convex. Thus, by [34, Theorem 4.12], the iterates  $x_k$  converge to a point  $\bar{x}$  at which the gradient of  $f_{\varepsilon,\mu}$  is 0. Since this is only the case for  $\bar{x} = \bar{x}_{\varepsilon,\mu}$  (see Section 3.3), convergence to  $\bar{x}_{\varepsilon,\mu}$  is obtained. The other statements now follow directly from [34, Theorem 4.12] and [34, Remark 4.13].  $\square$

*Remark.* The parameter  $\nu$  describes the rate of the linear convergence of Algorithm 4.2. It holds  $\nu < 1$  since all step sizes  $\alpha_k$  are bounded away from 0 when obtained by backtracking line-search [34, page 20], and the norms  $\|B_k\|_2$  are bounded as was discussed above and will further be quantified in Lemma 4.6.

These results prove that the algorithm of the inner solver converges for any parameters  $\varepsilon, \mu > 0$ , and that the stopping criterion  $\|\nabla f_{\varepsilon,\mu}(x_k)\|_2 \leq C\varepsilon^2$  is satisfied after finitely many iterations. Therefore, as the overall Tikhonov-regularised IPM described in Algorithm 3.2 requires only finitely many outer steps (see Theorem 3.12), it also terminates after finitely many inner / l-BFGS iterations. Furthermore, the convergence of each inner procedure in the residual gradient norm is r-linear.

It should be pointed out that these convergence results are independent of the actual choice of  $S_k$ , as long as the latter remains  $\varepsilon$ -strongly positive definite. This especially means that the results do not take into account that the choice (4.9) aims at capturing the spectrum of the Hessian of  $f_{\varepsilon,\mu}$  more properly. The use of  $S_k$  as defined in (4.9) is therefore solely a practical consideration to improve the numerical convergence speed of the inner iterations.

So far, the convergence analysis for the inner iterations was carried out for the isolated consideration of one outer iteration  $j$ , that is, for fixed parameters  $\varepsilon_j, \mu_j > 0$ . However, as Algorithm 4.2 is applied iteratively with decreasing parameters  $\varepsilon_j$  and  $\mu_j$ , the question arises how this inner convergence evolves with progressing outer iterations. For this purpose, in the next section, estimates for the inner convergence will be derived which globally depend on the parameters  $\varepsilon_j$  and  $\mu_j$ , ultimately allowing to prove the polynomial complexity result for the total number of l-BFGS iterations which is derived in Section 4.6 and presented in Theorem 4.12.

## 4.5 A bound on the number of required inner iterations

In the last section, it was shown that each inner iteration procedure described in Section 4.3 terminates after finitely many steps, and that the convergence in the residual gradient norm, which is relevant for the inner stopping criterion, is r-linear, cf. Theorem 4.5. However, these results are limited to an isolated inner loop, that is, to fixed parameters  $\varepsilon$  and  $\mu$ . Based on that, in this section, it will be analysed how the convergence and especially the number of required iterations evolves during the outer method, that is, for iteratively decreasing parameters. For this, global convergence estimates in terms of the parameters  $\varepsilon$  and  $\mu$  are needed.

#### 4.5.1 Bounds on the l-BFGS matrices

Unfortunately, the estimates in the convergence result of Theorem 4.5 are not only very conservative as can be verified in the proofs in [34], but they are also rather unspecific, as the constants  $\nu$  and  $\hat{\beta}$  include suprema over the step sizes  $\alpha_k$  and the norms of  $B_k$  and  $B_k^{-1}$ . Especially when a comparison of the cost for the respective inner method during the progressing outer iterations is sought, these unspecific terms cause an issue. Therefore, at first, specific estimates for the norm of  $B_k$  and  $B_k^{-1}$  are derived.

**Lemma 4.6.** *Let  $\varepsilon, \mu > 0$  with  $\varepsilon \leq 1$ , and let the corresponding Lipschitz constant  $L_{\varepsilon, \mu} \geq 1$ . For the l-BFGS matrices  $B_k = H_k^{-1}$  occurring in the iterations of Algorithm 4.2, it holds for all  $k \in \mathbb{N}_0$  that*

$$\|B_k\|_2 \leq 2\|S_k\|_2 + L_{\varepsilon, \mu} + \frac{\ell L_{\varepsilon, \mu}^2}{\varepsilon}, \quad (4.13)$$

$$\|B_k^{-1}\|_2 \leq \frac{5^\ell}{\varepsilon} \left( \frac{L_{\varepsilon, \mu}}{\varepsilon} \right)^{2\ell}, \quad (4.14)$$

where  $\ell := \min\{k, \ell_{\max}\}$  denotes the current memory length of the l-BFGS iteration.

*Proof.* The derivation of the estimates mainly follows the computations in [34, Lemma 4.1 and Lemma 4.6] and applies the specific case given here to those results.

- a) To obtain the estimate for  $B_k$ , it will first be reduced to its seed matrix  $B_k^{(0)}$  by analysing the BFGS update formula (4.4). For this, verify that by the  $\frac{\varepsilon}{2}$ -strong convexity and the Lipschitz continuous gradient of  $f_{\varepsilon, \mu}$  it holds

$$\frac{y_k^T s_k}{\|s_k\|_2^2} \geq \varepsilon, \quad \frac{s_k^T y_k}{\|y_k\|_2^2} \geq \frac{s_k^T y_k}{L_{\varepsilon, \mu}^2 \|s_k\|_2^2} \geq \frac{\varepsilon}{L_{\varepsilon, \mu}^2}.$$

Now, for any  $v \in \mathbb{R}^n$  with  $\|v\|_2 = 1$  it holds with (4.4), the positive-definiteness of  $B_k$ , and the Cauchy-Schwarz inequality [37, page 600] that

$$v^T B_{k+1} v = v^T B_k v - \frac{(v^T B_k s_k)^2}{s_k^T B_k s_k} + \frac{(v^T y_k)^2}{s_k^T y_k} \leq v^T B_k v + \frac{\|v\|_2^2 \|y_k\|_2^2}{s_k^T y_k},$$

where  $B_{k-\ell} := B_k^{(0)}$  is defined as the seed matrix here. Hence,

$$\|B_{k+1}\|_2 = \lambda_{\max}(B_{k+1}) \leq \lambda_{\max}(B_k) + \frac{\|y_k\|_2^2}{s_k^T y_k} \leq \|B_k\|_2 + \frac{L_{\varepsilon, \mu}^2}{\varepsilon},$$

and inductively it follows after the  $\ell$  undertaken l-BFGS update steps that

$$\|B_k\|_2 \leq \left\| B_k^{(0)} \right\|_2 + \frac{\ell L_{\varepsilon, \mu}^2}{\varepsilon}.$$

For the seed matrix  $B_k^{(0)} := S_k + \tau_k I$ , on the other hand, the choice  $\tau_k = \tau_k^g$ , defined in Lemma 4.1, and the definition  $z_{k-1} := y_k - S_k s_{k-1}$  in (4.8) yield

$$\tau_k^g = \frac{\|z_{k-1}\|_2}{\|s_{k-1}\|_2} \leq \frac{\|y_{k-1}\|_2 + \|S_k s_{k-1}\|_2}{\|s_{k-1}\|_2} \leq L_{\varepsilon, \mu} + \|S_k\|_2,$$

so it follows

$$\|B_k^{(0)}\|_2 = \|S_k\|_2 + \tau_k \leq 2\|S_k\|_2 + L_{\varepsilon,\mu}.$$

Taking all estimates together, the statement (4.13) is obtained.

- b) Similarly, the following estimate for the norm of  $B_k^{-1}$  can be derived, cf. [34, Lemma 4.1]:

$$\|B_k^{-1}\|_2 \leq 5^\ell \max \left\{ 1, \|(B_k^{(0)})^{-1}\|_2 \right\} \max \left\{ 1, \left(\frac{1}{\varepsilon}\right)^\ell, \left(\frac{L_{\varepsilon,\mu}}{\varepsilon}\right)^{2\ell} \right\}.$$

Now, the fact that  $S_k$  is  $\varepsilon$ -strongly positive definite yields with  $\tau_k \geq 0$  that

$$\lambda_{\min}(B_k^{(0)}) \geq \lambda_{\min}(S_k) + \tau_k \geq \varepsilon + \tau_k \geq \varepsilon,$$

so it follows

$$\|(B_k^{(0)})^{-1}\|_2 = \frac{1}{\lambda_{\min}(B_k^{(0)})} \leq \frac{1}{\varepsilon}.$$

With the assumptions  $\varepsilon \leq 1$  and  $L_{\varepsilon,\mu} \geq 1$  which imply  $\left(\frac{L_{\varepsilon,\mu}}{\varepsilon}\right)^2 \geq \frac{L_{\varepsilon,\mu}}{\varepsilon} \geq \frac{1}{\varepsilon} \geq 1$ , one finally obtains

$$\|B_k^{-1}\|_2 \leq 5^\ell \max \left\{ 1, \frac{1}{\varepsilon} \right\} \max \left\{ 1, \left(\frac{1}{\varepsilon}\right)^\ell, \left(\frac{L_{\varepsilon,\mu}}{\varepsilon}\right)^{2\ell} \right\} \leq 5^\ell \frac{1}{\varepsilon} \left(\frac{L_{\varepsilon,\mu}}{\varepsilon}\right)^{2\ell},$$

which is precisely (4.14). □

*Remark.* The assumptions  $\varepsilon \leq 1$  and  $L_{\varepsilon,\mu} \geq 1$  are not restrictive: During the method,  $\varepsilon_j$  is strictly reduced with every outer iteration and ultimately converges to zero, so from some outer iteration on it will lie below 1. The Lipschitz constant  $L_{\varepsilon,\mu}$ , on the other hand, is only an upper estimate and hence could always be chosen greater than 1. Moreover, the assumptions are only necessary to obtain the bound (4.14) without the need to distinguish the cases in the maximum terms.

Lemma 4.6 yields upper estimates for the norms of  $B_k$  and  $B_k^{-1}$ , but in order to establish a result for the overall convergence with respect to  $\varepsilon$  and  $\mu$ , they still require a bound on the norm of  $S_k$  as well as specific dependencies of the Lipschitz constant  $L_{\varepsilon,\mu}$  in the parameters  $\varepsilon$  and  $\mu$ .

#### 4.5.2 Assumptions required for the quantitative analysis

As will be seen later in this section, two assumptions are needed to obtain bounds on  $\|S_k\|_2$  and  $L_{\varepsilon,\mu}$  which hold for all inner iterations throughout the overall method. In addition to that, a third assumption is required to bound the step sizes away from zero.

Firstly, it is necessary to globally bound the size of all iterates  $x_k^j$  and of the gradients of all  $g_i$ , as well as to derive global Lipschitz constants for the gradients of  $f$  and all  $g_i$ . In contrast to the isolated case considered in the last section, however, the overall method does not intrinsically yield that it is globally limited to a compact set.

Therefore, it must be assumed for the further analysis that the entire method actually operates on a compact set, that is, that the size of all iterates  $x_k^j$  is globally bounded by some constant  $C^x$ . Overall, this assumption appears reasonable and not too restrictive, since the following is known from the discussions above:

- In each outer iteration  $j$ , the inner iterates  $(x_k^j)_{k \in \mathbb{N}_0}$  start at the last outer iterate  $x^{j-1}$  and converge r-linearly towards  $\bar{x}_{\varepsilon_j, \mu_j}$ , until the iteration is stopped when the residual gradient norm is sufficiently small, yielding the next outer iterate  $x^j$  (cf. Sections 4.3 and 4.4).
- The outer iterates  $(x^j)_{j \in \mathbb{N}_0}$  themselves are bounded as their distance to the respective current minimiser  $\bar{x}_{\varepsilon_j, \mu_j}$  decreases with  $j$ , and the sequence  $(\bar{x}_{\varepsilon_j, \mu_j})_{j \in \mathbb{N}_0}$  is bounded (cf. Sections 3.3 and 3.4).
- Furthermore, the outer iterates ultimately converge to  $x_M^*$ .

Thus, the only way how the conglomeration of all iterates  $x_k^j$  could not be bounded, is that, with increasing parameter  $j$ , there were some inner iterates  $x_k^j$  which, on their trajectory from the bounded  $x^{j-1}$  to the bounded  $x^j$ , would grow unboundedly large. This seems rather unlikely and did also not occur in the numerical tests discussed in Chapter 5, but the nature of the l-BFGS method and the objective  $f_{\varepsilon_j, \mu_j}$  which changes with every outer iteration did not allow us to actually prove such a bound on all iterates; hence the assumption.

The second assumption requires that the gradient of the respective regularised barrier function  $f_{\varepsilon_j, \mu_j}$  is bounded globally in all inner iterates  $x_k^j$ , that is, that it holds  $\|\nabla f_{\varepsilon_j, \mu_j}(x_k^j)\|_2 \leq C'$  for all  $j, k$  and some constant  $C'$ . For this assumption, it is less straight forward to argue whether it is reasonably satisfied, since the function  $f_{\varepsilon_j, \mu_j}$  changes with every outer iteration  $j$ , and especially its gradient might generally be unbounded when approaching the boundary of  $\mathcal{F}_+$ : For example, the one-dimensional constraint  $g(x) := x \leq 0$  corresponds to the gradient  $\frac{d}{dx}(-\ln(-x)) = -\frac{1}{x}$  which clearly grows unboundedly as  $x \rightarrow 0^-$ , yielding that the gradient of  $f_{\varepsilon_j, \mu_j}$  is not bounded on  $\mathcal{F}_+ = (-\infty, 0)$  for any  $j \in \mathbb{N}$ .

However, for the later derived convergence results it is only necessary that the gradient norm remains bounded at all inner iterates  $x_k^j$  of the respective outer iteration  $j$ . Although it could not be shown that this requirement is always met, similarly to above, the design of the overall method suggests that the assumption should usually not be problematic or represent a great restriction, as the following holds:

- In each outer iteration  $j$ , the residual gradient norm at the inner iterates  $x_k^j$  converges r-linearly to 0 (cf. Theorem 4.5).
- Each inner iteration loop is started at the last outer iterate  $x^{j-1}$  which satisfies  $\|\nabla f_{\varepsilon_{j-1}, \mu_{j-1}}(x^{j-1})\|_2 \leq C\varepsilon_{j-1}^2$ .
- The gradient of the regularised barrier function in a current outer iteration  $j$  differs from the one in the previous step only by the updated parameters:

$$\nabla f_{\varepsilon_j, \mu_j}(x) = \nabla f_{\varepsilon_{j-1}, \mu_{j-1}}(x) + (\varepsilon_j - \varepsilon_{j-1})x - (\mu_j - \mu_{j-1}) \sum_{i=1}^m \frac{1}{g_i(x)} \nabla g_i(x).$$



Thus, since the gradient norm of  $f_{\varepsilon_{j-1}, \mu_{j-1}}$  was reduced in the previous step until it lied below  $C\varepsilon_{j-1}^2$ , it is reasonable to expect that it does neither explode with the parameter updates, nor during the inner iterations. This expectation is also supported by the numerical results presented in Chapter 5.

The third assumption, which is necessary for bounding the steps sizes  $\alpha_k$  away from zero, can be considered as an extension of the Lipschitz continuity of  $\nabla f_{\varepsilon_j, \mu_j}$  to certain points along the search directions: It requires that, whenever the step size  $\alpha_k = 1$  is not accepted as it violates the Armijo condition (4.11), there exists some smaller step  $\delta_k \in (0, 1)$  for which the Armijo condition is still violated and the gradient of  $f_{\varepsilon_j, \mu_j}$  is Lipschitz continuous on the slice towards the latest iterate. In this Lipschitz estimate, the corresponding Lipschitz constant must not exceed  $C^L L_{\varepsilon_j, \mu_j}$  for some global constant  $C^L > 0$ .

As well as the other two assumptions, also this third assumption seems reasonable by the following considerations:

- As stated in Section 4.3, there always exists a step size  $\alpha_k > 0$  which satisfies the Armijo condition. Hence, whenever  $\alpha_k = 1$  is not accepted, the condition is violated down to a certain step size.
- In the next iterate  $x_{k+1}^j = x_k^j + \alpha_k p_k$ , the gradient of  $f_{\varepsilon_j, \mu_j}$  is Lipschitz continuous with constant  $L_{\varepsilon_j, \mu_j}$ .
- By definition of the backtracking line-search,  $\alpha_k$  is given as the first power  $\rho_{\text{ls}}^i$  for which the Armijo condition holds. Hence, the condition is still violated for  $\rho_{\text{ls}}^{i-1}$ , and the sought step size  $\delta_k$  can be taken from the interval  $(\rho_{\text{ls}}^i, \rho_{\text{ls}}^{i-1}]$ .
- Thus,  $\delta_k$  can be chosen such that  $x_k^j + \delta_k p_k$  lies in the proximity of  $x_{k+1}^j$ . By the continuity of  $\nabla f_{\varepsilon_j, \mu_j}$ , it is not expected that the Lipschitz estimate changes more between  $x_k^j$  and  $x_k^j + \delta_k p_k$  than what can be absorbed by some global constant  $C^L$ .

Together, this yields the following three assumptions which are required for the further quantitative convergence analysis of the overall method.

**Assumption 4.7.** Consider the Tikhonov-regularised IPM described in Algorithm 3.2 with the structured l-BFGS method of Algorithm 4.2 applied as inner solver. It is assumed that

- a) all iterates  $x_k^j$  produced by the method remain bounded, that is, there exists a constant  $C^x > 0$  such that for all outer iterations  $j \in \mathbb{N}$  and for all corresponding inner iterations  $k \in \mathbb{N}_0$  it holds

$$\left\| x_k^j \right\|_2 \leq C^x,$$

- b) at all iterates  $x_k^j$ , the gradient of the respective regularised barrier function  $f_{\varepsilon_j, \mu_j}$  is bounded globally, that is, there exists a constant  $C' > 0$  such that in all outer iterations  $j \in \mathbb{N}$  it holds for all corresponding inner iterations  $k \in \mathbb{N}_0$  that

$$\left\| \nabla f_{\varepsilon_j, \mu_j}(x_k^j) \right\|_2 \leq C',$$

- c) there exists a constant  $C^L > 0$  such that in all outer iterations  $j \in \mathbb{N}$  and all corresponding inner iterations  $k \in \mathbb{N}_0$  always one of the following holds:

- i. Either the step size  $\alpha_k = 1$  satisfies the Armijo condition (4.11), or
- ii. there exists a step size  $\delta_k \in (0, 1)$  such that the step  $x_k^j + \delta_k p_k$  does not satisfy the Armijo condition (4.11) and it holds

$$\left\| \nabla f_{\varepsilon_j, \mu_j}(x_k^j + \theta \delta_k p_k) - \nabla f_{\varepsilon_j, \mu_j}(x_k^j) \right\|_2 \leq C^L L_{\varepsilon_j, \mu_j} \theta \delta_k \|p_k\|_2 \quad \forall \theta \in (0, 1),$$

where  $x_k^j$  denotes the last iterate and  $p_k$  the current search direction.

For clarification it is noted that the constants  $C^x$ ,  $C'$  and  $C^L$  are assumed to be given globally. In part c), on the other hand, the distinction whether case i. or ii. holds as well as the step size  $\delta_k$  may differ with every iteration. Without loss of generality, the assumption implies that  $x_k^j + \delta_k p_k$  is strictly feasible.

The set of all points which are bounded by the constant  $C^x$  will from here on be denoted by

$$\Omega := \{x \in \mathbb{R}^n \mid \|x\|_2 \leq C^x\}. \quad (4.15)$$

Moreover, for any parameters  $\varepsilon, \mu > 0$ , the set of all points at which further the gradient norm of  $f_{\varepsilon, \mu}$  is bounded by  $C'$  will be denoted by

$$\mathcal{N}_{\varepsilon, \mu} := \{x \in \mathcal{F}_+ \cap \Omega \mid \|\nabla f_{\varepsilon, \mu}(x)\|_2 \leq C'\}. \quad (4.16)$$

By construction, all iterates  $x_k^j$  lie in the respective set  $\mathcal{N}_{\varepsilon_j, \mu_j}$ , and the entire Tikhonov-regularised IPM operates on  $\Omega$ .

### 4.5.3 Derivation of the quantitative estimates

With these assumptions and definitions, the quantitative estimates which are necessary for globally analysing the inner convergence rate can be derived. Since the set  $\Omega$  is compact, the extreme value theorem of continuous functions on compact sets [29, page 31] yields the existence of Lipschitz constants for the gradients of  $f$  and all  $g_i$  as well as a global bound on the gradient of all  $g_i$ .

**Lemma 4.8.** *Let  $\Omega$  be defined as above. There exist constants  $L^{\nabla f}, L^{\nabla g}, C^{\nabla g} > 0$  such that for all  $x, y \in \Omega$  it holds:*

$$\begin{aligned} \|\nabla f(x) - \nabla f(y)\|_2 &\leq L^{\nabla f} \|x - y\|_2, \\ \|\nabla g_i(x) - \nabla g_i(y)\|_2 &\leq L^{\nabla g} \|x - y\|_2 && \text{for } i = 1, \dots, m, \\ \|\nabla g_i(x)\|_2 &\leq C^{\nabla g} && \text{for } i = 1, \dots, m. \end{aligned}$$

*Proof.* Similarly to Lemma 4.4, the existence of the constants follows directly from the compactness of  $\Omega$ , the extreme value theorem [29, page 31] applied to the continuous functions  $\|\nabla g_i(\cdot)\|_2$ ,  $\|\nabla^2 g_i(\cdot)\|_2$  and  $\|\nabla^2 f(\cdot)\|_2$ , and from Lemma 4.3.  $\square$

These bounds now allow for the derivation of two key estimates which are needed for the quantification of the convergence: The first limits how close to the boundary of  $\mathcal{F}_+$  a point  $x$  can at most be as long as it lies in  $\mathcal{N}_{\varepsilon, \mu}$ , while the second one is an upper estimate on the Lipschitz constant  $L_{\varepsilon, \mu}$  and states that the latter does not grow faster than  $O(\frac{1}{\mu})$ .

**Lemma 4.9.** *Let Assumption 4.7 hold, and for each outer iteration  $j \in \mathbb{N}$ , let  $\mathcal{N}_{\varepsilon_j, \mu_j}$  be defined as above. Then it holds:*

a) *There exists a constant  $C^b > 0$  such that for any  $j \in \mathbb{N}$  and  $x \in \mathcal{N}_{\varepsilon_j, \mu_j}$  it holds:*

$$\sum_{i=1}^m \frac{1}{-g_i(x)} \leq C^b \frac{1}{\mu_j}. \quad (4.17)$$

b) *For any  $j \in \mathbb{N}$ , the gradient of  $f_{\varepsilon_j, \mu_j}$  is Lipschitz continuous on the set  $\mathcal{N}_{\varepsilon_j, \mu_j}$  with Lipschitz constant*

$$L_{\varepsilon_j, \mu_j} := L^{\nabla f} + C^b L^{\nabla g} + \varepsilon_j + (C^b)^2 (C^{\nabla g})^2 \frac{1}{\mu_j}, \quad (4.18)$$

where  $C^b$  is the constant from statement a), and  $L^{\nabla f}, L^{\nabla g}, C^{\nabla g}$  are the constants from Lemma 4.8.

*Proof.*

a) Firstly, if the set of strictly feasible points  $\mathcal{F}_+$  is empty, then  $\mathcal{N}_{\varepsilon_j, \mu_j} = \emptyset$  for all  $j$ , and the statement is true. Otherwise, the Slater CQ is satisfied, so there exist some Slater point  $x^{\text{Slat}} \in \mathcal{F}_+$  and, by definition of  $\mathcal{F}_+$ , some constant  $\theta > 0$  such that  $-g_i(x^{\text{Slat}}) \geq \theta$  for all  $i = 1, \dots, m$ .

Now, let  $x \in \mathcal{N}_{\varepsilon_j, \mu_j}$ . It holds  $-g_i(x) > 0$  for all  $i$ , but in general, it is unclear how close to 0 the term  $g_i(x)$  might be. The idea for deriving a bound on this ‘‘closeness’’ is to reduce it to the gradient of  $f_{\varepsilon_j, \mu_j}$ , which is bounded on  $\mathcal{N}_{\varepsilon_j, \mu_j}$ . For this, it will be distinguished between those constraints which are more than twice as close to 0 as  $\theta$ , and those which are not. Therefore, let  $I := \{i = 1, \dots, m \mid -g_i(x) \leq \frac{\theta}{2}\}$ . For any  $i \in I$  it holds with the convexity of  $g_i$  that

$$\frac{\theta}{2} \leq \underbrace{g_i(x)}_{\geq -\frac{\theta}{2}} - \underbrace{g_i(x^{\text{Slat}})}_{\leq -\theta} \stackrel{\text{Lemma 2.3 b)}}{\leq} \nabla g_i(x)^T (x - x^{\text{Slat}}),$$

so

$$\sum_{i \in I} \frac{\nabla g_i(x)^T (x - x^{\text{Slat}})}{-g_i(x)} \geq \frac{\theta}{2} \sum_{i \in I} \frac{1}{-g_i(x)}.$$

Thus, the ‘‘smaller’’  $g_i$  can be bounded away from 0 by some term dependent on the gradient of the respective barriers, cf. Section 2.3. For  $i \notin I$ , on the other hand, it is  $-g_i(x) > \frac{\theta}{2}$ , so since  $|\{1, \dots, m\} \setminus I| \leq m$ , it holds

$$\sum_{i \notin I} \frac{1}{-g_i(x)} < \frac{2m}{\theta}.$$

Together, this yields

$$\begin{aligned} \sum_{i=1}^m \frac{1}{-g_i(x)} &= \sum_{i \notin I} \frac{1}{-g_i(x)} + \sum_{i \in I} \frac{1}{-g_i(x)} \\ &< \frac{2m}{\theta} + \frac{2}{\theta} \sum_{i \in I} \frac{\nabla g_i(x)^T (x - x^{\text{Slat}})}{-g_i(x)} \\ &= \frac{2m}{\theta} + \frac{2}{\theta} \sum_{i=1}^m \frac{\nabla g_i(x)^T (x - x^{\text{Slat}})}{-g_i(x)} - \frac{2}{\theta} \sum_{i \notin I} \frac{\nabla g_i(x)^T (x - x^{\text{Slat}})}{-g_i(x)}. \end{aligned}$$

For the last term, it follows from the proof of Lemma 2.16 that

$$- \sum_{i \notin I} \frac{\nabla g_i(x)^T (x - x^{\text{Slat}})}{-g_i(x)} \leq m - |I| \leq m,$$

while the central term includes all constraints and thereby is precisely the inner product of  $(x - x^{\text{Slat}})$  and the gradient of the barrier term  $b(x) = -\sum_{i=1}^m \ln(-g_i(x))$ , which is given by  $\nabla b(x) = -\sum_{i=1}^m \frac{1}{g_i(x)} \nabla g_i(x)$ . Hence,

$$\sum_{i=1}^m \frac{1}{-g_i(x)} < \frac{4m}{\theta} + \frac{2}{\theta} \nabla b(x)^T (x - x^{\text{Slat}}).$$

With  $f_{\varepsilon_j, \mu_j}(x) = f_{\varepsilon_j}(x) + \mu_j b(x)$  and the Cauchy-Schwarz inequality [37, page 600] it now follows

$$\begin{aligned} \sum_{i=1}^m \frac{1}{-g_i(x)} &\leq \frac{4m}{\theta} + \frac{2}{\theta \mu_j} (\nabla f_{\varepsilon_j, \mu_j}(x) - \nabla f_{\varepsilon_j}(x))^T (x - x^{\text{Slat}}) \\ &\leq \frac{4m}{\theta} + \frac{2}{\theta \mu_j} (\|\nabla f_{\varepsilon_j, \mu_j}(x)\|_2 + \|\nabla f_{\varepsilon_j}(x)\|_2) \|x - x^{\text{Slat}}\|_2. \end{aligned}$$

Finally, Assumption 4.7 comes into play: Since  $x \in \mathcal{N}_{\varepsilon_j, \mu_j}$ , it implies directly that  $\|\nabla f_{\varepsilon_j, \mu_j}(x)\|_2 \leq C'$ , and furthermore that there exist global constants  $\tilde{C}_1, \tilde{C}_2 > 0$  such that

$$\|\nabla f_{\varepsilon_j}(x)\|_2 \leq \tilde{C}_1 \quad \text{and} \quad \|x - x^{\text{Slat}}\|_2 \leq \tilde{C}_2$$

for any  $j \in \mathbb{N}$ . These bounds exist as  $\|x\|_2 \leq C^x$ , and as  $\|\nabla f(\cdot)\|_2$  is bounded on the compact set  $\Omega$ , so both terms

$$\|\nabla f_{\varepsilon_j}(x)\|_2 \leq \|\nabla f(x)\|_2 + \varepsilon_j \|x\|_2 \leq \|\nabla f(x)\|_2 + \varepsilon_0 \sqrt{C^x}$$

and

$$\|x - x^{\text{Slat}}\|_2 \leq \|x\|_2 + \|x^{\text{Slat}}\|_2 \leq \sqrt{C^x} + \|x^{\text{Slat}}\|_2$$

are bounded on  $\Omega$  independent of the parameters  $\varepsilon_j, \mu_j$ .

With  $\mu_0 > \mu_1 > \dots > 0$ , it finally follows that there exists a global constant  $C^b > 0$  such that

$$\sum_{i=1}^m \frac{1}{-g_i(x)} \leq \frac{4m}{\theta} + \frac{2}{\theta \mu_j} (C' + \tilde{C}_1) \tilde{C}_2 \leq C^b \frac{1}{\mu_j}$$

for all  $j \in \mathbb{N}$  and  $x \in \mathcal{N}_{\varepsilon_j, \mu_j}$ , so (4.17) is shown.

- b) Let  $x, y \in \mathcal{N}_{\varepsilon_j, \mu_j}$ . To derive the Lipschitz constant for the gradient of  $f_{\varepsilon_j, \mu_j}$ , the latter will be split into its components: By the triangle inequality [37, page 600] and the Lipschitz continuity of  $\nabla f$ , given by Lemma 4.8, it holds:

$$\begin{aligned} \|\nabla f_{\varepsilon_j, \mu_j}(x) - \nabla f_{\varepsilon_j, \mu_j}(y)\|_2 &= \|\nabla f(x) - \nabla f(y) + \varepsilon_j(x - y) + \mu_j(\nabla b(x) - \nabla b(y))\|_2 \\ &\leq \underbrace{\|\nabla f(x) - \nabla f(y)\|_2}_{L^{\nabla f} \|x - y\|_2} + \varepsilon_j \|x - y\|_2 + \mu_j \|\nabla b(x) - \nabla b(y)\|_2. \end{aligned}$$

For the remaining term  $\|\nabla b(x) - \nabla b(y)\|_2$ , a bound on the Hessian of the barrier term  $b$  will be derived and afterwards Lemma 4.3 used to transfer that bound to the Lipschitz constant for the gradient of  $b$ . For any  $x \in \mathcal{N}_{\varepsilon_j, \mu_j}$  it holds:

$$\begin{aligned} \|\nabla^2 b(x)\|_2 &= \left\| \sum_{i=1}^m \left( \frac{1}{g_i(x)^2} \nabla g_i(x) \nabla g_i(x)^T - \frac{1}{g_i(x)} \nabla^2 g_i(x) \right) \right\|_2 \\ &\leq \sum_{i=1}^m \left\| \frac{1}{g_i(x)^2} \nabla g_i(x) \nabla g_i(x)^T \right\|_2 + \sum_{i=1}^m \left\| \frac{1}{g_i(x)} \nabla^2 g_i(x) \right\|_2 \\ &\leq \sum_{i=1}^m \left( \frac{1}{-g_i(x)} \|\nabla g_i(x)\|_2 \right)^2 + \sum_{i=1}^m \frac{1}{-g_i(x)} \|\nabla^2 g_i(x)\|_2. \end{aligned}$$

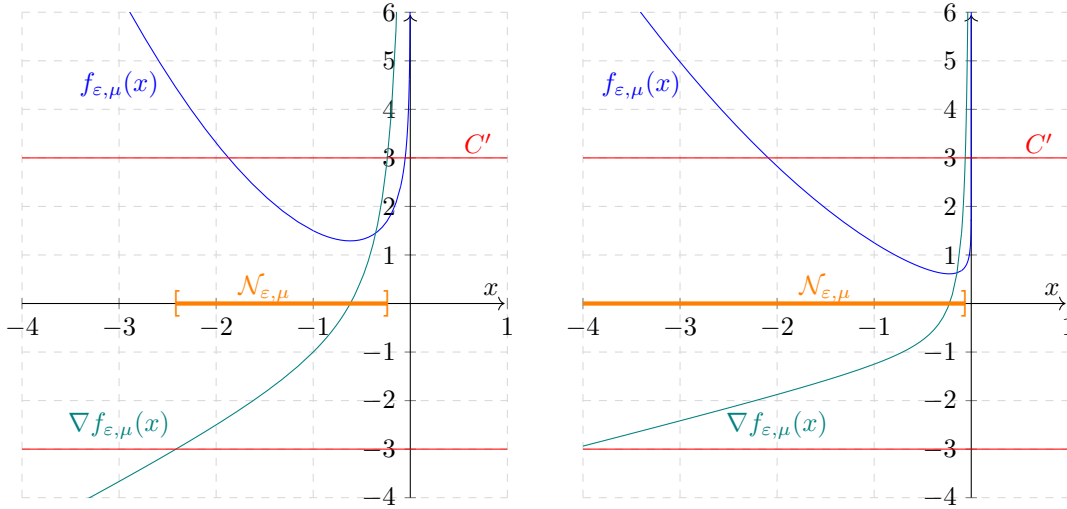
Now, Lemma 4.8 and its proof yield the bounds  $\|\nabla g_i(x)\|_2 \leq C^{\nabla g}$  and  $\|\nabla^2 g_i(x)\|_2 \leq L^{\nabla g}$  for all  $i = 1, \dots, m$ , so together with the estimate from part a) of this lemma it follows

$$\begin{aligned} \|\nabla^2 b(x)\|_2 &\leq \sum_{i=1}^m \left( \frac{1}{-g_i(x)} \underbrace{\|\nabla g_i(x)\|_2}_{\leq C^{\nabla g}} \right)^2 + \sum_{i=1}^m \frac{1}{-g_i(x)} \underbrace{\|\nabla^2 g_i(x)\|_2}_{\leq L^{\nabla g}} \\ &\leq (C^{\nabla g})^2 \sum_{i=1}^m \left( \frac{1}{-g_i(x)} \right)^2 + L^{\nabla g} \sum_{i=1}^m \frac{1}{-g_i(x)} \\ &\leq (C^{\nabla g})^2 \underbrace{\left( \sum_{i=1}^m \frac{1}{-g_i(x)} \right)^2}_{\leq \frac{C^b}{\mu_j}} + L^{\nabla g} \underbrace{\sum_{i=1}^m \frac{1}{-g_i(x)}}_{\leq \frac{C^b}{\mu_j}} \\ &\leq (C^{\nabla g})^2 (C^b)^2 \frac{1}{\mu_j^2} + L^{\nabla g} C^b \frac{1}{\mu_j}. \end{aligned}$$

Hence, Lemma 4.3 yields that the gradient of  $b$  is Lipschitz continuous on  $\mathcal{N}_{\varepsilon_j, \mu_j}$  with Lipschitz constant  $(C^b)^2 (C^{\nabla g})^2 \frac{1}{\mu_j^2} + C^b L^{\nabla g} \frac{1}{\mu_j}$ . Altogether, this ultimately gives

$$\|\nabla f_{\varepsilon_j, \mu_j}(x) - \nabla f_{\varepsilon_j, \mu_j}(y)\|_2 \leq \left( L^{\nabla f} + \varepsilon_j + (C^b)^2 (C^{\nabla g})^2 \frac{1}{\mu_j} + C^b L^{\nabla g} \right) \|x - y\|_2,$$

from which precisely the Lipschitz constant  $L_{\varepsilon_j, \mu_j}$  specified in (4.18) follows.  $\square$



**Figure 4.2:** Visualisation of the Tikhonov-regularised barrier function  $f_{\epsilon, \mu}$  (blue), its gradient (teal) and the resulting set  $\mathcal{N}_{\epsilon, \mu}$  (orange) for  $f(x) = -x$ ,  $g(x) = x$ , and parameters  $\epsilon = 1$ ,  $\mu = 1$  (left) and  $\epsilon = \frac{1}{2}$ ,  $\mu = \frac{1}{4}$  (right). When  $\epsilon$  and  $\mu$  are decreased, then the minimiser of  $f_{\epsilon, \mu}$  approaches 0, the boundary of the feasible set. Analogously, also the set  $\mathcal{N}_{\epsilon, \mu}$  gets closer to 0, as the global constant  $C'$  is fixed to 3. Lemma 4.9 a) states that  $\frac{1}{-x} \leq \frac{C^b}{\mu}$  for all  $x \in \mathcal{N}_{\epsilon, \mu}$ . Thus, the smaller  $\mu$ , the closer might  $\mathcal{N}_{\epsilon, \mu}$  be to the boundary, which is precisely the case here.

*Remark.* The dependencies of both estimates in  $\mu_j$  are not too surprising: The statements are limited to the set  $\mathcal{N}_{\epsilon_j, \mu_j}$  which enforces a certain proximity to the minimiser  $\bar{x}_{\epsilon_j, \mu_j}$ . The latter, in turn, lies in the strictly feasible set  $\mathcal{F}_+$ , and the gradient of  $f_{\epsilon_j, \mu_j}$  is zero in it. Now, when approaching the boundary of  $\mathcal{F}_+$ , this is accompanied by an increase in the gradient of  $f_{\epsilon_j, \mu_j}$  and thus only possible within  $\mathcal{N}_{\epsilon_j, \mu_j}$  to a certain degree. A larger barrier parameter  $\mu_j$  thereby puts more weight on the barrier term, so a greater distance to the boundary of  $\mathcal{F}_+$  is ensured. Conversely, when reducing  $\mu_j$ , the constraints  $g_i$  are allowed to get closer to 0, yielding the in 4.17 observed increase of the upper bound.

Similarly, the Lipschitz constant  $L_{\epsilon_j, \mu_j}$  might increase with smaller barrier parameter  $\mu_j$ , as it bounds how much curvature the function  $f_{\epsilon_j, \mu_j}$  can at most have on  $\mathcal{N}_{\epsilon_j, \mu_j}$ , and this curvature might increase rapidly near to the boundary of  $\mathcal{F}_+$ . This behaviour is illustrated in Figure 4.2.

The same computations as in the proof of Lemma 4.9 b) furthermore yield that, when Assumption 4.7 holds, the norm of the structured part of the seed matrix  $S_k$  defined by (4.9) is in each outer iteration  $j \in \mathbb{N}$  bounded for all corresponding inner iterations  $k \in \mathbb{N}_0$  by

$$\begin{aligned} \|S_k\|_2 &\leq \|\epsilon_j I\|_2 + \left\| \mu_j \sum_{i=1}^m \frac{1}{g_i(x_k^j)^2} \nabla g_i(x_k^j) \nabla g_i(x_k^j)^T \right\|_2 \\ &\leq \epsilon_j + (C^b)^2 (C^{\nabla g})^2 \frac{1}{\mu_j}. \end{aligned} \quad (4.19)$$

Here, it was used that  $x_k^j \in \mathcal{N}_{\epsilon_j, \mu_j}$  by definition.

Lemma 4.9 a) also yields another interesting observation: In Assumption 4.7, it is assumed that the norm of the gradient of  $f_{\varepsilon_j, \mu_j}$  evaluated at the respective iterates  $x_k^j$  remains bounded throughout the entire iteration. And in fact, the equation

$$\nabla f_{\varepsilon_j, \mu_j}(x) = \nabla f_{\varepsilon_{j-1}, \mu_{j-1}}(x) + (\varepsilon_j - \varepsilon_{j-1})x - (\mu_j - \mu_{j-1}) \sum_{i=1}^m \frac{1}{g_i(x)} \nabla g_i(x)$$

implies that, for  $x \in \mathcal{N}_{\varepsilon_{j-1}, \mu_{j-1}}$ , the norm of the next gradient will at most constitute of

$$\|\nabla f_{\varepsilon_j, \mu_j}(x)\|_2 \leq \|\nabla f_{\varepsilon_{j-1}, \mu_{j-1}}(x)\|_2 + (\varepsilon_{j-1} - \varepsilon_j)C^x + (\mu_{j-1} - \mu_j)C^b C^{\nabla g} \frac{1}{\mu_{j-1}}.$$

Hence, when running the previous inner loop until  $\|\nabla f_{\varepsilon_{j-1}, \mu_{j-1}}(x^{j-1})\|_2 \leq C\varepsilon_{j-1}^2$  and using the linear update procedure  $\varepsilon_j \leftarrow \beta\varepsilon_{j-1}$ ,  $\mu_j \leftarrow \beta^\gamma\mu_{j-1}$  as in line 4 of Algorithm 3.2, it holds at the first iterate  $x_j^0 := x^{j-1}$  of the new inner loop that

$$\|\nabla f_{\varepsilon_j, \mu_j}(x_j^0)\|_2 \leq \underbrace{\|\nabla f_{\varepsilon_{j-1}, \mu_{j-1}}(x_j^0)\|_2}_{\leq C\varepsilon_{j-1}^2} + (1 - \beta)C^x\varepsilon_{j-1} + (1 - \beta^\gamma)C^b C^{\nabla g}, \quad (4.20)$$

which lies in  $O(1)$  as  $j \rightarrow \infty$ .

With these estimates, the only missing parts for estimating the constants  $\hat{\beta}$  and  $\nu$  are an upper and lower bound on the step sizes  $\alpha_k$ . As is proven in the following lemma, these can be derived using Assumption 4.7 c) and the bound for the norm of  $B_k^{-1}$  derived in Lemma 4.6.

**Lemma 4.10.** *Let Assumption 4.7 hold, let  $\varepsilon_j \leq 1$ , and let the corresponding Lipschitz constant  $L_{\varepsilon_j, \mu_j} \geq 1$ . The step sizes  $\alpha_k$  computed during the iterations of Algorithm 4.2 with line-search parameters  $c_{\text{ls}}, \rho_{\text{ls}} \in (0, 1)$  are for all  $k \in \mathbb{N}_0$  bounded by*

$$1 \geq \alpha_k \geq \frac{(1 - c_{\text{ls}})\rho_{\text{ls}}}{C^L L_{\varepsilon_j, \mu_j} \|B_k^{-1}\|_2} \geq \frac{(1 - c_{\text{ls}})\rho_{\text{ls}}}{C^L 5^\ell} \left( \frac{\varepsilon_j}{L_{\varepsilon_j, \mu_j}} \right)^{2\ell+1}. \quad (4.21)$$

*Proof.* First note that, by Lemma 4.6, the norm of  $B_k^{-1}$  is bounded by (4.14), so the last inequality holds. By the concept of the Armijo backtracking line-search described in Section 4.3, the step size  $\alpha_k$  is the largest, i.e. first, of the values  $1, \rho_{\text{ls}}, \rho_{\text{ls}}^2, \rho_{\text{ls}}^3, \dots$  for which the Armijo condition (4.11) is satisfied. This already implies  $\alpha_k \leq 1$ .

For the remaining lower bound, the idea now is to use the fact that  $\alpha_k$  is precisely the smallest possible power of  $\rho_{\text{ls}}$ , so it either is 1 or its predecessor has still violated the Armijo condition. If  $\alpha_k = 1$  already satisfies the Armijo condition, then the statement (4.21) trivially follows. Otherwise, the step size is given by  $\alpha_k = \rho_{\text{ls}}^i$  with exponent  $i \in \mathbb{N}$  for which (4.11) is satisfied, while for  $\rho_{\text{ls}}^{i-1} \leq 1$  it is not.

By Assumption 4.7 c), there exists some step size  $\delta_k \in (0, 1)$  such that the Armijo condition is violated at the point  $x_k^j + \delta_k p_k \in \mathcal{F}_+$  and it furthermore holds

$$\left\| \nabla f_{\varepsilon_j, \mu_j}(x_k^j + \theta \delta_k p_k) - \nabla f_{\varepsilon_j, \mu_j}(x_k^j) \right\|_2 \leq C^L L_{\varepsilon_j, \mu_j} \|\theta \delta_k p_k\|_2 \quad \forall \theta \in (0, 1). \quad (4.22)$$

Since  $f_{\varepsilon_j, \mu_j}$  is convex and the Armijo condition becomes satisfied somewhere on the interval  $[\rho_{\text{ls}}^i, \rho_{\text{ls}}^{i-1})$ , the step size  $\delta_k$  can be chosen from the interval  $(\rho_{\text{ls}}^i, \rho_{\text{ls}}^{i-1}]$ . As the Armijo condition (4.11) is violated for  $\delta_k$ , it holds

$$f_{\varepsilon_j, \mu_j}(x_k^j + \delta_k p_k) > f_{\varepsilon_j, \mu_j}(x_k^j) + c_{\text{ls}} \delta_k \nabla f_{\varepsilon_j, \mu_j}(x_k^j)^T p_k.$$

Following the computations of [33, page 11], this can be rewritten with the mean value theorem [37, page 629] as

$$-c_{\text{ls}} \delta_k \nabla f_{\varepsilon_j, \mu_j}(x_k^j)^T p_k > f_{\varepsilon_j, \mu_j}(x_k^j) - f_{\varepsilon_j, \mu_j}(x_k^j + \delta_k p_k) = -\delta_k \nabla f_{\varepsilon_j, \mu_j}(x_k^j + \theta \delta_k p_k)^T p_k$$

for some  $\theta \in (0, 1)$ . Dividing by  $-\delta_k$  and subtracting  $\nabla f_{\varepsilon_j, \mu_j}(x_k^j)^T p_k$  yields

$$(c_{\text{ls}} - 1) \nabla f_{\varepsilon_j, \mu_j}(x_k^j)^T p_k < \left( \nabla f_{\varepsilon_j, \mu_j}(x_k^j + \theta \delta_k p_k) - \nabla f_{\varepsilon_j, \mu_j}(x_k^j) \right)^T p_k,$$

which, by the Cauchy-Schwarz inequality [37, page 600] and (4.22), can be estimated by

$$\begin{aligned} (c_{\text{ls}} - 1) \nabla f_{\varepsilon_j, \mu_j}(x_k^j)^T p_k &< \left\| \nabla f_{\varepsilon_j, \mu_j}(x_k^j + \theta \delta_k p_k) - \nabla f_{\varepsilon_j, \mu_j}(x_k^j) \right\|_2 \|p_k\|_2 \\ &\leq \theta \delta_k C^L L_{\varepsilon_j, \mu_j} \|p_k\|_2^2. \end{aligned}$$

Finally, with the inverse norm of  $B_k^{-1}$  and the search direction  $p_k = -B_k^{-1} \nabla f_{\varepsilon_j, \mu_j}(x_k^j)$  used in the l-BFGS method (cf. Section 4.1), it follows that

$$\frac{1 - c_{\text{ls}}}{\|B_k^{-1}\|_2} \|p_k\|_2^2 = (1 - c_{\text{ls}}) \lambda_{\min}(B_k) \|p_k\|_2^2 \leq (1 - c_{\text{ls}}) p_k^T B_k p_k < \theta \delta_k C^L L_{\varepsilon_j, \mu_j} \|p_k\|_2^2,$$

which yields

$$\frac{1 - c_{\text{ls}}}{C^L L_{\varepsilon_j, \mu_j} \|B_k^{-1}\|_2} < \theta \delta_k < \delta_k \leq \rho_{\text{ls}}^{i-1}.$$

Therefore, it holds

$$\alpha_k = \rho_{\text{ls}}^i = \rho_{\text{ls}} \rho_{\text{ls}}^{i-1} > \frac{(1 - c_{\text{ls}}) \rho_{\text{ls}}}{C^L L_{\varepsilon_j, \mu_j} \|B_k^{-1}\|_2},$$

and statement (4.21) is proven.  $\square$

#### 4.5.4 A global rate for the number of inner iterations

With these estimates, now, all components are prepared which are necessary for the global quantification of the convergence rate of the inner loop. In Theorem 4.5, it was stated that, for given parameters  $\varepsilon_j, \mu_j > 0$ , the inner iterates converge r-linearly in the residual gradient norm with the upper bound

$$\left\| \nabla f_{\varepsilon_j, \mu_j}(x_k^j) \right\|_2^2 \leq \nu^k \left( \frac{\hat{\beta}}{1 - \sqrt{\nu}} \right)^2 \frac{L_{\varepsilon_j, \mu_j}^2}{\varepsilon_j} \left\| \nabla f_{\varepsilon_j, \mu_j}(x_0^j) \right\|_2^2. \quad (4.23)$$

To quantify the convergence rate and especially to derive a rate for the number of inner iterations which are at most necessary in the entire outer loop, the constants

$$\hat{\beta} := \sup_{k \in \mathbb{N}_0} \sqrt{\frac{\alpha_k \|B_k^{-1}\|_2}{c_{\text{ls}}}} \quad \text{and} \quad \nu := \sup_{k \in \mathbb{N}_0} \left( 1 - \frac{c_{\text{ls}} \alpha_k \varepsilon}{\|B_k\|_2} \right)$$



must be bounded with respect to the parameters  $\varepsilon_j$  and  $\mu_j$ .

For this, the upper bounds

$$\|B_k\|_2 \leq 2 \|S_k\|_2 + L_{\varepsilon_j, \mu_j} + \frac{\ell L_{\varepsilon_j, \mu_j}^2}{\varepsilon_j} \leq 2\varepsilon_j + 2(C^b)^2 (C^{\nabla g})^2 \frac{1}{\mu_j} + L_{\varepsilon_j, \mu_j} + \frac{\ell L_{\varepsilon_j, \mu_j}^2}{\varepsilon_j}$$

and

$$\|B_k^{-1}\|_2 \leq \frac{5^\ell}{\varepsilon_j} \left( \frac{L_{\varepsilon_j, \mu_j}}{\varepsilon_j} \right)^{2\ell},$$

given by Lemma 4.6 and (4.19), and the bounds on the step size

$$\frac{(1 - c_{\text{ls}})\rho_{\text{ls}}}{CL5^\ell} \left( \frac{\varepsilon_j}{L_{\varepsilon_j, \mu_j}} \right)^{2\ell+1} \leq \alpha_k \leq 1,$$

derived in Lemma 4.10, yield the estimates

$$\hat{\beta} = \sup_{k \in \mathbb{N}_0} \sqrt{\frac{\alpha_k \|B_k^{-1}\|_2}{c_{\text{ls}}}} \leq \sqrt{\frac{5^\ell}{\varepsilon_j c_{\text{ls}}} \left( \frac{L_{\varepsilon_j, \mu_j}}{\varepsilon_j} \right)^{2\ell}}$$

and

$$\nu = \sup_{k \in \mathbb{N}_0} \left( 1 - \frac{c_{\text{ls}} \alpha_k \varepsilon_j}{\|B_k\|_2} \right) \leq 1 - \frac{c_{\text{ls}}(1 - c_{\text{ls}})\rho_{\text{ls}}\varepsilon_j^{2\ell+2}}{CL5^\ell L_{\varepsilon_j, \mu_j}^{2\ell+1} \left( 2\varepsilon_j + 2(C^b)^2 (C^{\nabla g})^2 \frac{1}{\mu_j} + L_{\varepsilon_j, \mu_j} + \frac{\ell L_{\varepsilon_j, \mu_j}^2}{\varepsilon_j} \right)}.$$

These estimates hold for the entire inner loop of one outer iteration  $j$ .

The goal behind the global quantification now is to express these results and thereby ultimately the entire convergence as a rate in the parameters  $\varepsilon_j$  and  $\mu_j$  which decrease during the method. For this, the upper estimate of the Lipschitz constant  $L_{\varepsilon_j, \mu_j}$  which was derived in Lemma 4.9 b) yields that it is at most increasing at a rate of

$$L_{\varepsilon_j, \mu_j} = L^{\nabla f} + C^b L^{\nabla g} + \varepsilon_j + (C^b)^2 (C^{\nabla g})^2 \frac{1}{\mu_j} = O\left(\frac{1}{\mu_j}\right). \quad (4.24)$$

Hence, the convergence parameter  $\hat{\beta}$  which is part of the prefactor in the estimate (4.23) increases during the outer iterations at a rate of at most

$$\hat{\beta} \leq \sqrt{\frac{5^\ell}{\varepsilon_j c_{\text{ls}}} \left( \frac{L_{\varepsilon_j, \mu_j}}{\varepsilon_j} \right)^{2\ell}} = O\left(\frac{1}{\varepsilon_j^{\ell+0.5} \mu_j^\ell}\right) \quad (4.25)$$

as  $j \rightarrow \infty$ , whereas the parameter  $\nu$  which determines the minimal convergence speed in each inner iteration loop and furthermore also occurs in the prefactor, satisfies that  $1 - \nu$  decreases to 0 at most at a rate of

$$\begin{aligned} 1 - \nu &\geq \frac{c_{\text{ls}}(1 - c_{\text{ls}})\rho_{\text{ls}}\varepsilon_j^{2\ell+2}}{CL5^\ell L_{\varepsilon_j, \mu_j}^{2\ell+1} \left( 2\varepsilon_j + 2(C^b)^2 (C^{\nabla g})^2 \frac{1}{\mu_j} + L_{\varepsilon_j, \mu_j} + \frac{\ell L_{\varepsilon_j, \mu_j}^2}{\varepsilon_j} \right)} \\ &= \Omega\left(\frac{\varepsilon_j^{2\ell+2}}{\frac{1}{\mu_j^{2\ell+1}} \left( \varepsilon_j + \frac{1}{\mu_j} + \frac{1}{\mu_j} + \frac{1}{\varepsilon_j \mu_j^2} \right)}\right) \\ &= \Omega\left(\varepsilon_j^{2\ell+3} \mu_j^{2\ell+3}\right), \end{aligned} \quad (4.26)$$

and correspondingly

$$1 - \sqrt{\nu} = \Omega\left(\varepsilon_j^{\ell+1.5} \mu_j^{\ell+1.5}\right) \quad (4.27)$$

as  $j \rightarrow \infty$ .

This can now finally be used to quantify the rate at which the number of iterations required in each inner loop evolves throughout the overall method. As motivated in Section 3.4 and used as stopping criterion in Algorithm 4.2, the objective of each inner iteration procedure is to reduce the residual gradient norm of  $f_{\varepsilon_j, \mu_j}$  until it lies below  $C\varepsilon_j^2$ . The estimate (4.23) given by the convergence result in Theorem 4.5 implies that this is the case when

$$\nu^k \left( \frac{\hat{\beta}}{1 - \sqrt{\nu}} \right)^2 \frac{L_{\varepsilon_j, \mu_j}^2}{\varepsilon_j} \left\| \nabla f_{\varepsilon_j, \mu_j}(x_0^j) \right\|_2^2 \stackrel{!}{\leq} C^2 \varepsilon_j^4.$$

Altogether, the estimates (4.20), (4.24), (4.25) and (4.27) yield that the prefactor on the left-hand side increases at most at a rate of

$$\begin{aligned} \left( \frac{\hat{\beta}}{1 - \sqrt{\nu}} \right)^2 \frac{L_{\varepsilon_j, \mu_j}^2}{\varepsilon_j} \left\| \nabla f_{\varepsilon_j, \mu_j}(x_0^j) \right\|_2^2 &= O\left( \frac{1}{\varepsilon_j^{2\ell+1}} \frac{1}{\mu_j^{2\ell}} \frac{1}{\varepsilon_j^{2\ell+3}} \frac{1}{\mu_j^{2\ell+3}} \frac{1}{\varepsilon_j \mu_j^2} \right) \\ &= O\left( \frac{1}{\varepsilon_j^{4\ell+5} \mu_j^{4\ell+5}} \right). \end{aligned} \quad (4.28)$$

Therefore, the stopping criterion  $\left\| \nabla f_{\varepsilon_j, \mu_j}(x_k^j) \right\|_2 \leq C\varepsilon_j^2$  is guaranteed to be satisfied when it holds for the number of inner iterations  $k$  that

$$\nu^k = O\left(\varepsilon_j^{4\ell+9} \mu_j^{4\ell+5}\right)$$

as  $j \rightarrow \infty$ .

If the inner convergence rate  $\nu$  was constant throughout the entire method, then this would yield that the number of inner iterations which are required at most in one outer step increases at most at a rate of  $k_{\max} = O(|\ln \varepsilon_j| + |\ln \mu_j|)$ . However, in (4.26) it was only shown that the rate at which  $1 - \nu$  decreases to 0 can be bounded below by  $\Omega\left(\varepsilon_j^{2\ell+3} \mu_j^{2\ell+3}\right)$ , not that it is globally bounded away from 0. Therefore, “only” the following estimate on the rate at which the number of inner iterations increases at most can be derived.

**Theorem 4.11.** *Consider the Tikhonov-regularised IPM described in Algorithm 3.2 with the structured l-BFGS method of Algorithm 4.2 applied as inner solver, and assume that Assumption 4.7 holds. Then the maximum number of inner iterations  $k_{\max}$  required at each outer step  $j$  satisfies*

$$k_{\max} = O\left( \frac{|\ln \varepsilon_j| + |\ln \mu_j|}{\varepsilon_j^{2\ell+3} \mu_j^{2\ell+3}} \right) \quad (4.29)$$

as  $j \rightarrow \infty$ .

*Proof.* As stated above, Theorem 4.5 yields that in each outer iteration  $j$ , the inner stopping criterion  $\left\| \nabla f_{\varepsilon_j, \mu_j}(x_k^j) \right\|_2 \leq C\varepsilon_j^2$  of line 3 in Algorithm 4.2 is satisfied when

$$\underbrace{(\nu_j)^k \left( \frac{\hat{\beta}}{1 - \sqrt{\nu_j}} \right)^2 \frac{L_{\varepsilon_j, \mu_j}^2}{\varepsilon_j} \left\| \nabla f_{\varepsilon_j, \mu_j}(x_0^j) \right\|_2^2}_{=: A_j} \leq C^2 \varepsilon_j^4. \quad (4.30)$$

Note that  $\nu$  is denoted with an index  $j$  here to emphasise that it changes during the outer iterations and thereby affects the convergence rate. Defining the prefactor in (4.30) as  $A_j$  and solving the inequality for  $k$  gives

$$\begin{aligned} & (\nu_j)^k A_j \leq C \varepsilon_j^4 \\ \Leftrightarrow & \underbrace{k \ln \nu_j + \ln A_j}_{<0} \leq \ln C + 4 \ln \varepsilon_j \\ \Leftrightarrow & k \geq \frac{\ln C + 4 \ln \varepsilon_j - \ln A_j}{\ln \nu_j}. \end{aligned}$$

By (4.28),  $A_j$  lies in the order  $O\left(\frac{1}{\varepsilon_j^{4\ell+5} \mu_j^{4\ell+5}}\right)$ , so for sufficiently large  $j$  it holds

$$A_j \leq \frac{C_1}{\varepsilon_j^{4\ell+5} \mu_j^{4\ell+5}}$$

for some constant  $C_1 > 0$ . Furthermore, (4.26) implies that

$$\nu_j \leq 1 - C_2 \varepsilon_j^{2\ell+3} \mu_j^{2\ell+3}$$

for some  $C_2 > 0$ .

Thus, for sufficiently large  $j$ , (4.30) and thereby the inner stopping criterion are satisfied when

$$k \geq \frac{\ln C + 4 \ln \varepsilon_j - \ln C_1 + (4\ell + 5) \ln \varepsilon_j + (4\ell + 5) \ln \mu_j}{\ln\left(1 - C_2 \varepsilon_j^{2\ell+3} \mu_j^{2\ell+3}\right)}.$$

The term in the numerator lies in  $O(|\ln \varepsilon_j| + |\ln \mu_j|)$  (note that  $\varepsilon_j$  and  $\mu_j$  converge to 0, so the logarithms become negative), while the term in the denominator yields an order of  $O\left(\frac{1}{\varepsilon_j^{2\ell+3} \mu_j^{2\ell+3}}\right)$  since it holds  $-\frac{1}{\ln(1-x)} = O\left(\frac{1}{x}\right)$  for  $x \rightarrow 0^+$ .

By taking both rates together, the order in (4.29) is obtained.  $\square$

*Remark.* As can be seen in Theorem 4.11, the number of l-BFGS iterations required in each outer step heavily depends on the update procedure for the parameters  $\varepsilon_j$  and  $\mu_j$  (line 4 in Algorithm 3.2). Intuitively, the greater these updates, the more inner iterations will be needed as they go along with a greater distance between the previous optimum  $\bar{x}_{\varepsilon_{j-1}, \mu_{j-1}}$  and the new inner minimiser  $\bar{x}_{\varepsilon_j, \mu_j}$ . It especially follows that the barrier parameter  $\mu_j$  should not be updated much faster than the regularisation parameter  $\varepsilon_j$ , leading to the recommendation for a rather small update parameter  $\gamma > 1$  in Algorithm 3.2.

The dependency on the l-BFGS memory length  $\ell$  arises from the problem that the  $\varepsilon$ -strong positive definiteness of  $B_k^{(0)}$  cannot be ensured to be maintained during the l-BFGS updates, hence yielding only the estimate (4.14). However, since  $\ell = \min\{k, \ell_{\max}\}$ , it is limited by the chosen maximum memory length  $\ell_{\max}$ .

Theorem 4.11 yields that there is some constant  $\tilde{C} > 0$  such that in each outer iteration  $j \in \mathbb{N}$ , the number of required inner iterations lies below  $\tilde{C} \frac{|\ln \varepsilon_j| + |\ln \mu_j|}{\varepsilon_j^{2\ell+3} \mu_j^{2\ell+3}}$ . In a final step, this upper estimate can now be used to establish a bound on the rate at which the total number of l-BFGS iterations in the overall Tikhonov-regularised IPM of Algorithm 3.2 increases. This rate will be derived in the next section, concluding the discussion in this chapter with the polynomial complexity result Theorem 4.12.

## 4.6 Polynomial complexity of the Tikhonov-regularised interior-point method

As final part of the convergence analysis conducted in this chapter, in this section, a polynomial complexity result for the Tikhonov-regularised IPM of Algorithm 3.2 is derived. It states that the total number of l-BFGS iterations required to reach a desired tolerance grows at most polynomially in this tolerance as it is reduced. The result is based on the upper estimate for the number of inner iterations required in each outer step which is given in Theorem 4.11.

The implications of Theorem 4.11 for the convergence analysis of the overall method are as follows: When the linear update procedure  $\varepsilon_j \leftarrow \beta\varepsilon_{j-1}$ ,  $\mu_j \leftarrow \beta^\gamma\mu_{j-1}$  is employed as proposed in Algorithm 3.2, then the overall method converges r-linearly in the objective with respect to the number of outer iterations (see Theorem 3.12), but in worst case, the number of necessary inner iterations at each outer step might grow exponentially with increasing outer iteration index  $j$ ; or at least, it is not shown otherwise here. This can be verified by inserting the corresponding parameters  $\varepsilon_j = \beta^j\varepsilon_0$  and  $\mu_j = \beta^{\gamma j}\mu_0$  into the upper iteration bound (4.29), by which one obtains

$$k_{\max} = O\left(\frac{|j \ln \beta| + |\gamma j \ln \beta|}{\beta^{j(2\ell+3)}\beta^{\gamma j(2\ell+3)}}\right) = O\left(je^{-j \ln \beta(1+\gamma)(2\ell+3)}\right) \quad (4.31)$$

as  $j \rightarrow \infty$ . As  $\beta < 1$  implies  $\ln \beta < 0$ , a positive exponent is given in the last term.

Hence, an exponential growth in the number of inner iterations cannot be ruled out with Theorem 4.11 for the linear update procedure. However, this exponential growth is only observed when considering the convergence with respect to  $j$ , that is, with the update choice given here, with respect to an “exponential improvement” in the accuracy of the solution, such as an accuracy improvement by one decimal place.

On the other hand, when analysing the convergence in terms of the actually imposed solution accuracy, that is, in terms of the tolerance  $\tau > 0$ , or, equivalently, in terms of a parameter  $\varepsilon_*$  which must be reached in the outer stopping criterion (line 3 of Algorithm 3.2), then this corresponds only to a logarithmic increase in the number of outer iterations. Hence, the overall number of necessary l-BFGS iterations is polynomial in  $\varepsilon_*$  and thereby in  $\tau$ .

**Theorem 4.12.** *Consider the Tikhonov-regularised IPM described in Algorithm 3.2 with the structured l-BFGS method of Algorithm 4.2 applied as inner solver, and assume that Assumption 4.7 holds. Then the total number of inner iterations  $K$  which are at most required to reach a tolerance  $\tau > 0$  increases at most polynomially in  $\tau$ . Specifically, it holds*

$$K = O\left((\ln \tau)^2 \left(\frac{1}{\tau}\right)^{(1+\gamma)(2\ell+3)}\right) \quad (4.32)$$

for  $\tau \rightarrow 0^+$ .

*Proof.* First note analogously to the proof of Theorem 3.12 that the outer iterates  $x^j$  produced by Algorithm 3.2 are bounded. Thereby, also the norm of the gradient of  $f$  evaluated in these iterates remains bounded. Hence, with  $\mu_j = o(\varepsilon_j)$  it follows that the regularisation parameter  $\varepsilon_*$  at which the left term in the stopping criterion in line 3 falls below the tolerance  $\tau$ , is linear in the latter.

Now let  $J \in \mathbb{N}$  denote the smallest number of outer iterations after which this parameter  $\varepsilon_*$  is reached. With above argument, it holds

$$\varepsilon_J = \Omega(\tau).$$

By (4.31), the number of inner iterations  $k_J$  required in the  $J$ -th outer step is bounded at a rate

$$k_J = O\left(Je^{-J \ln \beta(1+\gamma)(2\ell+3)}\right). \quad (4.33)$$

The total number of inner iterations  $K$  is given as the sum over the number of all inner iterations  $k_1, \dots, k_J$ . But since, in their limit, all those  $k_j$  are bounded by (4.31), and the term inside the brackets and thereby the maximum number of iterations is monotonically increasing with  $j$ , all  $k_1, \dots, k_J$  are in particular bounded by (4.33). Hence,

$$K = \sum_{j=1}^J k_j = J \cdot O\left(Je^{-J \ln \beta(1+\gamma)(2\ell+3)}\right) = O\left(J^2 e^{-J \ln \beta(1+\gamma)(2\ell+3)}\right). \quad (4.34)$$

The update rule  $\varepsilon_{j+1} \leftarrow \beta \varepsilon_j$  yields that  $\varepsilon_J = \beta^J \varepsilon_0$ , so it follows

$$J = \frac{\ln \frac{\varepsilon_J}{\varepsilon_0}}{\ln \beta}.$$

Inserting this formula for  $J$  into (4.34) yields

$$\begin{aligned} K &= O\left(J^2 e^{-J \ln \beta(1+\gamma)(2\ell+3)}\right) \\ &= O\left(\left(\frac{\ln \frac{\varepsilon_J}{\varepsilon_0}}{\ln \beta}\right)^2 e^{-\frac{\ln \frac{\varepsilon_J}{\varepsilon_0}}{\ln \beta} \ln \beta(1+\gamma)(2\ell+3)}\right) \\ &= O\left(\left(\frac{\varepsilon_J}{\varepsilon_0}\right)^2 \left(\frac{\varepsilon_J}{\varepsilon_0}\right)^{-(1+\gamma)(2\ell+3)}\right) \\ &= O\left((\ln \varepsilon_J)^2 \left(\frac{1}{\varepsilon_J}\right)^{(1+\gamma)(2\ell+3)}\right). \end{aligned}$$

Here, the order is considered for  $J \rightarrow \infty$  or, equivalently, for  $\varepsilon_J \rightarrow 0^+$ .

With  $\varepsilon_J = \Omega(\tau)$ , it finally follows

$$K = O\left((\ln \tau)^2 \left(\frac{1}{\tau}\right)^{(1+\gamma)(2\ell+3)}\right),$$

and the polynomial complexity result (4.32) is proven.  $\square$

*Remark.* A more elaborate rate could be obtained when taking the monotone increase of  $j$  into account in (4.34) instead of simply estimating it by the maximal value  $J$ . But for the concept of the polynomial complexity result (4.32), this is not necessary and therefore omitted here.

As alternative to the linear update procedure considered in Algorithm 3.2 and the discussion here, also a sublinear update rule such as the harmonic choice  $\varepsilon_j \leftarrow \frac{1}{j}\varepsilon_0$ ,  $\mu_j \leftarrow \frac{1}{j^\gamma}\mu_0$  with  $\gamma > 1$  could be used. It still satisfies the requirement  $\mu_j = o(\varepsilon_j)$  as  $j \rightarrow \infty$ , so all results except for the linear convergence rate in Theorem 3.12 remain valid. With this choice, the increase in the number of inner iterations is by (4.29) then restricted to

$$k = O\left(\frac{\left|\ln \frac{\varepsilon_0}{j}\right| + \left|\ln \frac{\mu_0}{j^\gamma}\right|}{\frac{\varepsilon_0^{2\ell+3}}{j^{2\ell+3}} \frac{\mu_0^{2\ell+3}}{j^{\gamma(2\ell+3)}}}\right) = O\left(\ln j j^{(1+\gamma)(2\ell+3)}\right),$$

which, in difference to the exponential growth in the linear update, is polynomial in the outer iteration index  $j$ .

However, this sublinear update scheme only yields a sublinear convergence rate in the objective  $f$  with respect to the outer iterations, as Theorem 3.11 b) states that the rate of this convergence lies in  $O(\varepsilon_j)$ . Therefore, an "exponential accuracy improvement" such as by one decimal place goes along with an exponential increase in the number of necessary outer iterations, so overall, the number of l-BFGS iterations would again grow exponentially in this case.

On the other hand, when considering the improvement directly in terms of the imposed tolerance  $\tau$  as it is done in Theorem 4.12, then the number of outer iterations is now linear in  $\tau$ , and again, a polynomial number of total inner iterations is obtained. This polynomial rate is similar to the rate (4.32) derived for the linear update procedure, and can easily be verified when inserting  $J = \frac{\varepsilon_*}{\varepsilon_0}$  into the estimate above and continuing as in the proof of Theorem 4.12.

# 5

## Numerical results

In the context of this thesis, the proposed and analysed Tikhonov-regularised IPM of Algorithm 3.2 was implemented and numerically tested. It is written in Python and published on GitHub [22]. As final part of the thesis, in this chapter, numerical results of the method are presented for nonlinear inequality-constrained problems, structured as follows.

In Section 5.1, details on the implementation are given. Afterwards in Section 5.2, the used test problems are described in detail. In Section 5.3, numerical results are presented for low-dimensional problems. These validate the theoretical convergence results of Chapters 3 and 4 and are used to discuss the numerical convergence behaviour of the method in detail. Finally, results for large-scale problems are presented and discussed in Section 5.4.

### 5.1 Implementation details

The code mainly consists of two files: The file “qnregipm.py” (Quasi-Newton based regularised interior-point method) contains a class with a solver for purely inequality constrained problems as given in  $(\mathcal{P})$ , whereas the file “qnregipmm.py” (Quasi-Newton based regularised interior-point method of multipliers) contains an extension of this class for problems with additional linear equality constraints. There, the Augmented Lagrangian method described in Section 3.5 is used.

Both classes support the use of Tikhonov regularisation as well as of the proximal-point regularisation outlined in Section 3.5, and for the inner solver it can be chosen between the structured l-BFGS method described in Section 4.3 and an l-BFGS version in which the seed matrix is only computed as scaled identity, cf. Section 4.1. If the structured l-BFGS method is used, then it optionally allows to use a matrix-free version of the seed matrix, which can be more suitable especially for large-scale problems. Furthermore, it can be decided whether the linear system in the l-BFGS two loop recursion (cf. line 6 in Algorithm 4.1) is solved exactly, or only computed approximately up to a specified tolerance using MINRES [38].

The Augmented Lagrangian method in the equality constrained case is implemented in a way that the barrier parameter  $\rho$  is initialised with  $\varepsilon_0$  and afterwards increased at the same speed as the reciprocal decrease of the barrier parameter  $\mu$ ; for details, see Section 3.5. In parts, this follows the idea of [41]. However,  $\rho$  is only updated here as long as the residual  $\|Ax - b\|_2^2$  does not lie below the tolerance  $\tau$ , in order to reduce the ill-conditioning of the overall method. It should be noted that in the equality constrained case, the upper estimate (3.16) used as stopping criterion does not in general hold as the iterates  $x^j$  will likely not satisfy the equality constraints exactly. However, the method is not terminated

until  $\|Ax - b\|_2^2$  lies below  $\tau$ .

To avoid issues with machine accuracy, the line-search breaks when it reaches  $\alpha \leq 10^{-12}$ . Similarly, the inner iteration is stopped prematurely when the size of the latest update in the iterate is too small, specifically, when  $\|\alpha_k p_k\|_2 < 10^{-12}$ . Furthermore, the factor  $\tau_{k+1}$  is only computed as long as the corresponding denominator lies above  $10^{-13}$ ; otherwise, the previous factor  $\tau_{k+1}$  is used. Finally, the method only imposes an inner approximation accuracy, i.e.  $\|\nabla f_{\varepsilon_j, \mu_j}(x^j)\|_2$ , of at most  $10^{-8}$ . Once the inner tolerance  $C\varepsilon_j^2$  falls below this threshold, it is kept at  $10^{-8}$ .

In addition to the two files mentioned above, the GitHub repository [22] also includes the code by which the numerical results and plots presented in Sections 5.3 and 5.4 were generated, as well as the programs used to illustrate Example 3.7.

## 5.2 Test problems

The results presented in this chapter were obtained by applying the Tikhonov-regularised IPM to the following three problems: Firstly, to the non-convex two-dimensional *Rosenbrock function* [37]

$$f(x_1, x_2) := 100(x_2 - x_1^2)^2 + (1 - x_1)^2,$$

which was augmented with the inequality constraints

$$g(x_1, x_2) := \begin{pmatrix} -x_1 \\ x_2^2 - 3 \\ x_2 - 1 \\ (x_1 - 1)^2 + (x_2 + 1)^2 - 4 \end{pmatrix} \leq 0.$$

Since the unique minimiser  $x^* = (1, 1)^T$  of the unconstrained problem lies on the boundary of the feasible set, it is also the unique solution for the constrained problem. The corresponding minimum is given by  $f^* := f(x^*) = 0$ . As strictly feasible starting point,  $x^0 = (1.5, 0.5)^T$  was used.

Secondly, following [8, page 573], the method was applied to inequality-constrained *geometric programs* (GP). Specifically, problems of the form

$$\begin{aligned} \inf_{x \in \mathbb{R}^n} \ln \left( \sum_{l=1}^{L_{\text{obj}}} \exp(F_l x + b_l) \right) \quad \text{s.t.} \quad & \ln \left( \sum_{l=1}^{L_{\text{obj}}} \exp(-F_l x - b_l + 1 - 2 \ln L_{\text{obj}}) \right) \leq 0 \\ & \ln \left( \sum_{l=1}^{L_{\text{ineq}}} \exp(G_{i,l} x + c_{i,l}) \right) \leq 0 \quad \text{for } i = 2, \dots, m \end{aligned}$$

were considered, where  $n, m, L_{\text{obj}}, L_{\text{ineq}} \in \mathbb{N}$ ,  $F_l, G_{i,l} \in \mathbb{R}^{1 \times n}$ , and  $b_l, c_l \in \mathbb{R}$  for all  $i$  and  $l$ .

The inequality constraints were specifically designed in the context of this thesis to suit the considered setting. By their choice, it is ensured that, on the feasible set, the objective function always lies above 1. Thus, the optimisation problem is bounded below, and for the minimal value it holds  $f^* \geq 1$ . In fact, it is  $f^* = 1$  if there exists  $x \in \mathcal{F}$  such



that  $F_l x + b_l = 1 - \ln L_{\text{obj}}$  for all  $l$ . Whether this is the case, therefore depends especially on the number of “objective summands”  $L_{\text{obj}}$  and constraints  $m$  relative to the problem dimension  $n$ .

The geometric programs were generated randomly by choosing the parameters  $F_l \in [-2, -1)^n$ ,  $G_{i,l} \in [1, 2)^n$ ,  $b_l \in [0, 1)$  and  $c_l \in [4, 5)$  from a uniform distribution on the respective interval. As starting point,

$$x^0 = \min \left\{ -\frac{\max_{i,l} c_{i,l} - \ln L_{\text{ineq}}}{n \min_{i,l,\nu} G_{i,l,\nu}}, \frac{\max_l b_l + \ln L_{\text{obj}} - 1}{-n \max_{l,\nu} F_{l,\nu}} \right\} (1, \dots, 1)^T$$

was used. By its design, it is guaranteed to be strictly feasible, while avoiding numerical problems which occur when  $x$  is too small.

The third problem class for which results are presented consists of *quadratically-constrained quadratic programs* (QCQP) [8, Section 4.4]

$$\inf_{x \in \mathbb{R}^n} \frac{1}{2} x^T P_0 x + q_0^T x \quad \text{s.t.} \quad \frac{1}{2} x^T P_i x + q_i^T x + r_i \leq 0 \quad \text{for } i = 1, \dots, m,$$

where  $P_0, \dots, P_m \in \mathbb{R}^{n \times n}$  are symmetric positive definite,  $q_0, \dots, q_m \in \mathbb{R}^n$ , and the offsets satisfy  $r_1, \dots, r_m < 0$ . Again, the problems were built randomly: The matrices  $P_0, \dots, P_m$  were obtained by randomly generating sparse matrices  $A_i \in [0, 10)^{n \times n}$  and computing  $P_i := A_i^T A_i + I$ . The entries of the vectors  $q_0, \dots, q_m$  were chosen uniformly from the interval  $[0, 1)$ , and the offsets  $r_i$  uniformly from  $[-100, -1)$ . By this choice, the used starting point  $x^0 = (0, \dots, 0)^T$  is strictly feasible.

In some of the figures presented next, a comparison to the optimal function value  $f^*$  and the optimal solution  $x^*$  is given. As described above, these are known for the Rosenbrock function. In the other cases, the parameters  $f^*$  and  $x^*$  were set to the minimal objective function value obtained during the iterations and to the corresponding iterate, respectively. Furthermore, if possible, solutions to the same problems were computed with CVXOPT [3] for the GPs and with CVXPY [15] for the QCQPs, and the parameters  $f^*$  and  $x^*$  were correspondingly updated when a smaller objective value was achieved.

All results were obtained by employing the structured l-BFGS method presented in Algorithm 4.2 as inner solver, where the linear systems were solved approximately with tolerance  $10^{-10}$ . Thereby, accurate initial guesses were computed while allowing for the use of matrix-free techniques. As line-search parameters,  $c_{\text{ls}} = 10^{-5}$  and  $\rho_{\text{ls}} = 0.5$  were used, and the l-BFGS memory size was limited to  $\ell_{\text{max}} = 5$ .

### 5.3 Results for small-scale problems

In order to numerically analyse the convergence behaviour of the Tikhonov-regularised IPM and compare it to the theoretical results derived in Chapters 3 and 4, at first, numerical results are presented for all three problems with a small number of variables: For the problem based on the Rosenbrock function, it holds  $n = 2$  and  $m = 4$ , as described in Section 5.2; the geometric program was generated with  $n = 4$ ,  $m = 6$ ,  $L_{\text{obj}} = 5$  and  $L_{\text{ineq}} = 5$ , and the

QCQP was generated with  $n = 10$  and  $m = 5$ . For all problems, the method was applied with parameters  $\varepsilon_1 = \mu_1 = 1$ ,  $\beta = 0.9$ ,  $\gamma = 1.1$ . and  $C = 1$ . It was applied until a tolerance of  $10^{-6}$  for Rosenbrock and the QCQP, and  $10^{-3}$  for the GP was reached in the upper bound (5.1).

In Figure 5.1, the convergence results of all three problems with respect to the outer iterations  $j$  are presented. The first row shows the difference between the function value in the current iterate  $x^j$  and the optimal solution  $f^*$ , as well as the upper bound on the function value which was used as stopping criterion. The latter was computed in every iteration based on the estimate (3.14) by

$$\text{“upper bound”} := \frac{1}{\varepsilon_j} \|\nabla f(x^j)\|_2 \|\nabla f_{\varepsilon_j, \mu_j}(x^j)\|_2 + m\mu_j + 2\varepsilon_j \|x^j\|_2^2. \quad (5.1)$$

The plots validate that this estimate bounds the residual in the function value from above, and that both terms decrease at a linear rate with respect to the barrier parameter  $\varepsilon_j$ , visualised by the linear graph on the logarithmically scaled  $y$ -axis. For the GP, high fluctuations in the upper bound were observed which are likely caused by the precautions taken to avoid numerical problems, as is further discussed below.

The linear convergence is also clearly visible in the plots in the second row, which furthermore show a convergence rate of 0.9 for Rosenbrock and GP, and even of 0.8 for most of the QCQP iteration. In all three cases, an update rate  $\beta = 0.9$  was used, so the numerical observation is consistent with the theoretical result presented in Theorem 3.12 by which the  $r$ -linear convergence rate is given by at least  $\beta$ . In addition, the plots mostly even show a  $q$ -linear convergence in the iterate which is stronger than the theoretically proven  $r$ -linear result.

In case of the QCQP, the method reached an accuracy of  $10^{-9}$  after about 100 iterations. However, the method continued for 40 additional iterations until the upper bound used as stopping criterion fell below the imposed accuracy of  $10^{-6}$ . In these iterations, hardly any further improvement was observed which is likely caused by the aforementioned precautions.

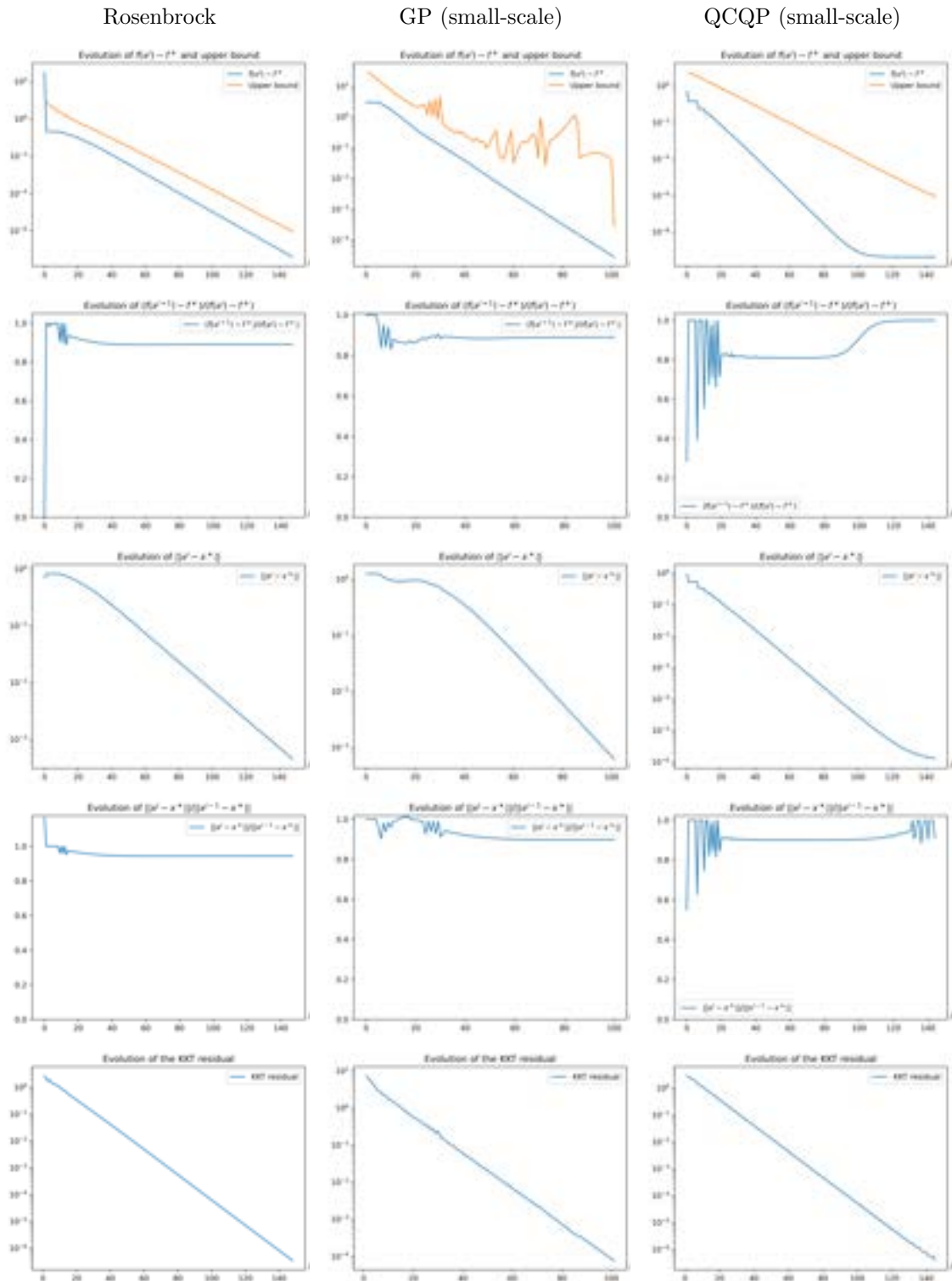
Similarly to the first two rows of Figure 5.1, the third and fourth row show the evolution of the distance between the outer iterates  $x^j$  and the optimal solution  $x^*$ . Although no theoretical results for the convergence rate in the iterates were derived in this thesis, the plots indicate that, at least for the problems considered here, the convergence is again  $q$ -linear, albeit with a slightly slower rate.

In the last row of Figure 5.1, an approximation to the KKT residual is presented for each outer iterate  $x^j$ . It is based on the KKT conditions of the original problem ( $\mathcal{P}$ ) which were defined in Theorem 2.11, and it is computed by

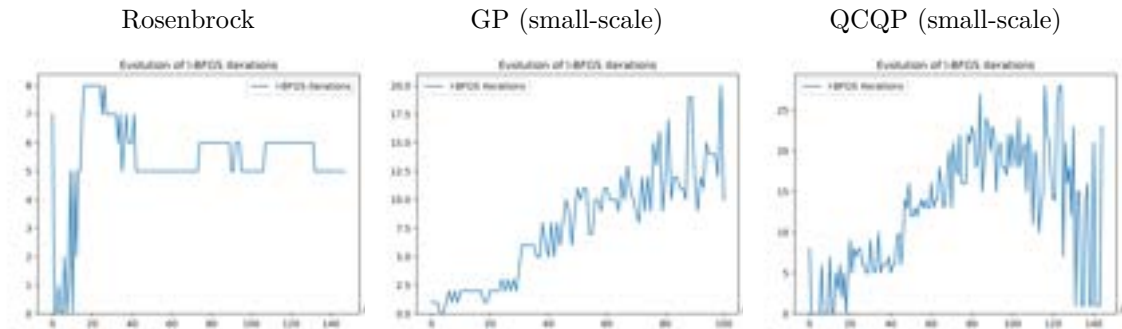
$$\begin{aligned} \text{“KKT residual”} &:= \left\| \nabla f(x^j) + \sum_{i=1}^m (\bar{\lambda}_{\mu_j})_i \nabla g_i(x^j) \right\|_2 + \left\| \left( (\bar{\lambda}_{\mu_j})_1 g_1(x^j), \dots, (\bar{\lambda}_{\mu_j})_m g_m(x^j) \right)^T \right\|_2 \\ &= \left\| \nabla f(x^j) - \mu_j \sum_{i=1}^m \frac{1}{g_i(x^j)} \nabla g_i(x^j) \right\|_2 + \mu_j \sqrt{m}, \end{aligned} \quad (5.2)$$

where the approximation  $(\bar{\lambda}_{\mu_j})_i := -\frac{\mu_j}{g_i(x^j)}$  is used for the Lagrange multiplier, cf. (2.7)–(2.8). As the first term corresponds to the gradient norm of  $f_{\varepsilon_j, \mu_j}$  after removing

## 5 Numerical results



**Figure 5.1:** Evolution of the Tikhonov-regularised IPM of Algorithm 3.2 in the outer iterations for a small number of variables. The first two rows show how the method converges linearly in the objective at a rate of  $\beta = 0.9$  for Rosenbrock and GP, and even of 0.8 for the QCQP. Similarly, the plots in the third and fourth row indicate that linear convergence is obtained in the iterates. In the last row, the decrease of the KKT residual is presented, which again is linear in the logarithmic  $y$ -axis.



**Figure 5.2:** Number of inner iterations required in each outer step of the Tikhonov-regularised IPM of Algorithm 3.2 for a small number of variables. In the GP and QCQP scenarios, a general trend for an approximately linear increase is observed, whereas for Rosenbrock, the number of iterations mainly remains constant.

the part  $\varepsilon_j x^j$ , and this gradient norm as well as the barrier parameter  $\mu_j$  are decreased linearly with rate  $\beta^\gamma$  and  $\beta^2$ , respectively, a linear decrease of the KKT residual with respect to the logarithmic  $y$ -axis is observed.

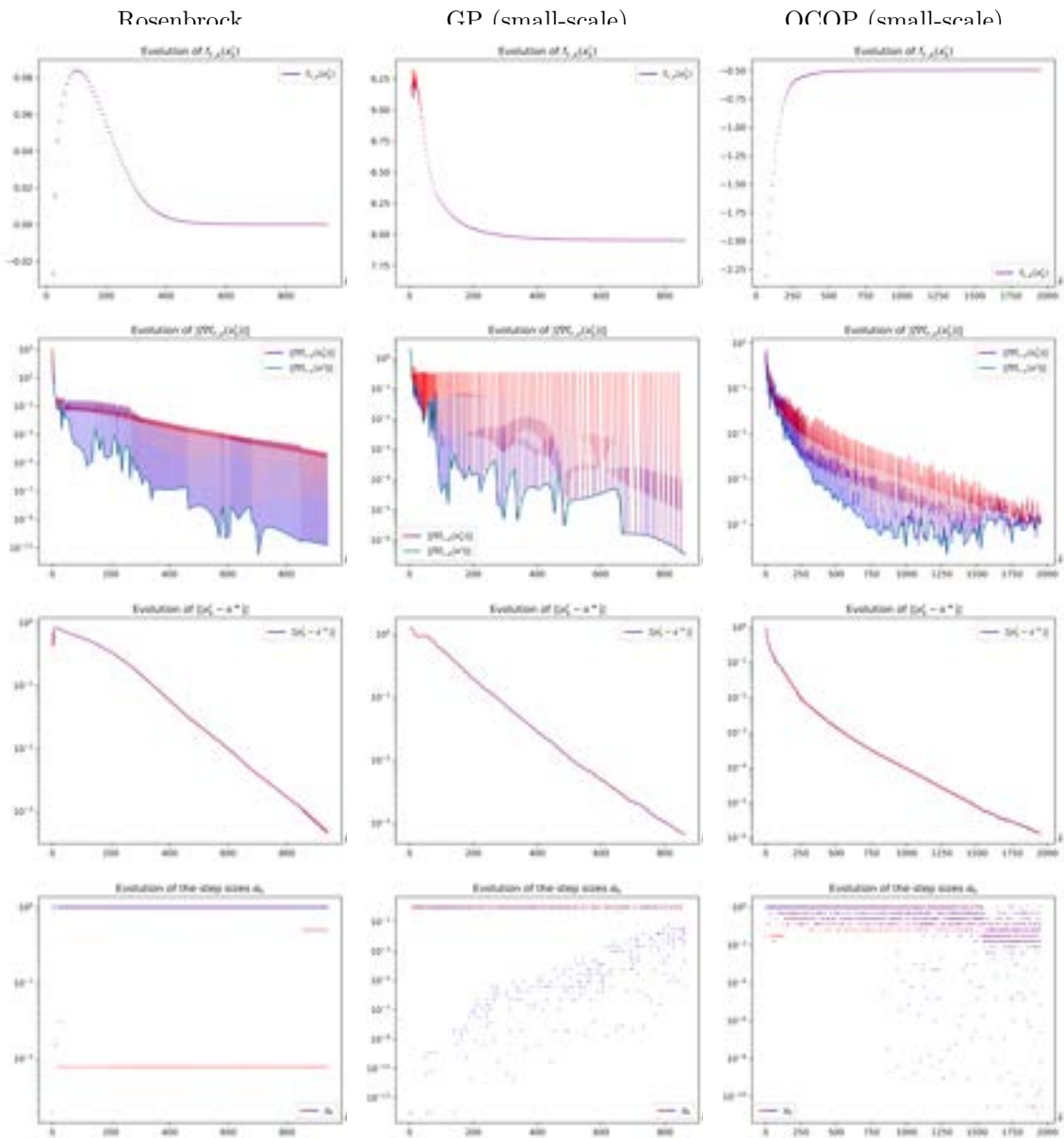
In Figure 5.2, the number of l-BFGS iterations required in each outer step is shown. Although it fluctuates over the steps, a general trend for an increasing number of required iterations was observed in case of the GP and QCQP. In the latter, this trend is at least given during the first 100 outer iterations in which the method did not interfere with the precautions regarding the machine accuracy.

Results on the inner iterations of the method are presented in Figure 5.3. It shows the inner iteration progress over the entire method. The  $x$ -axis therefore describes the accumulated number of inner iterations plus 1, denoted by  $\bar{k}$ . In particular,  $\bar{k} = 0, 1, 2, \dots$  corresponds to the inner iterate  $x_0^0, x_1^0, x_2^0, \dots$ , until, after  $k$  inner iterations, the iterate  $x_k^0$  is obtained for which the first inner loop is terminated. Then,  $\bar{k} = k + 1$  continues with  $x_0^1 = x^1 = x_k^0$ . This duplicate appearance of all outer iterates is purposely done to include their value in  $f_{\varepsilon_j, \mu_j}$  and the gradient norm of the latter both for the outer iteration which it had terminated, as well as for the following outer iteration in which it was used as starting point.

The plots are designed as follows. For each outer iteration, the values for the corresponding inner iterations are shown with a colour gradient: from red for the first inner iterate  $x_0^j$  to blue for the last inner iterate  $x_k^j$ . On the transition between two outer iterations, the connecting lines are omitted. Therefore, jumps are clearly visible whenever they occur.

In the first row of Figure 5.3, it is shown how the value of the regularised barrier function  $f_{\varepsilon_j, \mu_j}$  evolved during the iterations. As the function consists, on the one hand, of the objective  $f$  which decreases throughout the outer iterations, and on the other hand of the regularisation term and of the possibly negative barrier term, it might decrease as well as increase with progressing outer iterations, as the results for the GP and the QCQP illustrate. Within each inner loop, however, Theorem 4.2 guarantees that  $f_{\varepsilon_j, \mu_j}(x_k^j)$  is strictly decreasing.

## 5 Numerical results



**Figure 5.3:** Inner iteration behaviour of the Tikhonov-regularised IPM of Algorithm 3.2 for a small number of variables. The results for each inner loop are shown with a colour gradient: from red in the first inner iterate to blue in the last one. By construction of  $f_{\varepsilon_j, \mu_j}$ , its values can increase as well as decrease with progressing outer iterations; in each inner loop, however, they strictly decrease. The overall method demands that each inner loop reduces the gradient norm of  $f_{\varepsilon_j, \mu_j}$  below  $C\varepsilon_j^2$ . With the subsequent update of  $\varepsilon_j$  and  $\mu_j$ , a sudden increase in the gradient norm is observed, ultimately resulting in an overall increasing margin between the initial and final gradient norm in each inner loop. The inner iterates mainly follow the overall trend of the outer iterates observed in Figure 5.1. Their trajectories furthermore indicate that they follow a consistent pattern in each inner loop. In the last row, the step sizes are presented which were obtained in the inner iterations.

Below this, in the second row of Figure 5.3, the evolution of the gradient norm of  $f_{\varepsilon_j, \mu_j}$  is plotted. As each inner loop was terminated once the gradient norm fell below the tolerance  $C\varepsilon_j^2$ , a considerable reduction is observed for each loop. Furthermore, the final gradient norm  $\|\nabla f_{\varepsilon_j, \mu_j}(x^j)\|_2$ , additionally plotted in teal, decreases, with some fluctuation, mainly linearly on the logarithmic  $y$ -scale. Note that the generally increasing number of inner iterations increasingly stretches the  $x$ -axis with respect to the outer iterations, so the decrease appears more damped in the plot.

Following every inner loop, the parameters  $\varepsilon_j$  and  $\mu_j$  were updated, resulting in a sudden increase in the norm of the new gradient. In case of Rosenbrock and the QCQP, the initial gradient norm also decreased with progressing outer iterations. However, in case of the GP, it consistently returned to nearly the same value. This is consistent with the worst-case rate derived in Section 4.5 by which it holds  $\|\nabla f_{\varepsilon_j, \mu_j}(x^{j-1})\|_2 = O(1)$ .

In the first inner loops of the Rosenbrock scenario, it is shown how the gradient norm at first increased, before the curvature information gathered by the l-BFGS update yielded a search direction along which a decrease was obtained. As discussed above, after around 100 outer iterations in the QCQP, corresponding to  $\bar{k} = 1250$ , no further improvement was observed in the function value. The plot in Figure 5.3 indicates that this correlates with the residual gradient norm not being reduced below  $10^{-7}$ .

The reason for this is likely given by the precautions which were taken to avoid problems with the machine accuracy: Once the latest step  $\alpha_k p_k$  was too small, then the inner loop was terminated prematurely. This did not only affect the second half of the outer iterations in the QCQP case, but also occurred throughout most of the GP scenario, resulting in the fluctuations in the gradient norm as well as in the upper bound depicted in Figure 5.1.

In the third row, the evolution of the distance between the current iterate  $x_k^j$  and the solution  $x^*$  is shown—analogously to the third row of Figure 5.1. They follow the general trend of the linear convergence observed in the outer iterates, but the progression within each inner loop shows a certain pattern: In Rosenbrock and the GP, each inner trajectory bends upwards, indicating that the first steps were almost orthogonal to the direction towards  $x^*$ , and improved later with the inclusion of more curvature information in the l-BFGS update. In contrast to that, the decrease observed in the QCQP bends downwards, indicating that the initial steps pointed more towards the overall minimiser  $x^*$  than towards the current regularised barrier minimiser  $\bar{x}_{\varepsilon_j, \mu_j}$ .

The step sizes  $\alpha_k$  which were obtained in the inner iterations by backtracking line-search are presented in the last row of Figure 5.3. Whereas in most iterations the step size  $\alpha_k = 1$  was accepted, much smaller steps could sometimes be observed throughout the GP and in the second half of the QCQP scenario. This likely caused the inner loop to terminate prematurely, as the blue colour of the step sizes indicates. In order to include the cases when the line-search returned  $\alpha_k = 0$ , as the Armijo condition (4.11) was not satisfied for a step size greater than  $10^{-12}$ , these are represented in the plot at the level  $10^{-14}$ . As can be seen in Figure 5.3, this was only the case for a few l-BFGS iterations in the GP scenario.

In summary, the numerical results presented in Figures 5.1 to 5.3 are consistent with the convergence results derived in Chapters 3 and 4. Overall, the Tikhonov-regularised



IPM converges to a solution with the theoretically expected linear convergence rate in the objective, and also the outer iterates appear to converge linearly, at least for the three considered problems. However, it can be observed that the actual deviation from the optimal function value is way below its upper bound which is used as stopping criterion, so the method might be continued much longer than necessary. Hence, it would be worthwhile to develop a more accurate stopping criterion. Furthermore, the method might struggle with too small updates, causing issues with the machine accuracy.

## 5.4 Results for large-scale problems

As presented and discussed in the last section, the theoretical convergence results for the Tikhonov-regularised IPM can also be observed numerically, at least for low-dimensional problems. In this section, numerical results for larger-scale problems are presented: on the one hand, a GP with  $n = 5000$  variables,  $m = 501$  inequality constraints, and  $L_{\text{obj}} = 50$  and  $L_{\text{ineq}} = 5$  summands in the objective and constraint functions. On the other hand, a QCQP was generated with  $n = 10000$  variables and  $m = 500$  inequality-constraints.

For the GP, the method was applied with parameters  $\varepsilon_1 = \mu_1 = 1$ ,  $\beta = 0.95$ ,  $\gamma = 1.1$  and  $C = 100$ ; and for the QCQP, the parameters  $\varepsilon_1 = \mu_1 = 1$ ,  $\beta = 0.9$ ,  $\gamma = 1.2$  and  $C = 1$  were used. In order to illustrate the evolution of the method more extensively, in both scenarios, the tolerance  $\tau$  was chosen sufficiently small so that it was never reached by the stopping criterion (5.1), as is discussed below.

Analogously to Figures 5.1 to 5.3 for the small-scale problems, the results in this section are presented in the two figures 5.4 and 5.5, separated between the outer and inner iteration behaviour.

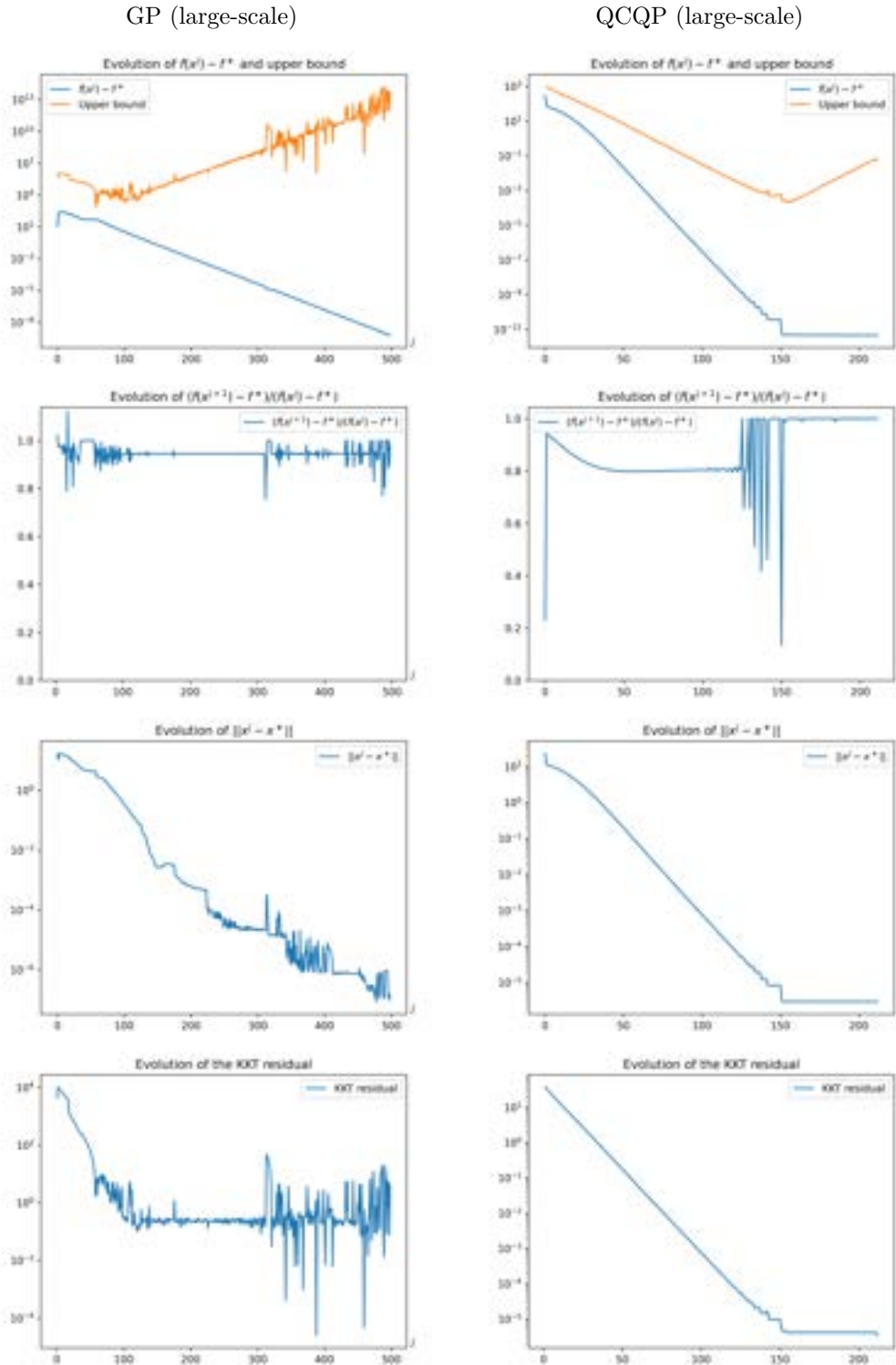
In general, there are many similarities between the results for the large-scale problems and the ones presented in the last section. As is shown in the first two rows of Figure 5.4, mainly q-linear convergence was obtained in the objective. For the GP, the convergence rate coincided with the update parameter  $\beta = 0.95$  as is expected by Theorem 3.12, and in the first 140 iterations, the QCQP again converged at a rate of even 0.8 instead of  $\beta = 0.9$ . After about 150 iterations, an approximation to a solution was obtained for the QCQP for which further improvements were impeded by the machine accuracy.

In both cases, however, the upper bound on the residual began to increase linearly in the logarithmic  $y$ -axis after an initial decrease. The reason for this was the inability of the inner solver to decrease the gradient norm of  $f_{\varepsilon_j, \mu_j}$  to the required tolerance  $C\varepsilon_j^2$ . Instead, it always terminated once the gradient norm had reached a certain level. As can be seen in Figure 5.5, this appears to have been the case around  $10^{-1}$  for the GP. Thus, after about 70 outer iterations, corresponding to  $\bar{k} = 1000$ , the achieved gradient norm started deviating from the imposed inner tolerance.

The inner iterations of the QCQP, on the other hand, terminated prematurely once a tolerance around  $10^{-5}$  was reached, which was the case after about 50 iterations, or, correspondingly,  $\bar{k} = 5000$ . However, the upper bound still decreased for around 100 additional iterations, likely because the terms  $m\mu_j$  and  $2\varepsilon_j \|x^j\|_2^2$  were dominant in (5.1).

The results in Figure 5.4 show that, even though the achieved “inner

## 5 Numerical results



**Figure 5.4:** Evolution of the Tikhonov-regularised IPM of Algorithm 3.2 in the outer iterations for large-scale problems. The objective converges linearly at a rate of  $\beta = 0.95$  (GP) and 0.8 (QCQP). Furthermore, the plots in the third row indicate a linear convergence in the iterates. After about 70 and 150 iterations for the GP and QCQP, respectively, the upper bound starts to increase in  $O(1/\varepsilon_j)$ , as the inner solver is not able to further reduce the gradient norm of  $f_{\varepsilon_j, \mu_j}$ . The KKT residual is affected similarly by this.



accuracy”  $\|\nabla f_{\varepsilon_j, \mu_j}(x^j)\|_2$  was not improved further, the linear convergence in the objective was still maintained, as well as the approximately linear convergence in the iterates.

The impact of the premature inner termination on the KKT residual (5.2) is shown in the last row of Figure 5.4. Whereas it stagnated correspondingly to the gradient norm in the GP scenario, it continued to decrease in case of the QCQP, until the method interfered with the machine accuracy. This different behaviour is likely related to the question whether the norm or the term  $\mu_j\sqrt{m}$  were dominant in the KKT residual (5.2).

In the first row of Figure 5.5, the evolution of the number of required inner iterations is shown. It differs notably between both scenarios: For the GP, the maximum number of iterations which was set 100 there was exhausted without reaching the required accuracy in the first inner loops. Afterwards, it always terminated after around 10 inner iterations due to an insufficiently large update. The number of inner iterations for the QCQP, on the other hand, increased almost linearly until the inner iterations started terminating prematurely.

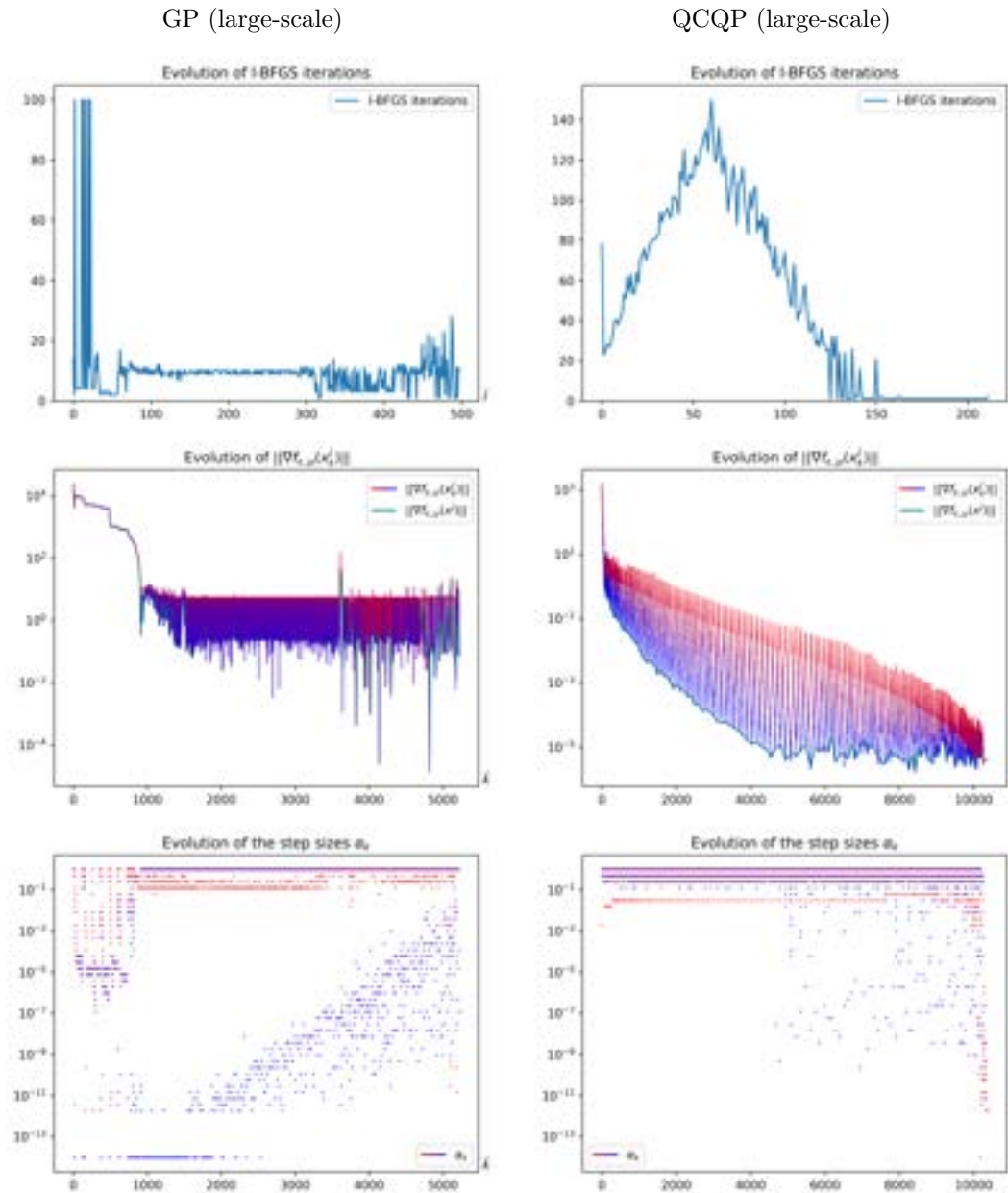
This also corresponds to the evolution of the gradient norm, presented in the third row of Figure 5.5. As, in case of the GP, it always returned to the same value of around  $10^1$  with each outer update, and was then reduced until around  $10^{-1}$ , the inner loop always required about 10 iterations. For the QCQP, on the other hand, both the initial and the ultimately reached gradient norm decreased linearly. But as the latter decreased faster, the difference between both grew, leading to the almost linear increase in the number of inner iterations. Once the premature terminations began, the difference in the gradient norm and thereby also the number of required inner iterations decreased with progressing outer steps.

The premature termination of the inner iterations was caused by the updates  $\alpha_k p_k$  becoming too small. This is related to the step sizes obtained in each inner iteration, presented in the last row of Figure 5.5. Especially in the GP scenario, the zero step sizes, represented by an entry at  $10^{-14}$ , and the generally smaller step sizes correspond to the termination of an inner loop, as their blue colour indicates. The increasing trend in the lower right corner of the plot thereby further indicates that, with progressing outer iterations, the search directions became smaller and smaller, so that even larger step sizes did not yield a sufficiently large update.

In general, it can be expected that higher problem dimensions lead to a larger, i.e. earlier obtained, threshold at which the inner iterations start to terminate prematurely. The reason for this lies in the nature of the Euclidean norm by which even small deviations in each component of the gradient add up to a large gradient norm.

The major problem with this premature inner termination is that, once the residual gradient norm is stagnating instead of being reduced in  $O(\varepsilon_j^2)$ , the definition of the upper bound (5.1) yields an increase of it in  $O(1/\varepsilon_j)$ . As the upper bound is used as stopping criterion for the overall method, the latter might therefore never terminate in practice. This, however, could be avoided by choosing a different stopping criterion for the overall method.

Nonetheless, the results presented in this section indicate that the Tikhonov regularised IPM proposed in this thesis still converges linearly to a solution for large-scale problems in practice, even when the inner loops are terminated prematurely.



**Figure 5.5:** Inner iteration behaviour of the Tikhonov-regularised IPM of Algorithm 3.2 for large-scale problems. In the plots in the second and third row, the results for each inner loop are shown with a colour gradient from red to blue. In the GP scenario, the inner loop is at first either terminated after few iterations or when their maximum number is reached. After around 70 outer steps, corresponding to  $\bar{k} = 1000$ , the inner solver is not able to sufficiently reduce the gradient norm any more. Instead, it is terminated constantly after around 10 iterations due to an insufficiently large update, always reaching a gradient norm of around  $10^{-1}$ . In case of the QCQP, the number of inner iterations first grows linearly, correlating with the increasing margin between initial and final gradient norm. After about 50 outer steps, corresponding to  $\bar{k} = 5000$ , the inner solver reaches its limit for the reduction of the gradient norm. As the margin subsequently decreases, also the number of employed inner iterations decreases.

# 6

## Conclusion

In this thesis, an l-BFGS based Tikhonov-regularised primal interior-point method was proposed, analysed, and numerically validated. The method aims at solving a general inequality-constrained convex problem, and its analysis is based on the mild and for feasible interior-point methods common assumptions that the problem has a minimiser and that a Slater point exists, cf. Assumption 2.10.

The method combines a Tikhonov regularisation with the barrier approach of interior-point methods, thereby allowing for unconstrained optimisation techniques while ensuring strong convexity in each sub-problem. During the iteration, the weight of both regularisation and barrier term is simultaneously reduced. The regularisation parameter is the central parameter in the method, and the barrier parameter as well as the inner approximation accuracy are forced to decrease faster than it. The precise design of the method was presented in Algorithm 3.2.

It was shown that each sub-problem in the method has a unique solution, and that these solutions converge to the minimal-norm solution of the original problem with a convergence rate linear in the regularisation parameter, cf. Theorem 3.6 and Lemma 3.8. Furthermore, conditions on the update procedure for the parameters and the inner approximation accuracy were derived (Section 3.4), and it was proven that also the outer iterates of the method converge to the minimal-norm solution, cf. Theorem 3.12. The convergence was further quantified with an r-linear convergence rate in the objective which was also observed in the numerical results presented in Chapter 5.

The convergence to the minimal-norm solution stands in contrast to other IPMs which typically converge to the analytic centre of the solution set, cf. Section 2.4. An advantage of this different behaviour is the clear characterisation of the limit point and the possibility to influence the optimisation and the obtained solution by shifting the regularisation centre, cf. Section 3.5. Furthermore, convergence of the method was proven in Theorem 3.12 without the additional assumption that the solution set of the problem is bounded, which is otherwise typical for IPMs [4, 18, 46, 48]. As one can easily construct examples, including Linear Programs, for which this is not the case, this is a notable advantage of the method proposed in this thesis.

The integration of a Tikhonov regularisation was mainly motivated with the goal to reduce the issue of increasing nonlinearity and ill-conditionedness which is well-known for primal IPMs, thereby reasonably allowing for the use of Quasi-Newton methods to solve the inner problems. But simultaneously, the different convergence behaviour caused by the regularisation might lead to a faster approach of the iterates to the boundary of the feasible

## 6 Conclusion

set. Thereby, it might also increase the nonlinearity and ill-conditionedness more drastically than other IPMs would, as the latter intrinsically strive to remain as far in the interior of the feasible set as possible.

To make the method suitable for large-scale nonlinear convex problems, an l-BFGS method was employed as inner solver. In contrast to Newton’s method, it requires neither knowledge on the Hessian of the objective and constraint functions, nor the solution of linear systems of equations. As the latter might become prohibitive when a large number of variables or constraints is given [4, 14]—a problem which also affects Quasi-Newton based primal-dual IPMs [5]—the here proposed method aims particularly at large-scale problems.

It was proven in this thesis that the overall method reaches an imposed accuracy in the function value after finitely many iterations, and that, under the requirements stated in Assumption 4.7, the total number of required l-BFGS iterations grows at most polynomially in the accuracy, cf. Theorems 3.12 and 4.12. To the best of our knowledge, this polynomial iteration complexity result is the first of its kind for purely Quasi-Newton based IPMs in nonlinear convex optimisation.

Although the derived rate (4.32) excludes an exponential growth, it still allows for an increase in the number of iterations which might not be suitable in practice. The numerical results presented in Section 5.4, however, indicate that the actual increase in the number of total l-BFGS iterations remains acceptable. Furthermore, at least for the considered test problems, the linear convergence rate was still observed even when the inner problems were not solved as accurately as theoretically required, limiting the number of inner iterations even further in practice.

In contrast to Newton-based IPMs [35, 43], the result does not explicitly estimate the required number of iterations, but only bounds its growth. Furthermore, the constants used in Section 4.5 might differ from problem to problem. Thereby, they especially do not allow to estimate the total number of l-BFGS iterations with respect to the variable and constraint dimensions  $n$  and  $m$ .

To deal with the increasing ill-conditionedness and thereby improve the convergence speed of the inner iterations, a structured l-BFGS method was proposed which aims at capturing the spectrum of the Hessian more properly by including its likely most problematic part directly in the seed matrix. Although the convergence and complexity results derived and presented in Chapter 4 do not cover the actual choice of the seed matrix, but only require that it contains  $\varepsilon I$ , it is reasonable to believe that the choice proposed in Section 4.3 improves the performance considerably.

The main drawback of the proposed structured inner solver is that, in each l-BFGS iteration, it requires the solution of a linear system of equations. This contrasts the concept of the overall method, as the latter explicitly aims at working without linear systems. But since the resultant solution is only used as initial guess for the next search direction, it may be computed only approximately, cf. Section 4.3. Furthermore, with the proposed choice for the seed matrix, a matrix-free implementation can be employed by which the matrix neither needs to be computed nor stored.

The convergence and complexity results derived in this thesis are not explicitly based on the structured l-BFGS method. Therefore, they still hold when the seed matrix is only given

as scaled identity, by which the entire method does not require the solution of linear systems.

In the context of general convex optimisation, the overall method has mainly two limitations: Firstly, it is a feasible IPM and therefore requires a strictly feasible starting point. If this is not given a priori, then a *Phase I method* [8, Section 11.4] can be employed to compute such a point. However, infeasible IPMs might be advantageous in these situations. As the method discussed in this thesis avoids linear systems, it cannot simply be converted to an infeasible one as it is done in [5] or [37, Chapter 19].

Secondly, if linear equality constraints are given in addition, then the convergence results of this thesis only remain valid if the optimisation is carried out on a corresponding reduced subspace. An alternative which is well-suited for the structured l-BFGS method is the integration of linear equality constraints via an Augmented Lagrangian method, cf. Section 3.5. However, deriving theoretical results for this approach is likely challenging.

With the extensive convergence analysis and the polynomial iteration complexity result, this thesis establishes a theoretical basis for the integration of Tikhonov regularisation into Quasi-Newton based primal interior-point methods. The proposed method may be refined with further modifications on the parameter and update choices, the stopping criteria, and the memory management, and it allows for additional extensions, such as in the design of the inner solver and the integration of linear equality constraints. Therefore, this thesis can serve as starting point for further research on similar methods in the context of large-scale constrained convex optimisation.

# Bibliography

- [1] H. O. Aggrawal and J. Modersitzki. Hessian initialization strategies for l-BFGS solving son-linear inverse problems. In: *International Conference on Scale Space and Variational Methods in Computer Vision*. Springer, 2021, pp. 216–228.
- [2] M. Andersen, J. Dahl, Z. Liu, and L. Vandenberghe. Interior-point methods for large-scale cone programming. In: *Optimization for Machine Learning* 5583, 2011.
- [3] M. S. Andersen, J. Dahl, and L. Vandenberghe. CVXOPT: A Python package for convex optimization, version 1.3.2. Available at [cvxopt.org](http://cvxopt.org). In: 2023.
- [4] P. Armand, J. C. Gilbert, and S. Jan-Jégou. A feasible BFGS interior point algorithm for solving convex minimization problems. In: *SIAM Journal on Optimization* 11(1):199–222, 2000.
- [5] P. Armand and Ségalat. A quasi-Newton penalty barrier method for convex minimization problems. In: *Computational Optimization and Applications* 26:5–34, 2003.
- [6] P. Armand and P. Ségalat. A limited memory algorithm for inequality constrained minimization. In: *Technical Report 2003–08, University of Limoges (France)*, 2003.
- [7] L. Bergamaschi, J. Gondzio, Á. Martínez, J. W. Pearson, and S. Pougkakiotis. A new preconditioning approach for an interior point-proximal method of multipliers for linear and convex quadratic programming. In: *Numerical Linear Algebra with Applications* 28(4):e2361, 2021.
- [8] S. P. Boyd and L. Vandenberghe. *Convex optimization*. Seventh printing 2009. Cambridge University Press, 2004. ISBN: 978-0-521-83378-3.
- [9] C. G. Broyden. The convergence of a class of double-rank minimization algorithms 1. general considerations. In: *IMA Journal of Applied Mathematics* 6(1):76–90, 1970.
- [10] D. Calvetti, B. Lewis, L. Reichel, and F. Sgallari. Tikhonov regularization with non-negativity constraint. In: *Electronic Transactions on Numerical Analysis* 18:153–173, 2004.
- [11] S. M. S. de Carvalho and A. R. L. de Oliveira. Application of the Quasi-Newton in interior point methods for solving the predispatch problem. In: *Revista Sergipana de Matemática e Educação Matemática* 8(1):63–80, 2023.
- [12] E. Chouzenoux, M.-C. Corbineau, and J.-C. Pesquet. A proximal interior point algorithm with applications to image processing. In: *Journal of Mathematical Imaging and Vision* 62:919–940, 2020.
- [13] M.-C. Corbineau, E. Chouzenoux, and J.-C. Pesquet. PIPA: A new proximal interior point algorithm for large-scale convex optimization. In: *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2018)*. 2018, pp. 1343–1347.
- [14] V. De Simone, D. di Serafino, J. Gondzio, S. Pougkakiotis, and M. Viola. Sparse approximations with interior point methods. In: *Siam Review* 64(4):954–988, 2022.

## Bibliography

- [15] S. Diamond and S. Boyd. CVXPY: A Python-embedded modeling language for convex optimization. In: *The Journal of Machine Learning Research* 17(1):2909–2913, 2016.
- [16] A. V. Fiacco and G. P. McCormick. *Nonlinear programming: sequential unconstrained minimization techniques*. Reprinted by SIAM Publications, 1990. Wiley, 1968.
- [17] R. Fletcher. A new approach to variable metric algorithms. In: *The computer journal* 13(3):317–322, 1970.
- [18] A. Forsgren, P. E. Gill, and M. H. Wright. Interior methods for nonlinear optimization. In: *SIAM review* 44(4):525–597, 2002.
- [19] K. R. Frisch. The logarithmic potential method of convex programming. In: *Memo-  
randum, University Institute of Economics, Oslo* 5(6), 1955.
- [20] C. Geiger and C. Kanzow. *Numerische Verfahren zur Lösung unrestringierter Opti-  
mierungsaufgaben*. Springer-Verlag, 2013. ISBN: 978-3-642-58582-1.
- [21] C. Geiger and C. Kanzow. *Theorie und Numerik restringierter Optimierungsaufgaben*. Springer-Verlag, 2013. ISBN: 978-3-642-56004-0.
- [22] O. Gildemeister. *code-master-thesis*. GitHub repository. 2024. URL: <https://github.com/gildol22/code-master-thesis>.
- [23] D. Goldfarb. A family of variable-metric methods derived by variational means. In: *Mathematics of computation* 24(109):23–26, 1970.
- [24] J. Gondzio. Interior point methods 25 years later. In: *European Journal of Operational  
Research* 218(3):587–601, 2012.
- [25] J. Gondzio and F. Sobral. Quasi-Newton approaches to interior point methods for  
quadratic problems. In: *Computational Optimization and Applications* 74(1):93–120,  
2019.
- [26] J. Gondzio and F. N. Sobral. Polynomial worst-case iteration complexity of quasi-  
Newton primal-dual interior point algorithms for linear programming. In: *arXiv  
preprint arXiv:2208.08771*, 2022.
- [27] A. Kaplan and R. Tichatschke. Proximal methods in view of interior-point strategies.  
In: *Journal of Optimization Theory and Applications* 98(2):399–429, 1998.
- [28] N. Karmarkar. A new polynomial-time algorithm for linear programming. In: *Pro-  
ceedings of the Sixteenth Annual ACM Symposium on Theory of Computing*. 1984,  
pp. 302–311.
- [29] K. Königsberger. *Analysis 2*. 4th ed. Springer-Verlag, 2013.
- [30] F. Kruse. Interior point methods for optimal control problems with pointwise state  
constraints. PhD thesis. Technische Universität München, 2014.
- [31] R. Kuhlmann. A primal-dual augmented lagrangian penalty-interior-point algorithm  
for nonlinear programming. PhD thesis. Universität Bremen, 2018.
- [32] D. C. Liu and J. Nocedal. On the limited memory BFGS method for large scale  
optimization. In: *Mathematical Programming* 45(1-3):503–528, 1989.
- [33] F. Mannel. A globalization of L-BFGS for nonconvex unconstrained optimization. In:  
*arXiv preprint arXiv:2401.03805*, 2024.



## Bibliography

- [34] F. Mannel, H. O. Aggrawal, and J. Modersitzki. A structured L-BFGS method and its application to inverse problems. In: *Inverse Problems*, 2024. DOI: 10.1088/1361-6420/ad2c31. URL: <http://iopscience.iop.org/article/10.1088/1361-6420/ad2c31>.
- [35] Y. Nesterov and A. Nemirovskii. *Interior-point polynomial algorithms in convex programming*. SIAM, 1994.
- [36] J. Nocedal. Updating quasi-Newton matrices with limited storage. In: *Mathematics of Computation* 35(151):773–782, 1980.
- [37] J. Nocedal and S. Wright. *Numerical optimization*. 2nd ed. Springer Science & Business Media, 2006. ISBN: 978-0-387-30303-1.
- [38] C. C. Paige and M. A. Saunders. Solution of sparse indefinite systems of linear equations. In: *SIAM Journal on Numerical Analysis* 12(4):617–629, 1975.
- [39] C. G. Petra. A memory-distributed quasi-Newton solver for nonlinear programming problems with a small number of general constraints. In: *Journal of Parallel and Distributed Computing* 133:337–348, 2019.
- [40] C. G. Petra, M. Salazar De Troya, N. Petra, Y. Choi, G. M. Oxberry, and D. Tortorelli. A quasi-Newton interior-point method for optimization in Hilbert spaces. In: *Proposed Journal Article, unpublished* (LLNL-JRNL-764097), 2019.
- [41] S. Pougkakiotis and J. Gondzio. An interior point-proximal method of multipliers for convex quadratic programming. In: *Computational Optimization and Applications* 78:307–351, 2021.
- [42] S. Pougkakiotis and J. Gondzio. An interior point-proximal method of multipliers for linear positive semi-definite programming. In: *Journal of Optimization Theory and Applications*:192:97–129, 2022.
- [43] J. Renegar. *A mathematical view of interior-point methods in convex optimization*. SIAM, 2001.
- [44] M. A. Saunders and J. A. Tomlin. *Solving regularized linear programs using barrier methods and KKT systems*. Vol. 10064. IBM Thomas J. Watson Research Center, 1996.
- [45] D. F. Shanno. Conditioning of quasi-Newton methods for function minimization. In: *Mathematics of computation* 24(111):647–656, 1970.
- [46] M. Ulbrich and S. Ulbrich. *Nichtlineare Optimierung*. Springer-Verlag, 2012.
- [47] A. Wächter and L. T. Biegler. On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. In: *Mathematical programming* 106:25–57, 2006.
- [48] M. H. Wright. Interior methods for constrained optimization. In: *Acta numerica* 1:341–407, 1992.
- [49] M. H. Wright. Some properties of the Hessian of the logarithmic barrier function. In: *Mathematical Programming* 67(1-3):265–295, 1994.