Technische Universität München
Lehrstuhl für Optimalsteuerung

Adaptive regularization and discretization for nonlinear inverse problems with PDEs

Alana Rena Kirchner

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

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Univ.–Prof. Dr. Martin Brokate
Prüfer der Dissertation:	1.	Univ.–Prof. Dr. Boris Vexler
	2.	Univ.–Prof. Dr. Barbara Kaltenbacher (Alpen–Adria–Universität Klagenfurt, Österreich)
	3.	Univ.–Prof. Dr. Arnd Rösch (Universität Duisburg–Essen)

Die Dissertation wurde am 08.01.2014 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 20.02.2014 angenommen.

Abstract

In this thesis, efficient methods for the solution of inverse problems, combining adaptive regularization and discretization are proposed. For the computation of a Tikhonov regularization parameter, we consider an inexact Newton method based on Morozov's discrepancy principle for noisy data. In each step, a regularized problem is solved on a different discretization level, which we control using dual-weighted residual error estimators. In the second part of this thesis, we combine this method with iteratively regularized Gauss-Newton methods. For both approaches, we provide a convergence analysis as well as numerical results.

Zusammenfassung

In dieser Arbeit werden effiziente Verfahren zur Lösung inverser Probleme vorgestellt, die adaptive Regularisierungs- und Diskretisierungsmethoden kombinieren. Zur Bestimmung eines Tichonov-Regularisierungsparameters betrachten wir ein inexaktes Newton-Verfahren basierend auf Morozov's Diskrepanzprinzip für gestörte Daten. In jedem Schritt wird ein regularisiertes Problem auf einer anderen Diskretisierungsebene gelöst, welche wir mittels residuenbasierter Fehlerschätzer steuern. Im zweiten Teil der Arbeit kombinieren dies mit iterativ regularisierten Gauß-Newton-Methoden. Für beide Ansätze liefern wir Konvergenzresultate sowie numerische Ergebnisse.

Contents

1. Introduction

Many complex processes in the field of natural sciences, medicine, and engineering are described by mathematical models with nonlinear partial differential equations (PDEs). The mentioned systems of PDEs mostly contain unknown parameter, e.g., space-dependent coefficient functions, source terms, initial or boundary data, whose determination leads to high-dimensional *inverse problems*. Many inverse problems are *ill-posed* in the sense that their solution does not depend continuously on the data, which in case of noisy data can lead to significant misinterpretations of the solution. This instability requires the use of appropriate *regularization* techniques, one of them being *Tikhonov regularization* which this thesis focusses on. The intensity of the regularization, i.e., how much the problem is changed to obtain a stable approximation of the solution, is controlled via a *regularization parameter*. The question is, how the regularization parameter should be chosen. A well-established method for the determination of a regularization parameter is *Morozov's discrepancy principle* where the residual norm, considered as a function of the regularization parameter, should equal an appropriate multiple of the noise level. Alternatively to applying Tikhonov regularization to the nonlinear problem of parameter identification, we consider *Newton type methods* [69]. For approaching the solution to the inverse problem in an iterative manner, there, in each iteration step, a linearization of the operator equation $F(q) = g$ around some approximate solution is solved. In particular, we investigate the *iteratively regularized Gauss-Newton method*, where in each Newton step Tikhonov regularization is applied to the linearized problem. Also here, the discrepancy principle plays a crucial role for selecting the regularization parameter in each iteration and to appropriately terminate the overall iteration.

The numerical effort for solving inverse problems with PDEs is usually much higher than for the numerical simulation of the underlying process with a given data set. Large potential for the construction of efficient algorithms for the solution of such inverse problems lies in *adaptive discretizations*. While the use of adaptive concepts for the choice of the discretization for numerical simulation has become state of the art in the last years, adaptivity in the context of inverse problems presents a new and highly relevant topic.

The central subject of this thesis consists in finding as generally applicable and analytically established methods as possible for the adaptive discretization of inverse problems, combining discretization and regularization. In this process, the main focus is on the efficiency of the constructed algorithms on the one hand and on the rigorous convergence analysis on the other hand.

This thesis is organized as follows:

Chapter 2: Basic concepts for the analysis and solution of inverse problems

In this chapter we give an overview of the existing methods for the numerical solution of ill-posed inverse problems, starting with the definition of a *nonlinear inverse problem* in the sense of Hadamard [41]. For a full introduction to the general theory of inverse problems we refer to [6, 21, 27, 107]. A classical reference on uniqueness results for parameter identification in PDEs is [53]. We continue by introducing *Tikhonov regularization*, which has been discussed, i.a., in [57, 59, 62, 80, 105, 107], and *Morozov's discrepancy principle* as *a posteriori parameter choice rule*, cf. [42, 89, 90]. Next, we build a bridge between inverse problems and optimal control, i.e., we formulate the inverse problem as an optimization problem constrained by the underlying PDE. We rely on the basic theory for PDEs [31, 40], and their numerical solution via finite element methods [20, 25, 76]. The monographs [19, 34, 46, 78, 108] represent a profound introduction to optimal control. We formulate the discretized problem and explain the use of *dual-weighted-residual error estimators* (DWR) for adaptive discretizations, cf. [9–11]. In the last section of Chapter 2 we briefly present the optimization methods considered in this thesis, namely *Newton's method*, *iteratively regularized Gauss-Newton methods*, *SQP methods* and *Penalty methods*, which are presented in more detail, e.g., in [19, 36, 95].

Chapter 3: Computation of a Tikhonov regularization parameter with adaptive discretizations

In this chapter we derive an algorithm for the efficient computation of a Tikhonov regularization parameter with adaptive mesh refinement.

To our knowledge, parameter choice rules and convergence analysis for uniform refinement go back to Natterer [92] and Engl, Neubauer [28]. Adaptive discretizations for parameter estimation have been considered by Chavent, Bisselll in [24] and Ben Ameur, Chavent, and Jaffré in [1], where the mesh is refined according to refinement and coarsening indicators. Later, Neubauer suggested an adaptive method for inverse problems which refines at discontinuities by means of appropriate weight factors, cf. [94]. In the context of adaptivity for inverse problems, we would also like to mention the publications [2] by Benameur, Kaltenbacher and [13] by Beilina, Klibanov as well as the recent paper by Wachsmuth [110], where an adaptive regularization and discretization method is presented for problems with control constraints.

We focus on Morozov's discrepancy principle as a posteriori parameter choice and suggest to solve the resulting scalar nonlinear equation by an inexact Newton method, where in each iteration step, a regularized problem is solved at a different discretization level. Dual-weighted-residual error estimators are used for adaptive refinement with respect to two quantities of interest, in order to keep the discretization level as coarse as possible but fine enough to guarantee global convergence of the inexact Newton method. The proposed algorithm is an extension of the method suggested in [39] for linear inverse problems. This concept leads to a highly efficient method for determining the Tikhonov regularization parameter for nonlinear ill-posed problems. Moreover it is shown that with the so obtained regularization parameter and an also adaptively discretized Tikhonov minimizer, usual convergence and regularization results from the continuous setting can be recovered. For the convergence analysis, we extend the ideas of Potra [98], using the existence of a lower and an upper bound to the second derivative of some quantity of interest. The efficiency of the proposed method is demonstrated by means of numerical experiments.

In the second part of Chapter 3 we extend the proposed method to parabolic problems. A detailed introduction to the theory of parabolic PDEs and their discretization can be found for instance in the textbook of Eriksson, Estep, Hansbo, and Johnson [29]. For the time discretization, we adopt the results obtained by Meidner and Vexler from [85–87], and use Galerkin finite element methods as well as an equilibrium criterion for an adaptive refinement strategy. Finally, we provide numerical results for a parabolic example to show the efficiency of the proposed method.

Chapter 4: Iteratively regularized Gauss-Newton methods

In this chapter we combine the method proposed in Chapter 3 with iteratively regularized Gauss-Newton methods.

A detailed overview of iterative regularization methods for ill-posed problems is given in the textbook [69] by Kaltenbacher, Neubauer, and Scherzer. We would also like to mention the publications [42, 51, 52, 60, 67, 91], which deal with iterative regularization methods. To our knowledge, the iteratively regularized Gauss-Newton method (IRGNM) has been introduced by Bakushinskii [4], together with a local convergence result in the continuous setting. Logarithmic rates have been shown by Hohage [50], and the results from [4] were extended to more general regularization methods by Kaltenbacher in [60]. Besides, IRGNM and its convergence in different settings has been considered by Kaltenbacher, Neubauer, and Ramm in [68], as well as Jin in [54, 55]. Generalized or related versions of IRGNM can be found in [18, 22, 23, 81]. To our knowledge, there exists no convergence analysis of IRGNM for adaptive regularization and discretization in the current literature.

We study adaptive discretizations of the iteratively regularized Gauss-Newton method IRGNM with a discrepancy principle choice of the regularization parameter in each Newton step. The stopping index plays the role of an additional regularization parameter, which is also chosen according to the discrepancy principle.

We first of all prove convergence and convergence rates under some accuracy requirements formulated in terms of four quantities of interest. Then computation of error estimators for these quantities based on a dual-weighted-residual method is discussed, which results in an algorithm for adaptive refinement.

Finally we extend the results from the Hilbert space setting with quadratic penalty to Banach spaces and general Tikhonov functionals for the regularization of each Newton step, where we base our analysis on the results obtained by Hofmann, Kaltenbacher, Pöschl, Scherzer [49], Kaltenbacher, Schöpfer, Schuster [71], Hofmann, Kaltenbacher [48], and Hohage, Werner [52].

In the second part of Chapter 4 we investigate all-at-once formulations of the iteratively regularized Gauss-Newton method. All-at-once formulations considering the PDE and the measurement equation simultaneously allow to avoid (approximate) solution of a potentially nonlinear PDE in each Newton step as compared to the reduced form. We analyze a least squares and a Generalized Gauss-Newton formulation and in both cases prove convergence and convergence rates with a posteriori choice of the regularization parameter in each Newton step and of the stopping index under certain accuracy requirements on four quantities of interest. Estimation of the error in these quantities by means of a dual-weighted-residual method is discussed, which leads to an algorithm for adaptive mesh refinement. Numerical experiments with an implementation of this algorithm show the numerical efficiency of this approach, which, in some aspects, outperforms the nonlinear Tikhonov regularization considered in Chapter 3.

Chapter 5: Conclusion and perspectives

In this chapter we summarize the results from the previous chapters and discuss possible extensions and future work.

Appendix

In the appendix the reader can find some additional material that is not included in the previous chapters in order to not distract the reader unnecessarily from the main presentation of this thesis.

2. Basic concepts for the analysis and solution of inverse problems

An inverse problem in general is a problem where the effect is known, but the source is not, in contrast to a *direct problem*, where we deduce the *effect* from the *source*. In abstract form an inverse problem can be written as

$$F(q) = g\,,$$

where q is the source, g is the data (effect) and F is some operator describing the underlying (physical, chemical, etc.) process. Depending on the linearity or nonlinearity of F, we speak of a *linear inverse problem* or *nonlinear inverse problem*. A typical example for a linear inverse problem is computer tomography. X-rays are sent through the patient's body and a detector measures the remaining outcoming rays on the other side. From these measurements (attenuation of X-rays) it is possible to infer the density distribution (source of the attenuation) and hence the different types of tissue. Another medical application and an example for a nonlinear inverse problem would be electrical impedance tomography, where electrodes on the patient's body measure the electric potentials on the skin while one of the electrodes plays the role of a current source. From these data one can conclude the conductivity (and implicitly the type) of the particular tissue inside the body. There are many more applications that could be named, e.g., inverse scattering, inverse seismic problems, inverse gravimetry and many more. Details on the mathematical modelling can be found, e.g., in [6, 21, 27, 53, 107]. Although we will keep the modelling PDE rather general, our numerical experiments will concentrate on parameter and coefficient identification, where we want to determine the unknown parameter standing for, e.g., unknown material properties.

2.1. Ill-Posedness, Tikhonov regularization, and Morozov's discrepancy principle

2.1.1. Inverse problem

In 1902, J. Hadamard established the notion of *well-posedness* [41] through the following characterization: A problem is *well-posed* if and only if

1. for all admissible data there exists a solution, (*existence*)
2. the solution is unique, (*uniqueness*) and
3. the solution depends continuously on the data. (*stability*).

Throughout this thesis, we consider the nonlinear *ill-posed* operator equation

$$F(q) = g\,, \tag{2.1}$$

where $F\colon \mathcal{D} \subseteq Q \to G$ is a nonlinear operator between Hilbert spaces Q and G, and $\mathcal{D}$ is some subset of Q, for instance the domain of F. We are interested in the situation that the solution of (2.1) does not depend continuously on the data g, i.e., where the ill-posedness is caused by violation of 3. Due to this lack of continuity (and possibly even non-existence) of F^{-1} equation (2.1) is not stably solvable. In addition, we assume that we are only given noisy data g^δ with (maximum) noise level δ, i.e.,

$$\|g^\delta - g\|_G \le \delta\,. \tag{2.2}$$

So we aim to find $q \in \mathcal{D}$, such that

$$F(q) \approx g^\delta\,. \tag{2.3}$$

(There may be no q fulfilling (2.3) with equality.)

2.1.2. Tikhonov regularization

In order to get a stable approximation of the exact solution $q^\dagger$ of (2.1), we need to apply regularization methods. A well-known method for doing so is *Tikhonov regularization* [107], where an approximate solution is obtained by solving the minimization problem

$$\min_{q\in\mathcal{D}} \|F(q) - g^\delta\|_G^2 + \frac{1}{\beta}\|q - q_0\|_Q^2 \tag{2.4}$$

for some given intial guess $q_0 \in \mathcal{D}$. We call $\beta > 0$ *regularization parameter*, which is crucial for the balance between minimizing the missfit term $\|F(q) - g^\delta\|_G^2$ and keeping the penalization term $\|q - q_0\|_Q^2$ small.

Remark 2.1. *We use the unusual notation $\frac{1}{\beta}$ with $\beta > 0$ called "regularization parameter" instead of the more common use of a regularization parameter $\alpha = \frac{1}{\beta}$, which we adopt from [39]. The reason behind this is that the resulting problem becomes "a little less nonlinear" with respect to β. In case of a linear inverse problem $Tq = g$ with $T\colon Q \to G$ linear, this can be seen very easily: A stationary point q of* (2.4) *(with $F = T$) satisfies*

$$q = q(\beta) = \left(T^*T + \frac{1}{\beta}\,\mathrm{id}\right)^{-1} T^* g^\delta\,,$$

where id *denotes the identity operator. For small singular values of T there holds*

$$\left(T^*T + \frac{1}{\beta}\,\mathrm{id}\right)^{-1} \approx \beta\,\mathrm{id}\,,$$

such that the map $\beta \mapsto q(\beta)$ is not "too nonlinear".

Existence of a global minimizer of the Tikhonov functional (and thereby of a stationary point, provided that the minimizer is an interior point of the domain $\mathcal{D}$) can be guaranteed if F is weakly sequentially closed, i.e.,

$$(q_n \rightharpoonup q \;\wedge\; F(q_n) \rightharpoonup g) \;\Rightarrow\; (q \in \mathcal{D} \;\wedge\; F(q) = g) \tag{2.5}$$

for all $(q_n)_{n\in\mathbb{N}} \subseteq \mathcal{D}$ cf., e.g., [105, Theorem 1].

2.1.3. Regularization parameter choice

We are interested in the (subsequential) convergence of a solution q_β^δ of the Tikhonov minimization problem (2.4) to an exact solution $q^\dagger$ of the inverse problem equation (2.1) for vanishing noise, i.e., $q_\beta^\delta \to q^\dagger$ as $\delta \to 0$. It is obvious that this cannot hold for arbitrary regularization parameter $\beta > 0$, but that it is reasonable to choose β depending on the noise δ. Rules to do so are called *parameter choice rules*. They are divided into two classes: *a priori* and *a posteriori* choice rules. A priori choice rules pick a regularization parameter β only in dependence of δ, i.e., $\beta = \beta(\delta)$ and are fixed a priori, whereas a posteriori choice rules choose β also in dependence of the noisy data g^δ, i.e., $\beta = \beta(\delta, g^\delta)$.

A priori parameter choice

It is evident that, the less noise, the weaker the regularization should be. In line with this, a reasonable and intuitive *a priori* choice for β would be such that $\beta = \beta(\delta) \to \infty$ as $\delta \to 0$. If this convergence holds and is not too fast, i.e., $\delta^2\beta(\delta) \to 0$ as $\delta \to 0$, it can in fact be shown that the optimization problem (2.4) is stable in the sense that solutions of (2.4) depend continuously on the data g^δ, cf. [27]. It can also be shown that a solution $q_{\beta(\delta)}^\delta$ of (2.4) converges subsequentially to a solution $q^\dagger$ of (2.1), in the sense that it has a convergent subsequence and each convergent subsequence converges (strongly) to a solution of (2.1) as $\delta \to 0$, cf. [27, 105].

Unfortunately, the convergence $q_\beta^\delta \to q^\dagger$ can be arbitrarily slow. To obtain convergence rates, most parameter choice rules require a *source condition* of the form

$$q^\dagger - q_0 \in \mathcal{R}((F'(q^\dagger)^* F'(q^\dagger))^\nu)\,, \tag{2.6}$$

for some $\nu \in [0, 1]$, where $\mathcal{R}$ denotes the range and we have assumed that F is continuously Fréchet-differentiable with derivative F'. These kinds of Hölder type source conditions in case $\nu < \frac{1}{2}$ in (2.6) or logarithmic source conditions (cf. [62]) represent a strong restriction on the exact solution, but the alternative tools for showing convergence rates need strong structural assumptions on the operator F (cf. [67]). In some (indeed practically relevant) cases those operator assumptions may be violated while some kind of source condition can still be fulfilled for moderately ill-posed problems (cf. [61, 63, 73]).

In this context the main disadvantage of a priori parameter choice rules arises: to obtain optimal rates, they usually need knowledge about ν (or $q^\dagger$), which makes them not applicable in most practically motivated settings.

For an appropriate a priori choice using this knowledge (cf., e.g., [27]) the optimal convergence rate

$$\|q_\beta^\delta - q^\dagger\|_Q = O\left(\delta^{\frac{2\nu}{2\nu+1}}\right) \tag{2.7}$$

for a solution q_β^δ to (2.4) can be obtained under the source condition (2.6). Appropriate a priori parameter choice rules for iterative regularization methods (instead of Tikhonov regularization) yielding optimal rates can be found in [69].

A posteriori parameter choice, Morozov's discrepancy principle

However, due to the frequent lack of a priori information, we will concentrate on an a posteriori choice rule for β, namely the famous *Morozov's discrepancy principle* [89, 90].

An overview and analysis of the most popular a posteriori choice rules, as for instance, the discrepancy principle, the generalized cross-validation, the quasi optimality criterion and the L-curve criterion, is given in [45, 51, 57, 58], where the latter two do not require the knowledge of the noise level. Other a posteriori noise-free choice rules are the balancing principle analyzed, e.g., in [77], [96] or the choice rule presented in [103].

Morozov's discrepancy principle is motivated by the idea of not solving the given ill-posed equation more accurately than the given data. We therefore seek a regularization parameter $\beta = \beta(\delta)$ depending on δ, such that there exists a solution q_β^δ of (2.4) fulfilling

$$\|F(q_\beta^\delta) - g^\delta\|_G = \tau\delta \tag{2.8}$$

for some $\tau \geq 1$. If such a parameter β exists, the optimal rates (2.7) are shown for $\nu < \frac{1}{2}$ and $\tau = 1$ in [27].

Since the residual norm will play the role of a quantity of interest later on when it comes to error estimation and mesh refinement (for a discretized version of (2.4), cf. Section 2.3), we introduce the following notation:

$$\iota\colon Q \to \mathbb{R}\,, \quad \iota(q) := \|F(q) - g^\delta\|_G^2 \tag{2.9}$$

$$i\colon \mathbb{R}^+ \to \mathbb{R}\,, \quad i(\beta) := \iota(q_\beta^\delta) = \|F(q_\beta^\delta) - g^\delta\|_G^2\,, \tag{2.10}$$

where q_β^δ is a (global) minimizer (or a stationary point) of (2.4).

Remark 2.2. *In fact, this definition is a bit sloppy in case the global minimum of* (2.4) *is not unique with respect to β, since then, i is not a proper function. In that case* (2.10) *is taken to mean picking one of the minimizers/stationary points.*

The discrepancy principle then reads: Find $\beta_* > 0$ fulfilling

$$i(\beta_*) = \tau^2\delta^2\,. \tag{2.11}$$

We will see that we don't have to solve (2.11) exactly. Namely, it will be shown that for obtaining convergence and optimal convergence rates it suffices to have a regularization parameter β that satisfies

$$\underline{\tau}^2\delta^2 \leq i(\beta) \leq \overline{\tau}^2\delta^2 \tag{2.12}$$

for some constants $0 < \underline{\tau} \leq \tau \leq \overline{\tau}$.

2.2. Inverse problem as optimal control problem

Our study is motivated by inverse problems for partial differential equations such as parameter identification problems, where F is the composition of a parameter-to-solution map

$$\begin{aligned} S\colon Q &\to V \\ q &\mapsto u \end{aligned} \tag{2.13}$$

with some measurement operator

$$\begin{aligned} C\colon V &\to G \\ u &\mapsto g\,, \end{aligned} \tag{2.14}$$

where V is an appropriate Hilbert space. (A typical choice for Q and G would be, e.g., the Lebesgue space $L^2(\Omega)$ and for V the Sobolev space $H^1(\Omega)$ or $H^1_0(\Omega)$, see Example 2.1)

The underlying (nonlinear) stationary PDE in a weak form is written as

$$A(q,u) = f \quad \text{in } W^* \tag{2.15}$$

for some second order differential operator $A\colon Q \times V \to W^*$, and $f \in W^*$ is some given right-hand side in the dual of a Hilbert space W. (Often the test space W is chosen as the state space, i.e., a typical choice would be $W = H^1(\Omega)$ or $W = H^1_0(\Omega)$, see Example 2.1).

Let $(\cdot,\cdot)_Q$ denote the inner product in the control (parameter) space Q, let $(\cdot,\cdot)_G$ be the inner product in the observation space G and let $\langle\cdot,\cdot\rangle_{V^*,V}$ be the duality pairing between the Hilbert spaces V and its dual V^* (analogously for W). We introduce the semi-linear form $a\colon Q \times V \times W \to \mathbb{R}$ which is linear in its third argument and defined by

$$a(q,u)(\varphi) \coloneqq \langle A(q,u), \varphi\rangle_{W^*,W} \tag{2.16}$$

and the notation $f(\varphi) \coloneqq \langle f, \varphi\rangle_{W^*,W}$. Correspondingly, if A is Fréchet-differentiable we write

$$a'_q(q,u)(\delta q,\varphi) \coloneqq \langle A'_q(q,u)(\delta q), \varphi\rangle_{W^*,W} \quad \text{and} \quad a'_u(q,u)(\delta u,\varphi) \coloneqq \langle A'_u(q,u)(\delta u), \varphi\rangle_{W^*,W}$$

for $\delta q \in Q$, $\delta u \in V$ (for higher derivatives analogously).

With this, we reformulate the state equation (2.15): Find for given control $q \in Q$ the state variable $u \in V$ such that

$$a(q,u)(\varphi) = f(\varphi) \quad \forall \varphi \in W\,. \tag{2.17}$$

There are several sets of assumptions on the nonlinearity in $a(\cdot,\cdot)(\cdot)$ and its dependence on the control variable q allowing the state equation (2.17) to be well-posed. Due to the fact that the analyis of this paper does not depend on the particular structure of the nonlinearity in a, we do not specify a set of assumptions on it, but assume throughout this thesis:

Assumption 2.1. *All arising state, dual, and tangent equations are uniquely stably solvable.*

In this setting Tikhonov regularization can be written in the form:

$$\begin{aligned} &\min_{(q,u)\in Q\times V} \quad J_\beta(q,u) \\ \text{s.t.} \quad &a(q,u)(\varphi) = f(\varphi) \quad \forall \varphi \in W\,, \\ &q \in \mathcal{D} \subset Q \end{aligned} \tag{2.18}$$

where the functional J_β is defined as

$$J_\beta\colon Q \times V \to \mathbb{R}\,, \quad J_\beta(q,u) := I(u) + \frac{1}{\beta}\|q - q_0\|_Q^2$$

with

$$I(u) := \|C(u) - g^\delta\|_G^2 \tag{2.19}$$

and some initial guess q_0.

A sample setting for (2.18) could be the following example.

Example 2.1. *Let $\Omega \subset \mathbb{R}^3$ be a Lipschitz domain. We consider*

$$\min_{q\in L^2(\Omega),v\in H_0^1(\Omega)} \|u - g^\delta\|_{L^2(\Omega)}^2 + \frac{1}{\beta}\,\|q - q_0\|_{L^2(\Omega)}^2$$

$$s.t. \quad \begin{cases} \Delta u + u^3 = q & \text{in } \Omega \\ \qquad\quad u = 0 & \text{on } \partial\Omega \end{cases}$$

(in strong form), which corresponds to (2.18) *with the choice $Q = L^2(\Omega)$, $V = W = H_0^1(\Omega)$, $a(q,u)(\varphi) = \int_\Omega \nabla u(x)\nabla\varphi(x) + u^3\varphi - q\varphi\, dx$, $G = L^2(\Omega)$, $C = \mathrm{id}\colon V \to G$.*

Existence of a solution to this problem as well as existence and differentiability of the solution operator S of the PDE is shown, e.g., in [108, Theorem 4.15, 4.16, 4.17].

Given the existence of the control-to-state mapping $S\colon Q \to V,\ q \mapsto u = S(q)$, we can define the reduced cost functional $j_\beta\colon Q \to \mathbb{R}$ by

$$j_\beta(q) := J_\beta(q, S(q)) = \|F(q) - g^\delta\|_G^2 + \frac{1}{\beta}\|q - q_0\|_Q^2\,, \tag{2.20}$$

which is consistent with (2.4) for $F = C \circ S$.

This definition allows to reformulate problem (2.18) as

$$\min_{q\in\mathcal{D}\subset Q} j_\beta(q)\,. \tag{2.21}$$

Remark 2.3. *In the theoretical analysis we consider the constraint $q \in \mathcal{D}$ to show the generality of the developed results, since in many examples (e.g., coefficient identification problems) the domain of F is a proper subset of Q or certain applications offer a priori knowledge (e.g., nonnegativity, monotonicity, convexity) about the searched-for parameter. Furthermore, some assumptions (cf. Assumption 3.3 and 3.6) are required to hold in $\mathcal{D}$, which in general would make them too strong assumptions on F for the choice $\mathcal{D} = Q$.*

In Chapter 4, we can even restrict these assumptions to a neighborhood of q_0*, and show that every minimizer/stationary point of an unconstrained version of* (2.21) *lies in this neighborhood (cf. Theorem 3.2, 4.2, 4.4, 4.8 and 4.11 as well as Assumption 3.1), such that (depending on the used optimization algorithm) the constraint* $q \in \mathcal{D}$ *plays no role.*

For these reasons we will neglect this constraint in Chapter 4, and for simplicity also when it comes to the actual solution process in the numerics of (2.20)*, i.e., in Section 3.6, and consider the unconstrained problem*

$$\min_{q \in Q} j_\beta(q) \tag{2.22}$$

instead.

Throughout this thesis we assume sufficient smoothness, i.e.,

Assumption 2.2. *J_β, A, S, j_β, C and F are three times continuously Fréchet-differentiable. (For some results, two times Fréchet-differentiability will suffice.)*

To find a stationary point of j_β in $\mathcal{D}$ we seek $q \in \mathcal{D}$ solving

$$j'_\beta(q)(\delta q) = 0 \quad \forall \delta q \in Q\,, \tag{2.23}$$

where $j'_\beta(q)(\delta q)$ denotes the directional derivative of j_β with respect to its component q in the direction δq.

Remark 2.4. *Throughout this thesis we will switch between* q_β^δ *being either a stationary point, i.e., a solution to* (2.23)*, or a global minimizer.*

In Chapter 3 we will consider stationary points most of the time and make use of the stationarity condition (2.23)*. At some points, we will provide an alternative proof for global minimizers and use minimality instead of stationarity. In order to transfer the results for stationary points to minimizers directly, we could claim that* $\mathcal{D}$ *has a nonempty interior and that the concerned minimizer lies in the interior of* $\mathcal{D}$*. This condition however is quite strong and can not be guaranteed in some cases, see Section 3.4, which is why we go without this condition in Chapter 3.*

In Chapter 4 we require indeed that $\mathcal{D}$ *contains an open ball, i.e., has nonempty interior, but in return, we get rid of the constraint* $\mathcal{D}$ *in a way, since we will show that, starting with an approximation (to the exact solution* $q^\dagger$*) which lies in a ball around* q_0 *that is contained in* $\mathcal{D}$*, the subsequent iterates/minimizers of the unconstrained problem don't leave this ball. Hence, the minimizers lie in the interior of* $\mathcal{D}$*. Together with the fact that we deal with convex optimization problems in Chapter 4, this implies that we don't have to distinguish between minimizers and stationary points.*

Summarizing, in Chapter 3 we don't need the assumption that $\mathcal{D}$ *has a nonempty interior, but we have to assume that there exists a stationary point in* $\mathcal{D}$*. In Chapter 4, on the contrary, we assume that* $\mathcal{D}$ *has a nonempty interior, but in return we automatically get that the iterates/minimizers lie in* $\mathcal{D}$*.*

To construct suitable expressions for the first and second derivatives of the reduced cost functional j and in view of Remark 2.4 as well as the upcoming discussion of error estimation, we introduce the Lagrangian $\mathcal{L}\colon Q \times V \times W \to \mathbb{R}$, defined as

$$\mathcal{L}(q,u,z) \coloneqq J_\beta(q,u) + f(z) - a(q,u)(z). \tag{2.24}$$

With its aid, we obtain the following standard representation of the first derivative $j'_\beta(q)(\delta q)$ [9]: If for given $q \in \mathcal{D}$ the state $u \in V$ fulfills the *state equation*

$$\mathcal{L}'_z(q,u,z)(\varphi) = 0, \quad \forall \varphi \in W \tag{2.25}$$

(i.e., (2.17)), and if additionally $z \in W$ is chosen as solution of the *adjoint state/dual equation*

$$\mathcal{L}'_u(q,u,z)(\varphi) = 0, \quad \forall \varphi \in V, \tag{2.26}$$

then the following expression of the first derivative of the reduced cost functional holds for given $\delta q \in Q$:

$$j'_\beta(q)(\delta q) = \mathcal{L}'_q(q,u,z)(\delta q). \tag{2.27}$$

Finding $x = (q,u,z) \in \mathcal{D} \times V \times W$ such that

$$\mathcal{L}'(x)(\varphi) = 0 \quad \forall \varphi \in Q \times V \times W \tag{2.28}$$

is equivalent to finding a stationary point to the optimization problem (2.21).

In the following we denote such a stationary point for some fixed β by $q_\beta^\delta \in \mathcal{D}$.

In the same manner one can obtain representations of the second derivatives of j in terms of the Lagrangian. It is possible to derive two different kinds of expressions: one is preferable for building up the whole Hessian, the other one for computing matrix-vector products of the Hessian times a given vector. Details on this can be found, e.g., in [9]. If the state equation (2.25) and the adjoint equation (2.26) are fulfilled, for given directions $\delta q, \tau q \in Q$ there holds:

(a) If $\delta u \in V$ satisfies the *linearized state/tangent equation*

$$\mathcal{L}''_{qz}(q,u,z)(\delta q,\varphi) + \mathcal{L}''_{uz}(q,u,z)(\delta u,\varphi) = 0 \quad \forall \varphi \in W, \tag{2.29}$$

and if $\delta z \in W$ is chosen as the solution of the *additional adjoint/dual-for-hessian equation*

$$\mathcal{L}''_{qu}(q,u,z)(\delta q,\varphi) + \mathcal{L}''_{uu}(q,u,z)(\delta u,\varphi) + \mathcal{L}''_{zu}(q,u,z)(\delta z,\varphi) = 0 \quad \forall \varphi \in V, \tag{2.30}$$

then the second derivative of the reduced cost functional is given as:

$$j''_\beta(q)(\delta q,\tau q) = \mathcal{L}''_{qq}(q,u,z)(\delta q,\tau q) + \mathcal{L}''_{uq}(q,u,z)(\delta u,\tau q) + \mathcal{L}''_{zq}(q,u,z)(\delta z,\tau q).$$

(b) If $\delta u, \tau u \in V$ satisfy the *linearized state/tangent equations*

$$\begin{aligned} \mathcal{L}''_{qz}(q,u,z)(\delta q,\varphi) + \mathcal{L}''_{uz}(q,u,z)(\delta u,\varphi) = 0 \quad \forall \varphi \in W, \\ \mathcal{L}''_{qz}(q,u,z)(\tau q,\varphi) + \mathcal{L}''_{uz}(q,u,z)(\tau u,\varphi) = 0 \quad \forall \varphi \in W, \end{aligned} \tag{2.31}$$

then the second derivative of the reduced cost functional is given as:

$$\begin{aligned} j''_\beta(q)(\delta q,\tau q) = \mathcal{L}''_{qq}(q,u,z)(\delta q,\tau q) + \mathcal{L}''_{qu}(q,u,z)(\delta q,\tau u) \\ + \mathcal{L}''_{uq}(q,u,z)(\delta u,\tau q) + \mathcal{L}''_{uu}(q,u,z)(\delta u,\tau u). \end{aligned}$$

Whether (a) or (b) is preferable depends on the dimension of the (discretized) control space. If it is low dimensional it is cheaper to build up the whole Hessian, thus choosing alternative (b), whereas for a large control space alternative (a) is less expensive.

For a profound introduction to the concepts of optimal control we refer to the monographs [19, 34, 46, 78, 108].

2.3. Discretization

In the main parts of this thesis, i.e., in Chapter 3 and 4 (except for the sections containing numerical experiments) we work with unspecified finite-dimensional spaces V_h, W_h, and Q_h. However, we assume the discretization to be such that the approaches *discretize-then-optimize* and *optimize-then-discretize* are equivalent (cf., e.g., [46, chapter 3]).

In this section, we briefly discuss a possible discretization of the optimization problem (2.18) (cf. [72]), which we use for our numerical experiments. We apply the Galerkin finite element method to discretize the state equation. This allows us to give a natural computable representation of the discrete gradient and Hessian similar to the continuous problem. The discretization of the control space Q is kept rather abstract by choosing a finite dimensional subspace $Q_h \subset Q$. A possible concrete choice is given below.

To describe the finite element discretization of the state space, we consider two- or three-dimensional shape-regular meshes, see, e.g., [25]. A mesh consists of quadrilateral or hexahedral cells K, which constitute a non-overlapping cover of the computational domain $\Omega \subset \mathbb{R}^n$, $n \in \{2, 3\}$. We allow hanging nodes (see, e.g., [84]) with the following restriction: We allow at most one hanging node on each face and this node has to lie on the midpoint of the face.

The corresponding mesh is denoted by $\mathcal{T}_h = \{K\}$, where we define the discretization parameter h as a cellwise constant function by setting $h\big|_K = h_K$ with the diameter h_K of the cell K.

On the mesh $\mathcal{T}_h$ we construct a conforming finite element space $V_h \subset V$ in a standard way:

$$V_h = \left\{ v \in V \cap C(\bar{\Omega}) \;\middle|\; v\big|_K \in \mathcal{Q}_s(K) \text{ for } K \in \mathcal{T}_h \right\},$$

where $\mathcal{Q}_s$ denotes a suitable space of polynomial-like functions of degree s on the cell K [84]. The Hilbert space W is discretized the same way, such that we get a discretized version W_h of W. (For the typical case $W = V$ one usually chooses $W_h = V_h$.)

Then, the discretization of the state equation (2.17) can be stated as: Find for given control $q \in Q$ a state $u_h = S_h(q) \in V_h$ with $S_h \colon Q \to V_h$ such that

$$a(q, u_h)(\varphi) = f(\varphi) \quad \forall \varphi \in W_h. \tag{2.32}$$

This discrete state equation is assumed to possess a unique solution for each $q \in Q$.

To discretize the control, we distinguish two cases: the control space is finite-dimensional, i.e., $Q = \mathbb{R}^d$ for some $d \in \mathbb{N}$, or the control space is infinite-dimensional, e.g., $Q = L^2(\Omega)$. In the first case, the control space does not have to be discretized. Hence, we may choose here $Q_h = Q$. In the latter case, we consider a nodewise discretization (i.e., a representation by nodal basis functions; cf., e.g., [84, 85, 99]) of the control variable on the same mesh as the state variable.

The optimization problem with discrete state and control variables is then given as follows:

$$\begin{aligned} &\min_{(q_h,u_h)\in(\mathcal{D}\cap Q_h)\times V_h} J_\beta(q_h,u_h) \\ \text{s.t.} \quad & a(q_h,u_h)(\varphi) = f(\varphi) \qquad \forall \varphi \in W_h\,. \end{aligned} \tag{2.33}$$

Similar to the continuous case, we introduce the discrete reduced cost functional $j_{\beta,h}\colon Q \to \mathbb{R}$ by

$$j_{\beta,h}(q) := J_\beta(q, S_h(q)) = \|C(S_h(q)) - g^\delta\|_G^2 + \frac{1}{\beta}\|q-q_0\|_Q^2 = \|F_h(q) - g^\delta\|_G^2 + \frac{1}{\beta}\|q-q_0\|_Q^2$$

with $F_h = C \circ S_h$. We assume that the observation operator C is already finite-dimensional or does not need to be discretized for other reasons, e.g., if $C = \mathrm{id}$. Then the discrete optimal control problem (2.33) can be reformulated as

$$\min_{q\in\mathcal{D}\cap Q_h} j_{\beta,h}(q)\,. \tag{2.34}$$

Similar to the previous section a stationary point $q^\delta_{\beta,h} \in \mathcal{D}\cap Q_h$ of $j_{\beta,h}$ in $\mathcal{D}$ is defined by

$$j'_{\beta,h}(q^\delta_{\beta,h})(\delta q) = 0 \quad \forall \delta q \in Q_h\,, \tag{2.35}$$

In the same manner we get discretized versions of $\iota(q)$ (cf. (2.9)) and $i(\beta)$ (cf. (2.10)):

$$\iota_h(q) := \|F_h(q) - g^\delta\|_G^2 \tag{2.36}$$

$$i_h(\beta) := \iota_h(q^\delta_{\beta,h}) = \|F(q^\delta_{\beta,h}) - g^\delta\|_G^2\,. \tag{2.37}$$

Here we have the same situation as in the continuous case concerning uniqueness of $q_{\beta,h}$, see Remark 2.2.

Since the state discretization as well as the control discretization are conforming Galerkin methods, the representation formulas for the first and second derivatives of j from Section 2.2 can directly be transferred to the discrete level for representing the derivatives of $j_{\beta,h}$ [9], i.e., in our setting *discretize-then-optimize* and *optimize-then-discretize* lead to the same result. A discretized form of (2.28) is defined by inserting the finite dimensional spaces $Q_h \subset Q$, $V_h \subset V$, and $W_h \subset W$ in place of Q, V, and W, i.e., we search for $x_h = (q_h, u_h, z_h) \in (\mathcal{D}\cap Q_h \times V_h \times W_h)$ such that

$$\mathcal{L}'(x_h)(\varphi) = 0 \quad \forall \varphi \in X_h = Q_h \times V_h \times W_h \tag{2.38}$$

to get a discretized stationary point $q_h = q^\delta_{\beta,h} \in \mathcal{D}\cap Q_h$ of $j_{\beta,h}$.

Analogously, the corresponding equations for the second derivatives in (2.29), (2.30) and (2.31) from Section 2.2 are considered in the discretized spaces Q_h, V_h, and W_h, and the optimization algorithms from Section 2.5 can be applied directly to the discretized versions $j_{\beta,h}$, $j'_{\beta,h}$, and $j''_{\beta,h}$ of j_β, j'_β, and j''_β.

Throughout this thesis we assume the following:

Assumption 2.3. *The operator C and the semilinear form $a\colon Q \times V \times W \to \mathbb{R}$ as well as the norms in G and Q are evaluated exactly, i.e.,*

$$\|\cdot\|_{Q_h} = \|\cdot\|_Q \,, \qquad \|\cdot\|_{G_h} = \|\cdot\|_G \,.$$

(In Chapter 4 the fact that the norm in the dual space W^* of W is not evaluated exactly will play an important role).

2.4. Adaptivity and error estimation

In order to solve PDEs numerically efficiently, it is reasonable to use local mesh refinement, e.g., in order to save computational effort or for the resolution of nonlinearities/singularities. Having computed an approximate (discrete) solution u_h, we are interested in estimating the error between this approximation and the exact (continuous) solution u and adapting the finite element mesh accordingly. There, a posteriori error estimators come into play. In particular we will use goal-oriented error estimators, which can serve to estimate the mentioned error with respect to some quantities of interest. The *dual-weighted residual (dwr) method* is a well-known and, due to its efficiency, very popular method to deduce such error estimators in the context of PDEs and optimal control, developed by Becker and Rannacher in [10] for the error in the cost functional J_β and extended to the case of a general quantity of interest by Becker and Vexler in [11], [12]. The naming may already tell that the concerned error can be estimated by weighted residuals of imposed adjoint/dual problems.

We aim to estimate the error with respect to a given quantity of interest $E\colon Q \times V \to \mathbb{R}$, i.e., we seek $\eta^E \in \mathbb{R}$ such that

$$E(q,u) - E(q_h,u_h) = \eta^E + \mathbf{R}$$

with some negligible higher order remainder term $\mathbf{R}$. For this we define an auxiliary Lagrangian

$$\mathcal{M}\colon X^2 \to \mathbb{R}\,, \quad \mathcal{M}(x,\xi) := E(q,u) + \mathcal{L}'(x)(\xi)\,, \tag{2.39}$$

where $x = (q,u,z) \in X = Q \times V \times W$ and $\mathcal{L}$ is the Lagrangian from (2.24). According to [11], for stationary points $y := (x,\xi) \in X^2$ of $\mathcal{M}$ in X and $y_h := (x_h,\xi_h) \in X_h^2 = (Q_h \times V_h \times W_h)^2$ of $\mathcal{M}$ in X_h there holds the error respresentation

$$E(q,u) - E(q_h,u_h) = \frac{1}{2}\mathcal{M}'(y_h)(y - \widehat{y}_h) + \mathbf{R}\,,$$

where $\widehat{y}_h \in X_h^2$ is arbitrary and

$$\mathbf{R} = \frac{1}{2}\int_0^1 \mathcal{M}'''(y_h + se)(e,e,e)s(s-1)\,ds \tag{2.40}$$

with $e = y - y_h$. The term $y - \widehat{y}_h$ will be computed approximately by using a suitable interpolation operator $\mathrm{IO}\colon X^2 \to X_h^2$ [7] and approximating the resulting interpolation error using an operator $\pi\colon X_h^2 \to \widehat{X}_h^2$, with $\widehat{X}_h \neq X_h$ such that $y - \pi y_h$ has a better local asymptotical behavior than $y - \mathrm{IO}\, y$ [39]. Then we compute

$$\eta^E = \frac{1}{2}\mathcal{M}'(y_h)(\pi y_h - y_h)\,.$$

A stationary point $y = (x, \xi)$ of $\mathcal{M}$ can be computed in a very efficient manner (similarly to Section 2.2), especially in combination with an optimization algorithm (see Section 3.3 for more details). In this context, note that a stationary point of $\mathcal{L}$ represents the first part x of a stationary point of $\mathcal{M}$.

2.5. Optimization methods

As mentioned at the end of Section 2.1, in order to solve the discrepancy principle equation (2.11) iteratively, we need to solve the optimization problem (2.34) in each iteration. There are plenty of ways to do so. We will give a short overview of some of them which will play a role later in this thesis: *Newton*'s method, *(Generalized) Gauss-Newton* methods, *Sequential Quadratic Programming (SQP)* methods, and *Penalty* methods.

Details can be found in [19, 36, 95].

2.5.1. Newton's method

Thanks to its fast local convergence properties, the method of choice for solving unconstrained (smooth) optimization problems is Newton's method (cf., e.g., [95]) as well as almost every textbook about optimization). The computation of a stationary point of the unconstrained problem (2.22) via Newton's method leads to the computation of a solution $\delta q \in Q$ to

$$j''_\beta(q)(\delta q, \varphi) = -j'_\beta(q)(\varphi) \quad \forall \varphi \in Q \tag{2.41}$$

in each iteration step, and updating q via line search (cf., e.g., [95, Section 3.5]). This means we solve (2.41) and set $q = q + \theta\delta q$ with some step size $\theta \in (0, 1]$ until $\|\nabla j_\beta(q)\|_Q$ is smaller than some given tolerance. The gradient $\nabla j_\beta(q) \in Q$ and the Hessian $\nabla^2 j_\beta(q) \colon Q \to Q$ are hereby defined via Riesz's representation as

$$\begin{aligned}(\nabla j_\beta(q), \delta q)_Q &= j'_\beta(q)(\delta q)\\ (\nabla^2 j_\beta(q)\delta q, \tau q)_Q &= j''_\beta(q)(\delta q, \tau q)\end{aligned}$$

for all $\delta q, \tau q \in Q$.

In our setting the gradient and the Hessian are explicitly given as

$$\begin{aligned}\nabla j_\beta(q) &= 2F'(q)^*(F(q) - g^\delta) + \tfrac{2}{\beta}(q - q_0)\\ \nabla^2 j_\beta(q) &= 2F''(q)^*(F(q) - g^\delta) + 2F'(q)^*F'(q) + \tfrac{2}{\beta}\,\mathrm{id}\ .\end{aligned}$$

The Newton equation (2.41) reads

$$\left[F''(q)^*(F(q) - g^\delta) + F'(q)^*F'(q) + \tfrac{1}{\beta}\,\mathrm{id}\right]\delta q = -F'(q)^*(F(q) - g^\delta) - \tfrac{1}{\beta}(q - q_0)\,, \tag{2.42}$$

and analogously in the discrete setting (cf. Section 2.3) with j_β, F, and Q replaced by $j_{\beta,h}$, F_h, and Q_h.

2.5.2. Iteratively regularized Gauss-Newton methods

As already mentioned in the two previous subsections, computing second order information can be very expensive, which motivates the idea of neglecting the second order term $F''(q)$ in (2.42), such that we get the following iteration rule: Solve

$$(F'(q)^* F'(q) + \tfrac{1}{\beta}\,\mathrm{id})\delta q = -F'(q)^*(F(q) - g^\delta) - \tfrac{1}{\beta}(q - q_0) \tag{2.43}$$

and set $q = q + \delta q$ for the next iteration.

The equation (2.43) is a necessary and (due to convexity) sufficient optimality condition for the minimization problem

$$\min_{\delta q \in Q} \|F(q) + F'(q)(\delta q) - g^\delta\|_G^2 + \tfrac{1}{\beta}\|q + \delta q - q_0\|_Q^2 \,. \tag{2.44}$$

Considering a different regularization parameter in each iteration, this leads to the iterative regularized Gauss-Newton method (IRGNM)

$$q^k = q^{k-1} - (F'(q^{k-1})^* F'(q^{k-1}) + \tfrac{1}{\beta_k}\,\mathrm{id})^{-1}\left[F'(q^{k-1})^*(F(q^{k-1}) - g^\delta) + \tfrac{1}{\beta}(q^{k-1} - q_0)\right] ,$$

which belongs to the class of Newton type methods [69].

This method was introduced by Bakushinskii in [4], where he showed local convergence on the continuous level under the assumption that F' is Lipschitz continuous and the source condition (2.6) holds for $\nu \geq 1$. For the noise free case (i.e., $\delta = 0$) he even proved the convergence rate

$$\|q^k - q^\dagger\|_Q = O\left(\tfrac{1}{\beta_k}\right) , \quad \text{as } \beta_k \to \infty \,,$$

under the following condition on the regularization parameter

$$\beta_k \to \infty \quad \text{as } k \to \infty \quad \text{and} \quad 1 \leq \frac{\beta_{k+1}}{\beta_k} \leq \vartheta$$

for some $\vartheta > 1$, which is satisfied, e.g., for the choice $\beta_k := \beta_0 \vartheta^n$, cf. [51].

This result has been extended in the last years by Kaltenbacher, Neubauer, Ramm, Scherzer, Schöpfer and Schuster in [17, 68, 69] to the case $\nu < 1$, where for $\nu < \frac{1}{2}$ restrictions on the nonlinearity of F, such as

$$\|F(q) - F(q^\dagger) - F'(q^\dagger)(q - q^\dagger)\|_G \leq C\|F(q) - F(q^\dagger)\|_G \|q - q^\dagger\|_Q \tag{2.45}$$

for all q from a neighborhood of $q^\dagger$ and some constant $C > 0$, are required.

Since (2.44) is equivalent to

$$\begin{aligned} &\min_{(\delta q, \delta u) \in Q \times V} \|C(u) + C'(u)(\delta u) - g^\delta\|_G^2 + \tfrac{1}{\beta}\|q + \delta q - q_0\|_Q^2 \\ \text{s.t.} \quad & a'_q(q,u)(\delta q, \varphi) + a'_u(q,u)(\delta u, \varphi) = 0 \quad \forall \varphi \in W \,, \end{aligned}$$

with $u \in V$ solving $a(q,u)(\varphi) = f(\varphi)$ for all $\varphi \in W$, this method is related to the *Generalized Gauss-Newton method* [18]: Solve

$$\begin{aligned} &\min_{(\delta q,\delta u)\in Q\times V} \|C(u) - g^\delta + C'(u)(\delta u)\|_G^2 + \tfrac{1}{\beta}\|q + \delta q - q_0\|_Q^2 \\ \text{s.t.} \quad & a'_q(q,u)(\delta q,\varphi) + a'_u(q,u)(\delta u,\varphi) + a(q,u)(\varphi) - f(\varphi) = 0 \quad \forall \varphi \in W\,, \end{aligned} \tag{2.46}$$

and set $q = q + \delta q$, $u = u + \delta u$ for the next iteration. The optimization problem (2.46) is formed by linearizing the functions $u \mapsto C(u) - g^\delta$, $(q,u) \mapsto a(q,u)(\varphi) - f(\varphi)$ (and $q \mapsto q - q_0$) in (2.18) separately, such that one circumvents the solution of the nonlinear PDE. We will discuss this method in more detail in Chapter 4 in combination with adaptivity and the choice of the regularization parameter β.

2.5.3. SQP methods

The idea of *Sequential Quadratic Programming* (SQP) is the solution of a constrained optimization problem by solving a sequence of linear-quadratic optimal control problems using the Lagrangian and the first order optimality condition. They can be used in combination with line search as well as trust region methods (cf. [95]), but we will focus on the local SQP method, because this section mainly serves as motivation for Chapter 4 and Section 4.2 in particular.

Since SQP methods are tailored for constrained optimization problems, in contrast to the previously considered optimization methods, here we do not consider the reduced problem (2.22), but the formulation (2.18), again neglecting the constraint $q \in \mathcal{D}$ (cf. Remark 2.3), i.e.,

$$\begin{aligned} &\min_{(q,u)\in Q\times V} \quad J_\beta(q,u) := \|C(u) - g^\delta\|_G^2 + \tfrac{1}{\beta}\|q - q_0\|_Q^2 \\ \text{s.t.} \quad & a(q,u)(\varphi) = f(\varphi) \quad \forall \varphi \in W\,. \end{aligned} \tag{2.47}$$

Applying the Lagrange-Newton method to the optimization problem (2.47) consists in using Newton's method to solve (2.38), which leads to the iteration rule

$$\begin{aligned} \mathcal{L}''(x)(\delta x,\varphi) &= -\mathcal{L}'(x)(\varphi) \quad \forall \varphi \in X = Q\times V\times W \\ x &= x + \delta x\,. \end{aligned}$$

This is equivalent to solving the quadratic program (SQP subproblem)

$$\begin{aligned} &\min_{(\delta q,\delta u)\in Q\times V} \varPhi(\delta q,\delta u) \\ \text{s.t.} \quad & b(\delta q,\delta u)(\varphi) := a'_q(q,u)(\delta q,\varphi) + a'_u(q,u)(\delta u,\varphi) + a(q,u)(\varphi) - f(\varphi) = 0 \quad \forall \varphi \in W\,, \end{aligned}$$

where

$$
\begin{aligned}
\Phi(\delta q, \delta u) &:= J_\beta(q,u) + (J_\beta)'_q(q,u)(\delta q) + (J_\beta)'_u(q,u)(\delta u)\\
&\quad + \tfrac{1}{2}(\mathcal{L})''_{qq}(q,u,z)(\delta q,\delta q) + \tfrac{1}{2}(\mathcal{L})''_{uu}(q,u,z)(\delta u,\delta u) + (\mathcal{L})''_{qu}(q,u,z)(\delta q,\delta u)\\
&= \|C(u) - g^\delta\|_G^2 + \tfrac{1}{\beta}\|q - q_0\|_Q^2\\
&\quad + \tfrac{2}{\beta}(q - q_0, \delta q)_Q + 2(C(u) - g^\delta, C'(u)(\delta u))_G\\
&\quad + \tfrac{1}{\beta}\|\delta q\|_Q^2 + \|C'(u)(\delta u)\|_G^2 + (C(u) - g^\delta, C''(u)(\delta u,\delta u))_G\\
&\quad - \tfrac{1}{2}a''_{qq}(q,u)(\delta q,\delta q,z) - \tfrac{1}{2}a''_{uu}(q,u)(\delta u,\delta u,z) - a''_{qu}(q,u)(\delta q,\delta u,z)\\
&= \|C(u) - g^\delta + C'(u)(\delta u)\|_G^2 + \tfrac{1}{\beta}\|q + \delta q - q_0\|_Q^2\\
&\quad + (C(u) - g^\delta, C''(u)(\delta u,\delta u))_G\\
&\quad - \tfrac{1}{2}a''_{qq}(q,u)(\delta q,\delta q,z) - \tfrac{1}{2}a''_{uu}(q,u)(\delta u,\delta u,z) - a''_{qu}(q,u)(\delta q,\delta u,z)\,,
\end{aligned}
$$

and setting $q = q + \delta q$, $u = u + \delta u$ and z as solution to the dual equation

$$
b'_{\delta u}(\delta q, \delta u)(\varphi, z) = \Phi'_{\delta u}(\delta q, \delta u)(\varphi) \quad \forall \varphi \in V
$$

for the next iteration. Neglecting second derivatives, in our setting, this turns into the *Generalized Gauss-Newton method* (2.46) (see also *Levenberg-Marquardt SQP* in [22], [23]).

Although we will not treat SQP methods as such, they present a basis for the ideas in Section 4.2.

2.5.4. Penalty methods

Another common way to treat constrained optimization problems ist via penalty methods (cf., e.g., [95]). The main idea is to couple the constraint and the objective functional in the following way (applied to (2.47)):

$$
\min_{(q,u)\in Q\times V} J_\beta(q,u) + \varrho\|A(q,u) - f\|_{W^*}^r
$$

(cf. the relation (2.16)) for $r = 1$ or $r = 2$ and a *penalty parameter* $\varrho > 0$. In fact, we won't apply a penalty method to the nonlinear optimization problem (2.18) directly, but to the linearized version (2.46) (see Section 4.2), which we rewrite in a condensed form as

$$
\min_{(q,u)\in Q\times V} \mathcal{T}(q,u) \quad \text{s.t. } B(q,u) = f \text{ in } W^* \tag{2.48}
$$

with a convex functional $\mathcal{T}\colon Q \times V \to \mathbb{R}$ and a linear operator $B\colon Q \times V \to W^*$, with corresponding penalty problem

$$
\min_{(q,u)\in Q\times V} \mathcal{P}(q,u) := \mathcal{T}(q,u) + \varrho\|B(q,u) - f\|_{W^*}^r\,. \tag{2.49}
$$

In Section 4.2.2 we will especially be interested in *exact penalty methods* (cf., e.g., [95, Section 17.2]), i.e., the case $r = 1$ and ϱ sufficiently large. More precisely, if ϱ ist chosen larger than the norm of the adjoint state (cf. (2.26)), a solution to the optimal control problem (2.48) solves the penalty problem (2.49):

Let $(\bar{q}, \bar{u})$ be a solution of (2.48) with corresponding adjoint state $\bar{z}$, i.e., let $(\bar{q}, \bar{u}, \bar{z})$ solve

$$L'(\bar{q}, \bar{u}, \bar{z})(dx) = 0 \quad \forall dx \in Q \times V \times W$$

for the corresponding Lagrangian $L(q, u, z) := \mathcal{T}(q, u) + f(z) - B(q, u)(z)$. Then there holds

$$\begin{aligned} \mathcal{P}(\bar{q}, \bar{u}) = \mathcal{T}(\bar{q}, \bar{u}) = L(\bar{q}, \bar{u}, \bar{z}) &\leq L(q, u, \bar{z}) \\ &\leq \mathcal{T}(q, u) + \|\bar{z}\|_W \, \|B(q, u) - f\|_{W^*} \\ &\leq \mathcal{T}(q, u) + \rho \|B(q, u) - f\|_{W^*} = \mathcal{P}(q, u) \end{aligned}$$

for all $(q, u) \in Q \times V$, where we have used convexity of the Lagrangian L.

Remark 2.5. *For $\rho > \|\bar{z}\|_W$ we even get the other direction: For a solution (q^*, u^*) of the penalty problem* (2.49) *there holds*

$$\begin{aligned} L(q^*, u^*, \bar{z}) &= \mathcal{T}(q^*, u^*) + \langle B(q^*, u^*) - f, \bar{z} \rangle_{W^*,W} \\ &\leq \mathcal{T}(q^*, u^*) + \|\bar{z}\|_W \, \|B(q^*, u^*) - f\|_{W^*} \\ &\leq \mathcal{T}(q^*, u^*) + \rho \|B(q^*, u^*) - f\|_{W^*} \\ &\leq \mathcal{T}(\bar{q}, \bar{u}) + \rho \|B(\bar{q}, \bar{u}) - f\|_{W^*} \\ &= \mathcal{T}(\bar{q}, \bar{u}) \\ &= \mathcal{T}(\bar{q}, \bar{u}) + \langle B(\bar{q}, \bar{u}) - f, \bar{z} \rangle_{W^*,W} \\ &= L(\bar{q}, \bar{u}, \bar{z}) \\ &\leq L(q^*, u^*, \bar{z}) \,, \end{aligned}$$

such that there holds in fact equality in the previous inequalities. This implies

$$\rho \|B(q^*, u^*) - f\|_{W^*} = \|\bar{z}\|_W \, \|B(q^*, u^*) - f\|_{W^*} \,,$$

which finally yields $\|B(q^*, u^*) - f\|_{W^*} = 0$, since $\rho > \|\bar{z}\|_W$.

3. Computation of a Tikhonov regularization parameter with adaptive discretizations

The aim of this chapter is to combine the choice of the Tikhonov regularization parameter β (cf. (2.11), (2.12)) with adaptive mesh refinement (cf. Section 2.3 and 2.4).

In [39] Griesbaum, Kaltenbacher, and Vexler already formulated an algorithm, which iteratively determines a regularization parameter β fulfilling the relaxed version (2.12) of the discrepancy principle in the discretized setting and adaptively refines the discretization by means of goal-oriented a posteriori error estimators (cf. Section 2.4) for linear inverse problems

$$Tq = g \tag{3.1}$$

with $T\colon Q \to G$ linear.

The presented algorithm for linear problems leads not only to a considerable saving of computational effort when solving a single regularized problem, but yields an even higher gain in CPU time and storage when used within a Newton method for determining the regularization parameter according to the discrepancy principle (2.11), which involves several regularized problems to be solved at different discretization levels.

As a matter of fact, many interesting parameter identification problems are governed by nonlinear PDEs, which motivated us to extend the results from [39] as much as possible to the nonlinear case. Like in [39] we apply an inexact Newton method to the equation (2.11). In this process, we have to solve (2.34) in each iteration in order to obtain a minimizer or rather (2.38) to get a stationary point $q_{\beta,h}^{\delta}$ of $j_{\beta,h}$ for the current value of β, which is required to evaluate $i_h(\beta) = \|F_h(q_{\beta,h}^{\delta}) - g^{\delta}\|_G^2$. In accordance with Section 2.3 and 2.4, we propose to do so on adaptively refined discretizations of the problem, which enables to save a considerable amount of computational effort.

Most of the results of this chapter have been published in [65].

3.1. Convergence of adaptively discretized stationary points to an exact solution for vanishing noise

We will first of all carry over the convergence analysis of Tikhonov regularization from the continuous setting to an adaptively discretized one. The crucial point here is that we do not impose accuracy in the sense of smallness of operator norms or closeness of Hilbert space elements, but only three scalar valued quantities have to be computed precisely enough, namely the value of i and its derivative i' at the current iterate β, and finally the value of the Tikhonov functional $j_\beta(q_\beta^\delta)$. In order to do so, throughout this thesis, we will need some assumptions on the operator F.

First, througout this chapter, we assume the existence of an exact solution to the inverse problem, i.e.,

Assumption 3.1. *There exists $q^\dagger \in \mathcal{D}$, such that $q^\dagger$ solves* (2.1)*.*

In order to obtain convergence $F(q_n) \to g$ (as $n \to \infty$) for some appropriate sequence q_n we need that F is weakly sequentially closed, which is formulated in the following assumption.

Assumption 3.2. *Let the reduced forward operator F be continuous and satisfy*

$$(q_n \rightharpoonup q \ \wedge \ F(q_n) \to g) \ \Rightarrow \ (q \in \mathcal{D} \ \wedge \ F(q) = g)$$

for all $(q_n)_{n\in\mathbb{N}} \subseteq \mathcal{D} \subseteq Q$

Note that Assumption 3.2 is less restrictive as compared to (2.5), since the premiss contains strong convergence of $(F(q_n))_{n\in\mathbb{N}}$.

Furthermore F can not be “too nonlinear”(cf. (2.45))

Assumption 3.3. *Let the* tangential cone condition

$$\|F(q) - F(\bar{q}) - F'(\bar{q})(q - \bar{q})\|_G \leq c_{tc}\|F(q) - F(\bar{q})\|_G$$

hold for all $q, \bar{q} \in \mathcal{D} \subseteq Q$ and for some $0 < c_{tc} < 1$.

These two assumptions (Assumption 3.2 and 3.3) are very typical means in the analysis of regularization methods. Some of the first ones who used the tangential cone condition Assumption 3.3, to show convergence of the Landweber iteration, were Hanke, Scherzer, and Neubauer [44] and although it is an open question whether this condition is fulfilled for some standard examples, it is the weakest condition on the operator F found so far to guarantee convergence of methods working with the first derivative of F (cf., e.g., [27], [69], [71], [102]). We will discuss the validity of Assumption 3.2 and 3.3 in more detail by means of some examples in Section 3.4.

Assumption 3.4. *Let $0 < \underline{\beta} \leq \overline{\beta} \in \mathbb{R}$. For all $\beta \in [\underline{\beta}, \overline{\beta}]$ there exist stationary points q_β^δ and $q_{\beta,h}^\delta$ in $\mathcal{D}$ of the Tikhonov functionals j_β and $j_{\beta,h}$ (cf.* (2.23) *and* (2.35)*).*

Since Assumption 3.1 and 3.4 constitute the basic prerequisites of this chapter, we will not list them again explicitly in the following theorems.

Theorem 3.1. *Let Assumption 3.2 and 3.3 hold and let $q_\beta^\delta \in \mathcal{D}$ and $q_{\beta,h}^\delta \in \mathcal{D} \cap Q_h$ be stationary points of j_β and $j_{\beta,h}$ respectively (cf.* (2.28), (2.38)*). Let further the regularization parameter $\beta = \beta(\delta, g^\delta)$ and the discretizations Q_h,V_h,W_h be chosen such that for i_h as in* (2.37) *(see also* (2.10)*)*

$$\underline{\underline{\tau}}^2 \delta^2 \leq i_h(\beta) \leq \overline{\overline{\tau}}^2 \delta^2 \tag{3.2}$$

for constants $\overline{\overline{\tau}} > \tau > \underline{\underline{\tau}} > \frac{1+c_{tc}}{1-c_{tc}}$ *as well as*

$$|i_h(\beta) - i(\beta)| \le \tilde{\tau}^2\delta^2 \tag{3.3}$$

for some constant $\tilde{\tau} > 0$ *with*

$$\tilde{\tau}^2 < \underline{\underline{\tau}}^2 - \left(\tfrac{1+c_{tc}}{1-c_{tc}}\right)^2 . \tag{3.4}$$

Then there holds:

(i) With

$$\underline{\tau}^2 = \underline{\underline{\tau}}^2 - \tilde{\tau}^2 , \quad \overline{\tau}^2 = \overline{\overline{\tau}}^2 + \tilde{\tau}^2 \tag{3.5}$$

(2.12) *is fulfilled.*

(ii) For any solution $q^\dagger \in \mathcal{D}$ *of* (2.1) *there holds the estimate*

$$\|q_\beta^\delta - q_0\|_Q \le \|q^\dagger - q_0\|_Q . \tag{3.6}$$

(iii) The sequence $q_\beta^\delta = q_{\beta(\delta)}^\delta$ *converges (weakly) subsequentially to a solution of* (2.1) *as* $\delta \to 0$ *in the sense that it has a weakly convergent subsequence and each weakly convergent subsequence converges strongly to a solution of* (2.1)*. If the solution* $q^\dagger$ *to* (2.1) *is unique, then* $q_{\beta(\delta)}^\delta$ *converges to* $q^\dagger$ *as* $\delta \to 0$.

Proof. (i): With $\underline{\tau}, \underline{\underline{\tau}}, \overline{\tau}, \overline{\overline{\tau}}, \tilde{\tau}$ as in (3.5) we can conclude from (3.2) that

$$\underline{\tau}^2\delta^2 + \tilde{\tau}^2\delta^2 \le i_h(\beta) \le \overline{\tau}^2\delta^2 - \tilde{\tau}^2\delta^2$$

holds, and using (3.3) we have

$$\underline{\tau}^2\delta^2 \le i_h(\beta) - \tilde{\tau}\delta^2 \le i(\beta) \le i_h(\beta) + \tilde{\tau}^2\delta^2 \le \overline{\tau}^2\delta^2 .$$

(ii): Since q_β^δ is a stationary point of j_β, there holds

$$\nabla j_\beta(q_\beta^\delta) = F'(q_\beta^\delta)^*(F(q_\beta^\delta) - g^\delta) + \tfrac{1}{\beta}(q_\beta^\delta - q_0) = 0 . \tag{3.7}$$

Forming the inner product with the error $q_\beta^\delta - q^\dagger$, we get

$$(F(q_\beta^\delta) - g^\delta, F'(q_\beta^\delta)(q_\beta^\delta - q^\dagger))_G + \tfrac{1}{\beta}\|q_\beta^\delta - q^\dagger\|_Q^2 = \tfrac{1}{\beta}(q_0 - q^\dagger, q_\beta^\delta - q^\dagger)_Q ,$$

hence for β satisfying (3.2), by (2.2), (2.12), Assumption 3.3, and (3.4)

$$\begin{aligned}
&\|F(q_\beta^\delta) - g^\delta\|_G^2 + \tfrac{1}{\beta}\|q_\beta^\delta - q^\dagger\|_Q^2 - \tfrac{1}{\beta}(q_0 - q^\dagger, q_\beta^\delta - q^\dagger)_Q \\
&= (F(q_\beta^\delta) - g^\delta, F(q_\beta^\delta) - g^\delta - F'(q_\beta^\delta)(q_\beta^\delta - q^\dagger))_G \\
&\le \|F(q_\beta^\delta) - g^\delta\|_G \|F(q_\beta^\delta) - g^\delta - F'(q_\beta^\delta)(q_\beta^\delta - q^\dagger)\|_G \\
&\le \|F(q_\beta^\delta) - g^\delta\|_G \left(\|F(q_\beta^\delta) - F(q^\dagger) - F'(q_\beta^\delta)(q_\beta^\delta - q^\dagger)\|_G + \delta\right) \\
&\le \|F(q_\beta^\delta) - g^\delta\|_G \left(c_{tc}\|F(q_\beta^\delta) - g\|_G + \delta\right) \\
&\le \|F(q_\beta^\delta) - g^\delta\|_G \left(c_{tc}\|F(q_\beta^\delta) - g^\delta\|_G + (1 + c_{tc})\delta\right) \\
&\le \|F(q_\beta^\delta) - g^\delta\|_G \left(c_{tc}\|F(q_\beta^\delta) - g^\delta\|_G + \tfrac{1+c_{tc}}{\underline{\tau}}\|F(q_\beta^\delta) - g^\delta\|_G\right) \\
&= \left(c_{tc} + \tfrac{1+c_{tc}}{\underline{\tau}}\right)\|F(q_\beta^\delta) - g^\delta\|_G^2 ,
\end{aligned}$$

and therewith

$$
\begin{aligned}
\|q_\beta^\delta - q_0\|_Q^2 &= \|q_\beta^\delta - q^\dagger\|_Q^2 - (q_0 - q^\dagger, q_\beta^\delta - q^\dagger)_Q + (q^\dagger - q_0, q_\beta^\delta - q_0)_Q \\
&\leq \beta \left(c_{tc} + \tfrac{1+c_{tc}}{\underline{\tau}} - 1\right) \|F(q_\beta^\delta) - g^\delta\|_G^2 + (q^\dagger - q_0, q_\beta^\delta - q_0)_Q \\
&\leq \beta \left(c_{tc} + \frac{1+c_{tc}}{\sqrt{\underline{\underline{\tau}}^2 - \tilde{\tau}^2}} - 1\right) \|F(q_\beta^\delta) - g^\delta\|_G^2 + \|q^\dagger - q_0\|_Q \|q_\beta^\delta - q_0\|_Q \qquad (3.8) \\
&\leq \beta \left(c_{tc} + (1 - c_{tc}) - 1\right) \|F(q_\beta^\delta) - g^\delta\|_G^2 + \|q^\dagger - q_0\|_Q \|q_\beta^\delta - q_0\|_Q \\
&= \|q^\dagger - q_0\|_Q \|q_\beta^\delta - q_0\|_Q
\end{aligned}
$$

which implies (3.6).

(iii): Due to (3.6) there exists a weakly convergent subsequence of $q_\beta^\delta = q_{\beta(\delta, g^\delta)}^\delta$. By

$$
\|F(q_\beta^\delta) - g\|_G \leq \|F(q_\beta^\delta) - g^\delta\|_G + \delta \leq (\overline{\overline{\tau}} + 1)\delta \to 0 \qquad \text{as } \delta \to 0
$$

(due to (3.2)) and the (weak) sequential closedness of F, the weak limit of any weakly convergent subsequence of q_β^δ is a solution to (2.1) and therefore can be inserted in place of $q^\dagger$ in (3.8), which implies even strong convergence by a standard argument (see, e.g, [27]):

Due to (3.6) and weak convergence we can conclude

$$
\begin{aligned}
\|q_\beta^\delta - q^\dagger\|_Q^2 &= \|q_\beta^\delta - q_0\|_Q^2 + \|q^\dagger - q_0\|_Q^2 - 2(q_\beta^\delta - q_0, q^\dagger - q_0)_Q \\
&\leq 2\|q^\dagger - q_0\|_Q^2 - 2(q_\beta^\delta - q_0, q^\dagger - q_0)_Q \\
&\to 0 \qquad \text{as } \delta \to 0\,.
\end{aligned}
$$

□

Remark 3.1. *In Chapter 4 we will show a similar convergence results (cf. Theorem 4.2). The difference to here will be that there, no tangential cone condition Assumption 3.3 is needed to show* (3.6)*.*

Remark 3.2. *Instead of* (3.3) *it suffices to require*

$$
|i_h(\beta) - i(\beta)| \leq \max\{c \cdot i_h(\beta), \tilde{\tau}^2 \delta^2\} \qquad (3.9)
$$

for some constant $c \leq \frac{\tilde{\tau}^2}{\overline{\overline{\tau}}^2}$*, since this less restrictive condition implies* (3.3)*, once* (3.2) *is satisfied.*

Remark 3.3. *In case* $q_\beta^\delta \in \mathcal{D}$ *is a global minimizer instead of a stationary point (cf. Remark 2.4), the tangential cone condition Assumption 3.3 is unnecessary,* (3.4) *can be changed to the weaker condition* $\tilde{\tau}^2 < \underline{\underline{\tau}}^2 - 1$ *and* $\overline{\overline{\tau}} > \tau > \underline{\underline{\tau}} > 1$*, and the proof of (ii) in Theorem 3.1* *can be simplified as follows:*

$$\begin{aligned}
\tfrac{1}{\beta}\|q_\beta^\delta - q_0\|_Q^2 &\leq (\underline{\tau}^2 - \tilde{\tau}^2 - 1)\delta^2 + \tfrac{1}{\beta}\|q_\beta^\delta - q_0\|_Q^2 \\
&\leq i_h(\beta) - \tilde{\tau}^2\delta^2 - \delta^2 + \tfrac{1}{\beta}\|q_\beta^\delta - q_0\|_Q^2 \\
&\leq \|F(q_\beta^\delta) - g^\delta\|_G^2 - \delta^2 + \tfrac{1}{\beta}\|q_\beta^\delta - q_0\|_Q^2 \\
&\leq \|F(q^\dagger) - g^\delta\|_G^2 - \delta^2 + \tfrac{1}{\beta}\|q^\dagger - q_0\|_Q^2 \\
&\leq \tfrac{1}{\beta}\|q^\dagger - q_0\|_Q^2
\end{aligned}$$

For proving convergence rates, we additionally need a source condition of the form (2.6):

Assumption 3.5. *Let*

$$q^\dagger - q_0 \in \mathcal{R}\left(\kappa\left(F'(q^\dagger)^* F'(q^\dagger)\right)\right)$$

hold with some $\kappa\colon \mathbb{R}^+ \to \mathbb{R}^+$*, such that* $\kappa^2\colon \lambda \mapsto \kappa(\lambda)^2$ *is strictly monotonically increasing on* $(0, \|F'(q^\dagger)\|_{Q\to G}^2]$*,* φ *defined by* $\varphi^{-1} = \kappa^2$ *is convex and* ψ *defined by* $\psi(\lambda) := \kappa(\lambda)\sqrt{\lambda}$ *is strictly monotonically increasing on* $(0, \|F'(q^\dagger)\|_{Q\to G}^2]$*.*

By $\|T\|_{X\to Y}$ we denote the operator norm $\|T\|_{X\to Y} := \sup_{x\in X, x\neq 0} \frac{\|Tx\|_Y}{\|x\|_X}$ for some operator $T\colon X \to Y$ and Hilbert spaces X, Y.

Remark 3.4. *For some operator* $T\colon Q \to G$ *the function* κ *in Assumption 3.5 should be understood as spectral representation* $\kappa(T^*T) = \int_0^a \kappa(\lambda)\, dE_\lambda$ *, where* $\{E_\lambda |\ 0 \leq \lambda \leq a\}$ *denotes the spectral family of (the nonnegative and self-adjoint operator)* T^*T *with* $a > 0$ *such that for the spectrum* $\sigma(T^*T)$ *of* T^*T *there holds* $\sigma(T^*T) \subset [0, a]$ *and* $\|T\|_{Q\to G}^2 \leq a$ *(cf., e.g., [27, 51, 69, 91]).*

Since the following result is not restricted to Tikhonov minimizers, we consider a quite general approximation $\tilde{q}$ of $q^\dagger$. Due to this universality the following theorem is of interest on its own and might serve as very useful tool for the convergence theory in the context of regularized problems in combination with the discrepancy principle (see also, e.g., the proof of Theorem 4.2).

Theorem 3.2. *Let Assumption 3.5 hold. Moreover let* F *satisfy the tangential cone condition Assumption 3.3 and let* $\tilde{q}$ *be a regularized approximation (not necessarily defined by Tikhonov regularization) of a solution* $q^\dagger$ *of* $F(q) = g$ *with* $\|g - g^\delta\|_G \leq \delta$ *such that*

$$\|\tilde{q} - q_0\|_Q \leq \|q^\dagger - q_0\|_Q\,, \tag{3.10}$$

$$\|F(\tilde{q}) - g^\delta\|_G \leq \hat{\tau}\delta \tag{3.11}$$

with some $\hat{\tau} > 0$ *independent of* δ*. Then the rate*

$$\|\tilde{q} - q^\dagger\|_Q = \mathcal{O}\left(\frac{\delta}{\sqrt{\psi^{-1}(C\delta)}}\right) \tag{3.12}$$

holds for some constant $C > 0$ *independent of* δ*.*

Proof. If $\|\tilde{q} - q^\dagger\|_Q$ vanishes we are trivially done. So we assume $\|\tilde{q} - q^\dagger\|_Q \neq 0$ for the rest of the proof.

By Assumption 3.5 there exists $\delta q \in Q$ such that

$$|(q^\dagger - q_0, q - q^\dagger)_Q| = |(\delta q, \kappa(F'(q^\dagger)^* F'(q^\dagger))(q - q^\dagger))_Q| \tag{3.13}$$

for any $q \in Q$. With Jensen's inequality (cf. [51, Lemma 3.5]), analogously to [51, Proposition 3.6]) for $a \geq \|F'(q^\dagger)\|^2_{Q\to G}$, we get

$$\begin{aligned}
&\varphi\left(\frac{\|\kappa(F'(q^\dagger)^* F'(q^\dagger))(q - q^\dagger)\|_G^2}{\|q - q^\dagger\|_Q^2}\right) \\
&= \varphi\left(\frac{\int_0^a \kappa(\lambda)^2 d\|E_\lambda(q - q^\dagger)\|^2}{\int_0^a d\|E_\lambda(q - q^\dagger)\|^2}\right) \leq \left(\frac{\int_0^a \varphi(\kappa(\lambda)^2) d\|E_\lambda(q - q^\dagger)\|^2}{\int_0^a d\|E_\lambda(q - q^\dagger)\|^2}\right) \\
&= \left(\frac{\int_0^a \lambda d\|E_\lambda(q - q^\dagger)\|^2}{\|q - q^\dagger\|_Q^2}\right) = \frac{\|(F'(q^\dagger)^* F'(q^\dagger))^{1/2}(q - q^\dagger)\|_G^2}{\|q - q^\dagger\|_Q^2} \\
&= \frac{\|F'(q^\dagger)(q - q^\dagger)\|_G^2}{\|q - q^\dagger\|_Q^2},
\end{aligned}$$

which implies

$$\|\kappa(F'(q^\dagger)^* F'(q^\dagger))(q - q^\dagger)\|_G^2 \leq \|q - q^\dagger\|_Q^2 \, \kappa^2\left(\frac{\|F'(q^\dagger)(q - q^\dagger)\|_G^2}{\|q - q^\dagger\|_Q^2}\right),$$

since $\varphi^{-1} = \kappa^2$ is strictly monotonically increasing. The used notation is the same as explained in Remark 3.4. For an introduction to functional calculus in this context, we refer to [27]. Together with (3.13) this yields

$$|(q^\dagger - q_0, q - q^\dagger)_Q| \leq \|\delta q\|_Q \, \|q - q^\dagger\|_Q \kappa\left(\frac{\|F'(q^\dagger)(q - q^\dagger)\|_G^2}{\|q - q^\dagger\|_Q^2}\right) \tag{3.14}$$

for any $q \in Q$. Combining (3.10) and (3.14) we obtain

$$\begin{aligned}
0 &\geq \|\tilde{q} - q_0\|_Q^2 - \|q^\dagger - q_0\|_Q^2 \\
&= \|\tilde{q} - q^\dagger\|_Q^2 + 2(q^\dagger - q_0, \tilde{q} - q^\dagger)_Q \\
&\geq \|\tilde{q} - q^\dagger\|_Q^2 - 2\|\delta q\|_Q \, \|\tilde{q} - q^\dagger\|_Q \kappa\left(\frac{\|F'(q^\dagger)(\tilde{q} - q^\dagger)\|_G^2}{\|\tilde{q} - q^\dagger\|_Q^2}\right).
\end{aligned} \tag{3.15}$$

With Assumption 3.3 and (3.11) we get the following estimate

$$\begin{aligned}
\|F'(q^\dagger)(\tilde{q} - q^\dagger)\|_G &\leq \|F(\tilde{q}) - F(q^\dagger)\|_G + \|F'(q^\dagger)(\tilde{q} - q^\dagger) + F(q^\dagger) - F(\tilde{q})\|_G \\
&\leq (1 + c_{tc})\|F(\tilde{q}) - F(q^\dagger)\|_G \\
&\leq (1 + c_{tc})(\|F(\tilde{q}) - g^\delta\|_G + \delta) \\
&\leq (1 + c_{tc})(\hat{\tau} + 1)\delta,
\end{aligned}$$

which together with (3.15) and by monotonicity of κ yields

$$\|\tilde{q} - q^\dagger\|_Q^2 \leq 2\|\delta q\|_Q \,\|\tilde{q} - q^\dagger\|_Q \kappa \left(\frac{\bar{C}^2\delta^2}{\|\tilde{q} - q^\dagger\|_Q^2} \right) \tag{3.16}$$

with $\bar{C} = (1 + c_{tc})(\hat{\tau} + 1)$. As we consider the case $\|\tilde{q} - q^\dagger\|_Q \neq 0$, we can multiply both sides of (3.16) by $\frac{\bar{C}\delta}{\|\tilde{q}-q^\dagger\|_Q^2}$ to obtain

$$\bar{C}\delta \leq 2\|\delta q\|_Q \frac{\bar{C}\delta}{\|\tilde{q} - q^\dagger\|_Q} \kappa \left(\frac{\bar{C}^2\delta^2}{\|\tilde{q} - q^\dagger\|_Q^2} \right) = 2\|\delta q\|_Q \psi \left(\frac{\bar{C}^2\delta^2}{\|\tilde{q} - q^\dagger\|_Q^2} \right) .$$

By strict monotonicity of ψ, this yields

$$\psi^{-1} \left(\frac{\bar{C}}{2\|\delta q\|_Q} \delta \right) \leq \frac{\bar{C}^2\delta^2}{\|\tilde{q} - q^\dagger\|_Q^2} ,$$

which proves (3.12). $\square$

Now, we come back to the specific setting from this thesis and prove convergence rates for a stationary point of the Tikhonov functional, and discretizations according to Theorem 3.1.

Corollary 3.3. *Let the source condition Assumption 3.5 and the assumptions from Theorem 3.1 be fulfilled. Then the following convergence rate is obtained:*

$$\|q_\beta^\delta - q^\dagger\|_Q = \mathcal{O} \left(\frac{\delta}{\sqrt{\psi^{-1}(C\delta)}} \right) . \tag{3.17}$$

for some constant $C > 0$ independent of δ.

Proof. Since (3.6) and (2.12) (see also Theorem 3.1 (i)) provide exactly the assumptions from Theorem 3.2, the proposition follows directly with Theorem 3.2. $\square$

Remark 3.5. *With $\kappa = \kappa_\nu \colon \lambda \mapsto \lambda^\nu$, $\nu \in (0, \frac{1}{2}]$, $\lambda > 0$ in Assumption 3.5, we can conclude the usual optimal Hölder type rates (cf. (2.7),[27])*

$$\|q_\beta^\delta - q^\dagger\|_Q = \mathcal{O} \left(\delta^{\frac{2\nu}{2\nu+1}} \right) :$$

Obviously $\kappa_\nu^2 \colon \lambda \mapsto \lambda^{2\nu}$ is strictly monotonically increasing. For $\varphi \colon \mu \mapsto \mu^{\frac{1}{2\nu}}$ there holds $\varphi^{-1} = \kappa^2$ and $\varphi''(\mu) = \frac{1}{2\nu}\left(\frac{1}{2\nu} - 1\right)\mu^{\frac{1}{2\nu}-2}$. Thus, φ is convex for $\nu \in (0, \frac{1}{2}]$. Further, we have $\psi(\delta) = \kappa_\nu(\delta)\sqrt{\delta} = \delta^{\nu+\frac{1}{2}}$, i.e., ψ is strictly monotonically increasing and there holds

$$\frac{\delta}{\sqrt{\psi^{-1}(\delta)}} = \delta^{1-\frac{1}{2(\nu+1/2)}} = \delta^{\frac{2\nu}{2\nu+1}} .$$

For a convergence (rates) proof under different conditions on the forward operator than those used here we refer to [58].

To obtain convergence rates for the discrete version $\|q^\delta_{\beta,h} - q^\dagger\|_Q$ we need another quantity of interest to be computed with sufficient precision, namely $j_{\beta,h}(q^\delta_{\beta,h})$, which is shown by means of the following proposition.

Theorem 3.4. *Let the conditions of Theorem 3.1 be fulfilled. Let moreover $\underline{\underline{\tau}}$, $\tilde{\tau}$ be chosen such that*

$$\sqrt{\underline{\underline{\tau}}^2 - \tilde{\tau}^2} > \frac{1 + c_{tc} + \sqrt{(1+c_{tc})^2 + 2(1-c_{tc})(1+\tilde{\tau}^2)}}{1 - c_{tc}} \tag{3.18}$$

is fulfilled (instead of the weaker condition (3.4)). If, for the discretization error with respect to the cost functional,

$$|j_\beta(q^\delta_\beta) - j_{\beta,h}(q^\delta_{\beta,h})| \leq \Theta^2\,\delta^2 \tag{3.19}$$

holds, where Θ is sufficiently small so that

$$\sqrt{\underline{\underline{\tau}}^2 - \tilde{\tau}^2} \geq \frac{1 + c_{tc} + \sqrt{(1+c_{tc})^2 + 2(1-c_{tc})(1+\Theta^2+\tilde{\tau}^2)}}{1 - c_{tc}}\,, \tag{3.20}$$

then a stationary point $q^\delta_{\beta,h} = q^\delta_{\beta(\delta),h(\delta)}$ converges (weakly) subsequentially to a solution of (2.1) in the same sense as in Theorem 3.1, and strongly in case $q^\dagger$ is unique as $\delta \to 0$.

If additionally the source condition Assumption 3.5 holds with f, ψ as in Theorem 3.2, then the convergence rate (3.17) is obtained for $q^\delta_{\beta,h}$ in place of q^δ_β.

Proof. There holds

$$\begin{aligned} j_\beta(q^\delta_\beta) - j_\beta(q^\dagger) &= \|F(q^\delta_\beta) - g^\delta\|_G^2 - \|F(q^\dagger) - g^\delta\|_G^2 + \tfrac{1}{\beta}\left(\|q^\delta_\beta - q_0\|_Q^2 - \|q^\dagger - q_0\|_Q^2\right) \\ &= \|F(q^\delta_\beta) - g^\delta\|_G^2 - \|F(q^\dagger) - g^\delta\|_G^2 + \tfrac{1}{\beta}(q^\delta_\beta + q^\dagger - 2q_0, q^\delta_\beta - q^\dagger)_Q \\ &= \|F(q^\delta_\beta) - g^\delta\|_G^2 - \|g - g^\delta\|_G^2 + \tfrac{2}{\beta}(q^\delta_\beta - q_0, q^\delta_\beta - q^\dagger)_Q - \tfrac{1}{\beta}\|q^\delta_\beta - q^\dagger\|_Q^2 \\ &\leq \|F(q^\delta_\beta) - g^\delta\|_G^2 + \tfrac{2}{\beta}(q^\delta_\beta - q_0, q^\delta_\beta - q^\dagger)_Q - \tfrac{1}{\beta}\|q^\delta_\beta - q^\dagger\|_Q^2\,. \end{aligned} \tag{3.21}$$

From the first order optimality condition (3.7) we get

$$\tfrac{1}{\beta}(q^\delta_\beta - q_0, q^\delta_\beta - q^\dagger)_Q = -(F(q^\delta_\beta) - g^\delta, F'(q^\delta_\beta)(q^\delta_\beta - q^\dagger))_G\,,$$

hence using the tangential cone condition Assumption 3.3

$$\begin{aligned} \|F(q^\delta_\beta) - g^\delta\|_G^2 + \tfrac{1}{\beta}(q^\delta_\beta - q_0, q^\delta_\beta - q^\dagger)_Q &= (F(q^\delta_\beta) - g^\delta, F(q^\delta_\beta) - g^\delta - F'(q^\delta_\beta)(q^\delta_\beta - q^\dagger))_G \\ &\leq \|F(q^\delta_\beta) - g^\delta\|_G \|F(q^\delta_\beta) - g^\delta - F'(q^\delta_\beta)(q^\delta_\beta - q^\dagger)\|_G \\ &\leq \|F(q^\delta_\beta) - g^\delta\|_G \left(c_{tc}\|F(q^\delta_\beta) - g\|_G + \delta\right) \\ &\leq \|F(q^\delta_\beta) - g^\delta\|_G \left(c_{tc}\|F(q^\delta_\beta) - g^\delta\|_G + (1+c_{tc})\delta\right) \\ &\leq c_{tc}\|F(q^\delta_\beta) - g^\delta\|_G^2 + (1+c_{tc})\delta\|F(q^\delta_\beta) - g^\delta\|_G \end{aligned}$$

and with (3.21)

$$\begin{aligned}
j_\beta(q_\beta^\delta) - j_\beta(q^\dagger) &\leq -\|F(q_\beta^\delta) - g^\delta\|_G^2 - \tfrac{1}{\beta}\|q_\beta^\delta - q^\dagger\|_Q^2 \\
&\quad + 2\left(c_{tc}\|F(q_\beta^\delta) - g^\delta\|_G^2 + (1+c_{tc})\delta\|F(q_\beta^\delta) - g^\delta\|_G\right) \\
&= -(1-2c_{tc})\|F(q_\beta^\delta) - g^\delta\|_G^2 + 2(1+c_{tc})\delta\|F(q_\beta^\delta) - g^\delta\|_G - \tfrac{1}{\beta}\|q_\beta^\delta - q^\dagger\|_Q^2 .
\end{aligned}$$

By (3.19) and Assumption 3.3, for the discretized quantity we further get

$$\begin{aligned}
j_{\beta,h}(q_{\beta,h}^\delta) &\leq j_\beta(q_\beta^\delta) + \Theta^2\delta^2 \\
&\leq j_\beta(q^\dagger) + \Theta^2\delta^2 - (1-2c_{tc})\|F(q_\beta^\delta) - g^\delta\|_G^2 + 2(1+c_{tc})\delta\|F(q_\beta^\delta) - g^\delta\|_G \\
&= \|g - g^\delta\|_G^2 + \tfrac{1}{\beta}\|q^\dagger - q_0\|_Q^2 + \Theta^2\delta^2 - (1-2c_{tc})\|F(q_\beta^\delta) - g^\delta\|_G^2 \\
&\qquad + 2(1+c_{tc})\delta\|F(q_\beta^\delta) - g^\delta\|_G \\
&\leq (1+\Theta^2)\delta^2 - (1-2c_{tc})\|F(q_\beta^\delta) - g^\delta\|_G^2 + 2(1+c_{tc})\delta\|F(q_\beta^\delta) - g^\delta\|_G \\
&\qquad + \tfrac{1}{\beta}\|q^\dagger - q_0\|_Q^2 ,
\end{aligned}$$

i.e., by (3.3)

$$\begin{aligned}
\|q_{\beta,h}^\delta - q_0\|_Q^2 - \|q^\dagger - q_0\|_Q^2 &= \beta\left(j_{h,\beta}(q_{\beta,h}^\delta) - \|F_h(q_{\beta,h}^\delta) - g^\delta\|_G^2\right) - \|q^\dagger - q_0\|_Q^2 \\
&\leq \beta\left(j_{h,\beta}(q_{\beta,h}^\delta) - \|F(q_\beta^\delta) - g^\delta\|_G^2 + \tilde{\tau}^2\delta^2\right) - \|q^\dagger - q_0\|_Q^2 \\
&\leq \beta\Big((1+\Theta^2)\delta^2 - (1-2c_{tc})\|F(q_\beta^\delta) - g^\delta\|_G^2 \\
&\qquad +2(1+c_{tc})\delta\|F(q_\beta^\delta) - g^\delta\|_G - \|F(q_\beta^\delta) - g^\delta\|_G^2 + \tilde{\tau}^2\delta^2\Big) \\
&\leq \beta\Big((1+\Theta^2+\tilde{\tau}^2)\delta^2 - 2(1-c_{tc})\|F(q_\beta^\delta) - g^\delta\|_G^2 \\
&\qquad +2(1+c_{tc})\delta\|F(q_\beta^\delta) - g^\delta\|_G\Big) .
\end{aligned} \tag{3.22}$$

From the fact that (2.12) holds with (3.5) (see Theorem 3.1 (i)) and by (3.20) we can conclude

$$\|F(q_\beta^\delta) - g^\delta\|_G \geq \sqrt{\underline{\tau}^2\delta^2} = \delta\sqrt{\underline{\underline{\tau}}^2 - \tilde{\tau}^2} \geq \delta\tfrac{1+c_{tc}+\sqrt{(1+c_{tc})^2+2(1-c_{tc})(1+\Theta^2+\tilde{\tau})}}{2(1-c_{tc})} ,$$

which implies

$$\beta\left((1+\Theta^2+\tilde{\tau})\delta^2 + 2(1+c_{tc})\delta\|F(q_\beta^\delta) - g^\delta\|_G - 2(1-c_{tc})\|F(q_\beta^\delta) - g^\delta\|_G^2\right) \leq 0 ,$$

since

$$x_{1/2} = \delta\tfrac{1+c_{tc}\pm\sqrt{(1+c_{tc})^2+2(1-c_{tc})(1+\Theta^2+\tilde{\tau})}}{2(1-c_{tc})}$$

are the only solutions to the quadratic equation

$$(1+\Theta^2+\tilde{\tau})\delta^2 + 2(1+c_{tc})\delta x - 2(1-c_{tc})x^2 = 0 ,$$

and $-2(1-c_{tc}) < 0$. By (3.22) this finally leads to

$$\|q_{\beta,h}^\delta - q_0\|_Q \leq \|q^\dagger - q_0\|_Q . \tag{3.23}$$

The rest of the proof follows the lines of the proof of Theorem 3.1 as well as Theorem 3.2, since (2.12) (see also Theorem 3.1 (i)) implies (3.11). □

Remark 3.6. *Again (like for Theorem 3.1, see Remark 3.3), the given proof can be simplified in case a global minimizer (and not only a stationary point) is available, insofar as the tangential cone condition Assumption 3.3 and* (3.18) *are not needed to show convergence (without convergence rates). Besides, instead of* (3.20) *we require*

$$\underline{\underline{\tau}}^2 \geq 1 + \Theta^2 .$$

Then the inequality (3.23) *is obtained as follows:*

$$\begin{aligned}
\tfrac{1}{\beta}\|q^\delta_{\beta,h} - q_0\|^2_Q &= j_{\beta,h}(q^\delta_{\beta,h}) - i_h(\beta) \\
&\leq j_\beta(q^\delta_\beta) + \Theta^2\delta^2 - \underline{\underline{\tau}}^2\delta^2 \\
&\leq \tfrac{1}{\beta}\|q^\dagger - q_0\|^2_Q + \|F(q^\dagger) - g^\delta\|^2_G + \Theta^2\delta^2 - \underline{\underline{\tau}}^2\delta^2 \\
&\leq \tfrac{1}{\beta}\|q^\dagger - q_0\|^2_Q + (1 + \Theta^2 - \underline{\underline{\tau}})\delta^2 \\
&\leq \tfrac{1}{\beta}\|q^\dagger - q_0\|^2_Q .
\end{aligned}$$

Remark 3.7. *Please note that* (3.18) *guarantees the existence of* $\Theta > 0$ *fulfilling* (3.20)*, e.g.,*

$$\Theta^2 := \frac{\left(\sqrt{\underline{\underline{\tau}}^2 - \tilde{\tau}^2}(1 - c_{tc}) - 1 - c_{tc}\right)^2 - (1 + c_{tc})^2}{2(1 - c_{tc})} - 1 - \tilde{\tau} ,$$

which is how Θ *will be chosen in the numerical test; cf. Section 3.6.*

3.2. Inexact Newton method for the computation of a regularization parameter

After having deduced conditions on the regularization parameter β and the discretizations, which guarantee convergence of the discretized control $q^\delta_{\beta,h} \to q^\dagger$ as $\delta \to 0$ (cf. Section 3.1), in this section, we will discuss how to determine such a β and discretizations Q_h, V_h, W_h fulfilling (3.2), (3.3) (or (3.9)) and (3.19) iteratively. We will do so by applying an inexact Newton method to the discrepancy principle equation (2.11) and deriving refinement criteria to guarantee convergence of the resulting sequence of regularization parameters.

As in [39], we use goal-oriented error estimators (see Section 3.3), where i and i' turn out to be appropriate quantities of interest. The main difficulty in transferring the results from the linear case [39] is the following: while $\beta \mapsto i(\beta)$ is convex and monotone in the linear case, which implies global monotone convergence of Newton's method for (2.11), these properties partially get lost in the nonlinear case. Still, monotonicity follows in a very general setting from minimality arguments.

First, we restrict ourselves to a closed interval $\beta \in [\underline{\beta}, \overline{\beta}]$, with some $\overline{\beta} > \underline{\beta} > 0$ and show later that our iterates do not leave this interval.

Lemma 3.5. *Provided, for any* $\beta \in [\underline{\beta}, \overline{\beta}]$ *there exists a minimizer of the Tikhonov functional, then the function* i *(defined by* (2.10) *with any selection* $q^\delta_\beta \in \operatorname{argmin}_{q \in \mathcal{D}} j_\beta(q)$*) is a monotonically decreasing function on* $[\underline{\beta}, \overline{\beta}]$

Proof. For any $\underline{\beta} \leq \beta_1 \leq \beta_2 \leq \overline{\beta}$ and any $q(\beta_i) = q^\delta_{\beta_i} \in \operatorname{argmin}_{q\in\mathcal{D}}(j_{\beta_i})$, $i = 1, 2$, we have by minimality

$$
\begin{aligned}
j_{\beta_2}(q(\beta_2)) \leq j_{\beta_2}(q(\beta_1)) &= j_{\beta_1}(q(\beta_1)) + \left(\tfrac{1}{\beta_2} - \tfrac{1}{\beta_1}\right) \|q(\beta_1) - q_0\|^2_Q \\
&\leq j_{\beta_1}(q(\beta_2)) + \left(\tfrac{1}{\beta_2} - \tfrac{1}{\beta_1}\right) \|q(\beta_1) - q_0\|^2_Q \\
&= j_{\beta_2}(q(\beta_2)) + \left(\tfrac{1}{\beta_2} - \tfrac{1}{\beta_1}\right) \left(\|q(\beta_1) - q_0\|^2_Q - \|q(\beta_2) - q_0\|^2_Q\right) ,
\end{aligned}
\tag{3.24}
$$

which implies $\left(\frac{1}{\beta_2} - \frac{1}{\beta_1}\right) (\|q(\beta_1) - q_0\|^2_Q - \|q(\beta_2) - q_0\|^2_Q) \geq 0$, thus monotone increase of the mapping $\beta \mapsto \|q(\beta) - q_0\|^2_Q$. This yields monotone decrease of i:

$$
\begin{aligned}
i(\beta_2) &= j_{\beta_2}(q(\beta_2)) - \frac{1}{\beta_2}\|q(\beta_2) - q_0\|^2_Q \\
&\leq j_{\beta_2}(q(\beta_2)) - \frac{1}{\beta_2}\|q(\beta_1) - q_0\|^2_Q \\
&\leq j_{\beta_2}(q(\beta_1)) - \frac{1}{\beta_2}\|q(\beta_1) - q_0\|^2_Q \\
&\leq i(\beta_1) + \frac{1}{\beta_2}\|q(\beta_1) - q_0\|^2_Q - \frac{1}{\beta_2}\|q(\beta_1) - q_0\|^2_Q \\
&= i(\beta_1).
\end{aligned}
$$

□

Two-times differentiability of i as well as bounds on the second derivative i'' will be shown in Lemma 3.6 under sufficient optimality conditions of second order

Assumption 3.6. *For every $\beta \in [\underline{\beta}, \overline{\beta}]$ let a stationary point q^δ_β of j_β exist, which fulfills the second order optimality condition*

$$
j''_\beta(q^\delta_\beta)[\delta q, \delta q] = 2\|F'(q^\delta_\beta)\delta q\|^2_G + 2(F(q^\delta_\beta)) - g^\delta, F''(q^\delta_\beta)(\delta q, \delta q))_G + \tfrac{2}{\beta}\|\delta q\|^2_Q \geq 2\eta\|\delta q\|^2_Q
$$

for all $\delta q \in Q$.

We are well aware of the fact that sufficient optimality conditions of second order are strong assumptions and in general hard to check. But at the same time, we would like to point out that in nonlinear optimization there are common means for extending results from a convex setting to non-convex objective functionals.

Lemma 3.6. *Let Assumption 3.3 and 3.6 be satisfied.*

(i) Then for every $\beta \in [\underline{\beta}, \overline{\beta}]$ there exists a neighborhood U_β of β and a choice of stationary points $q(\beta)$ of j_β such that the function i (defined by (2.10)*) is continuously differentiable and monotonically decreasing on U_β with*

$$
i'(\beta) \leq -2\underline{\beta}\eta\|q'(\beta)\|^2_Q \leq 0\,.
\tag{3.25}
$$

(ii) If additionally F''' is continuous, we obtain continuity of i'' on the compact interval $[\underline{\beta}, \overline{\beta}]$; hence, there exist $\underline{\gamma}, \overline{\gamma} \in \mathbb{R}$ such that

$$\underline{\gamma} \le i'' \le \overline{\gamma} \quad \textit{in } U_\beta\,. \tag{3.26}$$

(iii) Provided

$$\|F(q_\beta^\delta) - g^\delta\|_G > \tfrac{1+c_{tc}}{1-c_{tc}}\delta\,, \tag{3.27}$$

(which is fulfilled, e.g., by the choice (3.2)–(3.4)*) there even holds strict monotonicity in* (3.25)*, i.e.,*

$$i'(\beta) \le -2\underline{\beta}\eta\|q'(\beta)\|_Q^2 < 0\,. \tag{3.28}$$

Proof. (i): We define the function

$$\Psi\colon Q \times \mathbb{R} \to Q^*\,,\ \Psi(q,\beta) := \beta F'(q)^*(F(q) - g^\delta) + q - q_0\,,$$

or in terms of inner products and using Riesz's representation

$$\Psi(q,\beta)(\delta q) = \beta(F(q) - g^\delta, F'(q)(\delta q))_G + (q - q_0, \delta q)_Q\,. \tag{3.29}$$

Then a stationary point q_β^δ satisfies the equation $\Psi(q_\beta^\delta,\beta)(\delta q) = 0$ for all $\delta q \in Q$. Due to Assumption 3.6 there holds

$$\begin{aligned}(\Psi_q'(q_\beta^\delta,\beta)\delta q, \delta q)_Q &= \beta\left(\|F'(q_\beta^\delta)(\delta q)\|_G^2 + (F(q_\beta^\delta) - g^\delta, F''(q_\beta^\delta)(\delta q,\delta q))_G\right) + \|\delta q\|_Q^2 \\ &\ge \beta\eta\|\delta q\|_Q^2 - \|\delta q\|_Q^2 + \|\delta q\|_Q^2 \\ &= \beta\eta\,\|\delta q\|_Q^2\,.\end{aligned}$$

Hence the bilinear form $(\varphi,\Phi) \mapsto (\Psi_q'(q_\beta^\delta,\beta)\varphi,\Phi)_Q$ is coercive in $Q\times Q$ and therewith the mapping $\Psi_q'(q_\beta^\delta,\beta)\colon Q \to Q^*$ is continuously invertible. The implicit function theorem provides the existence of a neighboorhood U_β of β and a smooth path $q(\beta) = q_\beta^\delta$ for $\beta \in U_\beta$ with the claimed differentiability under the respective differentiability assumptions on F. This implies

$$i \in C^1(U_\beta) \text{ if } F \in C^2(Q) \text{ and } i \in C^2(U_\beta) \text{ if } F \in C^3(Q)\,. \tag{3.30}$$

To obtain (3.25), we differentiate the necessary first order optimality condition equation

$$\Psi(q(\beta),\beta)(\delta q) = 0 \qquad \forall \delta q \in Q \tag{3.31}$$

with respect to β:

$$\begin{aligned}&(F(q(\beta)) - g^\delta, F'(q(\beta))(\delta q))_G + \beta(F(q(\beta)) - g^\delta, F''(q(\beta))(q'(\beta),\delta q))_G \\ &\quad + \beta(F'(q(\beta))(q'(\beta)), F'(q(\beta))(\delta q))_G + (q'(\beta),\delta q)_Q = 0 \qquad \forall \delta q \in Q\,.\end{aligned} \tag{3.32}$$

Setting $\delta q = q'(\beta)$ yields

$$\begin{aligned}&(F(q(\beta)) - g^\delta, F'(q(\beta))(q'(\beta)))_G \\ &\quad = -\beta(F(q(\beta)) - g^\delta, F''(q(\beta))(q'(\beta),q'(\beta)))_G - \beta\|F'(q(\beta))(q'(\beta))\|_G^2 - \|q'(\beta)\|_Q^2\end{aligned}$$

and finally

$$\begin{aligned}
i'(\beta) &= \tfrac{d}{d\beta}\|F(q(\beta)) - g^\delta\|_G^2 \\
&= 2(F(q(\beta)) - g^\delta, F'(q(\beta))q'(\beta))_G \\
&= -2\beta\left((F(q(\beta)) - g^\delta, F''(q(\beta))(q'(\beta), q'(\beta)))_G + \|F'(q(\beta))q'(\beta)\|_G^2\right) - 2\|q'(\beta)\|_Q^2 \\
&\le -2\beta\eta\|q'(\beta)\|_Q^2 \\
&\le -2\underline{\beta}\eta\|q'(\beta)\|_Q^2 \,,
\end{aligned}$$

where we have used, once again, (3.31), Assumption 3.6, and $\beta \le \overline{\beta}$. This implies the first estimate in (3.28).

(ii): (3.30) implies (3.26).

(iii): It remains to show that $q'(\beta) \neq 0$. Let us assume that $q'(\beta) = 0$. Then (3.32) would simplify to

$$(F(q(\beta)) - g^\delta, F'(q(\beta))(\delta q))_G = 0 \qquad \forall \delta q \in Q\,.$$

Setting $\delta q = q(\beta) - q^\dagger$ we would then arrive at

$$\begin{aligned}
0 &= (F(q(\beta)) - g^\delta, F'(q(\beta))(q(\beta) - q^\dagger))_G \\
&\ge \|F(q(\beta)) - g^\delta\|_G^2 - (F(q(\beta)) - g^\delta, F(q(\beta)) - g^\delta - F'(q(\beta))(q(\beta) - q^\dagger))_G \\
&\ge \|F(q(\beta)) - g^\delta\|_G^2 - \|F(q(\beta)) - g^\delta\|_G\|F(q(\beta)) - g^\delta - F'(q(\beta))(q(\beta) - q^\dagger)\|_G \\
&\ge \|F(q(\beta)) - g^\delta\|_G^2 - \|F(q(\beta)) - g^\delta\|_G\left(\|F(q(\beta)) - g - F'(q(\beta))(q(\beta) - q^\dagger)\|_G + \delta\right) \\
&\ge \|F(q(\beta)) - g^\delta\|_G^2 - \|F(q(\beta)) - g^\delta\|_G\,(c_{tc}\|F(q(\beta)) - g\|_G + \delta) \\
&\ge \|F(q(\beta)) - g^\delta\|_G^2 - \|F(q(\beta)) - g^\delta\|_G\left(c_{tc}\|F(q(\beta)) - g^\delta\|_G + (1 + c_{tc})\delta\right) \\
&= \|F(q(\beta)) - g^\delta\|_G\left((1 - c_{tc})\|F(q(\beta)) - g^\delta\|_G - (1 + c_{tc})\delta\right)\,,
\end{aligned}$$

which gives a contradiction to (3.27).

□

Therewith, we are led to make use of globally and superlinearly convergent monotone modifications of Newton's method, cf., e.g., [98], where the solution is approached from above and below by two simultaneously computed sequences of iterates. In the simple one-dimensional situation we deal with here, we can use the fact that quadratic equations can be solved explicitly (see (3.52) below) to get rid of the necessity of computing two sequences. For this purpose we follow the ideas of [98], i.e., using the existence of a lower and an upper bound to the second derivative of the quantity of interest i to guarantee quadratic convergence of the sequence of regularization parameters β. But we will see that for showing linear convergence the existence of a lower bound suffices. An approximation i''_h to the second derivative i'' can be computed very efficiently in the context of goal oriented error estimation, see Section 3.3 and [39, Section 2].

Correspondingly we formulate the following convergence theorem for a sequence $(\beta^k)_{k\in\mathbb{N}}$ produced by an inexact Newton method for (2.11), which represents an extension of [39, Theorem 1] to nonlinear inverse problems.

We wish to point out that Theorem 3.7 is not restricted to the definition of i as in (2.10). In fact, the results hold for any function fulfilling

$$i \in C^2(\mathbb{R}^+)\,, \quad i'(\beta) < 0\,, \quad \text{and} \quad i'' \text{ is bounded from below.}$$

for all $\beta > 0$, such that Theorem 3.7 might also be of interest for a different setting outside of this thesis.

Theorem 3.7. *Let i be twice continuously differentiable and let $\underline{\gamma} < \gamma$ for some $\gamma \geq 0$ independent of β, denote a lower bound on i'', i.e.,*

$$\underline{\gamma} \leq i''(\beta) \quad \forall \beta \in \mathbb{R} \tag{3.33}$$

and assume that

$$i'(\beta) < 0 \quad \forall \beta \in \mathbb{R}\,. \tag{3.34}$$

Choose $\beta^0 > 0$ so that

$$i_h^0 = i_h(\beta^0) \geq \tau^2\delta^2 \tag{3.35}$$

for some $\underline{\underline{\tau}} < \tau$ and define

$$k_* = \min\{k \in \mathbb{N} \mid i_h^k - \overline{\overline{\tau}}^2\delta^2 \leq 0\} \tag{3.36}$$

for some $\overline{\overline{\tau}} > \tau$,

$$\begin{aligned} s_N^{k+1} &= -\frac{i_h^k - \tau^2\delta^2}{{i'}_h^k}\,, \quad \sigma^{k+1} = \frac{2}{1 + \sqrt{1 - 2\gamma s_N^{k+1}/{i'}_h^k}} \\ s^{k+1} &= \sigma^{k+1}\, s_N^{k+1}\,, \quad \beta^{k+1} = \beta^k + s^{k+1} \end{aligned} \tag{3.37}$$

with i_h^k, ${i'}_h^k$ satisfying

$$|i(\beta^k) - i_h^k| \leq c_1 |i_h^k - \tau^2\delta^2| \tag{3.38}$$

$$|i'(\beta^k) - {i'}_h^k| \leq c_2 |{i'}_h^k| \tag{3.39}$$

$$|i(\beta^k) - i_h^k| + |i'(\beta^k) - {i'}_h^k|\, |s^{k+1}| \leq \tfrac{\gamma+\underline{\gamma}}{2}|s^{k+1}|^2 \tag{3.40}$$

for all $k \in \mathbb{N}$ and

$$|i(\beta^{k_*}) - i_h^{k_*}| \leq (\tau^2 - \underline{\underline{\tau}}^2)\delta^2 \tag{3.41}$$

for some constants $c_1 \in (0,1)$, $c_2 \in (0, \frac{1}{2})$ independent of k and k_*, as well as for $k \geq k_* - 1$ additionally

$$|i(\beta^{k+1}) - i_h^{k+1}| + |i(\beta^k) - i_h^k| + |i'(\beta^k) - {i'}_h^k|\, |s^{k+1}| < \tfrac{\gamma+\underline{\gamma}}{2}|s^{k+1}|^2\,. \tag{3.42}$$

Then there holds

(i) negativity of the approximate derivative, i.e.,

$$i'^{k}_{h} < 0 \quad \forall k \in \mathbb{N}\,, \tag{3.43}$$

(ii) well-definedness, i.e.,

$$s_N^k \geq 0 \quad \forall k \in \mathbb{N} \quad \textit{and} \quad \sigma^k \in (0,1] \quad \forall k \in \mathbb{N}\,, \tag{3.44}$$

(iii) monotonicity

$$\beta^{k-1} \leq \beta^k \leq \beta_* \quad \forall k \in \mathbb{N}\,, \tag{3.45}$$

(iv) boundedness of i'^{k}_{h}, i.e., existence of constants $a, b > 0$, such that for all $k \in \mathbb{N}$

$$a \leq |i'^{k}_{h}| \leq b\,, \tag{3.46}$$

(v) finiteness of k_*,

(vi) convergence

$$\sigma^k \to 1 \quad \textit{as } k \to \infty\,, \tag{3.47}$$

(vii) convergence

$$\beta^k \to \beta_* \qquad \textit{as } k \to \infty\,, \tag{3.48}$$

(viii) boundedness of $i_h^{k_*}$ with

$$\underline{\underline{\tau}}^2\delta^2 \leq i_h^{k_*} \leq \overline{\overline{\tau}}^2\delta^2\,. \tag{3.49}$$

Remark 3.8. *At this point we give a short motivation for the strange-looking step size σ^{k+1} in (3.37) before this is analysed in more detail in the subsequent proof. Let $s = \beta_* - \beta^k$. Taylor expansion yields*

$$0 = i(\beta^k) - i(\beta_*) + i'(\beta^k)s + \frac{1}{2}i''(\beta^k)s^2 + O(s^3) \approx i_h^k - i(\beta_*) + i'^{k}_{h}s + \frac{\gamma^2}{2}s^2 + O(s^3)\,,$$

and it can be shown that for $s = s^{k+1}$ there holds $i_h^k - i(\beta_*) + i'^{k}_{h}s + \frac{\gamma^2}{2}s^2 = 0$.

Proof. (i): By (3.39) we have

$$|i'(\beta^k)| = -i'(\beta^k) \geq |i'^{k}_{h}| - |i'(\beta^k) - i'^{k}_{h}| \geq (1-c_2)|i'^{k}_{h}| \qquad \forall k \in \mathbb{N}\,,$$

and thus

$$\begin{aligned} i'^{k}_{h} &\leq i'(\beta^k) + |i'(\beta^k) - i'^{k}_{h}| \leq i'(\beta^k) + c_2|i'^{k}_{h}| \leq i'(\beta^k) + \tfrac{c_2}{1-c_2}|i'(\beta^k)| \\ &= i'(\beta^k) - \tfrac{c_2}{1-c_2}i'(\beta^k) = \tfrac{1-2c_2}{1-c_2}i'(\beta^k) < 0 \end{aligned} \tag{3.50}$$

for all $k \in \mathbb{N}$.

(ii): We show (3.44) by induction. By (3.35) together with (3.43) we have $s_N^1 \geq 0$ and $\sigma^1 \in (0,1]$. For the induction step we assume that $s_N^{k+1} \geq 0$ and $\sigma^{k+1} \in (0,1]$ for some $k \geq 0$.

Then there holds

$$\beta^k + ts^{k+1} \geq 0 \qquad \forall t \in (0,1]\,. \tag{3.51}$$

Observe that the roots of the quadratic equation

$$i_h^k - \tau^2\delta^2 + {i'}_h^k s - \tfrac{\gamma}{2}s^2 = 0 \tag{3.52}$$

are

$$s = \tfrac{1}{\gamma}\left({i'}_h^k \pm \sqrt{({i'}_h^k)^2 + 2\gamma(i_h^k - \tau^2\delta^2)}\right)\,.$$

Consequently

$$\begin{aligned}
s^{k+1} &= \sigma^{k+1}s_N^{k+1}\\
&= -\frac{2(i_h^k - \tau^2\delta^2)}{{i'}_h^k(1+\sqrt{1-2\gamma s_N^{k+1}/{i'}_h^k})}\\
&= -\frac{2(i_h^k - \tau^2\delta^2)}{{i'}_h^k + \sqrt{({i'}_h^k)^2 - 2\gamma s_N^{k+1}{i'}_h^k}}\\
&= -\frac{2(i_h^k - \tau^2\delta^2)}{{i'}_h^k + \sqrt{({i'}_h^k)^2 + 2\gamma(i_h^k - \tau^2\delta^2)}}\\
&= \frac{2(i_h^k - \tau^2\delta^2)({i'}_h^k - \sqrt{({i'}_h^k)^2 + 2\gamma(i_h^k - \tau^2\delta^2)})}{2\gamma(i_h^k - \tau^2\delta^2)}\\
&= \tfrac{1}{\gamma}\left({i'}_h^k - \sqrt{({i'}_h^k)^2 + 2\gamma(i_h^k - \tau^2\delta^2)}\right)
\end{aligned}$$

solves (3.52) and there holds

$$\begin{aligned}
i(\beta^{k+1}) - i(\beta_*) &= i(\beta^{k+1}) - \tau^2\delta^2\\
&= i(\beta^{k+1}) - i_h^k - {i'}_h^k s^{k+1} + \tfrac{\gamma}{2}(s^{k+1})^2\\
&= i(\beta^k) + i'(\beta^k)s^{k+1} + \int_0^1 (1-t)i''(\beta^k + ts^{k+1})\,dt\,(s^{k+1})^2\\
&\quad - i_h^k - {i'}_h^k s^{k+1} + \tfrac{\gamma}{2}(s^{k+1})^2\\
&\geq i(\beta^k) + i'(\beta^k)s^{k+1} + \tfrac{\underline{\gamma}}{2}(s^{k+1})^2 - i_h^k - {i'}_h^k s^{k+1} + \tfrac{\gamma}{2}(s^{k+1})^2\\
&\geq \tfrac{\gamma+\underline{\gamma}}{2}|s^{k+1}|^2 - |i(\beta^k) - i_h^k| - |i'(\beta^k) - {i'}_h^k||s^{k+1}|\,,
\end{aligned} \tag{3.53}$$

where we have used (3.51).

Now we distinguish two cases: $k < k_* - 1$ and $k \geq k_* - 1$. If $k < k_* - 1$, the definition of k_* (cf. (3.36)) implies $i_h^{k+1} \geq \overline{\overline{\tau}}^2\delta^2 \geq \tau^2\delta^2$, which together with (3.43) leads to the desired assertion $s_N^{k+2} \geq 0$ and $\sigma^{k+2} \in (0,1]$.

If $k \geq k_* - 1$, (3.53) together with (3.42) implies

$$i(\beta^{k+1}) - i(\beta_*) > |i(\beta^{k+1}) - i_h^{k+1}| \tag{3.54}$$

and finally

$$i_h^{k+1} > i(\beta_*) = \tau^2\delta^2\,,$$

which together with (3.43) again leads to $s_N^{k+2} \geq 0$ and $\sigma^{k+2} \in (0,1]$.

(iii): Since we already showed (3.44), we can conclude that (3.53) holds for all $k \in \mathbb{N}$. With (3.40) we have

$$i(\beta^k) \geq i(\beta_*) \qquad \forall k \in \mathbb{N}\,, \tag{3.55}$$

and with $i'(\beta) < 0$ (cf. (3.34)) there holds $\beta^k \leq \beta_*$ for all $k \in \mathbb{N}$, which gives the right inequality in (3.45). The left inequality in (3.45) follows directly from (3.44).

So $(\beta^k)_{k\in\mathbb{N}}$ is a monotonically increasing sequence which is bounded from above by β_* and consequently convergent. In the following we show that its limit is in fact β_*. In order to do so, we first prove the boundedness of $|{i'}_h^k|$.

(iv): To show (3.46), we use (3.45) with the fact that by our assumption $i' < 0$ (cf. (3.34)) we have

$$0 < \underline{c} \coloneqq \min_{\beta\in[\beta^0,\beta_*]} -i'(\beta) \leq -i'(\beta^k) \leq \max_{\beta\in[\beta^0,\beta_*]} -i'(\beta) =: \bar{c}$$

and therewith, by (3.50)

$$|{i'}_h^k| = -{i'}_h^k \geq \tfrac{1-2c_2}{1-c_2}\underline{c}$$

as well as by (3.39)

$$|{i'}_h^k| = -{i'}_h^k \leq -i'(\beta^k) + |i'(\beta^k) - {i'}_h^k| \leq -i'(\beta^k) + \tfrac{c_2}{1-c_2}|i'(\beta^k)| = \tfrac{1}{1-c_2}|i'(\beta^k)| \leq \tfrac{1}{1-c_2}\bar{c}$$

for all $k \in \mathbb{N}$. So with $a \coloneqq \frac{1-2c_2}{1-c_2}\underline{c}$ and $b \coloneqq \frac{1}{1-c_2}\bar{c}$, (3.46) is satisfied.

(v): By (3.45) the sequence $(\beta^k)_{k\in\mathbb{N}}$ is convergent, such that

$$\sigma^{k+1}s_N^{k+1} = \beta^{k+1} - \beta^k \to 0 \qquad \text{as } k \to \infty\,. \tag{3.56}$$

This leads to two cases:

1. case: $s_N^{k+1} \not\to 0$, hence there exists a subsequence $(s_N^{k_l+1})_{l\in\mathbb{N}}$ which is bounded away from zero, which by (3.56) implies $\sigma^{k_l+1} \to 0$. This, by the definition of σ^{k_l+1} and (3.43) implies

$$\frac{s_N^{k_l+1}}{|{i'}_h^{k_l}|} \to \infty\,,$$

which by (3.46)/(v) implies $s_N^{k_l+1} \to \infty$, so that

$$\sigma^{k_l+1}s_N^{k_l+1} \geq \frac{2s_N^{k_l+1}}{1+\sqrt{1+2\gamma\frac{s_N^{k_l+1}}{a}}} \to \infty\,,$$

which is a contradiction to (3.56).

2. case: $s_N^{k+1} \to 0$. Then

$$\frac{|i_h^k - \tau^2\delta^2|}{b} \leq \frac{|i_h^k - \tau^2\delta^2|}{|{i'}_h^k|} = s_N^{k+1} \to 0,$$

hence $|i_h^k - \tau^2\delta^2| \to 0$.

As the first case led to a contradiction we can conclude

$$s_N^{k+1} \to 0 \quad \text{and} \quad |i_h^k - \tau^2\delta^2| \to 0 \qquad \text{as } k \to \infty\,, \tag{3.57}$$

which implies that k_* is finite.

(vi): From (3.57) and (3.46) it follows directly that $\sigma^k \to 1$ as $k \to \infty$.

(vii): The convergence $\beta^k \to \beta_*$ can now be shown using the Mean Value Theorem for differential calculus

$$i(\beta_*) = i(\beta^k) + i'(\xi)(\beta_* - \beta^k)$$

for some $\xi \in (\beta^k, \beta_*)$ as well as (3.38) and (3.57):

$$\begin{aligned} \underline{c}|\beta^k - \beta_*| &\leq |i(\beta^k) - i(\beta_*)| \leq |i(\beta^k) - i_h^k| + |i_h^k - i(\beta_*)| \\ &\leq (c_1 + 1)|i_h^k - \tau^2\delta^2| \ \to 0 \quad \text{as } k \to \infty\,. \end{aligned}$$

(viii): To show (3.49), we use monotonicity (3.34), (3.45) and (3.41) to conclude the lower bound

$$i_h^{k_*} \geq i(\beta^{k_*}) - |i_h^{k_*} - i(\beta^{k_*})| \geq i(\beta_*) - |i_h^{k_*} - i(\beta^{k_*})| \geq \tau^2\delta^2 - (\tau^2 - \underline{\tau})\delta^2\,,$$

while the upper bound follows directly from the definition of k_*.

□

Please note that the condition (3.35) together with (3.38) implies $i(\beta^0) \geq \tau^2\delta^2$, since there holds $i_h^0 - i(\beta^0) \leq |i(\beta^0) - i_h^0| \leq c_1(i_h^0 - \tau^2\delta^2)$, which implies $i(\beta^0) - i_h^0 \geq -c_1(i_h^0 - \tau^2\delta^2)$ and finally

$$i(\beta^0) - \tau^2\delta^2 \geq i_h^0 - \tau^2\delta^2 - c_1(i_h^0 - \tau^2\delta^2) = (1 - c_1)(i_h^0 - \tau^2\delta^2) \geq 0\,.$$

With this inital choice a start with a too large regularization parameter β is avoided. The resulting monotonicity (3.45) implies that we approach the exact parameter β_* in fact from the stable side in the sense that a smaller β corresponds to a more stable Tikhonov problem.

Knowing that the sequence $(\beta^k)_k \in \mathbb{N}$ is monotonically increasing and that $\beta^0 \leq \beta^k \leq \beta_*$ $\forall k \in \mathbb{N}$ (cf. (3.28)), tracking the proof of Theorem 3.7, we note that we don't need i to be monotonically decreasing in all $\mathbb{R}$, but that $i' < 0$ in $[\beta^0, \beta_*]$ is sufficient, which is guaranteed by Lemma 3.6.

Please note that (3.42) is only a theoretical bound for proving finiteness of the stopping index k_* and does not get active, since in practice, we stop the iteration at the index k_*. That is because we only aim at achieving (3.49) (also see (2.12)), instead of $i_h^k = \tau^2\delta^2$ for some h, k.

By means of the results of Theorem 3.7 we will prove that the convergence $\beta^k \to \beta_*$ as $k \to \infty$ is in fact linear.

Proposition 3.1. *Let the assumptions of Theorem 3.7 hold. If additionally* $2c_1 + c_2 < 1$*, we get linear convergence, i.e., for all* $C \in (\frac{c_1+c_2}{1-c_1}, 1)$ *there exists* $k_0 \in \mathbb{N}$ *such that for all* $k \geq k_0$

$$|\beta^{k+1} - \beta_*| \leq C|\beta^k - \beta_*|\,. \tag{3.58}$$

Proof. There holds

$$\begin{aligned}
|\beta^{k+1} - \beta_*| &= \beta_* - \beta^{k+1} \\
&= \beta_* - \beta^k + \sigma^{k+1}\frac{i_h^k - \tau^2\delta^2}{{i'}_h^k} \\
&= \frac{1}{{i'}_h^k}\left({i'}_h^k(\beta_* - \beta^k) + \sigma^{k+1}(i_h^k - \tau^2\delta^2)\right) \\
&= \frac{1}{{i'}_h^k}\left((1-\sigma^{k+1}){i'}_h^k(\beta_* - \beta^k) + \sigma^{k+1}(i_h^k - \tau^2\delta^2 + {i'}_h^k(\beta_* - \beta^k))\right) \\
&= \frac{1}{|{i'}_h^k|}\left((1-\sigma^{k+1})|{i'}_h^k|(\beta_* - \beta^k) + \sigma^{k+1}(\tau^2\delta^2 - i_h^k - {i'}_h^k(\beta_* - \beta^k))\right) \\
&= (1-\sigma^{k+1})(\beta_* - \beta^k) + \frac{\sigma^{k+1}}{|{i'}_h^k|}(\tau^2\delta^2 - i_h^k - {i'}_h^k(\beta_* - \beta^k))\,.
\end{aligned}$$

Using the Taylor expansion/Mean value theorem for differential calculus

$$\tau^2\delta^2 = i(\beta_*) = i(\beta^k) + i'(\xi)(\beta_* - \beta^k) \tag{3.59}$$

for some $\xi \in [\beta^k, \beta_*]$, we further get

$$\begin{aligned}
|\beta^{k+1} - \beta_*| &= (1-\sigma^{k+1})(\beta_* - \beta^k) \\
&\quad + \frac{\sigma^{k+1}}{|{i'}_h^k|}\left(i(\beta^k) + i'(\xi)(\beta_* - \beta^k) - i_h^k + (i'(\beta^k) - {i'}_h^k)(\beta_* - \beta^k) - i'(\beta^k)(\beta_* - \beta^k)\right) \\
&= (1-\sigma^{k+1})(\beta_* - \beta^k) \\
&\quad + \frac{\sigma^{k+1}}{|{i'}_h^k|}\left((i'(\xi) - i'(\beta^k))(\beta_* - \beta^k) + (i(\beta_k) - i_h^k) + (i'(\beta_k) - {i'}_h^k)(\beta_* - \beta^k)\right) \\
&= \left[1 - \sigma^{k+1} + \frac{\sigma^{k+1}}{|{i'}_h^k|}\left(i'(\xi) - i'(\beta^k) + \frac{i(\beta_k) - i_h^k}{\beta_* - \beta^k} + (i'(\beta_k) - {i'}_h^k)\right)\right](\beta_* - \beta^k)\,.
\end{aligned}$$

The term $i(\beta_k) - i_h^k$ can be estimated via (3.38) and (3.59) as follows:

$$\begin{aligned}
|i(\beta^k) - i_h^k| &\leq c_1|\tau^2\delta^2 - i_h^k| \\
&\leq c_1\left(|{i'}_h^k||\beta_* - \beta^k| + |\tau^2\delta^2 - i_h^k - {i'}_h^k(\beta_* - \beta^k)|\right) \\
&\leq c_1\left(|{i'}_h^k||\beta_* - \beta^k| + |i(\beta^k) - i_h^k| + |i'(\xi) - {i'}_h^k||\beta_* - \beta^k|\right) \\
&\leq c_1\Big(|{i'}_h^k||\beta_* - \beta^k| + |i(\beta^k) - i_h^k| + |i'(\xi) - i'(\beta^k)||\beta_* - \beta^k| \\
&\qquad + |i'(\beta^k) - {i'}_h^k||\beta_* - \beta^k|\Big)
\end{aligned}$$

and further

$$(1-c_1)|i(\beta^k)-i_h^k| \le c_1\left(|{i'}_h^k||\beta_*-\beta^k| + |i'(\xi)-i'(\beta^k)||\beta_*-\beta^k| + |i'(\beta^k)-{i'}_h^k||\beta_*-\beta^k|\right),$$

which finally gives

$$\frac{|i(\beta^k)-i_h^k|}{|\beta_*-\beta^k|} \le \frac{c_1}{1-c_1}\left(|{i'}_h^k| + |i'(\xi)-i'(\beta^k)| + |i'(\beta^k)-{i'}_h^k|\right). \tag{3.60}$$

Going back to (3.60), using (3.38), (3.39), (3.44), (3.46), (3.59), and (3.60), we have

$$\begin{aligned}
\frac{|\beta^{k+1}-\beta_*|}{|\beta^k-\beta_*|} &\le 1-\sigma^{k+1} + \frac{\sigma^{k+1}}{|{i'}_h^k|}\Big[|i'(\xi)-i'(\beta^k)| \\
&\qquad + \frac{c_1}{1-c_1}\left(|{i'}_h^k| + |i'(\xi)-i'(\beta^k)| + |i'(\beta^k)-{i'}_h^k|\right) + |i'(\beta_k)-{i'}_h^k|\Big] \\
&= 1-\sigma^{k+1} + \frac{\sigma^{k+1}}{|{i'}_h^k|}\left[\frac{1}{1-c_1}|i'(\xi)-i'(\beta^k)| + \frac{c_1}{1-c_1}|{i'}_h^k| + \frac{1}{1-c_1}|i'(\beta^k)-{i'}_h^k|\right] \\
&\le 1-\sigma^{k+1} + \frac{1}{|{i'}_h^k|}\left[\frac{1}{1-c_1}|i'(\xi)-i'(\beta^k)| + \frac{c_1+c_2}{1-c_1}|{i'}_h^k|\right] \\
&\le 1-\sigma^{k+1} + \frac{1}{a(1-c_1)}|i'(\xi)-i'(\beta^k)| + \frac{c_1+c_2}{1-c_1}.
\end{aligned} \tag{3.61}$$

Since $|i'(\xi)-i'(\beta^k)| \to 0$ and $(1-\sigma^{k+1}) \to 0$ for $k\to\infty$, for any $C \in (\frac{c_1+c_2}{1-c_1}, 1)$, there exists $k_0 \in \mathbb{N}$, such that the right-hand side of (3.61) is smaller than C for all $k \ge k_0$, which shows the linear convergence result (3.58). □

Remark 3.9. *Note that we here also recover the superlinear convergence stated in the case of exact evaluation $c_1 = c_2 = 0$ in [98] with only a lower bound on i''. Nevertheless, we will even obtain quadratic convergence for our setting (with non-exact evaluation) in the following.*

Under additional quadratic conditions on the errors in i and i' as well as the assumption that i'' is bounded from above, there even holds quadratic convergence $\beta^k \to \beta_*$, which we will show in terms of the following proposition.

Proposition 3.2. *Let the assumptions of Theorem 3.7 be satisfied. If additionally $i''(\beta) \le \overline{\gamma}$ for some $\overline{\gamma} \in \mathbb{R}$ independent of β, and*

$$|i(\beta^k)-i_h^k| \le C_1\frac{|i_h^k-\tau^2\delta^2|^2}{|{i'}_h^k|^2} \tag{3.62}$$

$$|i'(\beta^k)-{i'}_h^k| \le C_2\frac{|i_h^k-\tau^2\delta^2|}{|{i'}_h^k|} \tag{3.63}$$

holds, then we obtain the quadratic convergence estimate

$$|\beta^{k+1}-\beta_*| \le C_k|\beta^k-\beta_*|^2 \tag{3.64}$$

with

$$C_k := \frac{1}{|{i'}_h^k|}\left(\frac{\max\{0,\overline{\gamma}\}}{2} + \frac{C_1}{(1-c_1)^2}\frac{|i'(\xi)|^2}{|{i'}_h^k|^2} + \frac{2C_2+\gamma}{2(1-c_1)}\frac{|i'(\xi)|}{|{i'}_h^k|}\right)$$

for some $\xi \in (\beta^k, \beta_*)$.

Proof. The quadratic convergence estimate (3.64) can be concluded by continuing the estimate (3.59) via the Taylor expansion

$$i(\beta_*) - i(\beta^k) = -(\beta^k-\beta_*)^2\int_0^1 (1-t)i''(\beta_* + t(\beta^k-\beta_*))\,dt$$

as follows:

$$\begin{aligned}
|\beta^{k+1}-\beta_*| &= (1-\sigma^{k+1})(\beta_*-\beta^k) + \frac{\sigma^{k+1}}{|{i'}_h^k|}(i(\beta_*) - i_h^k - {i'}_h^k(\beta_*-\beta^k))\\
&= (1-\sigma^{k+1})(\beta_*-\beta^k)\\
&\quad + \frac{\sigma^{k+1}}{|{i'}_h^k|}\Big(i(\beta_*) - i(\beta^k) + (i(\beta^k) - i_h^k) - (i'(\beta^k) - {i'}_h^k)(\beta_*-\beta^k) - i'(\beta^k)(\beta_*-\beta^k)\Big)\\
&= (1-\sigma^{k+1})(\beta_*-\beta^k) + \frac{\sigma^{k+1}}{|{i'}_h^k|}\Big(-\int_0^1(1-t)i''(\beta_*+t(\beta^k-\beta_*))\,dt\,(\beta^k-\beta_*)^2\\
&\quad +(i(\beta^k) - i_h^k) - (i'(\beta^k) - {i'}_h^k)(\beta_*-\beta^k) - i'(\beta^k)(\beta_*-\beta^k)\Big)\\
&= (1-\sigma^{k+1})(\beta_*-\beta^k) + \frac{\sigma^{k+1}}{|{i'}_h^k|}\Big(\frac{1}{2}\max\{0,\overline{\gamma}\}(\beta_*-\beta^k)^2\\
&\quad +(i(\beta^k) - i_h^k) - (i'(\beta^k) - {i'}_h^k)(\beta_*-\beta^k) - i'(\beta^k)(\beta_*-\beta^k)\Big)\\
&\le (1-\sigma^{k+1})(\beta_*-\beta^k) + \frac{\sigma^{k+1}}{|{i'}_h^k|}\Big(\frac{1}{2}\max\{0,\overline{\gamma}\}(\beta_*-\beta^k)^2\\
&\quad +|i(\beta^k) - i_h^k| + |i'(\beta^k) - {i'}_h^k|(\beta_*-\beta^k)\Big)\\
&\le \frac{1}{|{i'}_h^k|}\Bigg[|{i'}_h^k|\left(1 - \frac{2}{1+\sqrt{1-2\gamma s_N^{k+1}/{i'}_h^k}}\right)(\beta_*-\beta^k) + \frac{1}{2}\max\{0,\overline{\gamma}\}(\beta_*-\beta^k)^2\\
&\quad +|i(\beta^k) - i_h^k| + |i'(\beta^k) - {i'}_h^k|(\beta_*-\beta^k)\Bigg]\,.
\end{aligned} \tag{3.65}$$

It can easily be shown that $\frac{1-\frac{1}{2}\lambda}{1+\frac{1}{2}\lambda}\sqrt{1+2\lambda} \le 1\ \forall\lambda \ge 0$, which implies $1 - \frac{2}{1+\sqrt{1+2\lambda}} \le \frac{1}{2}\lambda\ \forall\lambda\ge 0$ and therefore

$$|{i'}_h^k|\left(1 - \frac{2}{1+\sqrt{1+2\gamma s_N^{k+1}/|{i'}_h^k|}}\right) \le \frac{\gamma}{2}s_N^{k+1} = \frac{\gamma}{2}\frac{|i_h^k - \tau^2\delta^2|}{|{i'}_h^k|}\,. \tag{3.66}$$

Besides, (with (3.59) there holds

$$|i_h^k - \tau^2\delta^2| \le \frac{1}{1-c_1}|i(\beta^k) - \tau^2\delta^2| \le \frac{1}{1-c_1}|i'(\xi)||\beta^k - \beta_*| \tag{3.67}$$

for some $\xi \in (\beta^k, \beta_*)$. With the estimates (3.66), (3.67) and using (3.62), (3.63) we can continue (3.65) as

$$
\begin{aligned}
|\beta^{k+1} - \beta_*| \leq \frac{1}{|{i'}^k_h|} \Bigg[& \frac{\gamma}{2} \frac{|i^k_h - \tau^2\delta^2|}{|{i'}^k_h|}(\beta_* - \beta^k) + \frac{1}{2}\max\{0, \overline{\gamma}\}(\beta_* - \beta^k)^2 \\
& + C_1 \frac{|i^k_h - \tau^2\delta^2|^2}{|{i'}^k_h|^2} + C_2 \frac{|i^k_h - \tau^2\delta^2|}{|{i'}^k_h|}(\beta_* - \beta^k) \Bigg] \\
= \frac{1}{|{i'}^k_h|} \Bigg[& \left(\frac{\gamma}{2} + C_2\right) \frac{|i^k_h - \tau^2\delta^2|}{|{i'}^k_h|}(\beta_* - \beta^k) + \frac{1}{2}\max\{0, \overline{\gamma}\}(\beta_* - \beta^k)^2 \\
& + C_1 \frac{|i^k_h - \tau^2\delta^2|^2}{|{i'}^k_h|^2} \Bigg] \\
\leq \frac{1}{|{i'}^k_h|} \Bigg[& \frac{\frac{\gamma}{2} + C_2}{1 - c_1} \frac{|i'(\xi)|}{|{i'}^k_h|} + \frac{1}{2}\max\{0, \overline{\gamma}\} + \frac{C_1}{(1 - c_1)^2} \frac{|i'(\xi)|^2}{|{i'}^k_h|^2} \Bigg] |\beta^k - \beta_*|^2
\end{aligned}
$$

for some $\xi \in (\beta^k, \beta_*)$. $\square$

Note that discrete versions $i'_h(\beta)$ and $i''_h(\beta)$ of $i'(\beta)$ and $i''(\beta)$ and therewith also a posteriori estimates for the lower and upper bounds on $i''_h \approx i''(\beta)$ can be computed with low numerical effort, once the error estimator for i' has been evaluated (see Section 3.3 and [39]).

In case upper and lower bounds on i'' are not known, bisection still provides a globally and R-linearly convergent method for obtaining β such that (3.49) holds.

To summarize the results from this section, for obtaining convergence of q^δ_β according to Theorem 3.1 and Theorem 3.4 and linear (and quadratic) convergence of the sequence of β's produced by the proposed inexact Newton method according to Proposition 3.1 and 3.2, we check if (3.9), (3.19), (3.38), (3.39), (3.40), (3.41) (and (3.62), (3.63)) hold in our search for the correct regularization parameter and refine the discretization according to the corresponding error estimator if one of these conditions is violated. Note that these bounds allow for a rather rough approximation as long as the regularization parameter β^k under consideration is still "far away" from the actual one in the sense that $i(\beta^k) - \tau^2\delta^2$ is not small yet.

We also wish to point out that the conditions (3.9), (3.19), (3.38), (3.39), (3.40), (3.41), (3.62), (3.63) are tailored to the use with goal oriented error estimators (see the next section). Indeed, adaptive refinement according to these estimators allows to enforce error bounds on terms of the type $I(q, F(q)) - I(q_h, F_h(q_h))$, whereas terms of the type $I(q, F(q)) - I(q_h, F(q_h))$ or $I(q, F(q)) - I(q, F_h(q))$ would not be directly tractable with this technique.

Vice versa, the same reasoning encourages the use of the inexact Newton method: Although we consider the scalar equation $i(\beta_*) = \tau^2\delta^2$, for which methods of first order (e.g., bisection) might seem to be the means of choice, the fact that second derivatives of i have to be computed anyway in the course of error estimation via Newton's method, we obtain better convergence without putting much more effort into the computations during the optimization process.

3.3. Use of goal oriented error estimators

For evaluating the error estimates needed for the determination of β according to Theorem 3.7 we employ the goal oriented error estimators (cf. Section 2.4) proposed in [11] and [12]. We make use of the results from [39], where concrete error estimators for j, i and i' and computation formulas for i' and i'' are derived.

We begin with an error estimator for j, which is needed to estimate (3.19). According to [39, Proposition 1] for continuous and discrete stationary points $x = (q, u, z) \in X = Q \times V \times W$ (cf. (2.28)) and $x_h = (q_h.u_h) \in X_h = Q_h \times V_h \times W_h$ (cf. (2.38)) of $\mathcal{L}$ and $\mathcal{L}_h$ respectively, an error representation is given by

$$j_\beta(q) - j_{\beta,h}(q_h) = J_\beta(q, u) - J_\beta(q_h, u_h) = \frac{1}{2}\mathcal{L}'(x_h)(x - \widehat{x}_h) + \mathbf{R}_1 \tag{3.68}$$

for arbitrary $\widehat{x}_h = (\widehat{q}_h, \widehat{u}_h, \widehat{z}_h) \in X_h$, where $\mathbf{R}_1$ is a third order remainder term.

For the purpose of formulating an error estimator for i (needed in (3.3), (3.38), (3.40), (3.41), (3.62) and (3.9)), we define the auxiliary functional

$$\mathcal{M}\colon X^2 \to \mathbb{R}, \quad \mathcal{M}(x, x_1) \coloneqq I(u) + \mathcal{L}'(x)(x_1)$$

(with I defined by (2.19)), where $x = (q, u, z)$ and $x_1 = (q_1, u_1, z_1)$.

Let $x = (q, u, z)$ and $x_h = (q_h, u_h, z_h)$ solve (2.28) and (2.38) respectively. According to [39, Proposition 3] we get continuous and discrete stationary points $y =: (x, x_1)$ and $y_h \coloneqq (x_h, x_{1,h}) = (q_h, u_h, z_h, q_{1,h}, u_{h,1}, z_{h,1})$ of $\mathcal{M}$ by solving

$$\begin{aligned} \mathcal{L}''(x)(\delta x, x_1) &= -I'(u)(\delta u) && \forall \delta x = (\delta q, \delta u, \delta z) \in X\,, \\ \mathcal{L}''(x_h)(\delta x_h, x_{1,h}) &= -I'(u_h)(\delta u_h) && \forall \delta x_h = (\delta q_h, \delta u_h, \delta z_h) \in X_h\,. \end{aligned} \tag{3.69}$$

For such stationary points the error representation

$$I(u) - I(u_h) = i(\beta) - i_h(\beta) = \frac{1}{2}\mathcal{M}'(y_h)(y - \widehat{y}_h) + \mathbf{R}_2 \tag{3.70}$$

holds for arbitrary $\widehat{y}_h \in X_h \times X_h$, where $\mathbf{R}_2$ is a third order remainder term.

By means of (x, x_1) and $(x_h, x_{1,h})$ the continuous and discrete version of the first derivative i' can be evaluated by

$$i'(\beta) = \frac{2}{\beta^2}(q, q_1)_Q \quad \text{and} \quad i_h'(\beta) = \frac{2}{\beta^2}(q_h, q_{1,h})_Q\,, \tag{3.71}$$

see [39, Proposition 4].

We define an additional auxiliary functional in order to express an error estimator for i' (needed in (3.39), (3.40), (3.63))

$$\mathcal{N}\colon X^2 \times X^2 \to \mathbb{R}, \quad \mathcal{N}(y, \mathbf{y}) \coloneqq K(q, q_1) + \mathcal{M}_x'(x, x_1)(x_2) + \mathcal{M}_{x_1}'(x, x_1)(x_3)\,,$$

where $x_1 = (q_1, u_1, z_1)$, $x_2 = (q_2, u_2, z_2)$, $x_3 = (q_3, u_3, z_3)$, $y = (x, x_1)$, $\mathbf{y} = (x_2, x_3)$, and K is defined by

$$K(q, q_1) = i'(\beta) = -\frac{2}{\beta^2}(q, q_1)_Q\,.$$

According to [39, Proposition 6] we obtain continuous and stationary points $(x, x_1, x_2, x_3) \in X^4$ and $(x_h, x_{1,h}, x_{2,h}, x_{3,h}) \in X_h^4$ of $\mathcal{N}$ by solving

$$\begin{aligned} \mathcal{L}''(x)(x_2, \delta x_1) &= -K'_{q_1}(q, q_1)(\delta q_1) \qquad \forall \delta x_1 = (\delta q_1, \delta u_1, \delta z_1) \in X\,, \\ \mathcal{L}''(x)(x_3, \delta x) &= -K'_q(q, q_1)(\delta q) - I''_{uu}(q, u)(u_2, \delta u) - \mathcal{L}'''(x)(x_1, x_2, \delta x) \\ &\qquad \forall \delta x = (\delta q, \delta u, \delta z) \in X\,, \end{aligned} \tag{3.72}$$

and

$$\begin{aligned} \mathcal{L}''(x_h)(x_{2,h}, \delta x_{1,h}) &= -K'_{q_{1,h}}(q_h, q_{1,h})(\delta q_{1,h}) \qquad \forall \delta x_{1,h} = (\delta q_{1,h}, \delta u_{1,h}, \delta z_{1,h}) \in X_h\,, \\ \mathcal{L}''(x_h)(x_{3,h}, \delta x_h) &= -K'_q(q_h, q_{1,h})(\delta q_h) - I''_{uu}(q_h, u_h)(u_{2,h}, \delta u_h) - \mathcal{L}'''(x_h)(x_{1,h}, x_{2,h}, \delta x_h) \\ &\qquad \forall \delta x_h = (\delta q_h, \delta u_h, \delta z_h) \in X_h\,. \end{aligned} \tag{3.73}$$

For these stationary points there holds the error representation

$$K(q, q_1) - K(q_h, q_{1,h}) = i'(\beta) - i'_h(\beta) = \frac{1}{2}\mathcal{N}'_y(y_h, \mathbf{y}_h)(y - \widehat{y}_h) + \frac{1}{2}\mathcal{N}'_{\mathbf{y}}(y_h, \mathbf{y}_h)(\mathbf{y} - \widehat{\mathbf{y}}_h) + \mathbf{R}_3 \tag{3.74}$$

for arbitrary $\widehat{y}_h, \widehat{\mathbf{y}}_h \in X_h^2$, where $\mathbf{R}_3$ is a third order remainder term, cf. [39, Proposition 5].

Finally, the second derivative i'' and its discrete equivalent can be evaluated by

$$i''(\beta) = \frac{4}{\beta^3}(q, q_1)_Q - \frac{2}{\beta^2}(q_2, q_1)_Q - \frac{2}{\beta^2}(q, q_3)_Q$$

and

$$i''_h(\beta) = \frac{4}{\beta^3}(q_h, q_{1,h})_Q - \frac{2}{\beta^2}(q_{2,h}, q_{1,h})_Q - \frac{2}{\beta^2}(q_h, q_{3,h})_Q\,, \tag{3.75}$$

see [39, Proposition 7].

The error estimators η^J, η^I, and η^K for J_β, I, K are computed by

$$\eta^J = \frac{1}{2}\mathcal{L}'(x_h)(e_h^x)\,, \quad \eta^I = \frac{1}{2}\mathcal{M}'(y_h)(e_h^y) \quad \text{and} \quad \eta^K = \frac{1}{2}\mathcal{N}'_y(y_h, \mathbf{y}_h)(e_h^y) + \frac{1}{2}\mathcal{N}'_{\mathbf{y}}(y_h, \mathbf{y}_h)(e_h^{\mathbf{y}})\,, \tag{3.76}$$

where e_h^x, e_h^y and $e_h^{\mathbf{y}}$ are approximations of interpolation errors (see Section 2.4) obtained by local averaging or higher order approximations, cf., e.g., [85]. When using spaces Q_h, V_h, W_h with locally supported basis functions, the estimators can be written as a sum of local contributions, which enables to implement a local refinement strategy based on the estimators, cf. [11]. Additionally to that, each local error can be decomposed into its components due to the discretization of Q on the one hand and of V on the other hand. Therewith the proposed method for determining β could also be applied when using different discretizations of Q and V. Then in each iteration step it can be decided whether to refine Q_h or V_h (or W_h), if necessary, cf. [85].

For further details on the evaluation of the error estimators we refer to [11, 39].

Remark 3.10. *Although these error estimators are known to work efficiently in practice (see, e.g., [11, 14, 39]), they are not reliable, i.e., the conditions* (3.38)–(3.40) *cannot be guaranteed in a strict sense in our computations since we have to neglect the remainder terms* $\mathbf{R}_1$, $\mathbf{R}_2$, *and* $\mathbf{R}_3$ *and use an approximation for* $x - \widehat{x}_h$, $y - \widehat{y}_h$, *and* $\mathbf{y} - \widehat{\mathbf{y}}_h$. *Since our analysis is kept rather general, it is not restricted to dual-weighted-residual error estimators. However, since they are based on residuals which are computed in the optimization process, the additional costs for estimation are very low, which makes the DWR error estimators tailored for our purposes.*

3.4. Test examples and verification of assumptions

For illustrating the performance of the presented method, we consider the following parameter/coefficient identification sample PDEs:

Let $\zeta \in \mathbb{R}^+$, $f \in L^2(\Omega)$, and Ω be a bounded smooth or polygonal and convex domain in $\mathbb{R}^d$, $d \in \{1, 2, 3\}$, and $C = \mathrm{id}$. The control space is chosen as $Q = L^2(\Omega)$ and the state and test spaces are $V = W = H_0^1(\Omega)$ in case of homogeneous Dirichlet boundary data and $V = W = H^1(\Omega)$ in case of homogeneous Neumann boundary data. The observation space is chosen as $G = L^2(\Omega)$.

Example 3.1. *For* $q \in L^2(\Omega)$ *find* $u \in H_0^1(\Omega)$ *such that*

$$\begin{cases} -\Delta u + \zeta u^3 = q & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega\,. \end{cases}$$

Example 3.2. *For* $q \in L^2(\Omega)$, $q \geq 0$ *almost everywhere in* Ω *find* $u \in H_0^1(\Omega)$ *such that*

$$\begin{cases} -\Delta u + \zeta q u = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega\,. \end{cases}$$

Example 3.3. *For* $q \in L^2(\Omega)$, $q \geq 0$ *almost everywhere in* Ω *find* $u \in H_0^1(\Omega)$ *such that*

$$\begin{cases} -\Delta u - \zeta(1+q)u = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega\,. \end{cases}$$

This example is also called "Helmholtz equation" with homogeneous Dirichlet boundary data.

Example 3.4. *For* $q \in L^2(\Omega)$, $q \geq 0$ *almost everywhere in* Ω *find* $u \in H^1(\Omega)$ *such that*

$$\begin{cases} -\Delta u - \zeta(1+q)u = f & \text{in } \Omega \\ \partial_n u = 0 & \text{on } \partial\Omega\,. \end{cases}$$

This example is also called "Helmholtz equation" with homogeneous Neumann boundary data.

Throughout this section let $(\cdot,\cdot) = (\cdot,\cdot)_{L^2(\Omega)}$ denote the inner product in $L^2(\Omega)$, i.e., $(\varphi, \psi) := \int_\Omega \varphi(x)\psi(x)\,dx$. Concerning the notation of scalar products (in $L^2(\Omega)$, $H^1(\Omega)$, etc.) we make no difference whether we deal with vector valued functions or not, for instance, $(\varphi, \psi) = (\varphi, \psi)_{L^2(\Omega)} = (\varphi, \psi)_{L^2(\Omega)^d} := \int_\Omega (\varphi(x), \psi(x))_{\mathbb{R}^d}\,dx$, where $(\cdot,\cdot)_{\mathbb{R}^d}$ denotes the euclidian scalar product.

Remark 3.11. *Note that, since $W = H^1(\Omega)$ is densely embedded in $L^2(\Omega)$, the duality pairing $\langle \cdot, \cdot \rangle_{W^*,W}$ can be interpreted as an inner product in the space $L^2(\Omega)$ in the following sense: Let inj denote the injection from W to $L^2(\Omega)$. Then its adjoint inj^* is the injection from $L^2(\Omega)^* = L^2(\Omega)$ to W^*. Let further $v \in W \subset L^2(\Omega)$. Then every element $v \in L^2(\Omega)$ can be understood as linear continuous functional from W to $\mathbb{R}$, in the sense that*

$$\langle inj^*(v), \varphi \rangle_{W^*,W} = (v, inj(\varphi)) \qquad \forall \varphi \in W\,.$$

Thus, for $f \in L^2(\Omega)$ there exists an element in W^*, which we also denote by f, such that $f(\varphi) = \langle f, \varphi \rangle_{W^*,W} = (f, \varphi)$ for all $\varphi \in W$.

Before taking a look at the results from applying the proposed method (cf. Algorithm 3.1 in the subsequent section) to these sample equations, we will discuss whether the imposed conditions on F of the last sections are satisfied for the considered examples. We will analyse whether

(i) the assumption that F is well–defined and bounded,

(ii) the weak sequentially closedness condition Assumption 3.2 (or (2.5)),

(iii) the tangential cone condition Assumption 3.3

are fulfilled for the considered examples.

We are well aware of the fact that Lemma 3.6 (and implicitly also Theorem 3.7) additionally requires Assumption 3.6, which represents the sufficient optimality condition of second order and is known to be hard to verify. We refer to [65], where the condition is discussed at least for Example 3.1.

We will see that the tangential cone condition Assumption 3.3 can only be assured if the domain $\mathcal{D}$ is sufficiently small. To put this in perspective, please note that provided that global minimizers (instead of stationary points) are available, the tangential cone condition Assumption 3.3 is only needed in Lemma 3.6 and for showing convergence rates (cf. Remark 3.3 and 3.6).

1. **Example 3.1:** We consider the weak formulation

$$(\nabla u, \nabla \varphi) + \zeta(u^3, \varphi) = (q, \varphi) \quad \forall \varphi \in H_0^1(\Omega)\,. \tag{3.77}$$

(i): Well–definedness and boundedness of F as a mapping from $L^2(\Omega)$ to $H_0^1(\Omega)$ together with the estimate

$$\|\nabla u\|_{L^2(\Omega)} + \|u\|_{C(\bar{\Omega})} \leq C\, \|q\|_{L^2(\Omega)} \tag{3.78}$$

for a constant $C > 0$ follows directly from [108, Theorem 4.7], for instance.

(ii): To verify Assumption 3.2, we consider an arbitrary sequence $(q_n)_{n\in\mathbb{N}} \subseteq \mathcal{D}$ with $q_n \rightharpoonup q$ in $L^2(\Omega)$. This implies that $(q_n)_{n\in\mathbb{N}}$ is bounded in $L^2(\Omega)$ and because of (3.78) also that $(u_n)_{n\in\mathbb{N}}$ is bounded in $H_0^1(\Omega) \cap C(\bar{\Omega})$. This implies that there exists a weakly convergence subsequence, which we also denote by $(u_n)_{n\in\mathbb{N}}$ with $u_n \rightharpoonup g$ in $H_0^1(\Omega)$ and that $\|u_n\|_{L^\infty(\Omega)} \leq M$ for some constant $M > 0$. Using the compact embedding of $H_0^1(\Omega)$ into $L^2(\Omega)$ we get $u_n \to g$ in $L^2(\Omega)$. Since the set $\{w \in L^2(\Omega)|\ \|w\|_{L^\infty(\Omega)} \le M\}$ is closed in $L^2(\Omega)$ we also have $\|g\|_{L^\infty(\Omega)} \le M$. Then, according to [108, Lemma 4.11], there exists a constant $L(M) > 0$ such that $\|u_n^3 - g^3\|_{L^2(\Omega)} \le L(M)\|u_n - g\|_{L^2(\Omega)}$. Finally, we get $0 = (\nabla u_n, \nabla\varphi) + \zeta(u_n^3, \varphi) - (q_n, \varphi) \to (\nabla g, \nabla\varphi) + \zeta(g^3, \varphi) - (q, \varphi)$ as $n \to \infty$, which implies $F(q) = g$. If we choose $\mathcal{D} \subset Q$ as a weakly sequentially closed set we also get $q \in \mathcal{D}$.

(iii): First of all we state that the necessary smoothness of F, a and all other upcoming semilinear forms follows from [108, Section 4.3].

Let $q, \bar{q} \in \mathcal{D}$. It can easily be checked that $v = F(q) - F(\bar{q}) - F'(\bar{q})(q - \bar{q})$ is a weak solution to the PDE

$$\begin{cases} -\Delta v + 3\zeta F(q)^2 v = \zeta(F(\bar{q}) + 2F(q))(F(\bar{q}) - F(q))^2 & \text{in } \Omega \\ v = 0 & \text{on } \partial\Omega\,. \end{cases}$$

Using the Hölder inequality, Sobolev's embedding theorem

$$\|\varphi\|_{L^p(\Omega)} \le c_s \|\varphi\|_{H^1(\Omega)} \quad \text{for all } p \le 6 \text{ for some constant } c_s > 0 \tag{3.79}$$

and Poincaré's inequality

$$\|\varphi\|_{L^2(\Omega)} \le c_p \|\nabla\varphi\|_{L^2(\Omega)} \tag{3.80}$$

for all $\varphi \in H_0^1(\Omega)$, we get

$$\begin{aligned}
&\|\nabla[F(q) - F(\bar{q}) - F'(\bar{q})(q - \bar{q})]\|_{L^2(\Omega)}^2 \\
&\quad = \zeta((F(\bar{q}) + 2F(q))(F(\bar{q}) - F(q))^2, F(q) - F(\bar{q}) - F'(\bar{q})(q - \bar{q})) \\
&\qquad - 3\zeta\|F(q)(F(q) - F(\bar{q}) - F'(\bar{q})(q - \bar{q}))\|_{L^2(\Omega)}^2 \\
&\quad \le \zeta\|F(\bar{q}) + 2F(q)\|_{L^6(\Omega)}\|F(\bar{q}) - F(q)\|_{L^6(\Omega)}\|F(\bar{q}) - F(q)\|_{L^2(\Omega)} \\
&\qquad \cdot \|F(q) - F(\bar{q}) - F'(\bar{q})(q - \bar{q})\|_{L^6(\Omega)} \\
&\quad \le \zeta c_{sp}^3\|\nabla[F(\bar{q}) + 2F(q)]\|_{L^2(\Omega)}\|\nabla[F(\bar{q}) - F(q)]\|_{L^2(\Omega)}\|F(\bar{q}) - F(q)\|_{L^2(\Omega)} \\
&\qquad \cdot \|\nabla[F(q) - F(\bar{q}) - F'(\bar{q})(q - \bar{q})]\|_{L^2(\Omega)}\,,
\end{aligned}$$

where $c_{sp} := c_s\sqrt{1 + c_p^2}$.

Division by $\|\nabla[F(q) - F(\bar{q}) - F'(\bar{q})(q - \bar{q})]\|_{L^2(\Omega)}$ leads to

$$\begin{aligned}
&\|\nabla[F(q) - F(\bar{q}) - F'(\bar{q})(q - \bar{q})]\|_{L^2(\Omega)} \\
&\quad \le \zeta c_{sp}^3\|\nabla[F(\bar{q}) + 2F(q)]\|_{L^2(\Omega)}\|\nabla[F(\bar{q}) - F(q)]\|_{L^2(\Omega)}\|F(\bar{q}) - F(q)\|_{L^2(\Omega)} \\
&\quad \le \zeta c_{sp}^3\left(\|\nabla F(\bar{q})\|_{L^2(\Omega)} + 2\|\nabla F(q)\|_{L^2(\Omega)}\right)\|\nabla[F(\bar{q}) - F(q)]\|_{L^2(\Omega)} \\
&\qquad \cdot \|F(\bar{q}) - F(q)\|_{L^2(\Omega)}\,.
\end{aligned} \tag{3.81}$$

Let $\bar{u} = F(\bar{q})$ and $u = F(q)$. It can easily be shown that $w = F(q) - F(\bar{q})$ is a weak solution to

$$\begin{cases} -\Delta w + \frac{1}{2}\zeta\left(\bar{u}^2 + u^2 + (\bar{u}+u)^2\right) w = q - \bar{q} & \text{in } \Omega \\ w = 0 & \text{on } \partial\Omega \,, \end{cases}$$

which yields the estimate

$$\begin{aligned} \|\nabla[F(q) - F(\bar{q})]\|^2_{L^2(\Omega)} + \tfrac{1}{2}\zeta \|\sqrt{F(\bar{q})^2 + F(q)^2 + (F(\bar{q}) + F(q))^2}(F(q) - F(\bar{q}))\|^2_{L^2(\Omega)} \\ \leq c_p \|\nabla[F(q) - F(\bar{q})]\|_{L^2(\Omega)} \|q - \bar{q}\|_{L^2(\Omega)} \,. \end{aligned}$$

Dividing this by $\|\nabla[F(q) - F(\bar{q})]\|_{L^2(\Omega)}$ we get

$$\|\nabla[F(q) - F(\bar{q})]\|_{L^2(\Omega)} \leq c_p \|q - \bar{q}\|_{L^2(\Omega)} \,.$$

Inserting this into (3.81) as well as using the fact that $F(q)$ (and $F(\bar{q})$) solve (3.77) (with q replaced by $\bar{q}$), we finally arrive at

$$\begin{aligned} \|F(q) - F(\bar{q}) - F'(\bar{q})(q - \bar{q})\|_{L^2(\Omega)} &\leq c_p \|\nabla[F(q) - F(\bar{q}) - F'(\bar{q})(q - \bar{q})]\|_{L^2(\Omega)} \\ &\leq \zeta c_{sp}^3 c_p C \left(\|\bar{q}\|_{L^2(\Omega)} + 2\|q\|_{L^2(\Omega)} \right) \|q - \bar{q}\|_{L^2(\Omega)} \\ &\quad \cdot \|F(\bar{q}) - F(q)\|_{L^2(\Omega)} \,, \end{aligned}$$

which implies Assumption 3.3 with small c_{tc} upon restriction of the domain $\mathcal{D}$ to a sufficiently small neighborhood of a solution $q^\dagger$.

2. **Example 3.2:** The validity of the claimed assumptions, except for Assumption 3.6, has been discussed by Kaltenbacher, Schöpfer, and Schuster in [71] in a Banach space setting. We will follow their approach for showing the same results for the choice $Q = L^2(\Omega)$, $V = W = H_0^1(\Omega)$, which has not been considered in [71].

   Let $\mathcal{D} = \{q \in Q = L^2(\Omega) |\; q \geq 0 \text{ a.e. in } \Omega\}$. The weak formulation reads

$$a(q,u)(\varphi) := (\nabla u, \nabla\varphi) + \zeta(qu, \varphi) = (f, \varphi) \quad \forall \varphi \in H_0^1(\Omega) \,. \tag{3.82}$$

   (i): Since with (3.79) we have

$$\begin{aligned} |a(q,u)(\varphi)| &\leq \|\nabla u\|_{L^2(\Omega)} \|\nabla\varphi\|_{L^2(\Omega)} + \zeta \|q\|_{L^2(\Omega)} \|u\|_{L^4(\Omega)} \|\varphi\|_{L^4(\Omega)} \\ &\leq \|\nabla u\|_{L^2(\Omega)} \|\nabla\varphi\|_{L^2(\Omega)} + \zeta \|q\|_{L^2(\Omega)} c_{sp}^2 \|\nabla u\|_{L^2(\Omega)} \|\nabla\varphi\|_{L^2(\Omega)} \\ &\leq (1 + c_{sp}^2 \zeta \|q\|_{L^2(\Omega)}) \|\nabla u\|_{L^2(\Omega)} \|\nabla\varphi\|_{L^2(\Omega)} \end{aligned}$$

   and

$$a(q,u)(u) = \|\nabla u\|^2_{L^2(\Omega)} + \zeta(qu, u) \geq \|\nabla u\|^2_{L^2(\Omega)}$$

   for all $q \in \mathcal{D}$, the semilinear form a is continuous and coercive, so that F is well–defined and bounded from $\mathcal{D}$ to $H_0^1(\Omega)$.

(ii): To verify Assumption 3.2, by the same argumentation as in the previous example, we get that for any sequence $(q_n)_{n\in\mathbb{N}} \subseteq \mathcal{D}$ with $q_n \rightharpoonup q$ in $L^2(\Omega)$ and $u_n := F(q_n) \to g$ in $L^2(\Omega)$ even $u_n \rightharpoonup g$ in $H^1(\Omega)$ and $u_n \to g$ in $L^4(\Omega)$ holds. Therewith, $g \in H_0^1(\Omega)$ and for any $\varphi \in H_0^1(\Omega)$ we have

$$\lim_{n\to\infty}(q_n(u_n - g), \varphi) \le \lim_{n\to\infty} \|q_n\|_{L^2(\Omega)} \|u_n - g\|_{L^4(\Omega)} \|\varphi\|_{L^4(\Omega)} = 0$$

and consequently

$$\begin{aligned}
&\lim_{n\to\infty}(\nabla u_n, \nabla\varphi) + \zeta(q_n u_n, \varphi) - (f, \varphi) \\
&= \lim_{n\to\infty}(\nabla u_n, \nabla\varphi) + \zeta(q_n(u_n - g), \varphi) + \zeta(q_n, g\varphi) - (f, \varphi) \\
&= \lim_{n\to\infty}(\nabla u_n, \nabla\varphi) + \zeta(q_n, g\varphi) - (f, \varphi) \\
&= (\nabla g, \nabla\varphi) + \zeta(q, g\varphi) - (f, \varphi) \\
&= 0\,,
\end{aligned}$$

i.e., $F(q) = g$, which implies Assumption 3.2.

(iii): The tangential cone condition has already been investigated by Hanke, Neubauer, and Scherzer in [44] for our example, but for the choice

$$\mathcal{D} = \{q \in L^2(\Omega)|\ \|q - q^\dagger\|_{L^2(\Omega)} \le \varepsilon \text{ and } q^\dagger \ge 0 \text{ almost everywhere in } \Omega\}\,,$$

which does not allow our argumentation in (i). In fact, well-definedness of F from $L^2(\Omega)$ to $H^2(\Omega)$ is shown by Colonius and Kunisch in [74], but only in 1D.

First we show that the PDE posesses a unique solution in $H^2(\Omega)$. According to [40, Theorem 4.4.3.7] (in 2D) and [83, Theorem 4.3.2] (in 3D) a solution to the Poisson equation

$$\begin{cases} -\Delta u = \tilde{f} & \text{in } \Omega \\ \quad\ u = 0 & \text{on } \partial\Omega\,. \end{cases}$$

is unique in $W^{2,p}(\Omega)$ with $\|u\|_{W^{2,p}(\Omega)} \le c_p\|\tilde{f}\|_{L^p}$, if there exists a $p_\Omega > 2$, such that the right-hand side $\tilde{f}$ lies in $L^p(\Omega)$ for all $1 < p < p_\Omega$. Thus, we have to show that $f - qu \in L^2(\Omega)$. We do this by a boot strapping argument: Since $u \in H_0^1(\Omega)$, which is continuously embedded in $L^6(\Omega)$ and $q \in L^2(\Omega)$, we get $qu \in L^{3/2}(\Omega)$ by Hölder's inequality. This implies $u \in W^{2,3/2}(\Omega)$. Using the continuous embedding from $W^{2,3/2}(\Omega)$ into $L^k(\Omega)$ for all $1 \le k < \infty$, i.e., in particular into $L^7(\Omega)$ we get $qu \in L^{14/9}$ and therewith $u \in W^{2,14/9}(\Omega)$ by the same argumentation. Since $\frac{14}{9} > \frac{3}{2}$, $W^{2,14/9}(\Omega)$ is continuously embedded in $L^\infty(\Omega)$. Finally we get that $qu \in L^2(\Omega)$ and therewith $u \in W^{2,2}(\Omega) = H^2(\Omega)$.

Let $q, \bar{q} \in \mathcal{D}$. With the abbreviations $u = F(q)$, $\bar{u} := F(\bar{q})$, $w := F'(q)(\bar{q} - q)$, w can be written as a weak solution of the tangent equation (2.29)

$$\begin{cases} -\Delta w + \zeta q w = -\zeta(\bar{q} - q)u & \text{in } \Omega \\ \qquad\qquad w = 0 & \text{on } \partial\Omega\,. \end{cases}$$

Since

$$\begin{aligned}
(\nabla(\bar{u}-u-w),\nabla\varphi)+\zeta(q(\bar{u}-u-w),\varphi)&\\
&=(\nabla\bar{u},\nabla\varphi)+\zeta(\bar{q}\bar{u},\varphi)+\zeta((q-\bar{q})\bar{u},\varphi)\\
&\qquad-[(\nabla u,\nabla\varphi)+\zeta(qu,\varphi)]-[(\nabla w,\nabla\varphi)+\zeta(qw,\varphi)]\\
&=\zeta((q-\bar{q})\bar{u},\varphi)-\zeta((q-\bar{q})u,\varphi)\\
&=-\zeta'((\bar{q}-q)(\bar{u}-u),\varphi)
\end{aligned}$$

for all $\varphi\in H_0^1(\Omega)$, $v=F(\bar{q})-F(q)-F'(q)(\bar{q}-q)=\bar{u}-u-w$ solves

$$\begin{cases}-\Delta v+\zeta qv=-\zeta(\bar{q}-q)(\bar{u}-u) & \text{in } \Omega\\ \qquad\qquad v=0 & \text{on } \partial\Omega\,.\end{cases}\tag{3.83}$$

With the same argumentation as before, we have $\bar{u},w,v\in H^2(\Omega)$. We introduce the dual variable y as solution to

$$\begin{cases}-\Delta y+\zeta qy=v & \text{in } \Omega\\ \qquad\qquad y=0 & \text{on } \partial\Omega\,,\end{cases}\tag{3.84}$$

which exists in $H^2(\Omega)$ with

$$\|y\|_{H^2(\Omega)}\le c\|v\|_{L^2(\Omega)}$$

for some constant $c>0$, again with the same argumentation. Testing (3.84) with v yields

$$\|v\|^2_{L^2(\Omega)}=(\nabla y,\nabla v)+\zeta(qy,v)\,,$$

and testing (3.83) with y yields

$$(\nabla v,\nabla y)+\zeta(qv,y)=(-\zeta(\bar{q}-q)(\bar{u}-u),y)\,.$$

Putting this together we get

$$\begin{aligned}
\|v\|^2_{L^2(\Omega)}&=-\zeta((\bar{q}-q)(\bar{u}-u),y)\\
&\le\zeta\|\bar{q}-q\|_{L^2(\Omega)}\|\bar{u}-u\|_{L^2(\Omega)}\|y\|_{L^\infty}\\
&\le\zeta C\|\bar{q}-q\|_{L^2(\Omega)}\|\bar{u}-u\|_{L^2(\Omega)}\|y\|_{H^2(\Omega)}\\
&\le\zeta C\|\bar{q}-q\|_{L^2(\Omega)}\|\bar{u}-u\|_{L^2(\Omega)}\|v\|_{L^2(\Omega)}
\end{aligned}$$

for some constant $C>0$, where we have used the fact that $H^2(\Omega)$ is continuously embedded in $L^\infty(\Omega)$. Finally we have

$$\|v\|_{L^2(\Omega)}\le\zeta C\|\bar{q}-q\|_{L^2(\Omega)}\|\bar{u}-u\|_{L^2(\Omega)}\,.$$

Restricting $\mathcal{D}$ to a sufficiently small L^2-neighborhood of $q^\dagger$, would keep the term $\zeta\|\bar{q}-q\|_{L^2(\Omega)}$ small, such that the tangential cone condition Assumption 3.3 would be satisfied, but this choice does not ensure that $q \geq 0$ almost everywhere in Ω for all $q \in \mathcal{D}$. This problem can be avoided by choosing

$$\mathcal{D} = \{q \in L^2(\Omega)|\ q \geq 0 \text{ almost everywhere in } \Omega \text{ and } \|q - q^\dagger\|_{L^2(\Omega)} \leq \varepsilon\} \tag{3.85}$$

for some $\varepsilon > 0$ sufficiently small.

At this point, we would like to refer to Remark 2.4, where we mentioned that it can happen that $\mathcal{D}$ has an empty interior. This would be the case here for the choice (3.85). We also mentioned that the assumption that $\mathcal{D}$ has a nonempty interior would allow to use stationarity for a minimizer. By means of this example, we can see, that this assumption is too strong in general for this purpose:

As usual for optimal control problems with control constraints, the constraint $q \geq 0$ almost everywhere leads to a variational inequality as necessary optimality condition, which can be expressed equivalently as a projection formula. For a minimizer $q_\beta^\delta > 0$ a.e., the projection operator becomes the identity, and the variational inequality turns into an equality. Together with the fact that the ball $\|q - q^\dagger\|_Q$ has a nonempty interior, a minimizer $q_\beta^\delta \in \{q \in L^2(\Omega)|\ q > 0$ almost everywhere in Ω and $\|q - q^\dagger\|_{L^2(\Omega)} < \varepsilon\}$ satisfies indeed the stationarity equation (2.28).

3. **Example 3.3**

We will only carry out the case that ζ is small (the notion "small" in this context will be specified in the following, cf. (3.86)), which corresponds to a small wave number in a scattering problem modeled by the Helmholtz equation (cf., e.g., [15]), and refer to [79] for different choices of boundary control. The main difficulty and difference to Example 3.2 is the lack of coercivity of the semilinear form in case that the wave number matches an eigenvalue of the Laplacian, such that existence and continuity of the solution operator F cannot be shown for arbitrary ζ (cf. [88]).

The weak formulation reads

$$a(q,u)(\varphi) := (\nabla u, \nabla\varphi) - \zeta((1+q)u, \varphi) = (f, \varphi) \quad \forall \varphi \in H_0^1(\Omega)\,.$$

(i): Similarly as in Example 3.2 (i) we get

$$|a(q,u)(\varphi)| \leq (1 + c_{sp}^2\zeta\ \|1+q\|_{L^2(\Omega)})\ \|\nabla u\|_{L^2(\Omega)}\ \|\nabla\varphi\|_{L^2(\Omega)}$$

and

$$\begin{aligned} a(q,u)(u) &= \|\nabla u\|_{L^2(\Omega)}^2 - (\zeta(1+q)u, u) \\ &\geq \|\nabla u\|_{L^2(\Omega)}^2 - \zeta\|1+q\|_{L^2(\Omega)}\|u\|_{L^4(\Omega)}^2 \\ &\geq \left(1 - \zeta c_{sp}^2\|1+q\|_{L^2(\Omega)}\right) \|\nabla u\|_{L^2(\Omega)}^2\,, \end{aligned}$$

which implies continuity and coercivity of a for

$$q \in \mathcal{D} := \{q \in L^2(\Omega)|\ \ \|1+q\|_{L^2(\Omega)} < \tfrac{1}{\zeta c_{sp}^2}\}\,, \tag{3.86}$$

hence well–definedness of $F\colon \mathcal{D} \to H_0^1(\Omega)$.

Further there holds

$$\|\nabla u\|^2_{L^2(\Omega)} \leq c_p \, \|f\|_{L^2(\Omega)} \, \|\nabla u\|_{L^2(\Omega)} + \zeta c_{sp}^2 \, \|1+q\|_{L^2(\Omega)} \, \|\nabla u\|^2_{L^2(\Omega)} \,,$$

which implies

$$\|\nabla u\|_{L^2(\Omega)} \leq \frac{c_p}{1 - \zeta c_{sp}^2 \, \|1+q\|_{L^2(\Omega)}} \, \|f\|_{L^2(\Omega)} \,.$$

So F is also bounded from $\mathcal{D}$ to $H_0^1(\Omega)$.

(ii),(iii): Assumption 3.2 and Assumption 3.3 follow analogously to Example 3.2.

4. **Example 3.4:**

In this case we cannot proceed as for Example 3.3, since then, 0 is an eigenvalue of the Laplacian due to the Neumann zero boundary condition (cf. [38]), such that for small ζ the given Helmholtz equation is not coercive. A possible setting, which avoids the mentioned difficulty, would be the use of impedance boundary conditions, i.e., $\partial_n u - i\sqrt{\zeta} u = \xi$, cf. [88], since the eigenvalues of the Laplacian with impedance boundary are not real.

Due to the lack of coercivity in our setting, we won't provide a theoretical analysis here, but discuss the numerical results in Section 3.6 anyway. We will also consider a linearization of the Helmholtz equation, which is obtained via the Born approximation cf. [15] and apply the algorithm from [39] to this simplified linear problem (see Appendix D). The claimed assumptions are trivially satisfied in that case, since the operator F is linear.

3.5. Algorithm and implementational realization

Summarizing the results from this and the previous section, the concrete algorithm for determining the regularization parameter β according to Theorem 3.7 and Proposition 3.1 can be seen in Algorithm 3.1.

The last loop (step 18– 20 in Algorithm 3.1) only serves as a refinement loop, with β fixed and represents the implementation of Theorem 3.4.

As the theoretical assumption $\underline{\gamma} \leq i''(\beta)$ for all $\beta > 0$ (see Theorem 3.7) is not exactly transferable to practice, we decided to choose $\underline{\gamma}$ as follows:

$$\underline{\gamma} = \underline{\gamma}_k := i''^k_h \,, \qquad \gamma = \gamma_k := \begin{cases} 0 \,, & \text{if } i''^k_h \geq 0 \\ -\frac{3}{2} i''^k_h \,, & \text{if } i''^k_h < 0 \,. \end{cases} \tag{3.87}$$

Note that only existence of an upper bound $\overline{\gamma}$ but not its explicit value is needed in order to compute the regularization parameter according to Theorem 3.7 and Proposition 3.1. For the tested examples, the numerical results of the version of the algorithm for quadratic convergence

Algorithm 3.1: Inexact Newton method for determining a regularization parameter for nonlinear inverse problems

---

Choose $c_{tc} < 1$, $\overline{\overline{\tau}} > \tau > \underline{\underline{\tau}} > 1$, $\tilde{\tau} < \underline{\underline{\tau}}$ such that (3.20) is fulfilled, $c_1 \in (0,1)$, $c_2 \in (0, \frac{1}{2})$ such that $2c_1 + c_2 < 1$ and $c \leq \frac{\tilde{\tau}^2}{\overline{\overline{\tau}}^2}$ and choose $\Theta > 0$ such that (3.20).
Choose initial guess $\beta^0 > 0$, initial discretizations Q_h, V_h, set $k = 0$.
Compute $x_h = (q_h, u_h, z_h)$ by solving (2.38) .
Evaluate i_h^k by (2.37).
while $i_h^k > \overline{\overline{\tau}}^2 \delta^2$ **do**
  Compute $x_{h,1} = (q_{h,1}, u_{h,1}, z_{h,1})$ by solving (3.69).
  Evaluate ${i'}_h^k$ by (3.71).
  Evaluate the error estimator η^I (for $i(\beta^k)$) according to (3.76).
  Compute $x_{h,2}$ and $x_{h,3}$ by solving (3.73).
  Evaluate the error estimator η^K (for $i'(\beta^k)$) according to (3.76).
  if (3.38), (3.39), (3.40), (3.41) and (3.9) are satisfied **then**
    Evaluate $\underline{\gamma}$ according to (3.87).
    Set $\beta^{k+1} = \beta^k + \sigma^{k+1} s^{k+1}$ and $k = k + 1$.
  else
    Refine with respect to the corresponding error estimator.
  Compute $x_h = (q_h, u_h, z_h)$ by solving (2.38) .
  Evaluate i_h^k by (2.37).
while (3.19) is violated **do**
  Refine with respect to the error estimator for j (see (3.19) and (3.68)).
  Compute $x_h = (q_h, u_h, z_h)$ by solving (2.38).

(cf. Proposition 3.2) did not differ significantly from the one given above for linear convergence, which may be caused by the fact that we stop at k_* and do not check whether (3.42) holds.

In Section 3.3 we discussed the evaluation of the error estimators needed for Algorithm 3.1 insofar as we formulated the equations, which have to be solved. We will now go into more detail and explain how these equations can be solved efficiently.

Although, of course, in practice we solve the discrete versions of the equations in Section 3.3, for notational simplicity, we consider the continuous equations in all following algorithms in this section. The discrete counterparts are obtained by replacing Q, V, W by Q_h, V_h, W_h, see also Section 2.3.

We start with the error estimator η^J for J_β, which is given by

$$J_\beta(q, u) - J_\beta(q_h, u_h) \approx \eta^J := \frac{1}{2} L'(x_h)(\pi x_h - x_h)$$

(cf. Section 3.3 and Section 2.4) and computed according to Algorithm 3.2 using the notation $e_u := \pi u_h - u_h$, $e_z := \pi z_h - z_h$, $e_q := \pi q_h - q_h$.

For obtaining x_1 according to [39, Proposition 3] we need to solve (3.69). We do this in the manner presented in Algorithm 3.3 (as in [85]).

Algorithm 3.2: Computation of η^J

1: Compute $\eta_{\text{state}} := \mathcal{L}'_z(x_h)(e_z)$
2: Compute $\eta_{\text{dual}} := \mathcal{L}'_u(x_h)(e_u)$
3: Compute $\eta_{\text{control}} := \mathcal{L}'_q(x_h)(e_q)$
4: Compute $\eta^J = \eta_{\text{state}} + \eta_{\text{dual}} + \eta_{\text{control}}$

Algorithm 3.3: Computation of x_1

1: Determine $z_1^{(0)} \in W$ as a solution to the dual equation "DualQI" (see Appendix A).

2: Determine $q_1 \in Q$ by solving the hessian equation"QIEq":

$$j''(q)(q_1, \varphi) = -\mathcal{L}''_{zq}(x)(z_1^{(0)}, \varphi) \qquad \forall \varphi \in Q\,. \tag{3.88}$$

We thereby evaluate the second derivative $j''(q)(\delta q)$ for given $\delta q \in Q$ as in Section 2.2

3: u_1 is a solution to (2.29) for given $\delta q = q_1$. It is already computed in the process of solving (3.88).

4: Set $z_1 = z_1^{(0)} + z_1^{(1)}$, where $z_1^{(1)}$ is the solution of (2.30) for given $\delta q = q_1$ and $\delta u = u_1$, which is already computed in the process of solving (3.88).

By means of x_1 we can then compute the error estimator η^I for I according to Section 3.3, i.e.,

$$\begin{aligned} I(u) - I(u_h) \approx \eta^I &:= \frac{1}{2}\mathcal{M}'_y(y_h)(y - \pi_h y) \\ &= \frac{1}{2}\Big[I'_u(q,u)(e_u) + \mathcal{L}''_{qq}(x)(q_1, e_q) + \mathcal{L}''_{qu}(x)(q_1, e_u) \\ &\quad + \mathcal{L}''_{qz}(x)(q_1, e_z) + \mathcal{L}''_{uq}(x)(u_1, e_q) + \mathcal{L}''_{uu}(x)(u_1, e_u) \\ &\quad + \mathcal{L}''_{zq}(x)(z_1, e_q) + \mathcal{L}''_{zu}(x)(z_1, e_u) + \mathcal{L}''_{uz}(x)(u_1, e_z) \\ &\quad + \mathcal{L}'_q(x)(e_{q_1}) + \mathcal{L}'_u(x)(e_{u_1}) + \mathcal{L}'_z(x)(e_{z_1})\Big]\,, \end{aligned}$$

which leads to Algorithm 3.4, where we use the notation $e_{u_1} := \pi u_{h,1} - u_{h,1}$, $e_{z_1} := \pi z_{h,1} - z_{h,1}$, $e_{q_1} := \pi q_{h,1} - q_{h,1}$.

Computing x_2, i.e., solving the first part of (3.72) is easier than the computation of x_1, since the right-hand side only contains q and q_1, which leads to Algorithm 3.5 for the computation of x_2.

The second equation in (3.72) can be solved efficiently as formulated in Algorithm 3.6 for the computation of x_3.

Finally we can compute the error estimator η^K according to Section 3.3, i.e.,

$$K(q, q_1) - K(q_h, q_{1,h}) \approx \eta^K := \frac{1}{2}N'_y(y_h, \mathbf{y}_h)(\pi y_h - y_h) + \frac{1}{2}N'_{\mathbf{y}}(y_h, \mathbf{y}_h)(\pi \mathbf{y}_h - \mathbf{y}_h)\,,$$

Algorithm 3.4: Computation of η^I

1: Compute $\eta_{\text{state}} := \mathcal{L}'_z(x_h)(e_{z_1})$
2: Compute $\eta_{\text{dual}} := \mathcal{L}'_u(x_h)(e_{u_1})$
3: Compute $\eta_{\text{control}} := \mathcal{L}'_q(x_h)(e_{q_1})$
4: Compute $\eta_{\text{tangent}} := \mathcal{L}''_{qz}(x_h)(q_{1,h}, e_z) + \mathcal{L}''_{uz}(x_h)(u_{1,h}, e_z)$
5: Compute

$$\eta_{\text{dualForHessianQI}} := \mathcal{L}''_{qu}(x_h)(q_{1,h}, e_u)+\mathcal{L}''_{uu}(x_h)(u_{1,h}, e_u)+\mathcal{L}''_{zu}(z_{1,h}, e_u)+I'(u_h)(e_u)$$

6: Compute $\eta^I = \eta_{\text{state}} + \eta_{\text{dual}} + \eta_{\text{control}} + \eta_{\text{tangent}} + \eta_{\text{dualForHessianQI}}$

Algorithm 3.5: Computation of x_2

1: Determine $q_2 \in Q$ by solving the hessian equation "K1Eq":

$$j''(q)(q_2, \varphi) = -K'_{q_1}(q, q_1)(\varphi) \qquad \forall \varphi \in Q\,. \tag{3.89}$$

We thereby evaluate the second derivative $j''(q)(\delta q)$ for given $\delta q \in Q$ as in Section 2.2
2: u_2 is a solution to (2.29) for given $\delta q = q_2$. It is already computed in the process of solving (3.89).
3: z_2 is a solution to (2.30) for given $\delta q = q_2$ and $\delta u = u_2$, which is already computed in the process of solving (3.89).

where

$$\begin{aligned}
N'_y(y_h, \mathbf{y}_h)(\pi y_h - y_h) &= K'_q(q_h, q_{1,h})(\pi q_h - q_h) + K'_{q_1}(q_h, q_{2,h})(\pi q_{h,1} - q_{h,1}) \\
&\quad + M''_{yy}(y_h)(\mathbf{y}_h, \pi y_h - y_h) \\
&= K'_q(q_h, q_{1,h})(\pi q_h - q_h) + +K'_{q_1}(q_h, q_{2,h})(\pi q_{h,1} - q_{h,1}) \\
&\quad + I''(u_h)(u_{2,h}, \pi u_h - u_h) + \mathcal{L}''_{xx}(x_h)(\pi x_{h,1} - x_{h,1}, x_{2,h}) \\
&\quad + \mathcal{L}''_{xx}(x_h)(x_{3,h}, \pi x_h - x_h) + \mathcal{L}'''_{xxx}(x_h)(x_{1,h}, x_{2,h}, \pi x_h - x_h)
\end{aligned}$$

$$\begin{aligned}
N'_{\mathbf{y}}(y_h, \mathbf{y}_h)(\pi \mathbf{y}_h - \mathbf{y}_h) &= \mathcal{M}''_{y\mathbf{y}}(y_h)(\mathbf{y}_h, \pi \mathbf{y}_h - \mathbf{y}_h) \\
&= I'(u_h)(\pi u_{2,h} - u_{2,h}) + \mathcal{L}''_{xx}(x_h)(x_{1,h}, \pi x_{2,h} - x_{2,h}) \\
&\quad + \mathcal{L}'_x(x_h)(\pi x_{3,h} - x_{3,h})\,.
\end{aligned}$$

With

$$\begin{aligned}
e_{u_2} &:= \pi u_{2,h} - u_{2,h} & e_{z_2} &:= \pi z_{2,h} - z_{2,h} & e_{q_2} &:= \pi q_{2,h} - q_{2,h} \\
e_{u_3} &:= \pi u_{3,h} - u_{3,h} & e_{z_3} &:= \pi z_{3,h} - z_{3,h} & e_{q_3} &:= \pi_h q_{3,h} - q_{3,h}
\end{aligned}$$

Algorithm 3.6: Computation of x_3

1: Determine $u_3^{(0)} \in V$ as a solution to the tangent equation "TangentK2For3rdDerivative" (see Appendix A).
2: Determine $z_3^{(0)} \in W$ as a solution to the dual equation "DualK2" (see Appendix A).
3: Determine $q_3 \in Q$ by solving the hessian equation "K2Eq" (see Appendix A). We thereby evaluate the second derivative $j''(q)(\delta q)$ for given $\delta q \in Q$ as explained in Section 2.2
4: Set $u_3 = u_3^{(0)} + u_3^{(1)}$, where $u_3^{(1)}$ is a solution to (2.29) for given $\delta q = q_3$. It is already computed in the process of solving "K2Eq".
5: Set $z_3 = z_3^{(0)} + z_3^{(1)}$, where $z_3^{(1)}$ is the solution of (2.30) for given $\delta q = q_3$ and $\delta u = u_3^{(1)}$, which is already computed in the process of solving "K2Eq".

this leads to the computation of η^K according to Algorithm 3.7, where

$$\begin{aligned}\eta_{\text{dualForHessianK2}} \coloneqq\ & \mathcal{L}''_{qu}(x_h)(q_{3.h}, e_u) + \mathcal{L}''_{uu}(x_h)(u_{3.h}, e_u) \\ & + \mathcal{L}''_{zu}(x_h)(z_{3.h}, e_u) + I''(u_h)(u_{2,h}, e_u) \\ & + \mathcal{L}'''_{quu}(x_h)(q_{1,h}, u_{2,h}, e_u) + \mathcal{L}'''_{quu}(x_h)(q_{2,h}, u_{1,h}, e_u) \\ & + \mathcal{L}'''_{qzu}(x_h)(q_{1,h}, z_{2,h}, e_u) + \mathcal{L}'''_{qzu}(x_h)(q_{2,h}, z_{1,h}, e_u) \\ & + \mathcal{L}'''_{zuu}(x_h)(z_{1,h}, u_{2,h}, e_u) + \mathcal{L}'''_{zuu}(x_h)(z_{2,h}, u_{1,h}, e_u) \\ & + \mathcal{L}'''_{uuu}(x_h)(u_{1,h}, u_{2,h}, e_u) + \mathcal{L}'''_{qqu}(x_h)(q_{1,h}, q_{2,h}, e_u)\end{aligned} \tag{3.90}$$

and

$$\begin{aligned}\eta_{\text{controlK2RHSO}} \coloneqq\ & \mathcal{L}''_{qq}(x_h)(q_{3,h}, e_q) + \mathcal{L}''_{uq}(x_h)(u_{3,h}, e_q) \\ & + \mathcal{L}''_{zq}(x_h)(z_{3,h}, e_q) + K'_q(q_h, q_{1,h})(e_q) \\ & + \mathcal{L}'''_{qzq}(x_h)(q_{1,h}, z_{2,h}, e_q) + \mathcal{L}'''_{qzq}(x_h)(q_{2,h}, z_{1,h}, e_q) \\ & + \mathcal{L}'''_{uzq}(x_h)(u_{1,h}, z_{2,h}, e_q) + \mathcal{L}'''_{uzq}(x_h)(u_{2,h}, z_{1,h}, e_q) \\ & + \mathcal{L}'''_{quq}(x_h)(q_{1,h}, u_{2,h}, e_q) + \mathcal{L}'''_{quq}(x_h)(q_{2,h}, u_{1,h}, e_q) \\ & + \mathcal{L}'''_{uuq}(x_h)(u_{1,h}, u_{2,h}, e_q) + \mathcal{L}'''_{qqq}(x_h)(q_{1,h}, q_{2,h}, e_q)\,.\end{aligned} \tag{3.91}$$

Remark 3.12. *Will will not go into more detail concerning the localization of the error estimators, but refer to [30, 84]. Also see [84, Remark 6.3], which states that in simple cases, e.g., when the control enters the right-hand side linearly, and in case we choose the same discrete space for the state and the control, via optimality conditions the terms containing the control error vanish.*

Algorithm 3.7: Computation of η^K

1: Compute $\eta_{\text{state}} \coloneqq \mathcal{L}'_z(x_h)(e_{z_3})$
2: Compute $\eta_{\text{dual}} \coloneqq \mathcal{L}'_u(x_h)(e_{u_3})$
3: Compute $\eta_{\text{control}} \coloneqq \mathcal{L}'_q(x_h)(e_{q_3})$
4: Compute $\eta_{\text{tangent}} \coloneqq \mathcal{L}''_{qz}(x_h)(q_{1,h}, e_{z_2}) + \mathcal{L}''_{uz}(x_h)(u_{1,h}, e_{z_2})$
5: Compute

$$\eta_{\text{dualForHessianQI}} \coloneqq \mathcal{L}''_{qu}(x_h)(q_{1,h}, e_{u_2}) + \mathcal{L}''_{uu}(x_h)(u_{1,h}, e_{u_2}) + \mathcal{L}''_{zu}(z_{1,h}, e_{u_2}) + I'(u_h)(e_{u_2})$$

6: Compute $\eta_{\text{HVO}} \coloneqq \mathcal{L}''_{qq}(x_h)(q_{1,h}, e_{q_2}) + \mathcal{L}''_{uq}(x_h)(u_{1,h}, e_{q_2}) + \mathcal{L}''_{zq}(x_h)(z_{1,h}, e_{q_2})$
7: Compute $\eta_{\text{tangentA1}} \coloneqq \mathcal{L}''_{qz}(x_h)(q_{2,h}, e_{z_1}) + \mathcal{L}''_{uz}(u_{2,h}, e_{z_1})$
8: Compute

$$\eta_{\text{dualForHessian}} \coloneqq \mathcal{L}''_{qu}(x_h)(q_{2.h}, e_{u_1}) + \mathcal{L}''_{uu}(x_h)(u_{2.h}, e_{u_1}) + \mathcal{L}''_{zu}(x_h)(z_{2.h}, e_{u_1})$$

9: Compute

$$\eta_{\text{HVOA1}} \coloneqq \mathcal{L}''_{qq}(x_h)(q_{2,h}, e_{q_1}) + \mathcal{L}''_{uq}(x_h)(u_{2,h}, e_{q_1}) + \mathcal{L}''_{zq}(x_h)(z_{2,h}, e_{q_1}) + K'_{q_1}(q_h, q_{1,h})(e_{q_1})$$

10: Compute

$$\begin{aligned} \eta_{\text{tangentA2}} \coloneqq{} & \mathcal{L}''_{qz}(x_h)(q_{3.h}, e_z) + \mathcal{L}''_{uz}(x_h)(u_{3.h}, e_z) + \mathcal{L}'''_{qqz}(x_h)(q_{1,h}, q_{2,h}, e_z) \\ & + \mathcal{L}'''_{uuz}(x_h)(u_{1,h}, u_{2,h}, e_z) + \mathcal{L}'''_{quz}(x_h)(q_{1,h}, u_{2,h}, e_z) + \mathcal{L}'''_{quz}(x_h)(q_{2,h}, u_{1,h}, e_z) \end{aligned}$$

11: Compute $\eta_{\text{dualForHessianK2}}$ according to (3.90)
12: Compute $\eta_{\text{controlK2RHSO}}$ according to (3.91)
13: Compute

$$\begin{aligned} \eta^K = {} & \eta_{\text{state}} + \eta_{\text{dual}} + \eta_{\text{control}} + \eta_{\text{tangent}} + \eta_{\text{dualForHessianQI}} + \eta_{\text{HVO}} + \eta_{\text{tangentA1}} \\ & + \eta_{\text{dualForHessian}} + \eta_{\text{controlHVOA1}} + \eta_{\text{tangentA2}} + \eta_{\text{dualForHessianK2}} + \eta_{\text{controlK2RHSO}} \end{aligned}$$

3.6. Numerical results

For illustrating the performance of the proposed method for the computation of a Tikhonov regularization parameter for nonlinear inverse problems with adaptive discretizations "(NT)" (for Nonlinear Tikhonov) according to the algorithms from Section 3.5 (especially Algorithm 3.1) we apply it to the example PDEs from Section 3.4.

All the computations in this thesis have been done using the optimization toolbox RoDoBo [101] in combination with the finite element toolbox Gascoigne [35] and the visualization software VisuSimple [109]. Some graphics have been created via MATLAB [82].

In our numerical computations of this section, step 3 and step 16 in Algorithm 3.1 are realized by applying Newton's method (cf. Section 2.5.1) to the optimization problem (2.22).

In all considered examples we aim to identify the parameter $q \in Q = L^2(\Omega)$ from noisy measurements $g^\delta \in G$ of the state $u \in H^1(\Omega)$ in $\Omega = (0,1)^2 \subset \mathbb{R}^2$. As for the measurements we consider two cases:

(i) via point functionals in $n_m = 100$ uniformly distributed points ξ_i, $i = 1, 2, \ldots, n_m$ and perturbed by uniformly distributed random noise of some percentage $p > 0$. Then the observation space is chosen as $G = \mathbb{R}^{n_m}$ and the observation operator is defined by $(C(v))_i = v(\xi_i)$ for $i = 1, \ldots, n_m$. The noisy data is created via $g_i^\delta = g(\xi_i)(1 + \varepsilon_i p)$ for $i = 1, \ldots, n_m$, where $\varepsilon_i \in (-1, 1)$ are random numbers and the exact data g is simulated as solution $u^\dagger$ of the given PDE on a very fine mesh with 1050625 nodes and equally sized quadratic cells.

(ii) via L^2-projection. Then we have $G = L^2(\Omega)$, $C = \mathrm{id}$, and

$$g^\delta = g + \delta \frac{r}{\|r\|_{L^2(\Omega)}} = g + p\, \|g\|_{L^2(\Omega)} \frac{r}{\|r\|_{L^2(\Omega)}},$$

where r denotes some uniformly distributed random noise and p the percentage of perturbation. The exact state $u^\dagger$ is simulated on a very fine mesh with 1050625 nodes and equally sized quadratic cells. We denote the corresponding finite element space by V_{h_L}. In order to evaluate $\|C(u) - g^\delta\|_{L^2(\Omega)} = \|u - g^\delta\|_{L^2(\Omega)}$ on coarser meshes and the corresponding finite element spaces V_{h_l} with $l = 0, 1, \ldots, L$ during the optimization algorithm, g^δ has to be transferred from V_{h_L} to the current grid V_{h_l}. As usual in the finite element context, this is done by the L^2-projection as the restriction operator.

We consider configurations with three different exact sources $q^\dagger$ (with $(x, y) \in \Omega$):

(a) A Gaussian distribution

$$q^\dagger = \frac{c}{2\pi\sigma^2} \exp\left(-\frac{1}{2}\left(\left(\frac{sx - \mu}{\sigma}\right)^2 + \left(\frac{sy - \mu}{\sigma}\right)^2\right)\right)$$

with $c = 10$, $\mu = 0.5$, $\sigma = 0.1$, and $s = 2$.

(b) Two Gaussian distributions added up to one distribution

$$q^\dagger = q_1 + q_2,$$

where

$$q_1 = \frac{c_1}{2\pi\sigma^2} \exp\left(-\frac{1}{2}\left(\left(\frac{s_1 x - \mu}{\sigma}\right)^2 + \left(\frac{s_1 y - \mu}{\sigma}\right)^2\right)\right),$$

$$q_2 = \frac{c_2}{2\pi\sigma^2} \exp\left(-\frac{1}{2}\left(\left(\frac{s_2 x - \mu}{\sigma}\right)^2 + \left(\frac{s_2 y - \mu}{\sigma}\right)^2\right)\right)$$

with $\sigma = 0.1$, $\mu = 0.5$, $s_1 = 2$, $s_2 = 0.8$, $c_1 = 1$, and $c_2 = 1$.

(c) The step function

$$q^{\dagger} = \begin{cases} 0 & \text{for } x \geq 0.5 \\ 1 & \text{for } x < 0.5\,. \end{cases}$$

Figure 3.1 shows the exact source distributions $q^{\dagger}$.

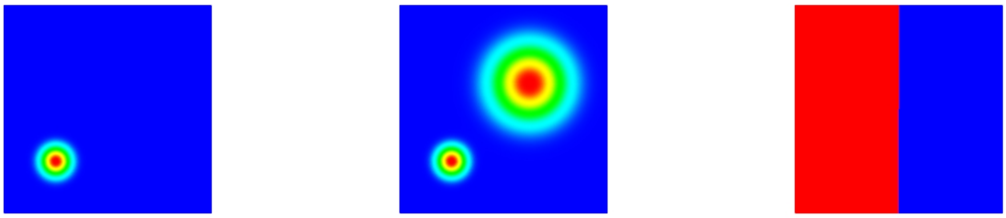

Figure 3.1.: Exact source distribution/control $q^{\dagger}$. FLTR: (a),(b),(c)

The concrete choice of the parameters for the numerical tests is the following: $\tilde{\tau} = 0.1$, $\underline{\underline{\tau}} = 3.1$, $\tau = 4$, $\overline{\overline{\tau}} = 5$, $c_{tc} = 10^{-7}$, $c_1 = 0.9$, $c_2 = 0.4$, $c_{tc} = 10^{-7}$ and

$$\Theta^2 := \frac{\left(\sqrt{\underline{\underline{\tau}}^2 - \tilde{\tau}^2}(1 - c_{tc}) - 1 - c_{tc}\right)^2 - (1 + c_{tc})^2}{2(1 - c_{tc})} - 1 - \tilde{\tau}\,,$$

which guarantees (3.20).

Remark 3.13. *In order to fulfill* (3.20)*, $\underline{\underline{\tau}}$ and $\tilde{\tau}$ have to be chosen such that* $\sqrt{\underline{\underline{\tau}}^2 - \tilde{\tau}^2} > 1 + \sqrt{3}$.

As explained in the previous section (Section 3.5), we restrict ourselves to linear convergence of the sequence of β's. That means all presented numerical results refer to Proposition 3.1, which guarantees (only) linear convergence of the produced sequence of β's, but uses less restrictive accuracy requirements, thus allowing for coarser grids than the quadratically convergent version.

We start with a regularization parameter $\beta^0 = 10$, an initial guess for the control $q_0 = 0 \in L^2(\Omega)$ and a mesh of 25 equally sized quadratic cells. In each refinement step the corresponding selected cell is divided into four equally sized quadratic cells.

Please note that throughout this thesis, the color maps for the different images are not the same, such that the presented images of the reconstructions only give evidence about the shape and not about absolute values. To make up for this we will mostly give the relative error of the reconstruction of the source, i.e.,

$$\frac{\|q^{\delta}_{\beta,h} - q^{\dagger}\|_Q}{\|q^{\dagger}\|_Q}\,. \tag{3.92}$$

3.6.1. Example 1

First we consider Example 3.1 from Section 3.4:

For $q \in L^2(\Omega)$ find $u \in H_0^1(\Omega)$ such that

$$\begin{cases} -\Delta u + \zeta u^3 = q & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \,. \end{cases}$$

For the choice $\zeta = 1000$, the corresponding exact states for the sources given in Figure 3.1 are displayed in Figure 3.2.

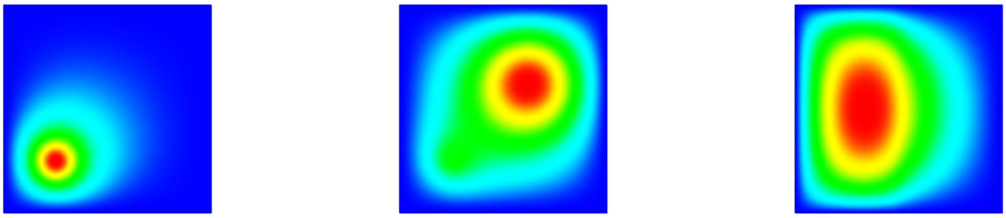

Figure 3.2.: Exact states $u^\dagger$ for Example 1 with $\zeta = 1000$. FLTR: configuration (a),(b),(c)

For the configuration (a) we also present the exact states for different choices of ζ in Figure 3.3 with the intention of showing that a different choice of ζ barely changes the shape of the state and that the upcoming images of the reconstructions of the state for different choices of ζ can be interpreted by comparing them to Figure 3.2.

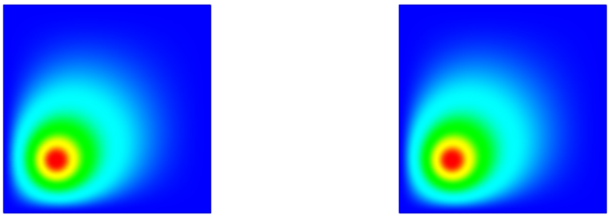

Figure 3.3.: Exact states $u^\dagger$ for Example 1 with configuration (a). FLTR: $\zeta = 10$, $\zeta = 1$

In Figure 3.4 we can see the reconstructed control, the reconstructed state and the adaptively refined mesh produced by Algorithm 3.1 for the choice $\zeta = 1$, 1% noise and point measurements (a). As we would expect, the algorithm refines mostly in the left lower corner, where the source is located. Also the control and the state seem to be reconstructed satisfactorily.

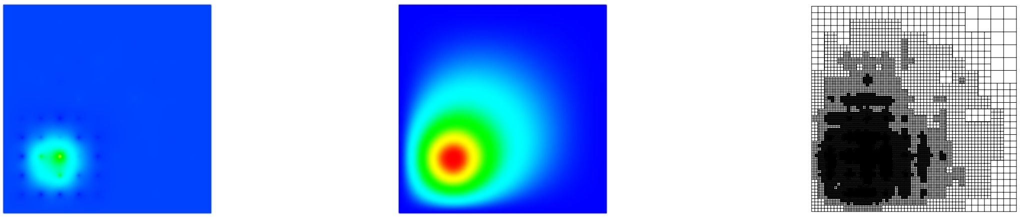

Figure 3.4.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 1 (a)(i), $\zeta = 1$, 1% noise using (NT)

In Table 3.1 we give a closer insight in the functioning of the presented algorithm and at the same time, compare the results of this adaptive strategy to uniform refinement. In the first column we indicate the iteration number of the inexact Newton method for β (see Theorem 3.7 and Algorithm 3.1), in the second and forth column we find the number of nodes in the adaptively or uniformly refined mesh respectively, and in the third and fifth column we display the size of β. Recapitulating Table 3.1, the algorithm first refines four times, then updates β seven times, until the "optimal" β is found. The subsequent four steps denote the second refinement loop, i.e., step 18–20 in Algorithm 3.1. Both versions, adaptive and uniform refinement, lead to the same regularization parameter, but with adaptivity we save about 96% of nodes.

Table 3.1.: Adaptive refinement using Algorithm 3.1 (NT) versus uniform refinement for Example 1 (a)(i) with $\zeta = 1$ and 1% noise

	Adaptive refinement		Uniform refinement	
k	# nodes	β^k	# nodes	β^k
0	25	10	25	10
0	41	10	81	10
0	137	10	289	10
0	413	10	1089	10
0	1309	10	4225	10
1	1309	30	4225	30
2	1309	71	4225	71
3	1309	154	4225	154
4	1309	320	4225	319
5	1309	652	4225	652
6	1309	1257	4225	1256
7	1309	2140	4225	2140
-	2473	2140	16641	2140
-	5327	2140	66049	2140
-	15081	2140	263169	2140
-	38665	2140	1050625	2140

In order to show the influence of the regularization parameter on the objective functional, we present some sort of L-curve for the same example (i.e., Example 1 (a)(i), $\zeta = 1$, 1%

noise) in Figure 3.5. The L-curve is a log-log graph with the misfit term (i.e., in our setting $\|F(q^\delta_{h,\beta}) - g^\delta\|^2_G$) on the x-axis and the regularization term without regularization parameter (i.e., in our setting $\|q^\delta_{h,\beta}\|^2_Q$) on the y-axis (see, e.g., [45]) for various choices of the regularization parameter β. In a continuous setting with respect to β, i.e., considering all $\beta \in (0, \infty)$ for a fixed noise level and a fixed discretization, such a graph usually has an L-shape. A "good" regularization parameter is obtained in its corner, where there is a balance between regularization and data fitting. What can be seen in Figure 3.5, is simply the horizontal line of this "L". We start from the stable side, i.e., with small β and large misfit term, then increase β iteratively, such that the misfit term is decreased and $\|q^\delta_{h,\beta}\|^2_Q$ is increased. Finally the algorithm stops before the "imaginary" corner is reached. Strictly speaking, the right part of the line and the green circles representing refinement steps do not belong to a proper L-curve graph, but only the red squares representing β updates. The four last refinement steps (see Table 3.1) are not included into the L-curve graph, since in this last refinement loop (cf. step 18–20 in Algorithm 3.1) β is fixed.

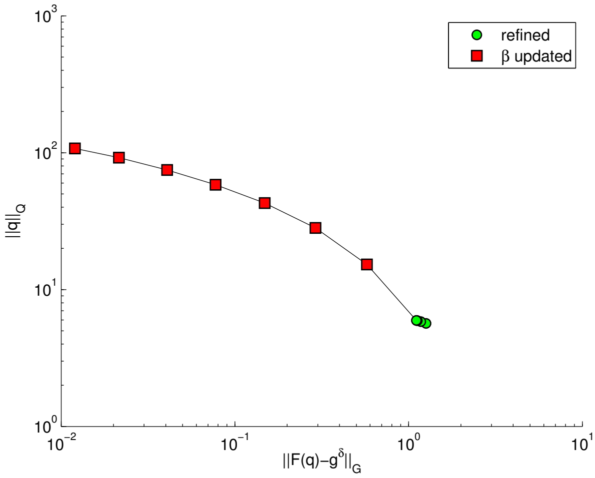

Figure 3.5.: L-Curve for Example 1 (a)(i) with $\zeta = 1$ and 1% noise

In Table 3.2 we display the results for different choices of the scalar factor ζ in the PDE. In the second and fifth column the reader can see the relative control error (3.92), in the third and sixth column the size of the computed regularization parameter, and in the forth and seventh column the number of nodes in the resulting refined mesh for adaptive and uniform refinement respectively. Finally, the last column shows the reduction of computation time (CTR) using adaptivity instead of uniform refinement, namely $69\% - 98\%$. Although a higher factor ζ, i.e., a boosted nonlinearity of the PDE, leads to higher computation times in general, the saving of computation time has not been affected much. The same holds for the number of nodes and the relative control error. Only the obtained regularization parameter seems slightly larger, the larger ζ. Besides, we wish to point out that the relative control error (≈ 0.456) as well as the obtained regularization parameter (≈ 5000) is about the same with adaptivity as with uniform refinement.

In order to get more representative results concerning the comparison of the measured computation times, we will consider $\zeta = 1000$ or $\zeta = 100$ in the following, because of the presumably higher computation times (probably caused by a larger β_*).

Table 3.2.: Adaptive refinement using Algorithm 3.1 (NT) versus uniform refinement for Example 1 (a)(i) for different choices of ζ with 1% noise). CTR: Computation time reduction using adaptivity

ζ	adaptive			uniform			CTR
	error	β	# nodes	error	β	# nodes	
1	0.456	3085	38665	0.456	3087	1050625	98%
10	0.452	3194	38377	0.452	3197	1050625	98%
100	0.443	4999	31967	0.443	4990	1050625	96%
1000	0.475	11463	44413	-	-	-	> 97%

We also tested the proposed method with noisy data produced via L^2-projection, i.e., alternative (ii). Figure 3.6 shows the reconstructed control, the reconstructed state, and the resulting adaptively refined mesh for the source (a), $\zeta = 1000$, and 1% noise. Here the refined mesh looks even more efficient than in Figure 3.4, because it refines less and more concentrated on the location of the source.

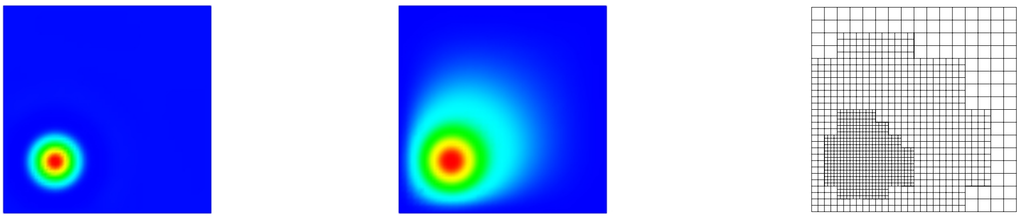

Figure 3.6.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 1 (a)(ii), $\zeta = 1000$, 1% noise using (NT)

Again, analogously to Table 3.1, we present the behavior of the algorithm concerning refinement and β updating steps in Table 3.3. Unfortunately, here we have the unusual case that the adaptively refined mesh has more nodes than the one obtained by uniform refinement, if only slightly. However, we interpret this fact as an accident, since this did not occur in any other tested example.

Please note that neither with adaptive nor with uniform refinement additional refinement according to step 18–20 in Algorithm 3.1 was needed here.

In Figure 3.7 one can see the L-curve diagram (cf. Figure 3.5) for the proposed algorithm applied to Example 1 (a)(ii), $\zeta = 1000$, 1% noise.

Next, we consider the source distribution (b) (see the beginning of Section 3.6). The reconstructed control, the reconstructed state, and the adaptively refined mesh for point measure-

Table 3.3.: Adaptive refinement using Algorithm 3.1 (NT) versus uniform refinement for Example 1 (a)(ii) with $\zeta = 1000$ and 1% noise

	Adaptive refinement		Uniform refinement	
k	# nodes	β^k	# nodes	β^k
0	25	10	25	10
0	69	10	81	10
1	69	462	81	460
1	137	462	289	460
1	413	462	1089	460
1	1305	462	-	-
2	1305	2232	1089	2232
3	1305	7956	1089	7947
4	1305	23397	1089	23387
5	1305	62004	1089	62053
6	1305	153697	1089	153648
7	1305	347474	1089	348885
8	1305	714429	1089	721494
9	1305	1344973	1089	1370237
10	1305	2211183	1089	2282419
11	1305	4368044	1089	4609382

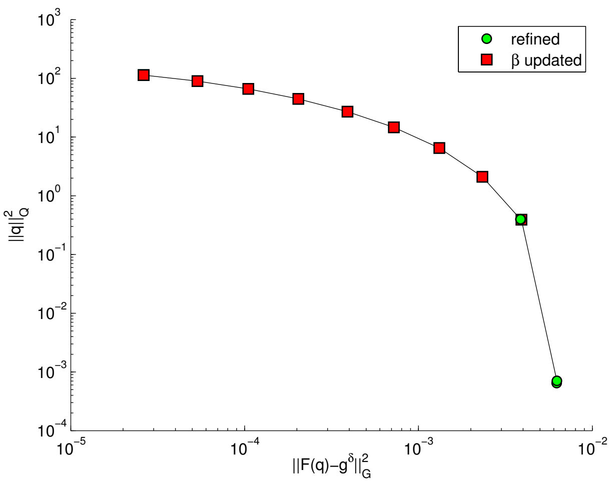

Figure 3.7.: L-Curve for Example 1 (a)(ii) with $\zeta = 1000$ and 1% noise using (NT)

ments (i) and $\zeta = 1000$ are displayed in Figure 3.8. Also for this configuration, the algorithm produces reasonable reconstructions and a reasonably refined mesh.

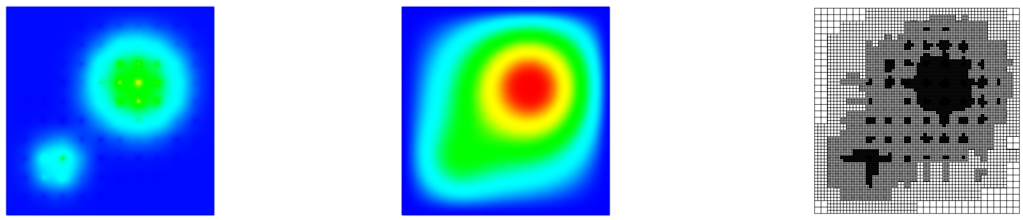

Figure 3.8.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 1 (b)(i), $\zeta = 1000$, 1% noise using (NT)

The refinement steps and β updating steps can be seen in Table 3.4 as well as in the L-curve Figure 3.9. Here we have four refinement steps, one β updating step, again one refinement step and from then on only β updating steps until the final β is found. The additional refinement loop only consists of one step in the adaptive case, whereas in the uniform case no additional refinement is needed. Again, both versions produce about the same β and the same relative error (3.92) of about 0.3, but using adaptivity we save about 69% of mesh nodes which leads to 54% saving of computation time.

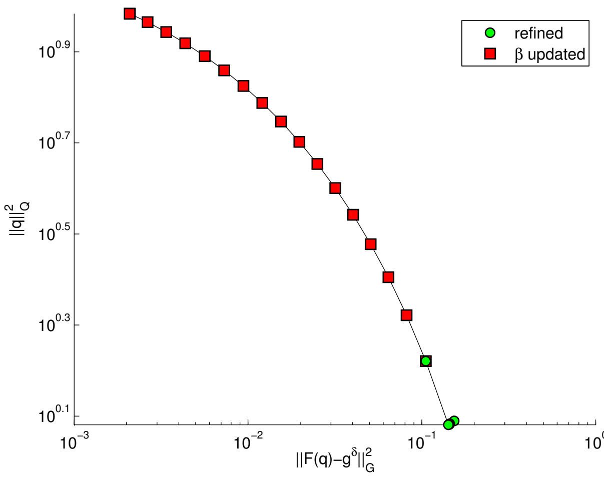

Figure 3.9.: L-Curve for Example 1 (b)(i) with $\zeta = 1000$ and 1% noise using (NT)

Considering the same example (Example 1 (b), $\zeta = 1000$, 1% noise) with noisy data (ii) (via L^2-projection) we obtain the reconstructions and the mesh from Figure 3.10.

As for Example 1 (a), the L^2-projection alternative (ii) produces a coarser mesh than the setting with point measurements (i). This is probably mainly because in case (i) the algorithm additionally refines at the measurement points, which are uniformly distributed in the domain. The reconstructions look even better comparing Figure 3.10 and Figure 3.8, which is also

Table 3.4.: Adaptive refinement using Algorithm 3.1 (NT) versus uniform refinement for Example 1 (b)(i) with $\zeta = 1000$ and 1% noise

	Adaptive refinement		Uniform refinement	
k	# nodes	β^k	# nodes	β^k
0	25	10	25	10
0	81	10	81	10
0	137	10	289	10
0	497	10	1089	10
0	1721	10	4225	10
1	1721	15	4225	16
1	4793	15	16641	16
2	4793	22	16641	22
3	4793	30	16641	30
4	4793	40	16641	40
5	4793	52	16641	52
6	4793	68	16641	68
7	4793	88	16641	88
8	4793	113	16641	113
9	4793	144	16641	144
10	4793	183	16641	183
11	4793	232	16641	231
12	4793	291	16641	290
13	4793	364	16641	363
14	4793	452	16641	450
15	4793	554	16641	552
16	4793	672	16641	669
17	4793	803	16641	799
17	10115	803	66049	799
18	10115	947	66049	943
-	20231	947	-	-

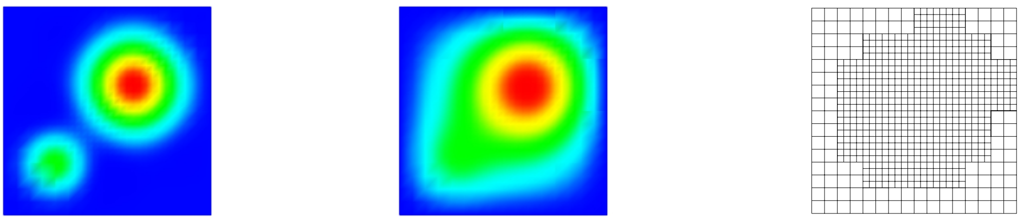

Figure 3.10.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 1 (b)(ii), $\zeta = 1000$, 1% noise using (NT)

confirmed by numbers considering Table 3.5, meaning that (ii) results in a slightly smaller relative control error. This is also plausible in view of the fact that (ii) provides information on u in all of Ω, whereas in (i) we only have measurements at 100 points. To put this in perspective we have to mention that the evaluation of this error (3.92) is done (numerically) on the final mesh of the particular example, which implies that this value is less accurate for (ii) than for (i) in this case here. With respect to uniform refinement we can see from the same table that again, uniform and adaptive refinement lead to about the same regularization parameter and that we save 73% of computation time due to 30% less nodes.

Table 3.5.: Adaptive refinement using Algorithm 3.1 (NT) versus uniform refinement for Example 1 (b) with 1% noise. CTR: Computation time reduction using adaptivity

noisy data	adaptive			uniform			CTR
	error	β	# nodes	error	β	# nodes	
(i)	0.300	947	20231	0.301	943	66049	54%
(ii)	0.276	151048	761	0.275	150342	1089	73%

Table 3.6 and the L-curve Figure 3.11 show about the same behavior of the proposed algorithm (Algorithm 3.1) as in the previous test cases.

Table 3.6.: Adaptive refinement using Algorithm 3.1 (NT) versus uniform refinement for Example 1 (b)(ii) with $\zeta = 1000$ and 1% noise

	Adaptive refinement		Uniform refinement	
k	# nodes	β^k	# nodes	β^k
0	25	10	25	10
0	81	10	81	10
0	253	10	289	10
1	253	229	289	229
1	761	229	1089	229
2	761	758	1089	757
3	761	2023	1089	2022
4	761	4798	1089	4796
5	761	10468	1089	10464
6	761	21246	1089	21238
7	761	40405	1089	40384
8	761	71171	1089	71095
9	761	110885	1089	110621
10	761	151048	1089	150341

Finally, we approach our last test configuration, namely the step function (c) (see the beginning of Section 3.6). The fact that, because of the different values of initial guess and exact solution on large parts of the boundary, only a very weak source condition (or possibly none at all) holds, makes this a challenging example. Additionally, the quadratic L^2-norm regularization term that we use here is the wrong type of regularization for such a piecewise constant function, which would require total variation regularization.

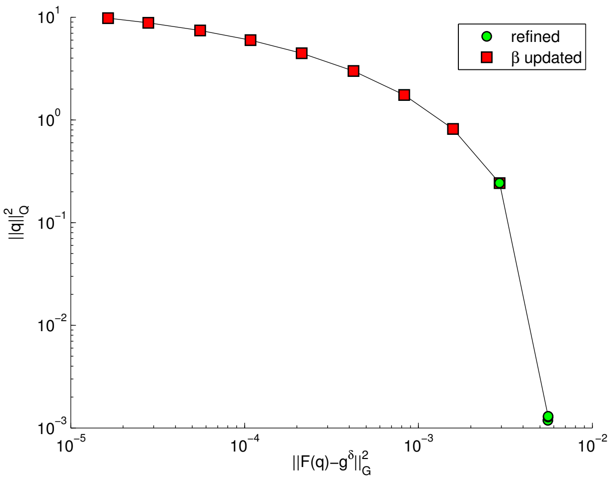

Figure 3.11.: L-Curve for Example 1 (b)(ii) with $\zeta = 1000$ and 1% noise using (NT)

Nevertheless, the shape of the control is reconstructed quite well by Algorithm 3.1, see Figure 3.12, and the corresponding control error of 0.450 is not larger than for the previous test configurations. Unfortunately, in the reconstruction of the source there appear some unwelcome peaks where the measurement points are located. Moreover, the mesh looks slightly overrefined, since it is refined on large parts of the domain where the solution is constant and nothing “is happening”. The corresponding control error amounts to 0.450. Still we save 29% of nodes and 61% of computation time.

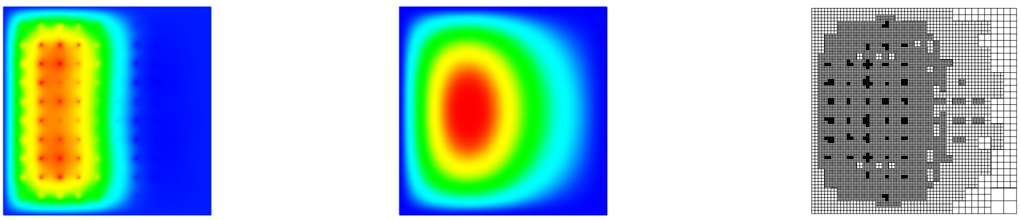

Figure 3.12.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 1 (c)(i), $\zeta = 1000$, 1% noise using (NT)

In the same manner as for the previous examples, the iterations and the resulting reduction of the misfit term can be read out of Table 3.7 and Figure 3.13.

The L^2-norm alternative (ii) for the same example (Example 1 (c), $\zeta = 1000$, 1% noise) leads to the reconstructions from Figure 3.14. The mesh, again, is coarser than in case of point measurements, but the jump is not detected very well either. This is not suprising however, since the L^2-projection is known to not perform well for the identification of discontinuities.

Table 3.8, Figure 3.15, and Table 3.9 confirm the general observations from the preceding

Table 3.7.: Adaptive refinement using Algorithm 3.1 (NT) versus uniform refinement for Example 1 (c)(i) with $\zeta = 1000$ and 1% noise

	Adaptive refinement		Uniform refinement	
k	# nodes	β^k	# nodes	β^k
0	25	10	25	10
0	55	10	81	10
0	189	10	289	10
0	673	10	1089	10
1	673	21	1089	21
1	1837	21	4225	21
1	5367	21	16641	21
2	5367	39	16641	39
3	5367	68	16641	68
4	5367	114	16641	114
5	5367	180	16641	180
6	5367	261	16641	260
-	11785	261	-	-

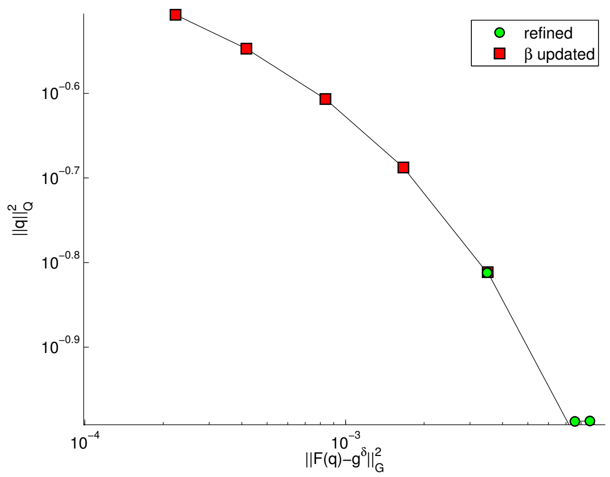

Figure 3.13.: L-Curve for Example 1 (c)(i) with $\zeta = 1000$ and 1% noise using (NT)

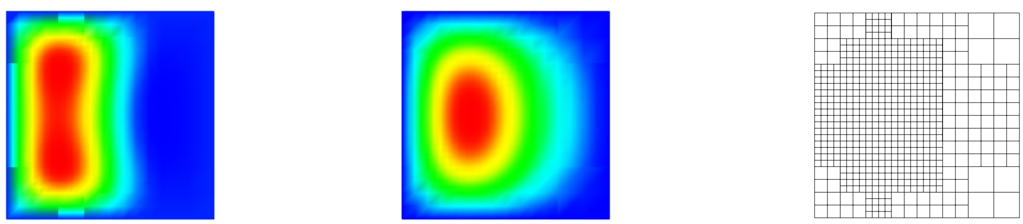

Figure 3.14.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 1 (c)(ii), $\zeta = 1000$, 1% noise using (NT)

test cases. Only the reduction of computation time with respect to uniform refinement is significantly smaller here (only 9%).

Table 3.8.: Adaptive refinement using Algorithm 3.1 (NT) versus uniform refinement for Example 1 (c)(ii) with $\zeta = 1000$ and 1% noise

	Adaptive refinement		Uniform refinement	
k	# nodes	β^k	# nodes	β^k
0	25	10	25	10
0	81	10	81	10
0	81	246	81	246
0	239	246	289	246
0	651	246	1089	246
0	651	672	1089	672
0	651	1486	1089	1486
0	651	3005	1089	3002
0	651	5679	1089	5669
0	651	10165	1089	10141
0	651	16867	1089	16798
0	651	27502	1089	27311
0	651	42541	1089	41946

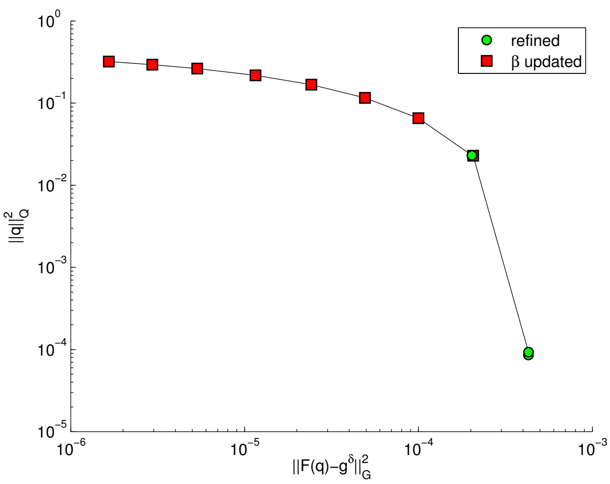

Figure 3.15.: L-Curve for Example 1 (c)(ii) with $\zeta = 1000$ and 1% noise using (NT)

In all the tests made so far, we restricted ourselves to 1% noise. We will now present the results for different noise levels by Example 1 (a) with $\zeta = 100$. In Figure 3.16 the reader finds the reconstructions of the control for 1%, 4%, and 8% noise. The intuitive expectations are indeed fulfilled: the larger the noise, the worse the reconstruction. The same holds for the reconstructions of the state in Figure 3.17.

Also when it comes to the adaptively refined meshes, the observations from Figure 3.18 satisfy the natural expectations, being: the more noise, the less refinement.

Table 3.10, where the relative control error, the obtained regularization parameter, as well

Table 3.9.: Adaptive refinement using Algorithm 3.1 (NT) versus uniform refinement for Example 1 (c) with 1% noise. CTR: Computation time reduction using adaptivity

noisy data	adaptive			uniform			CTR
	error	β	# nodes	error	β	# nodes	
(i)	0.450	261	11785	0.450	260	16641	61%
(ii)	0.461	42541	651	0.460	41946	1089	9%

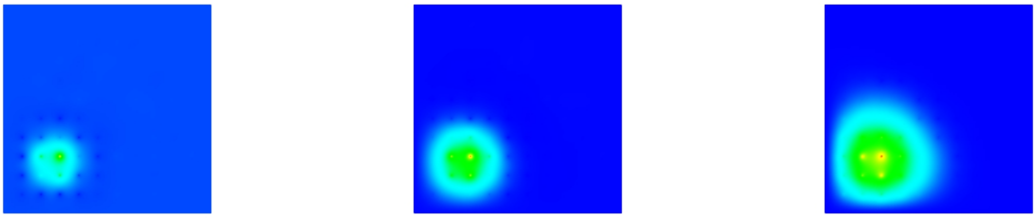

Figure 3.16.: Reconstructed control for Example 1 (a)(i), $\zeta = 100$ for different noise levels. FLTR: 1%, 4%, 8% noise using (NT)

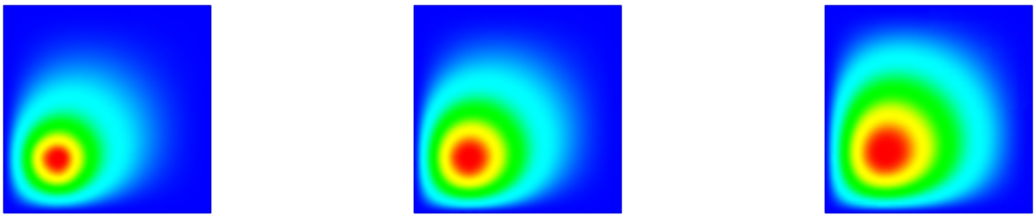

Figure 3.17.: Reconstructed state for Example 1 (a)(i), $\zeta = 100$ for different noise levels. FLTR: 1%, 4%, 8% noise using (NT)

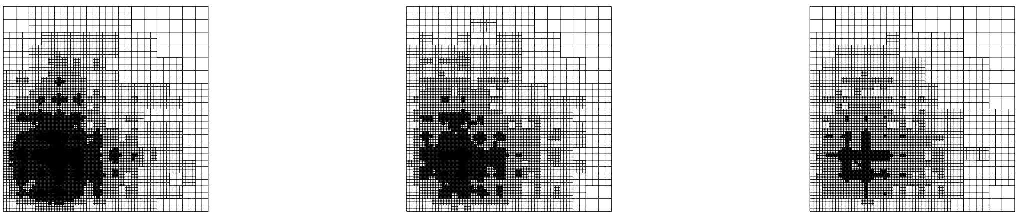

Figure 3.18.: Adaptively refined mesh for Example 1 (a)(i), $\zeta = 100$ for different noise levels. FLTR: 1%, 4%, 8% noise using (NT)

as the number of nodes in the obtained mesh is shown for 0.5%, 1%, 2%, 4%, and 8% noise, even reveals more coherences: the larger the noise, the larger the control error, the smaller the regularization parameter (i.e., the stronger the regularization), the coarser the mesh.

Table 3.10.: Different noise levels for Example 1 (a)(i), $\zeta = 100$ using (NT)

noise	error	β	# nodes
0.5%	0.387	13063	59151
1%	0.443	4999	31967
2%	0.588	1398	35995
4%	0.761	345	17065
8%	0.927	50	8191

3.6.2. Example 2

In this section we apply Algorithm 3.1 to Example 3.2 from Section 3.4, i.e., we consider the following PDE:

For $q \in L^2(\Omega)$, $q \geq 0$ almost everywhere in Ω find $u \in H_0^1(\Omega)$ such that

$$\begin{cases} -\Delta u + qu = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega\,. \end{cases}$$

The control constraint $q \geq 0$ is needed for theoretical purposes, but is neglected in the computations here. We try to make up for this by choosing the starting value for q appropriately, $q_0 = 0 \in L^2(\Omega)$.

We consider again the exact source distributions (a) and (b) cf. the beginning of Section 3.6. (Due to the possible lack of a source condition, which we mentioned in the previous example, we skip source distrbution (c) here.) The corresponding exact states simulated on a very fine mesh are shown in Figure 3.19.

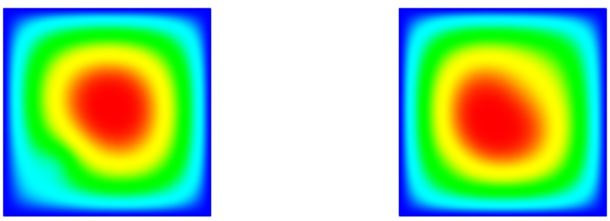

Figure 3.19.: Exact state $u^\dagger$ for Example 2. FLTR: configuration (a),(b)

The recontructions of the control and the state as well as the adaptively refined mesh obtained by Algorithm 3.1 (NT) can be seen in Figure 3.20 (for source (a)) and Figure 3.21 (for source

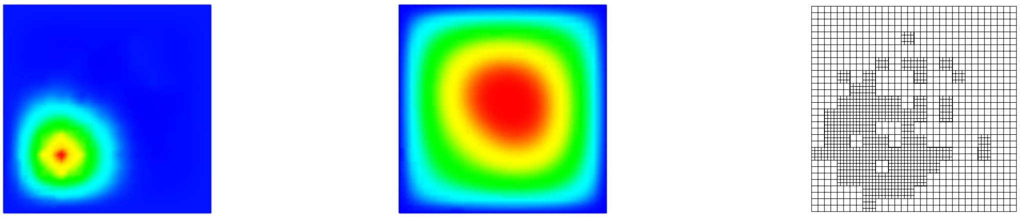

Figure 3.20.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 2 (a)(i), 1% noise using (NT)

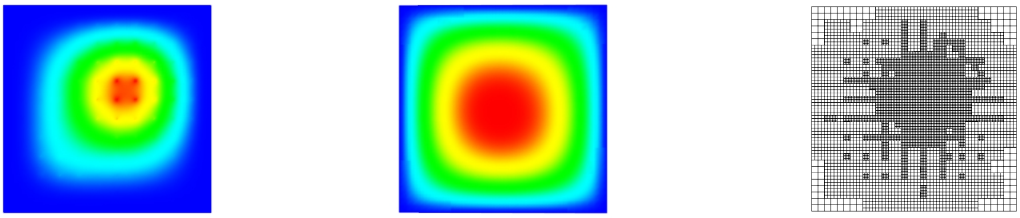

Figure 3.21.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 2 (b)(i), 1% noise using (NT)

(b)), both with 1% noise contained in 100 point measurements, see (i) at the beginning of Section 3.6.

The resulting relative control error (3.92), the number of mesh nodes, and the obtained regularization parameter β is indicated in Table 3.11 for adaptive and for comparison also uniform refinement. For the test configuration (a) we get 29% and for (b) 51% reduction of computation time using adaptivity. For (a) and also for (b), the obtained relative error and the regularization parameter are about the same. Unfortunately, we note that the control error is large, especially for (a).

Table 3.11.: Adaptive refinement using Algorithm 3.1 (NT) versus uniform refinement for Example 2 with 1% noise. CTR: Computation time reduction using adaptivity

Example	adaptive			uniform			CTR
	error	β	# nodes	error	β	# nodes	
(a)(i)	0.930	43504	2033	0.930	43154	1089	29%
(b)(i)	0.669	6027	7663	0.669	6022	16641	51%

To put this in perspective, we have to mention that for implementational reasons, in Example 2 (and also Example 3 in the following), we neglected the control error and only computed estimators for the state and adjoint errors (see Section 3.3 and Section 3.5). What makes the difference to Example 1 is the bilinearity of the PDE, meaning the coupled term qu. In cases, where the control enters linearly the right-hand side, the boundary conditions, or the intial condition and the control space is discretized in the same way as the state space, the error estimator for the error caused by discretization of the control vanishes. This is due to the optimality conditions, cf. [84, Remark 6.3].

3.6.3. Example 3

Finally, we consider the Helmholtz equation Example 3.3 and Example 3.4, where we set $\zeta = 1$.

For $q \in L^2(\Omega)$, $q \geq 0$ almost everywhere in Ω find $u \in H^1(\Omega)$ such that

$$-\Delta u - (1+q)u = f \quad \text{in } \Omega$$

under Dirichlet zero boundary conditions

$$u = 0 \quad \text{on } \partial\Omega\,,$$

or Neumann zero boundary conditions

$$\partial_n u = 0 \quad \text{on } \partial\Omega\,.$$

We consider two possible right-hand sides:

$$\begin{aligned}
f_1 &= -1 \\
f_2 &= 4\pi^2\left[\sin(2\pi x - \tfrac{\pi}{2})(\sin(2\pi y - \tfrac{\pi}{2}) + 1) + \sin(2\pi y - \tfrac{\pi}{2})(\sin(2\pi x - \tfrac{\pi}{2}) + 1)\right] \\
&\quad - (\sin(2\pi x - \tfrac{\pi}{2}) + 1)(\sin(2\pi y - \tfrac{\pi}{2}) + 1)\,.
\end{aligned}$$

Like for the previous example (Example 2), we choose the starting value $q_0 = 0 \in L^2(\Omega)$ for q and neglect the constraint $q \geq 0$.

We restrict ourselves to the test configuration (a), where we choose $c = 1$ (instead of $c = 10$ like in the previous examples) since from Section 3.4 we know that Example 3 is a challenging problem and that $\|q^\dagger\|_{L^2(\Omega)}$ should not be “too large”. For the same reason, we consider only a small perturbation of 0.01% noise.

In Figure 3.22 we show the exact states for $f = f_1$ and $f = f_2$ for Dirichlet and Neumann boundary conditions.

Figure 3.23 shows the recontructions of the control and the state as well as the adaptively refined mesh produced by Algorithm 3.1 (NT) for $f = f_1$ and Dirichlet zero boundary conditions. The relative control error (3.92) is 0.655. The corresponding results for Neumann boundary conditions are shown in Figure 3.24, where we have an error of 0.364. For $f = f_2$ and Dirichlet boundary conditions we obtain the reconstructions from Figure 3.26 with a relative error of 0.415 and finally, for $f = f_2$ and Neumann boundary conditions the proposed method yields the results from Figure 3.26 and a control error of 0.450.

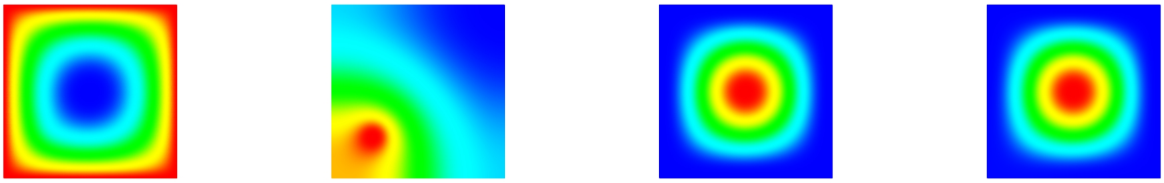

Figure 3.22.: Exact state $u^\dagger$ for Example 3 (a)(i) with 0.01% noise for different right-hand sides and boundary conditions. FLTR: $f = f_1$ with Dirichlet zero boundary conditions, $f = f_1$ with Neumann zero boundary conditions, $f = f_2$ with Dirichlet zero boundary conditions, $f = f_2$ with Neumann zero boundary conditions

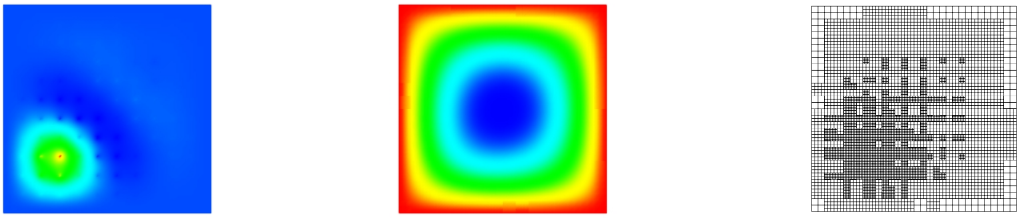

Figure 3.23.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 3 (a)(i), 0.01% noise, $f = f_1$, Dirichlet zero boundary conditions using (NT)

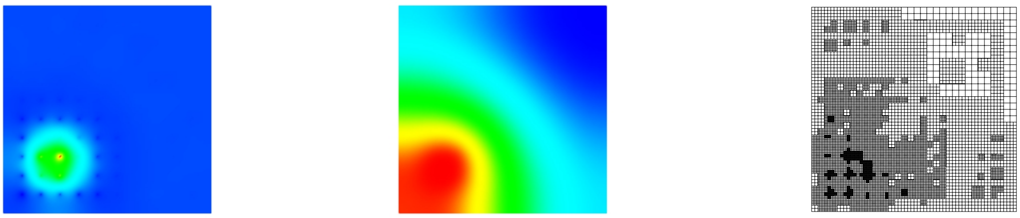

Figure 3.24.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 3 (a)(i), 0.01% noise, $f = f_1$, Neumann zero boundary conditions using (NT)

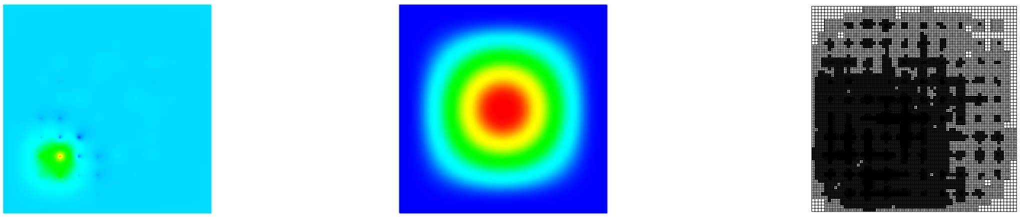

Figure 3.25.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 3 (a)(i), 0.01% noise, $f = f_2$, Dirichlet zero boundary conditions using (NT)

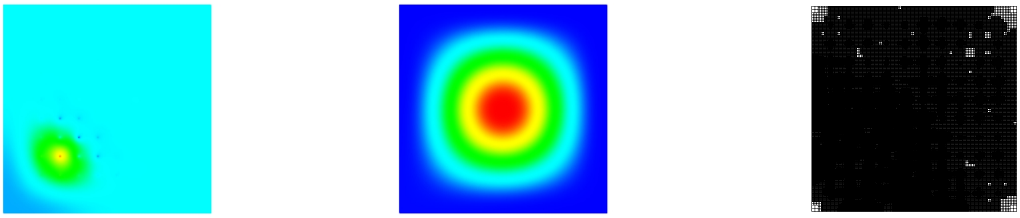

Figure 3.26.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 3 (a)(i), 0.01% noise, $f = f_2$, Neumann zero boundary conditions using (NT)

The sources seem to be reconstructed a little worse than for the two previous examples (Example 1 and 2) and the mesh refinement is not as concentrated on the source location (especially for the choice $f = f_2$). This is probably due to the smaller noise level on the one hand, which implies more iterations in general, and on the other hand due to the challenging Helmholtz equation itself.

3.7. Extension to parabolic equations

In practice, many inverse problems governed by partial differential equations additionally have a time component, which hails from time-dependence of the underlying PDE. We will concentrate on parabolic PDEs in the following. A detailed introduction to the theory of parabolic PDEs and their discretization can be found for instance in the textbook of Eriksson, Estep, Hansbo, and Johnson [29]. An introduction to inverse parabolic problems as well as examples and solution methods is given in [53, 112, 113]. Adaptive discretization and solution techniques as well as error estimates for parabolic PDEs and optimal control problems governed by parabolic PDEs have been investigated in [75, 84–87, 93, 104], for instance.

In this section we will extend the idea from Section 3.1–Section 3.5 to parabolic PDEs. This requires a reformulation of the PDE (2.15):

$$\begin{aligned} \partial_t u(t) + A(q(t), u(t)) &= f(t) \quad \text{for almost all } t \in I := (0, T) \\ u(0) &= u_0 \end{aligned} \tag{3.93}$$

with some end time point $T > 0$ and some given inital data $u_0 \in L^2(\Omega)$. For simplicity we choose $V = W = H_0^1(\Omega)$ and the "new" state space

$$\tilde{V} := \{\varphi \in L^2(I, H_0^1(\Omega)) |\ \partial_t \varphi \in L^2(I, H_0^1(\Omega)^*)\}\,.$$

This is a special case of the definition in [85]. The notation $\varphi \in L^2(I, H)$ connotes that φ lives in $L^2(I)$ in time with values in H. Note that the space $\tilde{V}$ is continuously embedded in $C(\bar{I}, L^2(\Omega))$ with $\bar{I} = [0, T]$. We will only consider examples where the control is constant in time, i.e., $q(t) = q$ such that the control space doesn't need to be changed. For setting up the weak formulation of the state equation we again denote the $L^2(\Omega)$ inner product by $(\cdot,\cdot) := (\cdot,\cdot)_{L^2(\Omega)}$ and, due to Remark 3.11, we introduce the inner product

$$(\varphi, \psi)_I := \int_I \langle \varphi(t), \psi(t)\rangle_{W^*,W}\ dt = \int_I (\varphi(t), \psi(t))_{L^2(\Omega)}\, dt\,. \tag{3.94}$$

Consistent with (2.16) we define the semilinear form

$$\tilde{a}(q,u)(\varphi) := (A(q,u), \varphi)_I = \int_I \langle A(q(t), u(t)), \varphi(t)\rangle_{W^*,W}\, dt = \int_I a(q(t), u(t))(\varphi(t))\, dt\,, \tag{3.95}$$

such that the weak formulation can be written as

$$(\partial_t u, \varphi)_I + \tilde{a}(q,u)(\varphi) + (u(0), \varphi(0)) = (f, \varphi)_I + (u_0, \varphi(0)) \qquad \forall \varphi \in \tilde{V}\,. \tag{3.96}$$

The cost functional $J(q, u)$ in this section is defined by

$$J(q,u) := I(u) + \frac{1}{\beta}\,\|q - q_0\|_Q^2 \quad \text{with} \quad I(u) = \int_I J_1(u(t))\, dt + J_2(u(T)) \tag{3.97}$$

and $J_1 : V = H_0^1(\Omega) \to \mathbb{R}$, $J_2 : L^2(\Omega) \to \mathbb{R}$. The Lagrange functional is consistently given by

$$\tilde{\mathcal{L}}(q,u,z) = J(q,u) + (f - \partial_t u, z)_I - \tilde{a}(q,u)(z) + (u_0 - u(0), z(0))\,. \tag{3.98}$$

The equations from Section 2.2 and Section 3.5 for $\mathcal{L} := \tilde{\mathcal{L}}$ then read as presented in Appendix B.

3.7.1. Space-time discretization

The numerical treatment of parabolic PDEs (3.93) requires a discretization in time additional to the space discretization presented in Section 2.3. As mentioned at the beginning of this chapter, the control is constant in time, such that we only need to discretize the state variables in time. To define such a semidiscretization, we partition the interval $\bar{I} = [0, T]$ into M subintervals I_m of size k_m ($m = 1, \dots M$):

$$[0, T] = \{0\} \cup I_1 \cup I_2 \cup \cdots \cup I_M\,,$$

where $I_m = (t_{m-1}, t_m]$ and $0 = t_0 < t_1 < \cdots < t_M = T$. We define the time discretization parameter k as a piecewise constant function $k|_{I_m} = k_m$ for all $m = 1, 2, \ldots, M$. Then we can define the semidiscretized space V_k^r by

$$\tilde{V}_k^r := \{v_k \in L^2(I, L^2(\Omega)) |\ v_k|_{I_m} \in \mathcal{P}_r(I_m, H_0^1(\Omega)),\ m = 1, 2, \ldots, M \text{ and } v_k(0) \in L^2(\Omega)\}$$

(cf. [84]), where $\mathcal{P}_r(I_m, V)$ denotes the space of polynomials up to order r defined on the interval I_m with values in V. By means of these spaces we formulate the discretized state equation using the Discontinuous Galerkin method “dG(r)” (cf. [9]) of order r as follows.

Let $v_{k,m}^+ := \lim_{t\to 0, t>0} v_k(t_m + t)$ be the limit “from the right-hand side”, $v_{k,m}^- := \lim_{t\to 0, t>0} v_k(t_m - t) = v_k(t_m)$ the limit “from the left-hand side” and $[v_k]_m := v_{k,m}^+ - v_{k,m}^-$ the jump in between. For given $q \in Q$ find $u_k \in \tilde{V}_k^r$ such that

$$\sum_{m=1}^{M} (\partial_t u_k, \varphi)_{I_m} + \tilde{a}(q, u_k)(\varphi) + \sum_{m=0}^{M-1} ([u_k]_m, \varphi_m^+) + (u_{k,0}^-, \varphi_0^-) = (f, \varphi)_I + (u_0, \varphi_0^-) \tag{3.99}$$

for all $\varphi \in \tilde{V}_k^r$. We consistently define the Lagrangian as

$$\begin{aligned}\tilde{\mathcal{L}}(q, u_k, z_k) := J(q, u_k) + (f, z_k)_I - \sum_{m=1}^{M} (\partial_t u_k, z_k)_{I_m} \\ - \tilde{a}(q, u_k)(z_k) - \sum_{m=0}^{M-1} ([u_k]_m, z_{k,m}^+) + (u_0 - u_{k,m}^-, z_{k,0}^-)\,.\end{aligned}$$

The time-discretized version of (2.21) can be formulated as

$$\min_{q \in \mathcal{D} \subset Q} j_k(q) := J(q, u_k)\,,$$

where u_k solves (3.99). Similarly to the space discretization (Section 2.3), semi-discretized versions of the equations (2.25), (2.26), (2.29), (2.30) and (2.31) from Section 2.2 and Section 3.5 are obtained by simply replacing $\mathcal{L}(q, u, z)$ by $\tilde{\mathcal{L}}(q, u_k, z_k)$ and the spaces V and W by $\tilde{V}_k^r$.

Now, we combine the Continuous Galerkin discretization “cG(s)” of order s from Section 2.3 with the dG(r) discretization in time such that the fully “cG(s)dG(r)” discretized state equation reads

$$\sum_{m=1}^{M} (\partial_t u_{kh}, \varphi)_{I_m} + \tilde{a}(q_h, u_{kh})(\varphi) + \sum_{m=0}^{M-1} ([u_{kh}]_m, \varphi_m^+) + (u_{kh,0}^-, \varphi_0^-) = (f, \varphi)_I + (u_0, \varphi_0^-)$$

for all $\varphi \in \tilde{V}_{k,h}^{r,s}$, where $\tilde{V}_{k,h}^{r,s}$ is the fully discrete space-time finite element space

$$\tilde{V}_{k,h}^{r,s} := \{v_{kh} \in L^2(I, L^2(\Omega)) |\ v_{kh}|_{I_m} \in \mathcal{P}_r(I_m, V_h^{s,m}),\ m = 1, 2, \ldots, M \text{ and } v_{kh}(0) \in V_h^{s,0}\} \subset \tilde{V}_k^r\,.$$

In here, $V_h^{s,m}$ denotes the finite element space from Section 2.3 associated with the time end point t_m.

Again, one gets the fully-discretized versions of j and the equations (2.25), (2.26), (2.27), (2.29), (2.30), (2.31) from Section 2.2 and Section 3.5 by replacing the index k by kh and $\tilde{V}_k^r$ by $\tilde{V}_{k,h}^{r,s}$ in the semidiscrete formulation.

As special case of cG(s)dG(r) discretization with $r = 0$, we consider the implicit Euler time stepping scheme (see, e.g., [106]). For a detailed presentation of the resulting state, dual and tangent equations from Section 3.5 we refer the reader to Appendix C. A single time step is solved by means of a Newton solver for treating the nonlinearities and by applying multigrid methods to the resulting linear subproblems. For details we refer to [8].

3.7.2. Space-time mesh refinement

Having established a space-time discretization for (3.93), the next step is the derivation of a posteriori error estimators (cf. Section 2.4 and 3.3), which allow us to estimate the error in space as well as the error in time, since our goal is to modify Algorithm 3.1, insofar as we either refine in space, or in time or in both in step 15 and 19.

DWR error estimators for parabolic problems

In order to extend the results from the stationary case (cf. Section 2.4 and 3.3) to parabolic PDEs, we use the results from [85, 104]. For a given quantity of interest functional E, there hold the same results as in Section 3.3 with h replaced by k, but in order to distinguish between time and space error, we split the total error as follows

$$E(q,u) - E(q_{hk}.u_{hk}) = E(q,u) - E(q_k,u_k) + E(q_k,u_k) - E(q_{kh},u_{kh}) \approx \eta_k^E + \eta_h^E \,,$$

and define analogously to (2.39), the auxiliary Lagrangian

$$\tilde{\mathcal{M}} \colon \tilde{X} \times \tilde{X} \to \mathbb{R} \,, \quad \tilde{\mathcal{M}}(x,\xi) \coloneqq E(q,u) + \mathcal{L}'(x)(\xi) \,,$$

where $\tilde{X} \coloneqq Q \times \tilde{V} \times \tilde{V}$. Adapted from [85], for continuous, semi-discrete and fully-discrete stationary points $(x,\xi) \in \tilde{X} \times \tilde{X}$, $(x_k,\xi_k) \in \tilde{X}_k \times \tilde{X}_k$ with $\tilde{X}_k \coloneqq Q \times \tilde{V}_k^r \times \tilde{V}_k^r$ and $(x_{kh},\xi_{kh}) \in \tilde{X}_{kh} \times \tilde{X}_{kh}$ with $\tilde{X}_{kh} \coloneqq Q \times \tilde{V}_{k,h}^{r,s} \times \tilde{V}_{k,h}^{r,s}$ of $\tilde{\mathcal{M}}$, there holds

$$\begin{aligned} E(q,u) - E(q_k,u_k) &= \frac{1}{2}\tilde{\mathcal{M}}'(x_k,\xi_k)(x - \widehat{x}_k, \xi - \widehat{\xi}_k) + \mathbf{R}_k \\ E(q,u) - E(q_{kh},u_{kh}) &= \frac{1}{2}\tilde{\mathcal{M}}'(x_{kh},\xi_{kh})(x - \widehat{x}_{kh}, \xi - \widehat{\xi}_{kh}) + \mathbf{R}_h \,, \end{aligned}$$

where $(\widehat{x}_k, \widehat{\xi}_k)$ and $(\widehat{x}_{kh}, \widehat{\xi}_{kh})$ are arbitrary and $\mathbf{R}_k$ and $\mathbf{R}_h$ are remainder terms of the form (2.40).

So for $E \coloneqq I$ we can estimate the errors $\eta_k^I \approx i(\beta) - i_k(\beta) = I(q,u) - I(q_k,u_k)$ and $\eta_h^I \approx i(\beta) - i_{kh}(\beta) = I(q,u) - I(q_{kh},u_{kh})$. An error estimator for the cost functional J (or j) can be derived in the same way (cf. Section 2.4, [85]) such that there holds

$$\begin{aligned} J(q,u) - J(q_k,u_k) &= \frac{1}{2}\tilde{\mathcal{L}}'(x_k)(x - \widehat{x}_k) + \mathbf{R}_k \approx \eta_k^J \\ J(q,u) - J(q_{kh},u_{kh}) &= \frac{1}{2}\tilde{\mathcal{L}}'(x_{kh})(x - \widehat{x}_{kh}) + \mathbf{R}_h \approx \eta_h^J \,. \end{aligned}$$

for continuous, semi-discrete and fully-discrete stationary points $x \in \tilde{X}$, $x_k \in \tilde{X}_k$ and $x_{kh} \in \tilde{X}_{kh}$ of $\tilde{\mathcal{L}}$. Since $i'(\beta) = K(q, q_1) = -\frac{2}{\beta^2}(q, q_1)_Q$ (cf. Section 3.3) and q, q_1 are constant in time there holds $q_k = q$, $q_{1,k} = q_1$, such that $i'(\beta) - i'_k(\beta) = K(q, q_1) - K(q_k, q_{1,k}) = 0$. This implies that we can set $\eta_k^K = 0$ and that the results from Section 3.3 concerning i' still hold if we replace the index h by hk.

For details on the numerical realization of the error estimators we refer to [85, 104].

Refinement criterion

As in Algorithm 3.1 we base our decision whether to refine or not on (3.9), (3.19) and (3.38)–(3.41). But the question is whether to refine in time, in space or in both. To answer this question, we adopt the concept from [85, 104] and aim to obtain a discretization fulfilling

$$|\eta_k^I| \approx |\eta_h^I| \quad \text{and} \quad |\eta_k^J| \approx |\eta_h^J| \,.$$

For this purpose we define an "equilibration factor" $e \in (1, 5)$ and formulate Algorithm 3.8, which replaces Step 15 and Step 19 in Algorithm 3.1. With $\eta = \eta_k + \eta_h$ being the error estimator corresponding to the violated condition (3.9), (3.19), (3.38), (3.39), (3.40) or (3.41) (i.e., $\eta \in \{\eta^I, \eta^J, \eta^K\}$).

Algorithm 3.8: Space-time refinement

1: **if** $|\eta_h| \geq e|\eta_k|$ **then**
2: Refine spatial discretization.
3: **else if** $|\eta_k| \geq e|\eta_h|$ **then**
4: Refine temporal discretization.
5: **else**
6: Refine spatial and temporal discretization.

3.7.3. Numerical results

Example 4

In order to test how the proposed method formulated in Algorithm 3.1 performs on parabolic problems, we consider the following initial-boundary value problem:

For given $q \in L^2(\Omega)$ find $u \in \tilde{V}$ such that

$$\begin{cases} \partial_t u - \Delta u + \zeta u^3 = q & \text{in } \Omega \times I \\ u = 0 & \text{on } \partial\Omega \times I \\ u = 0 & \text{on } \Omega \times \{0\} \,. \end{cases}$$

We choose the domain and the time interval as $\Omega = (0, 1)^2 \subset \mathbb{R}^2$ and $I = (0, 1)$. The noisy data is simulated via point functionals in $n_m = 100$ uniformly distributed points ξ_i, $i = 1, 2, \ldots, n_m$

and the end point, perturbed by some percentage p of uniformly distributed random noise. We set $J_1 = 0$ in (3.97), and $J_2(u(T)) = \|u(T) - g^\delta\|_G^2$ with observation space $G = \mathbb{R}^{n_m}$. Like in Section 3.6, the noisy data is created via $g_i^\delta = g(\xi_i)(1 + \varepsilon_i p)$ for $i = 1, \ldots, n_m$, where $\varepsilon_i \in (-1, 1)$ are random numbers and the exact data g is simulated as $g = u^\dagger(T)$, where $u^\dagger$ is the solution of the given PDE on a very fine mesh with 16641 space nodes (and equally sized quadratic cells) and 6000 equidistant time steps. The equilibration factor, affecting whether we refine in space or in time (see Algorithm 3.8), is chosen as $e = 20$.

We will use the same type of figures and tables as in Section 3.6. Therefore, we will not go into detail concerning a description of the means of presentation.

We consider the same exact source distributions (a), (b), and (c) as in Section 3.6 , which are visualized in Figure 3.1.

The corresponding exact states at the end point are presented in Figure 3.27.

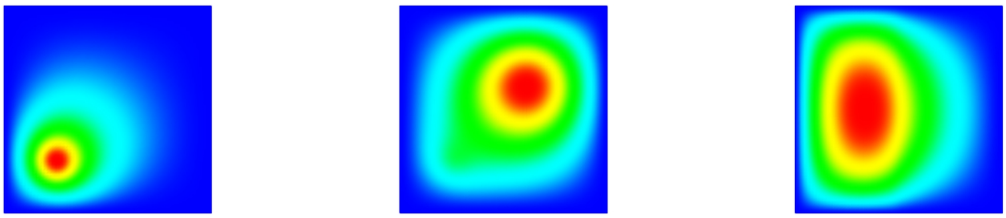

Figure 3.27.: Exact state $u^\dagger(T)$ for Example 4. FLTR: configuration (a),(b),(c)

We start with a space mesh of 25 nodes (like in Section 3.6) and a time mesh of 10 equally sized time steps. One refinement step in time looks as follows: a selected cell is divided into two cells of half the size.

For the source distribution (a) and 1% noise, Figure 3.28 shows the reconstructions of the source and of the state as well as the adaptively refined space mesh at the end point. In Figure 3.29 the reader can find the resulting time mesh.

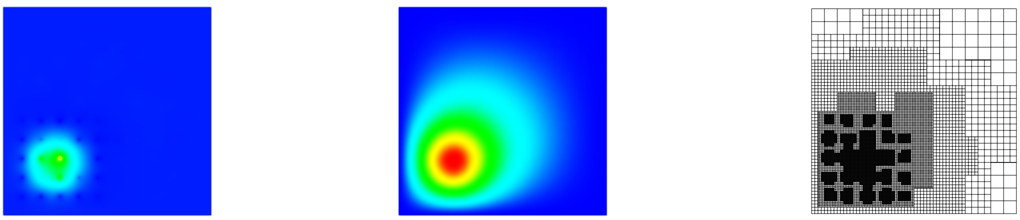

Figure 3.28.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 4 (a)(i), 1% noise using (NT)

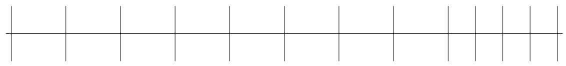

Figure 3.29.: Adaptively refined time mesh for Example 4 (a)(i) with 1% noise

The refinement and β iteration steps can be tracked by taking a look at Table 3.12 and Figure 3.30. We can see that the algorithm starts with a refinement step in space and time, then refines three times in space only, then updates β six times, refines in space twice again before it increases β three times and terminates. The time mesh is refined – as one would expect – at the end of the interval. The space mesh is refined where the source is located and at the measurement points, similar to the results from Section 3.6 .

Table 3.12.: Adaptive refinement using Algorithm 3.1 (NT) versus uniform refinement for Example 4 (a)(i) with 1% noise

	Adaptive refinement			Uniform refinement		
k	# nodes in space	# nodes in time	β^k	# nodes in space	# nodes in time	β^k
0	25	10	10	25	10	10
0	81	12	10	81	20	10
0	277	12	10	289	20	10
0	797	12	10	1089	20	10
1	2137	12	10	4225	20	29
2	2137	12	29	4225	20	10
3	2137	12	68	4225	20	68
4	2137	12	146	4225	20	146
5	2137	12	299	4225	20	299
6	2137	12	606	4225	20	606
6	5011	12	606	16641	20	606
6	12521	12	606	66049	20	606
7	12521	12	1157	66049	20	1157
8	12521	12	1964	66049	20	1964
9	12521	12	2825	66049	20	2825

Compared to uniform refinement, we save 81% of space nodes, 40% of time steps, while both alternatives produce the same regularization parameter. The L-curve graph Figure 3.30 (also see Section 3.6 for a more detailed description) for the parabolic Example 4 has a typical shape.

Beside the single Gauss distribution (a) we also tested Algorithm 3.1 together with Algorithm 3.8 by the example sources (b) and (c) from Section 3.6 (also with 1% noise). The results can be seen in Figure 3.31 and Figure 3.32. As for the time mesh, we only mention that the algorithm did not refine in time at all for the source (b), and for (c) it refined exactly like for (a), i.e., from 10 to 12 time steps at the end. Apart from that, the reconstructions and the space mesh look similar to the elliptic case (see Section 3.6 , Example 1).

In Table 3.13, we present a collection of computation time, regularization parameter, and space and time nodes for adaptive as well as uniform refinement for Example 4 with the three source distributions (a), (b), and (c). In all three cases, we save at least 31% of computation

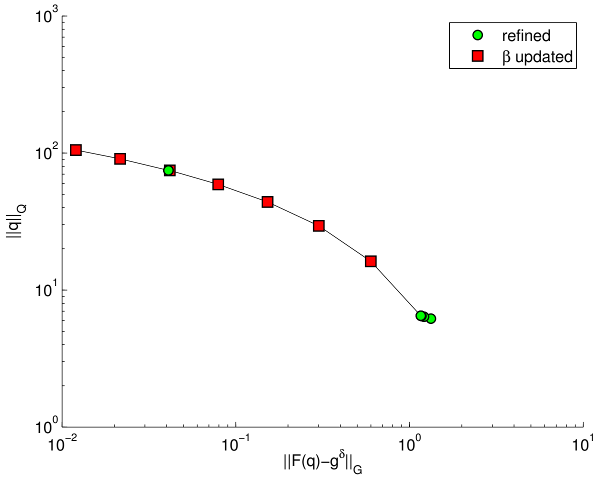

Figure 3.30.: L-Curve for Example 4 (a)(i) with 1% noise

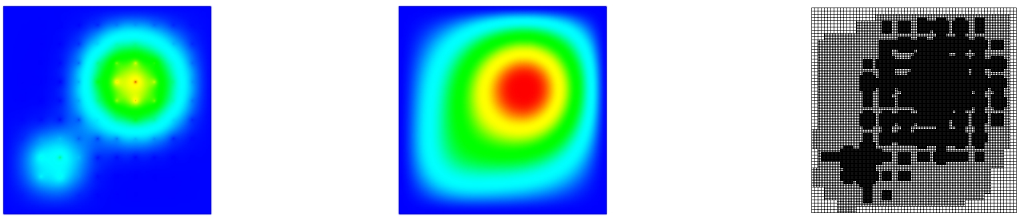

Figure 3.31.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 4 (b)(i), 1% noise using (NT)

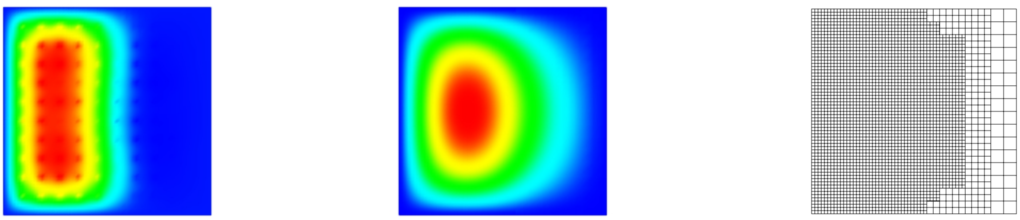

Figure 3.32.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 4 (c)(i), 1% noise using (NT)

time. Again, as for all the previous test setting, both, adaptive and uniform refinement yield the same β and a similar relative control error (3.92).

Table 3.13.: Adaptive refinement using Algorithm 3.1 (NT) versus uniform refinement for Example 4 with 1% noise. CTR: Computation time reduction using adaptivity

Example	adaptive				uniform				CTR
	error	β	space nodes	time nodes	error	β	space nodes	time nodes	
(a)(i)	0.466	2825	12521	12	0.466	2825	66049	20	84%
(b)(i)	0.350	272	35237	10	0.350	272	66049	20	31%
(c)(i)	0.452	251	3231	12	0.452	251	3231	20	49%

Testing configuration (a) again with a larger noise level $p = 4\%$, we obtain a worse reconstruction, as one can see at first glance by looking at Figure 3.33. The resulting error of 0.789 confirms this visual impression. It is also not surprising that – just like in the elliptic case – the algorithm refines much less (1913 space nodes) and stops with a smaller regularization parameter ($\beta = 175$) than with smaller noise.

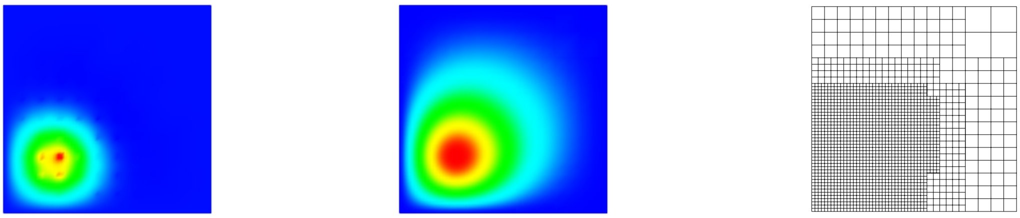

Figure 3.33.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 4 (a)(i), 4% noise using (NT)

4. Iteratively regularized Gauss-Newton methods

The implementation of variational regularization methods (i.e., regularization methods which approximate the exact solution by a (stable) solution of a minimization problem), such as Tikhonov regularization, requires an iterative algorithm for the minimization of the objective functional. Thus, it suggests itself to consider such iterative (mostly Newton type, cf. Section 2.5.2 and 2.5.3) methods directly and to make the regularization parameter β grow in the course of the iterations. Based on the existing literature about iterative regularization methods (see, e.g., [5, 69]) we will consider adaptive discretizations for these methods with the goal of saving computational effort compared to Algorithm 3.1. This is not an unrealistic hope due to weaker accuracy requirements and possibly associated coarser discretizations in the beginning of the iteration. The stopping index plays the role of an additional regularization parameter, which will also be chosen according to the discrepancy principle.

For the stable solution of (2.1) with noisy data, we will consider two types of iterative Newton type methods: a reduced form of the iteratively regularized Gauss-Newton method (IRGNM) (cf. Section 4.1) and all-at-once formulations, where we treat the PDE and the measurement equation at the same time. More specifically, as all-at-once formulations, we consider a least squares type (cf. Section 4.2.1) and a Generalized Gauss-Newton type [18] form of the IRGNM (cf. Section 4.2.2). The reduced formulation will not be implemented and serves mainly as theoretical basis for the all-at-once formulations. Since the analysis in Section 4.1 is done for a general forward operator, the obtained results hold for general inverse problems formulated as operator equations, and can also be applied to a system of equations as considered in the all-at-once formulations Section 4.2.

Most of the results from this chapter can be found in [64, 66].

4.1. Reduced form of the discretized IRGNM

In this section we consider the iteratively regularized Gauss-Newton method (IRGNM) (2.44) from Section 2.5.2. For a better understanding, we first discuss the method on the continuous level, i.e., in each iteration step we seek q_k^δ as the solution to

$$\min_{q\in Q} \|F'(q^{\delta,k-1})(q-q^{\delta,k-1})+F(q^{\delta,k-1})-g^\delta\|_G^2+\frac{1}{\beta_k}\|q-q_0\|_Q^2\,, \tag{4.1}$$

which leads to the iteration rule

$$
\begin{aligned}
q^{\delta,k} = q^{\delta,k-1} - \Big(F'(q^{\delta,k-1})^* F'(q^{\delta,k-1}) + \frac{1}{\beta_k}\,\mathrm{id}\Big)^{-1} \\
\cdot \Big(F'(q^{\delta,k-1})^*(F(q^{\delta,k-1}) - g^\delta) + \frac{1}{\beta_k}(q^{\delta,k-1} - q_0)\Big)
\end{aligned}
\tag{4.2}
$$

(see, e.g., [4, 5, 17, 50, 54, 55, 69]).

Remark 4.1. *In case the reader might wonder why the condition $q \in \mathcal{D}$ has disappeared in this formulation, the author would like to refer to Remark 2.3 and 2.4. In Theorem 4.2 we will show that all iterates stay in the ρ-neighborhood of q_0, where the tangential cone condition Assumption 4.4 holds. This implies $q^{\delta,k} \in \mathcal{D}$, provided that this neighborhood lies in $\mathcal{D}$, which we will claim in the following assumption.*

Througout this chapter, we assume the existence of an exact solution to the inverse problem in a neighborhood of the inital guess q_0, i.e.,

Assumption 4.1. *There exist $\rho > 0$ and $q^\dagger \in \mathcal{B}_\rho(q_0) \subset \mathcal{D}$, such that $q^\dagger$ solves* (2.1).

We would like to point out that the assumption $\mathcal{B}_\rho(q_0) \subset \mathcal{D}$ is nontrivial, since it implies that $\mathcal{D}$ contains an open set/has nonempty interior. Although this assumption is not necessary in all cases, depending on the PDE (see Section 3.4, Example 2), we keep this condition and refer, once again, to Remark 2.3 and Remark 2.4.

The regularization parameter β_k and the overall stopping index k_* have to be chosen in an appropriate way in order to guarantee convergence. We will here use an inexact Newton/discrepancy principle type strategy, as it has been shown to yield convergence of the IRGNM even in a Banach space setting in [71], see also [52] for a convergence analysis in an even more general setup but with different parameter choice strategies for β_k and k_*.

Again, as in Chapter 3, our aim is to consider adaptively discretized versions of the formulations (4.2) defined by replacing the spaces Q, V, W by finite dimensional counterparts Q_h, V_h, W_h (see Section 2.3). These should be sufficiently precise so that the convergence results from the continuous setting can be carried over, but we save computational effort by using degrees of freedom only where really necessary. For this purpose we will again make use of goal oriented error estimators (cf. Section 2.4 and 3.3), that control the error in some quantities of interest. Different from Chapter 3, we do not treat the nonlinear problem directly here, but use an iterative solution algorithm, the iteratively regularized Gauss-Newton method (4.2) and treat a sequence of linearized problems instead.

We start with a detailed description of a single iteration step (4.2) for fixed (discretized) previous iterate $q^{\delta,k-1} = q_{\text{old}} \in Q$ in the continuous setting. Afterwards, we transfer the formulation to the discretized setting actually used in the computations and specify the quantities of interest required in error estimation and adaptive refinement.

As in Section 2.2 we can formulate the optimization problem (4.1) as optimal control problem using the decomposition of the forward operator into solution and observation map.

$$
\min_{(q,u_{\text{old}},v)\in Q\times V\times V} \|C'(u_{\text{old}})(v) + C(u_{\text{old}}) - g^\delta\|_G^2 + \frac{1}{\beta_k}\|q - q_0\|_Q^2 \tag{4.3}
$$

$$\text{s.t.} \qquad a(q_{\text{old}}, u_{\text{old}})(\varphi) = f(\varphi) \qquad \forall \varphi \in W\,, \tag{4.4}$$
$$a'_u(q_{\text{old}}, u_{\text{old}})(v, \varphi) + a'_q(q_{\text{old}}, u_{\text{old}})(q - q_{\text{old}}, \varphi) = 0 \qquad \forall \varphi \in W\,, \tag{4.5}$$

since for a solution $q^{\delta,k}$ of (4.1) the triple $(q^{\delta,k}, S(q_{\text{old}}), S'(q_{\text{old}})(q^{\delta,k} - q_{\text{old}}))$ solves (4.3)–(4.5).

In most of this section we omit the superscript δ (denoting dependence on the noisy data) to be able to better indicate the difference between continuous and discretized quantities.

In order to apply Algorithm 3.1 to a sequence of problems of the type (4.1) we need other quantities of interest than $i(\beta)$, $i'(\beta)$, and j_β from Chapter 3, namely quantities with respect to the "old" iterate from step $k-1$ as well as the "current" iterate from step k. To this purpose, for fixed q_{old} (the previous iterate) we consider the following quantities of interest

$$\begin{aligned}
I_1\colon\ & Q \times Q \times \mathbb{R} \to \mathbb{R}\,, && (q_{\text{old}}, q, \beta) && \mapsto \|F'(q_{\text{old}})(q - q_{\text{old}}) + F(q_{\text{old}}) - g^\delta\|_G^2 + \frac{1}{\beta}\|q - q_0\|_Q^2 \\
I_2\colon\ & Q \times Q \to \mathbb{R}\,, && (q_{\text{old}}, q) && \mapsto \|F'(q_{\text{old}})(q - q_{\text{old}}) + F(q_{\text{old}}) - g^\delta\|_G^2 \\
I_3\colon\ & Q \to \mathbb{R}\,, && q_{\text{old}} && \mapsto \|F(q_{\text{old}}) - g^\delta\|_G^2 \\
I_4\colon\ & Q \to \mathbb{R}\,, && q && \mapsto \|F(q) - g^\delta\|_G^2\,,
\end{aligned} \tag{4.6}$$

which, for $u_{\text{old}} \in V$ and $v \in V$ solving (4.4) and (4.5), and $u \in V$ solving

$$a(q, u)(\varphi) = f(\varphi) \qquad \forall \varphi \in W\,, \tag{4.7}$$

satisfy the identities

$$\begin{aligned}
I_1(q_{\text{old}}, q, \beta) &= \|C'(u_{\text{old}})(v) + C(u_{\text{old}}) - g^\delta\|_G^2 + \frac{1}{\beta}\|q - q_0\|_Q^2 \\
I_2(q_{\text{old}}, q) &= \|C'(u_{\text{old}})(v) + C(u_{\text{old}}) - g^\delta\|_G^2 \\
I_3(q_{\text{old}}) &= \|C(u_{\text{old}}) - g^\delta\|_G^2 \\
I_4(q) &= \|C(u) - g^\delta\|_G^2\,.
\end{aligned}$$

The (continuous) quantities of interest in the k-th iteration step are then defined as follows: For a solution $(q^k, u^k_{\text{old}}, v^k)$ of (4.3)–(4.5) for given $q_{\text{old}} = q^k_{\text{old}}$ and $\beta = \beta_k$, and u^k fulfilling

$$a(q^k, u^k)(\varphi) = f(\varphi) \qquad \forall \varphi \in W$$

in the k-th iteration, let

$$\begin{aligned}
I_1^k &\coloneqq I_1(q^k_{\text{old}}, q^k, \beta_k) = \|F'(q^k_{\text{old}})(q^k - q^k_{\text{old}}) + F(q^k_{\text{old}}) - g^\delta\|_G^2 + \frac{1}{\beta_k}\|q^k - q_0\|_Q^2 \\
&= \|C'(u^k_{\text{old}})(v^k) + C(u^k_{\text{old}}) - g^\delta\|_G^2 + \frac{1}{\beta_k}\|q^k - q_0\|_Q^2 \\
I_2^k &\coloneqq I_2(q^k_{\text{old}}, q^k) = \|F'(q^k_{\text{old}})(q^k - q^k_{\text{old}}) + F(q^k_{\text{old}}) - g^\delta\|_G^2 = \|C'(u^k_{\text{old}})(v^k) + C(u^k_{\text{old}}) - g^\delta\|_G^2 \\
I_3^k &\coloneqq I_3(q^k_{\text{old}}) = \|F(q^k_{\text{old}}) - g^\delta\|_G^2 = \|C(u^k_{\text{old}}) - g^\delta\|_G^2 \\
I_4^k &\coloneqq I_4(q^k) = \|F(q^k) - g^\delta\|_G^2 = \|C(u^k) - g^\delta\|_G^2\,.
\end{aligned} \tag{4.8}$$

To formulate the quantities of interest (4.6)/(4.8) for a discrete setting, we consider finite element spaces Q_h, V_h, W_h as in Section 2.3 and formulate the discretized version of the optimal control problem (4.3) for given $q_{\text{old}} \in Q_h$.

$$\min_{(q,u_{\text{old}},v)\in Q_h\times V_h\times V_h} \|C'(u_{\text{old}})(v) + C(u_{\text{old}}) - g^\delta\|_G^2 + \frac{1}{\beta}\|q - q_0\|_Q^2 \tag{4.9}$$

subject to

$$a(q_{\text{old}}, u_{\text{old}})(\varphi) = f(\varphi) \qquad \forall \varphi \in W_h \tag{4.10}$$

$$a_u'(q_{\text{old}}, u_{\text{old}})(v,\varphi) + a_q'(q_{\text{old}}, u_{\text{old}})(q - q_{\text{old}},\varphi) = 0 \qquad \forall \varphi \in W_h\,. \tag{4.11}$$

According to (2.32), the equation (4.10) is equivalent to $u_{\text{old}} = S_h(q_{\text{old}})$ and (4.11) is equivalent to $v = S_h'(q_{\text{old}})(q - q_{\text{old}})$, such that the reduced form of (4.9)–(4.11) reads

$$\min_{q\in Q_h} \|F_h'(q_{\text{old}})(q - q_{\text{old}}) + F_h(q_{\text{old}}) - g^\delta\|_G^2 + \frac{1}{\beta}\|q - q_0\|_Q^2 \tag{4.12}$$

with $F_h = C \circ S_h$ (cf. Section 2.3).

Then the discrete counterparts to (4.6) are defined by

$$\begin{aligned}
&I_{1,h}\colon\ Q\times Q\times\mathbb{R}\to\mathbb{R}\,, && (q_{\text{old}}, q, \beta) && \mapsto \|F_h'(q_{\text{old}})(q - q_{\text{old}}) + F_h(q_{\text{old}}) - g^\delta\|_G^2 + \frac{1}{\beta}\|q - q_0\|_Q^2\\
&I_{2,h}\colon\ Q\times Q\to\mathbb{R}\,, && (q_{\text{old}}, q) && \mapsto \|F_h'(q_{\text{old}})(q - q_{\text{old}}) + F_h(q_{\text{old}}) - g^\delta\|_G^2\\
&I_{3,h}\colon\ Q\to\mathbb{R}\,, && q_{\text{old}} && \mapsto \|F_h(q_{\text{old}}) - g^\delta\|_G^2\\
&I_{4,h}\colon\ Q\to\mathbb{R}\,, && q && \mapsto \|F_h(q) - g^\delta\|_G^2\,.
\end{aligned} \tag{4.13}$$

Correspondingly, for a solution $(q_h^k, u_{\text{old},h}^k, v_h^k) = (q_{h_k}^k, u_{\text{old},h_k}^k, v_{h_k}^k)$of (4.9) for given $q_{\text{old}} = q_{\text{old}}^k \in Q_h$ and $u_h^k = u_{h_k}^k$ solving

$$a(q_h^k, u_h^k)(\varphi) = f(\varphi) \qquad \forall \varphi \in W_h\,,$$

the discrete quantities of interest in the k-th iteration step (i.e., the discrete counterparts to (4.8)) can be formulated as

$$\begin{aligned}
I_{1,h}^k &\coloneqq I_{1,h_k}^k(q_{\text{old}}^k, q_{h_k}^k, \beta_k) = \|F_{h_k}'(q_{\text{old}}^k)(q_{h_k}^k - q_{\text{old}}^k) + F_{h_k}(q_{\text{old}}^k) - g^\delta\|_G^2 + \frac{1}{\beta_k}\|q_{h_k}^k - q_0\|_Q^2\\
&= \|C(u_{\text{old},h_k}^k)(u_{h_k}^k - u_{\text{old},h_k}^k) + C(u_{\text{old},h_k}^k) - g^\delta\|_G^2 + \frac{1}{\beta_k}\|q_{h_k}^k - q_0\|_Q^2\\
I_{2,h}^k &\coloneqq I_{2,h_k}^k(q_{\text{old}}^k, q_{h_k}^k) = \|F_{h_k}'(q_{\text{old}}^k)(q_{h_k}^k - q_{\text{old}}^k) + F_{h_k}(q_{\text{old}}^k) - g^\delta\|_G^2\\
&= \|C(u_{\text{old},h_k}^k)(u_{h_k}^k - u_{\text{old},h_k}^k) + C(u_{\text{old},h_k}^k) - g^\delta\|_G^2\\
I_{3,h}^k &\coloneqq I_{3,h_k}^k(q_{\text{old}}^k) = \|F_{h_k}(q_{\text{old}}^k) - g^\delta\|_G^2 = \|C(u_{\text{old},h_k}^k) - g^\delta\|_G^2\\
I_{4,h}^k &\coloneqq I_{4,h_k}^k(q_{h_k}^k) = \|F_{h_k}(q_{h_k}^k) - g^\delta\|_G^2 = \|C(u_{h_k}^k) - g^\delta\|_G^2\,,
\end{aligned} \tag{4.14}$$

where we introduced the notation h_k (replacing h), denoting the discretization in step k, in order to distiguish between the possibly different discretizations during the iterative process in the following.

At the end of each iteration step we set

$$q_{\text{old}}^{k+1} := q_{h_k}^k . \tag{4.15}$$

Remark 4.2. *The sequence of iterates we actually consider is the discrete one $(q_{h_k}^k)_{k\in\mathbb{N}}$, which we also update according to* (4.15)*. Besides that, for theoretical purposes we keep a sequence of continuous iterates $(q^k)_{k\in\mathbb{N}}$, where each member q^k of this sequence emerges from a member $q_{\text{old}}^k = q_{h_{k-1}}^{k-1}$ of the sequence of discretized iterates $(q_{h_k}^k)_{k\in\mathbb{N}}$, but* not *from q^{k-1}, see Figure 4.1. One of the reasons for the necessity of considering this auxiliary continuous iterates is the*

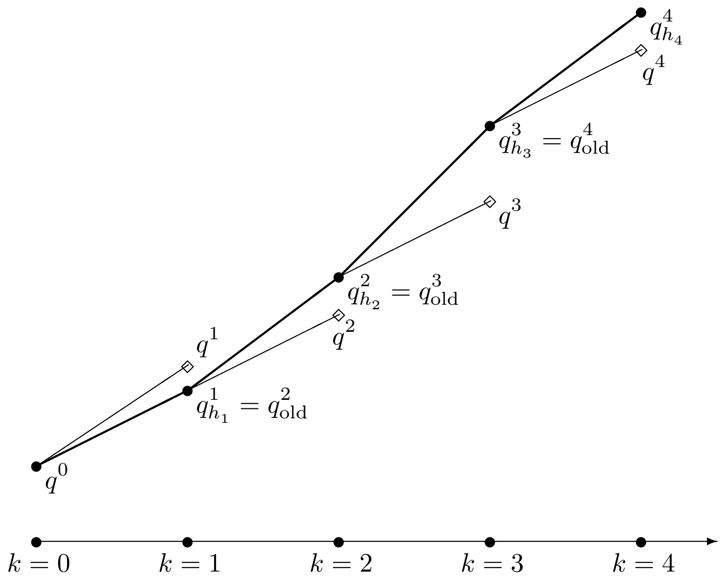

Figure 4.1.: Sequence of discretized iterates and auxiliary sequence of continuous iterates for the reduced form of IRGNM

key inequality (4.35) *in the proof of the convergence Theorem 4.2 below, which makes use of minimality of the iterate q^k in all of Q (and not only in the finite dimensional subspace Q_h), thus allowing for comparison to the infinite dimensional exact solution $q^\dagger$.*

We stress once more that the discretization may be different in each iteration, as indicated by the subscripts h_k, h_{k-1} here. In order to keep the notation readable we will suppress the iteration index k in the subscript h_k whenever this is possible without causing confusion.

Remark 4.3. *Please note that, in spite of* (4.15)*, $I_{3,h}^{k+1}$ and $I_{4,h}^k$ are not the same, since h_k and h_{k+1} may differ, such that*

$$I_{3,h}^{k+1} = \|F_{h_{k+1}}(q_{\text{old}}^{k+1}) - g^\delta\|_G^2 = \|C(S_{h_{k+1}}(q_{\text{old}}^{k+1})) - g^\delta\|_G^2 = \|C(u_{\text{old},h}^{k+1}) - g^\delta\|_G^2 ,$$

whereas

$$I_{4,h}^k = \|F_{h_k}(q_{\text{old}}^{k+1}) - g^\delta\|_G^2 = \|C(S_{h_k}(q_{\text{old}}^{k+1})) - g^\delta\|_G^2 = \|C(S_{h_k}(q_{h_k}^k)) - g^\delta\|_G^2 = \|C(u_k^{h_k}) - g^\delta\|_G^2 ,$$

i.e., $I_{3,h}^{k+1}$ contains $u_{\text{old},h}^{k+1} = S_{h_{k+1}}(q_{\text{old}}^{k+1})$, the solution of the state equation with respect to the discretization from step $k+1$, whereas in $I_{4,h}^{k}$ we have $u_{h_k}^{k} = S_{h_k}(q_{\text{old}}^{k+1})$, the solution of the state equation with respect to the discretization from step k.

Remark 4.4. *Also $I_3^{k+1} = \|F(q_{\text{old}}^{k+1}) - g^\delta\|_G^2$ and $I_4^k = \|F(q^k) - g^\delta\|_G^2$ are not the same, because*

$$q_{\text{old}}^{k+1} = q_{h_k}^{k} \neq q^k \,,$$

see Figure 4.1.

Remark 4.5. *Please note that even the discretizations h_k for fixed k can differ in the different quantities of interest during one Gauss-Newton iteration cf. Algorithm 4.1. Tracking the proof of the main convergence result Theorem 4.2 the reader can verify that only $I_{1,h}^k$ and $I_{2,h}^k$ have to be evaluated on the same mesh, since in the proof we will need the identity*

$$I_{1,h}^k = I_{2,h}^k + \frac{1}{\beta_k}\|q_{h_k}^k - q_0\|_Q^2 \,, \tag{4.16}$$

which is guaranteed by assuming exact evaluation of the Q-norm $\|q_h^k - q_0\|_Q$ (cf. Assumption 2.3).

In order to assess and – by adaptive refinement – control the differences

$$|I_{i,h}^k - I_i^k| \leq \eta_i^k \,, \quad i \in \{1,2,3,4\} \tag{4.17}$$

between the exact quantities of interest and their counterparts resulting from discretization, we will again make use of the goal oriented error estimators from the Section 2.4 and 3.3. In Section 4.1.2 and 4.2.1, we will explain in more detail what these estimators look like for the specific quantities of interest from this chapter.

We select the regularization parameter β_k according to an inexact Newton condition (cf. [42, 100]) which can be interpreted as a discrepancy principle with "noise level" $\tilde{\theta} I_{3,h}^k$

$$\underline{\tilde{\theta}} I_{3,h}^k \leq I_{2,h}^k \leq \tilde{\tilde{\theta}} I_{3,h}^k \,, \tag{4.18}$$

i.e.,

$$\underline{\tilde{\theta}}\|F_h(q_{\text{old}}^k) - g^\delta\|_G^2 \leq \|F_h'(q_{\text{old}}^k)(q_h^k - q_{\text{old}}^k) + F_h(q_{\text{old}}^k) - g^\delta\|_G^2 \leq \tilde{\tilde{\theta}}\|F_h(q_{\text{old}}^k) - g^\delta\|_G^2$$

for some $0 < \underline{\tilde{\theta}} \leq \tilde{\theta} \leq \tilde{\tilde{\theta}} < \frac{1}{2}$.

Note that we could compute this regularization parameter according to Algorithm 3.1 (as for (3.49) in Theorem 3.7, Chapter 3), but, as mentioned in the beginning of Chapter 3, Algorithm 3.1 is an extension of [39] to nonlinear inverse problems. Since (4.1)/(4.3) is in fact a linear-quadratic optimization problem and [39] is tailored for this case (e.g., by the use of explicit formulas of the quantity of interest with respect to β), we expect the algorithm from [39] to be more efficient and decide to use that one instead.

To this purpose we adopt the main result from [39, Theorem 1], whose counterpart for the nonlinear case is Theorem 3.7 with $i(\beta) = \|F'(q_{\text{old}}^k)(q_\beta^\delta - q_{\text{old}}^k) + F(q_{\text{old}}^k) - g^\delta\|_G^2$ and $\tau_\beta^2\delta_\beta^2 = \tilde{\tilde{\theta}}\|F(q_{\text{old}}^k) - g^\delta\|_G^2$.

Theorem 4.1. *Let $i \in C^3(\mathbb{R}^+)$, $i'(\beta) < 0$, $i''(\beta) > 0$, $i'''(\beta) \leq 0$ for all $\beta > 0$ and β_* solve*

$$i(\beta_*) = \tau_\beta^2 \delta_\beta^2 \,.$$

for some $\tau_\beta \geq 1$ and $\delta_\beta > 0$. Let moreover a sequence $(\beta^k)_{k\in\mathbb{N}}$ be defined by

$$\beta^{k+1} = \beta^k - \frac{i_h^k - \tau_\beta^2\delta_\beta^2}{{i'}_h^k}\,, \quad 0 < \beta^0 \leq \beta_*\,, \tag{4.19}$$

with some i_h^k, ${i'}_h^k \in \mathbb{R}$ satisfying

$$|i(\beta^k) - i_h^k| \leq \min\left\{c_{\beta,1}|i_h^k - \tau_\beta^2\delta_\beta^2|\,, \frac{C_{2,\beta}\|g^{\delta_\beta}\|_G^2}{|{i'}_h^k|^2(\beta^k)^2}|i_h^k - \tau_\beta^2\delta_\beta^2|^2\right\} \tag{4.20}$$

$$|i'(\beta^k) - {i'}_h^k| \leq \min\left\{C_{3,\beta}|{i'}_h^k|\,, \frac{C_{2,\beta}\|g^{\delta_\beta}\|_G^2}{|{i'}_h^k|^2(\beta^k)^2}|i_h^k - \tau_\beta^2\delta_\beta^2|\right\} \tag{4.21}$$

for some constants $C_{2,\beta}, C_{3,\beta} > 0$, $0 < c_{1,\beta} < 1$ independent of k. Let moreover k_* be given as

$$k_* = \min\left\{k \in \mathbb{N}|\ i_h^k \leq \left(\tau_\beta^2 + \frac{\tilde{\tau}_\beta^2}{2}\right)\delta_\beta^2\right\} \tag{4.22}$$

and the following conditions be fulfilled:

$$\begin{gathered} {i'}_h^k < 0 \quad \forall k \leq k_* - 1\,, \\ |i(\beta^{k_*-1}) - i_h^{k_*-1}| + \left|\frac{i_h^{k_*-1} - \tau_\beta^2\delta_\beta^2}{{i'}_h^{k_*-1}}\right| |i'(\beta^{k_*-1}) - {i'}_h^{k_*-1}| \leq \tilde{\tau}_\beta^2\,, \\ |i(\beta^{k_*}) - i_h^{k_*}| \leq \frac{\tilde{\tau}_\beta^2}{2}\delta_\beta^2 \end{gathered} \tag{4.23}$$

for some $\tilde{\tau}_\beta < \tau_\beta$.

Then k_* is finite and there holds

$$\beta^{k+1} \geq \beta^k \quad \textit{and} \quad \beta^k \leq \beta_* \quad \textit{for all } k \leq k_* - 1\,,$$

β^k satisfies the local quadratic convergence estimate

$$|\beta^{k+1} - \beta_*| \leq \frac{C\|g^{\delta_\beta}\|_G^2}{|i'(\beta^k)|(\beta^k)^2}(\beta^k - \beta_*)^2 + \mathcal{O}((\beta^k - \beta_*)^4) \quad \forall k \leq k_* - 1$$

for some $C > 0$ independent of β^k and k, and

$$(\tau_\beta^2 - \tilde{\tau}_\beta^2)\delta_\beta^2 \leq i(\beta^{k_*}) \leq (\tau_\beta^2 + \tilde{\tau}_\beta^2)\delta_\beta^2\,. \tag{4.24}$$

Proof. See [39, proof of Theorem 1]. □

Moreover in [39, Lemma 1] Griesbaum, Kaltenbacher and Vexler show that the choice $i(\beta) := \|Tq_\beta^{\delta_\beta} - t\|_G^2$ with some linear operator $T\colon Q \to G$ and some constant term t (with respect to $q_\beta^{\delta_\beta}$) fulfills all the requirements from Theorem 4.1. This includes our setting, since we can write $I_2(q_{\text{old}}, q) = \|F'(q_{\text{old}})(q - q_{\text{old}}) + F(q_{\text{old}}) - g^\delta\|_G^2$ as $\|Tq - t\|_G^2$ with $T := F'(q_{\text{old}})$ and $t := F'(q_{\text{old}})(q_{\text{old}}) + g^\delta - F(q_{\text{old}})$. The relation (4.24) would as well allow us to use the continuous version I_2^k in (4.18), but we prefer to formulate the condition with the discretized actually computed quantities.

The overall Newton iteration is stopped according to a discrepancy principle similar to (3.36)

$$k_* = \min\{k \in \mathbb{N} |\ I_{3,h}^k \le \tau^2\delta^2\}. \tag{4.25}$$

In our convergence analysis we will again (as in Section 3.1) make use of the weak sequential closedness Assumption 3.2 and a tangential cone condition similar to Assumption 3.3.

4.1.1. Convergence of adaptively discretized minimizers/stationary points to an exact solution for vanishing data noise

Similarly as in Section 3.1 (more specifically Theorem 3.4), in the following theorem we show convergence of the form $q_{\text{old}}^{k_*} = q_{\beta_{k_*-1},h_{k_*-1}}^{\delta,k_*-1} \to q^\dagger$ as $\delta \to 0$ for a sequence $(q_{\beta_k,h_k}^{\delta,k})_{k\in\mathbb{N}}$ produced by iteratively solving (4.9) with $\beta := \beta_k$ and updating according to (4.15).

Throughout Section 4.1 we assume that continuous and discrete solutions of the IRGNM subproblem exist, which means

Assumption 4.2. *For all $\beta_k \in [\underline{\beta}, \overline{\beta}]$ (for some $0 < \underline{\beta} < \overline{\beta}$) and all iterations k the subproblem (4.1)/(4.3)–(4.5) is solvable, i.e., there exists a solution $q^k = q^{\delta,k} = q_{\beta_k}^{\delta,k}$ to (4.1), where $q^{\delta,k-1} = q_{\text{old}}^k$ is chosen according to (4.15).*

Assumption 4.3. *For all $\beta_k \in [\underline{\beta}, \overline{\beta}]$ (for some $0 < \underline{\beta} < \overline{\beta}$) and all iterations k the discretized subproblem (4.12)/ (4.9)–(4.11) is solvable, i.e., there exists a solution $q_h^k = q_{h_k}^k = q_{h_k}^{\delta,k} = q_{\beta_k,h_k}^{\delta,k}$ to (4.12), where q_{old} is chosen according to (4.15).*

Remark 4.6. *Please note that different from Assumption 3.4 in Chapter 3, we do not assume that these minimizers lie in $\mathcal{D}$; cf. Remark 2.3, 2.4 and 3.1.*

Since the three assumptions Assumption 4.1, 4.2 and 4.3 constitute the basic prerequisites of this section, we will not list them again explicitly in the following theorems.

Different to Chapter 3 the tangential cone condition Assumption 3.3 doesn't need to be satisfied in the whole of $\mathcal{D}$, but only in a neighboorhood of q_0:

Assumption 4.4. *Let the* tangential cone condition *(cf. e.g., [27],[102])*

$$\|F(q) - F(\bar{q}) - F'(q)(q - \bar{q})\|_G \le c_{tc}\|F(q) - F(\bar{q})\|_G$$

hold for all $q, \bar{q} \in \mathcal{B}_\rho(q_0) \subseteq \mathcal{D} \subset Q$ (cf. Assumption 4.1) and for some $0 < c_{tc} < 1$ sufficiently small (in order to fulfill (4.26) in the following).

Assumption 4.5. *Let τ be chosen sufficiently large and $0 < \underline{\tilde{\theta}} < \tilde{\tilde{\theta}}$ sufficiently small (see* (4.18), (4.25)*), such that*

$$2\left(c_{tc}^2 + \frac{(1+c_{tc})^2}{\tau^2}\right) < \underline{\tilde{\theta}} \quad \text{and} \quad \frac{2\tilde{\tilde{\theta}} + 4c_{tc}^2}{1-4c_{tc}^2} < 1\,. \tag{4.26}$$

Finally, let for the discretization error with respect to the quantities of interest (4.17) *hold, where η_1^k, η_2^k, η_3^k, η_4^k are selected such that*

$$\eta_1^k + 2c_{tc}^2\eta_3^k \le \left(\underline{\tilde{\theta}} - 2\left(c_{tc}^2 + \frac{(1+c_{tc})^2}{\tau^2}\right)\right) I_{3,h}^k\,, \tag{4.27}$$

$$\eta_3^k \le c_1 I_{3,h}^k \quad \text{and} \quad \eta_2^k \to 0\,,\ \eta_3^k \to 0\,,\ \eta_4^k \to 0 \ \text{as}\ k \to \infty\,, \tag{4.28}$$

$$I_{3,h}^k \le (1+c_3) I_{4,h}^{k-1} + r^k\,, \quad \text{and} \quad (1+c_3)\frac{2\tilde{\tilde{\theta}} + 4c_{tc}^2}{1-4c_{tc}^2} \le c_2 < 1 \tag{4.29}$$

for some constants $c_1, c_2, c_3 > 0$, and a sequence $r^k \to 0$ as $k \to \infty$.

Remark 4.7. *Condition* (4.26) *can be satisfied fir c_{tc} sufficiently small, which (via the tangential cone condition Assumption 4.4) is a local restriction on the nonlinearity of F, and by choosing τ sufficiently large. Conditions* (4.27) *and* (4.28) *(where the right-hand side is always strictly positive by definition of the stopping index* (4.25)*) are smallness conditions on the η_i^k $(i = 1, 2, 3, 4)$, whereas the first condition in* (4.29) *links the discretizations of the forward operator at $q_{\text{old}}^k = q_h^{k-1}$ in the old and the new iteration. The second condition in* (4.29) *is enabled by the right inequality in* (4.26).

By means of these assumptions we can formulate the following convergence theorem (similar to Theorem 3.1).

For a better legibility, depending on which parameter will be of interest at a certain point, we will omit some indices and switch between the different notations $q^k = q^{\delta,k} = q^{\delta,k}_{\beta_k}$, $q_h^k = q_{h_k}^k = q_{h_k}^{\delta,k} = q_{\beta_k,h_k}^{\delta,k}$.

Theorem 4.2. *Let Assumption 3.2, 4.4 and 4.5 be satisfied. Further, let $q^\dagger \in \mathcal{B}_\rho(q_0) \subset \mathcal{D}$ be a solution to* (2.1) *and let the starting value be chosen such that $q_{\text{old}}^1 = q_{h_0}^0 \in \mathcal{B}_\rho(q_0)$ (which is obviously satisfied, for instance, by the choice $q_{\text{old}}^1 = q_0$).*

Then with β_k and $h = h_k$ fulfilling (4.18) *and k_* selected according to* (4.25)*, for a solution $q_{h_k}^k$ to* (4.12) *there holds:*

(o) for all $k < k_*$, provided

$$\|F_h'(q_{\text{old}}^k)(q_0 - q_{\text{old}}^k) + F_h(q_{\text{old}}^k) - g^\delta\|_G^2 \ge \tilde{\tilde{\theta}} \|F_h(q_{\text{old}}^k) - g^\delta\|_G^2\,, \tag{4.30}$$

there exists β_k satisfying (4.18),

(i) the estimate

$$\|q_{h_k}^k - q_0\|_Q \le \|q^\dagger - q_0\|_Q \quad \forall 1 \le k < k_* \tag{4.31}$$

holds, which implies that all iterates stay in the ball $\mathcal{B}_\rho(q_0)$, *(ii)* k_* *is finite,*

(iii) $q_{\text{old}}^{k_*} = q_{\text{old}}^{k_*(\delta)} = q_{\beta_{k_*(\delta)-1},h_{k_*(\delta)-1}}^{k_*(\delta)-1,\delta}$ *converges (weakly) subsequentially to a solution of* (2.1) *as* $\delta \to 0$ *in the sense that it has a weakly convergent subsequence* $\left(q_{\text{old}}^{k_*(\delta_l)}\right)_{l\in\mathbb{N}}$ *and each weakly convergent subsequence converges strongly to a solution of* (2.1)*. If the solution* $q^\dagger$ *to* (2.1) *is unique, then* $q_{\text{old}}^{k_*(\delta)}$ *converges strongly to* $q^\dagger$ *as* $\delta \to 0$.

Proof. (o): Let $I_1^k(\beta)$, $I_{1,h}^k(\beta)$, $I_2^k(\beta)$, $I_{2,h}^k(\beta)$ denote I_1^k, $I_{1,h}^k$, I_2^k, $I_{2,h}^k$ (see (4.14), (4.8)) with β_k replaced by β.

For all $1 \leq k < k_*$ and any solution $q^\dagger$ of (2.1), by (4.17) and minimality of q^k for the continuous Tikhonov functional, there holds

$$\begin{aligned} I_{2,h}^k(\beta) &\leq I_{1,h}^k(\beta) \\ &\leq I_1^k(\beta) + \eta_1^k \\ &\leq \|F'(q_{\text{old}}^k)(q^\dagger - q_{\text{old}}^k) + F(q_{\text{old}}^k) - g^\delta\|_G^2 + \eta_1^k + \frac{1}{\beta}\|q^\dagger - q_0\|_Q^2 \,. \end{aligned} \tag{4.32}$$

Using (2.2), Assumption 4.4, the inequality $(a+b)^2 \leq 2a^2 + 2b^2$ for arbitrary $a, b \in \mathbb{R}$, as well as (4.25), we can estimate as follows

$$\begin{aligned} \|F'(q_{\text{old}}^k)(q^\dagger - q_{\text{old}}^k) + F(q_{\text{old}}^k) - g^\delta\|_G^2 &\leq \left(\|F'(q_{\text{old}}^k)(q^\dagger - q_{\text{old}}^k) + F(q_{\text{old}}^k) - F(q^\dagger)\|_G + \delta\right)^2 \\ &\leq \left(c_{tc}\|F(q^\dagger) - F(q_{\text{old}}^k)\|_G + \delta\right)^2 \\ &\leq \left(c_{tc}\left(\|g^\delta - F(q_{\text{old}}^k)\|_G + \delta\right) + \delta\right)^2 \\ &\leq \left(c_{tc}\sqrt{I_3^k} + (1 + c_{tc})\delta\right)^2 \\ &\leq 2c_{tc}^2 I_3^k + 2(1 + c_{tc})^2\delta^2 + \eta_1^k \\ &\leq 2c_{tc}^2(I_{3,h}^k + \eta_3^k) + 2(1 + c_{tc})^2 \frac{I_{3,h}^k}{\tau^2} \\ &= 2\left(c_{tc}^2 + \frac{(1 + c_{tc})^2}{\tau^2}\right) I_{3,h}^k + 2c_{tc}^2\eta_3^k \,. \end{aligned} \tag{4.33}$$

Together with (4.32) and Assumption 4.5 this yields

$$I_{2,h}^k(\beta) \leq \underset{\sim}{\theta} I_{3,h}^k + \frac{1}{\beta}\|q^\dagger - q_0\|_Q^2 \,.$$

Hence, we have

$$\limsup_{\beta\to\infty} I_{2,h}^k(\beta) \leq \underset{\sim}{\theta} I_{3,h}^k \,.$$

On the other hand, since

$$\lim_{\beta\to 0} q_{\beta,h}^k = q_0 \,,$$

we have by (4.30)

$$\lim_{\beta\to 0} I_{2,h}^k(\beta) = \|F_h'(q_{\text{old}}^k)(q_0 - q_{\text{old}}^k) + F_h(q_{\text{old}}^k) - g^\delta\|_G^2 \geq \tilde{\bar{\theta}} I_{3,h}^k\,.$$

Thus, by continuity of the mapping $\beta \mapsto I_{2,h}^k(\beta)$ (see e.g., [39]) and the Intermediate Value Theorem the assertion follows.

(i): We will prove (4.31) as follows: We will show that, for fixed $k > 0$, provided that $q_{\text{old}}^k = q_{h_{k-1}}^{k-1} \in \mathcal{B}_\rho(q_0)$, there holds

$$\|q_{h_k}^k - q_0\|_Q \leq \|q^\dagger - q_0\|_Q \quad \forall 1 \leq k < k_*\,,$$

which implies $q_{\text{old}}^{k+1} = q_{h_k}^k \in \mathcal{B}_\rho(q_0)$. Then (4.31) follows recursively, since $q_{\text{old}}^1 = q_{h_0}^0 \in \mathcal{B}_\rho(q_0)$.

So we assume that $q_{\text{old}}^k \in \mathcal{B}_\rho(q_0)$ in the following.

From (4.16), (4.18) it follows that

$$I_{1,h}^k = I_{2,h}^k + \frac{1}{\beta_k}\|q_h^k - q_0\|_Q^2 \geq \underline{\tilde{\theta}} I_{3,h}^k + \frac{1}{\beta_k}\|q_h^k - q_0\|_Q^2\,. \tag{4.34}$$

As in (4.32), by (4.17) and minimality of q^k for (4.1) we have

$$I_{1,h}^k \leq I_1^k + \eta_1^k \leq \|F'(q_{\text{old}}^k)(q^\dagger - q_{\text{old}}^k) + F(q_{\text{old}}^k) - g^\delta\|_G^2 + \frac{1}{\beta_k}\|q^\dagger - q_0\|_Q^2 + \eta_1^k\,, \tag{4.35}$$

which together with (4.34), (4.33), and (4.27) gives

$$\begin{aligned}
\underline{\tilde{\theta}} I_{3,h}^k + \frac{1}{\beta_k}\|q_h^k - q_0\|_Q^2 &\leq \|F'(q_{\text{old}}^k)(q^\dagger - q_{\text{old}}^k) + F(q_{\text{old}}^k) - g^\delta\|_G^2 + \frac{1}{\beta_k}\|q^\dagger - q_0\|_Q^2 + \eta_1^k \\
&\leq 2\left(c_{tc}^2 + \frac{(1+c_{tc})^2}{\tau^2}\right) I_{3,h}^k + \frac{1}{\beta_k}\|q^\dagger - q_0\|_Q^2 + \eta_1^k + 2c_{tc}^2\eta_3^k \\
&\leq \underline{\tilde{\theta}} I_{3,h}^k + \frac{1}{\beta_k}\|q^\dagger - q_0\|_Q^2\,,
\end{aligned}$$

which implies (4.31).

(ii): Furthermore, for all $1 \leq k < k_*$ we have by the triangle inequality as well as Assumption 4.4 and (4.18)

$$\begin{aligned}
\sqrt{I_4^k} &= \|F(q^k) - g^\delta\|_G \\
&\leq \|F'(q_{\text{old}}^k)(q^k - q_{\text{old}}^k) + F(q_{\text{old}}^k) - g^\delta\|_G + \|F'(q_{\text{old}}^k)(q^k - q_{\text{old}}^k) - F(q^k) + F(q_{\text{old}}^k)\|_G \\
&\leq \sqrt{I_2^k} + c_{tc}\|F(q^k) - F(q_{\text{old}}^k)\|_G \\
&\leq \sqrt{\tilde{\bar{\theta}} I_{3,h}^k + \eta_2^k} + c_{tc}(\sqrt{I_4^k} + \sqrt{I_3^k})\,.
\end{aligned} \tag{4.36}$$

(Note that applying the tangential cone condition Assumption 4.4 to q^k here is allowed due to (4.31).) Hence, by $(a+b)^2 \leq 2a^2 + 2b^2$ for all $a, b \in \mathbb{R}$

$$I_4^k \leq 2(\tilde{\bar{\theta}} I_{3,h}^k + \eta_2^k) + 2c_{tc}^2(2I_4^k + 2I_3^k)\,,$$

which implies

$$I_4^k \leq \frac{1}{1-4c_{tc}^2}\left(2\tilde{\bar{\theta}} I_{3,h}^k + 2\eta_2^k + 4c_{tc}^2 I_3^k\right) \qquad \forall 1 \leq k < k_*\,.$$

With (4.17) and (4.29) we can further deduce

$$\begin{aligned}
I_{4,h}^k &\leq \frac{1}{1-4c_{tc}^2}\left((2\tilde{\bar{\theta}} + 4c_{tc}^2) I_{3,h}^k + 2\eta_2^k + 4c_{tc}^2\eta_3^k\right) + \eta_4^k \\
&\leq \frac{2\tilde{\bar{\theta}} + 4c_{tc}^2}{1-4c_{tc}^2}(1+c_3) I_{4,h}^{k-1} + \frac{1}{1-4c_{tc}^2}\left((2\tilde{\bar{\theta}} + 4c_{tc}^2) r^k + 2\eta_2^k + 4c_{tc}^2\eta_3^k\right) + \eta_4^k \\
&\leq c_2 I_{4,h}^{k-1} + \frac{1}{1-4c_{tc}^2}\left((2\tilde{\bar{\theta}} + 4c_{tc}^2) r^k + 2\eta_2^k + 4c_{tc}^2\eta_3^k\right) + \eta_4^k
\end{aligned}$$

for all $1 \leq k < k_*$. With the notation

$$a^i := \frac{1}{1-4c_{tc}^2}\left((2\tilde{\bar{\theta}} + 4c_{tc}^2) r^i + 2\eta_2^i + 4c_{tc}^2\eta_3^i\right) + \eta_4^i \quad \forall i \in \{1, 2, \ldots, k\} \tag{4.37}$$

there follows recursively

$$I_{4,h}^k \leq c_2^k I_{4,h}^0 + \sum_{j=0}^{k-1} c_2^j a^{k-j} \qquad \forall 1 \leq k < k_*\,. \tag{4.38}$$

Note that by the second part of (4.28), the second part of (4.29) and the fact that $r^k \to 0$ as $k \to \infty$ (by definition of r^k) , we have $c_2^k I_{4,h}^0 + \sum_{j=0}^{k-1} c_2^j a^{k-j} \to 0$ as $k \to \infty$. So, if the discrepancy principle never got active (i.e., $k_* = \infty$), the sequence $(I_{4,h}^k)_{k\in\mathbb{N}}$ and therewith by (4.28) also $(I_{3,h}^k)_{k\in\mathbb{N}}$ would be bounded by a sequence tending to zero as $k \to \infty$, which implies that $I_{3,h}^k$ would fall below $\tau^2\delta^2$ for k sufficiently large, thus yielding a contradiction. Hence the stopping index $k_* < \infty$ is well-defined and finite. (Note that here, we could not just argue $I_{3,h}^k \leq C I_{4,h}^k \to 0$, hence eventually smaller than $\tau\delta$, since the estimate (4.38) only holds for $k < k_*$.)

(iii): With (2.2), (4.17), (4.28) and the definition of k_*, we have

$$\|F(q_{\text{old}}^{k_*}) - g\|_G \leq \sqrt{I_3^{k_*}} + \delta \leq \sqrt{I_{3,h}^{k_*} + \eta_3^{k_*}} + \delta \leq \sqrt{(1+c_1) I_{3,h}^{k_*}} + \delta \leq (\tau\sqrt{1+c_1} + 1)\delta \to 0 \tag{4.39}$$

as $\delta \to 0$. Due to (4.31) $q_{\text{old}}^{k_*} = q_{h_{k_*-1}}^{k_*-1}$ has a weakly convergent subsequence $\left(q_{\text{old}}^{k_*(\delta_l)}\right)_{l\in\mathbb{N}}$ and due to the weak sequential closedness of F (Assumption 3.2) and (4.39) the limit of every weakly convergent subsequence is a solution to $F(q) = g$.

Strong convergence in the case of uniqueness of $q^\dagger$ follows by a standard argument as in (3.9).

□

Remark 4.8. *In view of the proof of Theorem 3.1 note that the estimate* (4.35) *can alternatively be shown using stationarity instead of minimality of q^k (which is equivalent by convexity together with the assumption that $\mathcal{D}$ has a nonempty interior; cf. Assumption 4.1): For all $\delta q \in Q$ stationarity means*

$$0 = (F'(q^k_{\text{old}})(q^k - q^k_{\text{old}}) + F(q^k_{\text{old}}) - g^\delta, F'(q^k_{\text{old}})(\delta q))_G + \frac{1}{\beta_k}(q^k - q_0, \delta q)_Q .$$

Setting $\delta q = q^k - q^\dagger$ this yields

$$\begin{aligned}
0 &= \|F'(q^k_{\text{old}})(q^k - q^k_{\text{old}}) + F(q^k_{\text{old}}) - g^\delta\|_G^2 \\
&\quad + (F'(q^k_{\text{old}})(q^k - q^k_{\text{old}}) + F(q^k_{\text{old}}) - g^\delta, F'(q^k_{\text{old}})(q^k - q^\dagger - (q^k - q^k_{\text{old}})) - F(q^k_{\text{old}}) + g^\delta)_G \\
&\quad + \frac{1}{\beta_k}(q^k - q_0, q^k - q^\dagger)_Q \\
&= \|F'(q^k_{\text{old}})(q^k - q^k_{\text{old}}) + F(q^k_{\text{old}}) - g^\delta\|_G^2 \\
&\quad - (F'(q^k_{\text{old}})(q^k - q^k_{\text{old}}) + F(q^k_{\text{old}}) - g^\delta, F'(q^k_{\text{old}})(q^\dagger - q^k_{\text{old}}) + F(q^k_{\text{old}}) - g^\delta)_G \\
&\quad + \frac{1}{\beta_k}\|q^k - q_0\|_Q^2 - \frac{1}{\beta_k}(q^k - q_0, q^\dagger - q_0)_Q ,
\end{aligned}$$

hence, by Cauchy-Schwarz and $ab \leq \frac{1}{2}a^2 + \frac{1}{2}b^2$ for all $a, b \in \mathbb{R}$,

$$\begin{aligned}
I_1^k &= \|F'(q^k_{\text{old}})(q^k - q^k_{\text{old}}) + F(q^k_{\text{old}}) - g^\delta\|_G^2 + \frac{1}{\beta_k}\|q^k - q_0\|_Q^2 \\
&= (F'(q^k_{\text{old}})(q^k - q^k_{\text{old}}) + F(q^k_{\text{old}}) - g^\delta, F'(q^k_{\text{old}})(q^\dagger - q^k_{\text{old}}) + F(q^k_{\text{old}}) - g^\delta)_G \\
&\quad + \frac{1}{\beta_k}(q^k - q_0, q^\dagger - q_0)_Q \\
&\leq \|F'(q^k_{\text{old}})(q^k - q^k_{\text{old}}) + F(q^k_{\text{old}}) - g^\delta\|_G \|F'(q^k_{\text{old}})(q^\dagger - q^k_{\text{old}}) + F(q^k_{\text{old}}) - g^\delta\|_G \\
&\quad + \frac{1}{\beta_k}\|q^k - q_0\|_Q \|q^\dagger - q_0\|_Q \\
&\leq \frac{1}{2}\|F'(q^k_{\text{old}})(q^k - q^k_{\text{old}}) + F(q^k_{\text{old}}) - g^\delta\|_G^2 + \frac{1}{2}\|F'(q^k_{\text{old}})(q^\dagger - q^k_{\text{old}}) + F(q^k_{\text{old}}) - g^\delta\|_G^2 \\
&\quad + \frac{1}{2\beta_k}\|q^k - q_0\|_Q^2 + \frac{1}{2\beta_k}\|q^\dagger - q_0\|_Q^2 \\
&= \frac{1}{2}I_1^k + \frac{1}{2}\|F'(q^k_{\text{old}})(q^\dagger - q^k_{\text{old}}) + F(q^k_{\text{old}}) - g^\delta\|_G^2 + \frac{1}{2\beta_k}\|q^\dagger - q_0\|_Q^2 .
\end{aligned}$$

This finally yields

$$I_1^k \leq \|F'(q^k_{\text{old}})(q^\dagger - q^k_{\text{old}}) + F(q^k_{\text{old}}) - g^\delta\|_G^2 + \frac{1}{\beta_k}\|q^\dagger - q_0\|_Q^2 .$$

To prove convergence rates we will again (like in Section 3.1) consider Hilbert space source conditions as in Assumption 3.5 and apply Theorem 3.2.

Corollary 4.3. *Let the conditions of Theorem 4.2 and additionally the source condition Assumption 3.5 be fulfilled. Then there exist $\bar{\delta} > 0$ and a constant $C > 0$ independent of δ such* *that for all* $\delta \in (0, \bar{\delta}]$

$$\left\| q_{\text{old}}^{k_*} - q^\dagger \right\|_Q = \left\| q_{\beta_{k_*-1}, h_{k*-1}}^{\delta, k_*-1} - q^\dagger \right\|_Q = \mathcal{O}\left(\frac{\delta}{\sqrt{\psi^{-1}(C\delta)}} \right)$$

Proof. The proposition follows directly from Theorem 3.2 due to (4.31) and (4.39). □

4.1.2. Computation of the error estimators

The computation of the error estimators η_1^k, η_2^k, η_3^k and η_4^k is done similarly to Section 3.3 and [39]. The only difference lies in the fact that in I_1^k we have three variables subject to discretization, namely q, u_{old} and v instead of only two (q and u) as usual, which leads to the following error estimators. In this section we omit the iteration index k for simplicity. The previous iterate q_{old} is fixed and not subject to new discretization and the dependence on β is not important for error estimation; so we neglect q_{old} and β as arguments.

Error estimator for I_1

We consider

$$I_1(q, u_{\text{old}}, v) = \|C'(u_{\text{old}})(v) + C(u_{\text{old}}) - g^\delta\|_G^2 + \frac{1}{\beta}\|q - q_0\|_Q^2 \,,$$

where the reader should not be confused about the ambiguity of I_1 as compared with (4.6), since both definitions represent the same thing and are just expressed by different arguments. We define the Lagrange functional

$$\begin{aligned} L(q, u_{\text{old}}, v, w, w_{\text{old}}) \coloneqq\ & I_1(q, u_{\text{old}}, v) \\ & + a'_u(q_{\text{old}}, u_{\text{old}})(v, w) + a'_q(q_{\text{old}}, u_{\text{old}})(q - q_{\text{old}}, w) \\ & + a(q_{\text{old}}, u_{\text{old}})(w_{\text{old}}) - f(w_{\text{old}}) \,. \end{aligned}$$

Proposition 4.1. *Let* $X = Q \times V \times V \times W \times W$ *and* $X_h = Q_h \times V_h \times V_h \times W_h \times W_h$. *Let* $x = (q, u_{\text{old}}, v, w, w_{\text{old}}) \in X$ *be a stationary point of* L, *i.e.*

$$L'(x)(dx) = 0 \qquad \forall dx \in X$$

and let $x_h = (q_h, u_{\text{old},h}, v_h, w_h, w_{\text{old},h}) \in X_h$ *be a discrete stationary point of* L, *i.e.*

$$L'(x_h)(dx) = 0 \qquad \forall dx \in X_h \,. \tag{4.40}$$

Then there holds

$$I_1(q, u_{\text{old}}, v) - I_1(q_h, u_{\text{old},h}, v_h) = \frac{1}{2} L'(x_h)(x - \widehat{x}_h) + R \,,$$

for an arbitrary $\widehat{x}_h \in X_h$ *and*

$$R = \frac{1}{2} \int_0^1 L'''(x + se_x)(e_x, e_x, e_x) s(s-1) \; ds$$

with $e_x \coloneqq x - x_h$.

Proof. cf. [39] and [10]. □

Explicitly, such stationary points can be computed by solving the equations

$$
\begin{aligned}
u_{\text{old}} \in V\colon \quad & a(q_{\text{old}}, u_{\text{old}})(dv_{\text{old}}) = f(dv_{\text{old}}) \quad \forall dv_{\text{old}} \in W && (4.41)\\
v \in V: \quad & a'_u(q_{\text{old}}, u_{\text{old}})(v, dw) = -a'_q(q_{\text{old}}, u_{\text{old}})(q - q_{\text{old}}, dw) \quad \forall dw \in W && (4.42)\\
w \in W: \quad & a'_u(q_{\text{old}}, u_{\text{old}})(dv, w) = -I'_{1,w}(q, u_{\text{old}}, v)(dv) \quad \forall dv \in V && (4.43)\\
w_{\text{old}} \in W: \quad & a'_u(q_{\text{old}}, u_{\text{old}})(du, w_{\text{old}}) = -I'_{1,u_{\text{old}}}(q, u_{\text{old}}, v)(du) && (4.44)\\
& \qquad - a''_{uu}(q_{\text{old}}, u_{\text{old}})(v, du, w) && (4.45)\\
& \qquad - a''_{qu}(q_{\text{old}}, u_{\text{old}})(q - q_{\text{old}}, du, w) \quad \forall du \in V && (4.46)\\
q \in Q: \quad & I'_{1,q}(q, u_{\text{old}}, v)(dq) = -a'_q(q_{\text{old}}, u_{\text{old}})(dq, w) \quad \forall dq \in Q\,. && (4.47)
\end{aligned}
$$

and their discrete counterparts.

Via interpolation and the operator π_h from Section 2.4 the error estimator η_1 for I_1 can be computed as

$$I_1 - I_{1,h} = I_1(q, u_{\text{old}}, v) - I_1(q_h, u_{\text{old},h}, v_h) \approx \frac{1}{2} L'(x_h)(\pi_h x_h - x_h) = \eta_1$$

(cf. [10]).

Remark 4.9. *Please note that the equations* (4.40) /(4.41)–(4.47) *(or their discrete counterparts) are solved anyway in the process of solving the optimization problem* (4.3)–(4.5) *(or its discrete equivalent* (4.9)–(4.11)*).*

Error estimator for I_2

We consider

$$I_2(u_{\text{old}}, v) = \|C'(u_{\text{old}})(v) + C(u_{\text{old}}) - g^\delta\|_G^2\,,$$

where the reader should not be confused about the ambiguity of I_2 as compared with (4.6), since both definitions represent the same thing and are just expressed by different arguments. For $x_1 := (q_1, u_{\text{old}_1}, v_1, w_1, w_{\text{old}_1}) \in X$ we define the Lagrange functional

$$M(x, x_1) := I_2(u_{\text{old}}, v) + L'(x)(x_1)\,.$$

By doing so, we combine information on the quantity of interest I_2, whose precision is to be assessed, with information on the underlying minimization problem (represented by the derivative of the Lagrangian L) into one functional. This approach is standard DWR technique and can be found, e.g., in [11] (also see Section 2.4). This allows us to conclude a similar result to Proposition 4.1 for the difference $I(u_{\text{old}}, v) - I(u_{\text{old},h}, v_h)$ for stationary points $y = (x, x_1) \in X \times X$ and $y_h = (x_h, x_{1,h}) \in X_h \times X_h$ of M (cf. [11]). Such a discrete stationary point y_h can be computed by solving the discrete counterparts of the equations (4.40)/(4.41)–(4.47) and

$$x_{1,h} \in X_h\colon \quad L''(x_h)(x_{1,h}, dx) = -I'_{2,v}(u_{\text{old},h}, v_h)(dv) - I'_{2,u_{\text{old}}}(u_{\text{old},h}, v_h)(du_{\text{old}}) \qquad (4.48)$$

for all $dx = (dq, du_{\text{old}}, dv, dw, dw_{\text{old}}) \in X_h$. The error estimator η_2 for I_2 can then be computed by

$$I_2 - I_{2,h} = I_2(u_{\text{old}}, v) - I_2(u_{\text{old},h}, v_h) \approx \frac{1}{2} M'(y_h)(\pi_h y_h - y_h) = \eta_2 \,.$$

Remark 4.10. *Once the optimality system* (4.40) /(4.41)–(4.47) *has been solved, computation of the auxiliary variable $x_{1,h}$ in $y_h = (x_h, x_{1,h})$ only requires the solution of the linear system* (4.48). *To avoid the computation of second order information in* (4.48) *the author would like to refer to [11], where* (4.48) *is replaced by an approximate equation of first order.*

Error estimator for I_3

For I_3 we again proceed similarly as for I_1 and I_2, i.e., we consider

$$I_3(u_{\text{old}}) = \|C(u_{\text{old}}) - g^\delta\|_G^2 \,,$$

where the reader should not be confused about the ambiguity of I_3 as compared with (4.6), since both definitions represent the same thing and are just expressed by different arguments. For $x_2 \in X$ we define the Lagrangian

$$N(x, x_2) := I_3(u_{\text{old}}) + L'(x)(x_2) \,.$$

As there holds again a similar results to Proposition 4.1, we compute a discrete stationary point $\chi_h = (x_h, x_{2,h}) \in X_h \times X_h$ of N by solving the equations (4.40)/(4.41)–(4.47) and

$$x_{2,h} \in X_h\colon \qquad L''(x_h)(x_{2,h}, dx) = -I_3'(u_{\text{old},h})(du_{\text{old}}) \qquad \forall dx \in X_h \,, \tag{4.49}$$

(where $dx = (dq, du_{\text{old}}, dv, dw, dw_{\text{old}})$) and compute the error estimator for I_3 as

$$I_3 - I_{3,h} = I_3(u_{\text{old}}) - I_3(u_{\text{old},h}) \approx \frac{1}{2} N'(\chi_h)(\pi_h \chi_h - \chi_h) = \eta_3 \,.$$

Remark 4.10 also holds for the computation of $x_{2,h}$, i.e., $x_{2,h}$ can be computed at low additional costs.

Error estimator for I_4

Different to the other error estimates, the bound on the error in I_4 only appears in connection with the very weak assumption $\eta_4^k \to 0$ as $k \to \infty$, which may be satisfied in practice without refining explicitly with respect to η_4, but simply by refining with respect to the other error estimators η_1, η_2, and especially η_3. Another way to make sure that $\eta_4^k \to 0$ as $k \to \infty$, is, of course, to refine globally every now and then, although this is admittedly, not a very efficient solution.

If one doesn't want to rely on such practically motivated speculations and actually wants to compute an error estimator for I_4, one has to include the decoupled constraint

$$A(q, u)(v) = f(v) \qquad \forall v \in W$$

in the definition of the Lagrangian L for I_1. In that case we redefine the Lagrange functional L as

$$\begin{aligned}L(q, u_{\text{old}}, v, w, w_{\text{old}}, u, z) \coloneqq{}& I_1(q, u_{\text{old}}, v)\\ &+ a'_u(q_{\text{old}}, u_{\text{old}})(v, w) + a'_q(q_{\text{old}}, u_{\text{old}})(q - q_{\text{old}}, w)\\ &+ a(q_{\text{old}}, u_{\text{old}})(w_{\text{old}}) - f(w_{\text{old}})\\ &+ a(q, u)(z) - f(z)\,.\end{aligned}$$

and the spaces $X \coloneqq Q \times V \times V \times W \times W \times V \times W$ and $X_h \coloneqq Q_h \times V_h \times V_h \times W_h \times W_h \times V_h \times W_h$. Then we consider

$$I_4(u) \coloneqq \|C(u) - g^\delta\|_G^2\,,$$

where the reader should not be confused about the ambiguity of I_4 as compared with (4.6), since both definitions represent the same thing and are just expressed by different arguments. For $x_3 \in X$ we define the auxiliary Lagrange functional

$$K(x, x_3) \coloneqq I_4(u) + L'(x)(x_3)$$

for $x, x_3 \in X$. Then we could estimate the difference $I_4(u) - I_4(u_h)$ by computing a discrete stationary point $\xi_h = (x_h, x_{3,h})$ of K; that means, we would solve the discrete counterparts of the equations (4.40)/(4.41)–(4.47) and

$$x_{3,h} \in X_h\colon \qquad L''(x_h)(x_{3,h}, dx) = -I'_4(u_h)(du) \quad \forall dx \in X_h\,,$$

(where $dx = (dq, du_{\text{old}}, dv, dw, dw_{\text{old}}, du, dz)$) and compute the error estimator η_4 for I_4 by

$$I_4 - I_{4,h} = I_4(u) - I_4(u_h) \approx \frac{1}{2}K'(\xi_h)(\pi_h\xi_h - \xi_h) = \eta_4\,.$$

Remark 4.10 also holds for the computation of $x_{3,h}$, i.e., $x_{3,h}$ can be computed at low additional costs.

4.1.3. Algorithm

By means of the error estimators from Section 4.1.2 and the results from [39], in particular Theorem 4.1, we formulate the following two algorithms. The main algorithm is given in Algorithm 4.1 and illustrated by means of a flowchart in Figure 4.2. It controls the Gauss-Newton iteration in which Algorithm 4.2 is used to produce a regularization parameter and a discretization by iteratively solving the IRGNM subproblem as in [39] and similar to Algorithm 3.1.

As mentioned and justified in Section 4.1.2, we neglect η_4^k and the conditions $\eta_2^k \to 0$ and $\eta_3^k \to 0$ as $k \to \infty$ from (4.28). In order to verify the condition (4.27) more easily, we split it into

$$\eta_1^k \le c_4 I_{3,h}^k \quad \text{with} \quad c_4 \coloneqq \frac{1}{2}\left(\underline{\tilde{\theta}} - 2\left(c_{tc}^2 + \frac{(1+c_{tc})^2}{\tau^2}\right)\right) \tag{4.50}$$

and

$$\eta_3^k \leq c_5 I_{3,h}^k \quad \text{with} \quad c_5 := \frac{1}{4c_{tc}^2}\left(\underline{\tilde{\theta}} - 2\left(c_{tc}^2 + \frac{(1+c_{tc})^2}{\tau^2}\right)\right) . \tag{4.51}$$

Additionally, we combine the inequality in (4.28), the first inequality in (4.29) and (4.51), since there holds

$$I_{3,h}^k \leq I_3^k + \eta_3^k \quad \text{and} \quad I_{4,h}^{k-1} \geq I_4^{k-1} - \eta_4^{k-1} ,$$

such that the condition

$$\eta_3^k + (1+c_3)\eta_4^{k-1} \leq (1+c_3)I_4^{k-1} - I_3^k + r^k \tag{4.52}$$

implies the first inequality in (4.29). As mentioned in Remark 4.4, I_3^k and I_4^{k-1} and $I_{3,h}^k$ and $I_{4,h}^{k-1}$ only differ in the discretization level, which motivates the assumption that for small h (i.e., for a fine discretization), we have $I_3^k \approx I_4^{k-1}$ and $\eta_4^{k-1} \approx \eta_3^k$, such that instead of (4.52) we check whether

$$\eta_3^k \leq \frac{c_3}{2(1+c_3)} I_{3,h}^k + \frac{r^k}{2(1+c_3)} . \tag{4.53}$$

Thus, as a combination of the inequalities in (4.28), (4.53) and (4.51), we formulate the condition

$$\eta_3^k \leq \min\left\{c_1, c_5, \frac{c_3}{2(1+c_3)}\right\} I_{3,h}^k , \tag{4.54}$$

which we check in each iteration step (see step 5 and 22 in Algorithm 4.1).

In the statement "Refine grids according to the error estimator ..." within the algorithms below, refinement of the spaces Q_h, V_h, and W_h is supposed to be done in principle separately according to the corresponding contributions within the error estimators. In fact, an error estimator of the form described in Section 4.1.2 comes as a sum of contributions corresponding to the different components of the variable that is subject to discretization $(x_h, x_{1,h}, x_{2,h}, x_{3,h})$, which would allow refinement with respect to one component. However, in our computations we set $V = W$ and use the same discretization for Q, V (and W).

Algorithm 4.1: Reduced form of discretized IRGNM

Choose τ, τ_β, $\tilde{\tau}_\beta$, $\underline{\tilde{\theta}}$, $\tilde{\tilde{\theta}}$ such that $0 < \underline{\tilde{\theta}} \leq \tilde{\tilde{\theta}} < 1$, $\max\{1\,, \tilde{\tau}_\beta\} < \tau_\beta \leq \tau$ and (4.26) holds, set $\tilde{\theta} = (\underline{\tilde{\theta}} + \tilde{\tilde{\theta}})/2$ and choose the constants c_1, c_2 and c_3, such that the second part of (4.29) is fulfilled.

Choose a discretization $h = h_0$ and a starting value $q_h^0 = q_{h_0}^0$ (not necessarily coinciding with q_0 in the regularization term) , and set $q_{\text{old}}^0 = q_{h_0}^0$.

Determine $u_{\text{old}}^0 = u_{\text{old},h_0}^0$, $I_{3,h}^0 = I_{3,h_0}^0$ and $\eta_3^0 = \eta_{3,h_0}^0$ by applying Algorithm 4.3 with $m = 0$ (and $h = h_0$).

Set $h_0^1 = h_0$.

while (4.54) is violated **do**

  Refine grids according to the error estimator η_3^0, such that we obtain a finer discretization h_0^1.

  Determine $u_{\text{old}}^0 = u_{\text{old},h_0^1}^0$, $I_{3,h}^0 = I_{3,h_0^1}^0$ and $\eta_3^0 = \eta_{3,h_0^1}^0$ by applying Algorithm 4.3 with $h = h_0^1$ and $m = 0$.

Set $k = 0$ and $h = h_0^1$ (possibly different from h_0).

while $I_{3,h}^k \geq \tau^2\delta^2$ or $k = 0$ **do**

  Set $h = h_k^1$.

  With q_{old}^k, $u_{\text{old},h}^k$ fixed, apply Algorithm 4.2 starting with the current mesh $h(= h_k^1)$ to obtain a regularization parameter β_k and a possibly different discretization h_k^2 such that (4.18) holds and the corresponding $v_h^k = v_{h_k^2}^k$, $q_h^k = q_{h_k^2}^k$.

  Set $h = h_k^2$.

  Evaluate error estimator $\eta_1^k = \eta_1^k(h_k^2)$.

  Set $h_k^3 = h_k^2$.

  while (4.50) is violated **do**

    Refine grids according to the error estimator η_1^k, such that we obtain a finer discretization h_k^3.

    Set $h = h_k^3$.

    With q_{old}^k and $u_{\text{old},h}^k$ fixed, determine $q_h^k = q_{h_k^3}^k$ and $v_h^k = v_{h_k^3}^k$ by solving (4.55)

  Set $q_{\text{old}}^{k+1} = q_h^k$

  Compute $u_{\text{old}}^{k+1} = u_{\text{old},h_k^3}^{k+1}$, $I_{3,h}^{k+1} = I_{3,h_k^3}^{k+1}$, and $\eta_3^{k+1} = \eta_{3,h_k^3}^{k+1}$ by applying Algorithm 4.3 with $m = k+1$ and $h = h_k^3$.

  Set $h_k^4 = h_k^3$.

  while (4.54) is violated **do**

    Refine grid according to the error estimator η_3^{k+1}, such that we obtain a finer discretization h_k^4.

    Determine $u_{\text{old},h}^{k+1} = u_{\text{old},h_k^4}^{k+1}$, $I_{3,h}^{k+1} = I_{3,h_k^4}^{k+1}$ and $\eta_3^{k+1} = \eta_{3,h_k^4}^{k+1}$ by applying Algorithm 4.3 with $m = k+1$ and $h = h_k^4$.

  Set $h_{k+1}^1 = h_k^4$ (i.e., use the current mesh as a starting mesh for the next iteration)

  Set $k = k+1$

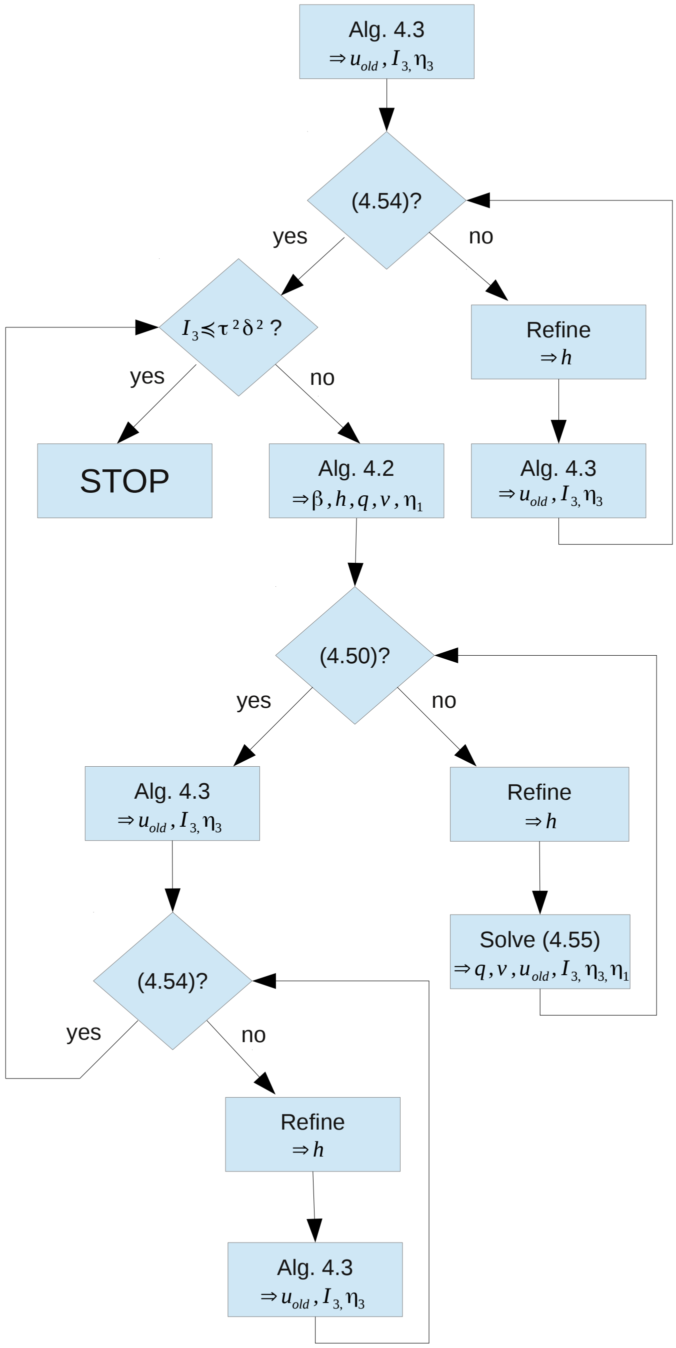

Figure 4.2.: Illustration of Algorithm 4.1. The arrow within the blocks point to variables that are computed in this step.

Algorithm 4.2: Inexact Newton method for the determination of a regularization parameter for the IRGNM subproblem

---

Set $\delta_\beta = \sqrt{\tilde{\theta} I_{3,h}^{k-1}}/\tau_\beta$.
Compute a Lagrange triple $x_h = (q_h, v_h, z_h)$ to the linear-quadratic optimization problem

$$\begin{cases} \min\limits_{(q,v)\in Q_h\times V_h} \|C'(u_{\text{old},h}^k)(v) + C(u_{\text{old},h}^k) - g^\delta\|_G^2 + \dfrac{1}{\beta_k}\|q - q_0\|_Q^2 \\ \text{s.t.} \quad a_u'(q_{\text{old}}^k, u_{\text{old},h}^k)(v,\varphi) + a_q'(q_{\text{old}}^k, u_{\text{old},h}^k)(q - q_{\text{old}}^k, \varphi) \\ \qquad + a(q_{\text{old}}^k, u_{\text{old},h}^k)(\varphi) - f(\varphi) = 0 \qquad \forall \varphi \in W_h \end{cases} \tag{4.55}$$

Evaluate $i_h = I_{2,h}^k = \|C'(u_{\text{old}}^k)(v_h) + C(u_{\text{old}}^k) - g^\delta\|_G^2$ (also see (2.37)).
while $i_h > (\tau_\beta^2 + \frac{\tilde{\tau}_\beta^2}{2})\delta_\beta^2$ **do**
  Evaluate i_h' (cf. [39], Chapter 3).
  Evaluate error estimator for $i(\beta) = I(v(\beta))$ with $I\colon v \mapsto I_2(u_{\text{old}}^k, v)$ (cf. [39]).
  Evaluate error estimator for $i'(\beta) = \frac{d}{d\beta} I(v(\beta))$ (cf. [39], Chapter 3).
  if accuracy requirements (cf. [39]) are violated **then**
    Refine with respect to the corresponding error estimator.
  else
    Update β according to an inexact Newton method (cf. [39]) $\beta \leftarrow \beta - \frac{i_h}{i_h'}$.
  Compute a Lagrange triple $x_h = (q_h, v_h, z_h)$ to (4.55).
  Evaluate $i_h = I_{2,h}^k = \|C'(u_{\text{old}}^k)(v_h) + C(u_{\text{old}}^k) - g^\delta\|_G^2$.

Remark 4.11. *Algorithm 4.2 is equivalent to the algorithm given in [39] with the following replacements*

in [39]	*here*
q	$q - q_0$
T	$F'(q_{\text{old}}^k)$
g^δ	$F'(q_{\text{old}}^k)(q_{\text{old}}^k) + g^\delta - F(q_{\text{old}}^k)$
$\tau^2\delta^2$	$\tilde{\theta} I_{3,h}^k$
$(\tau - \tilde{\tau})^2\delta^2$	$\underline{\tilde{\theta}} I_{3,h}^k$
$(\tau + \tilde{\tau})^2\delta^2$	$\bar{\tilde{\theta}} I_{3,h}^k$

With respect to loops and the solution of PDEs and optimization problems, the algorithm has the form sketched in Algorithm 4.4.

(We do not display the refinement loops of Algorithm 4.1 and Algorithm 4.2 but only the iteration loops.)

Algorithm 4.3: Evaluation of I_3^m

Determine

$$u_{\text{old}}^m \in V_h : \qquad a(q_{\text{old}}^m, u_{\text{old}}^m)(\varphi) = f(\varphi) \qquad \forall \varphi \in W_h \,.$$

Evaluate $I_{3,h}^m$ according to (4.14).
Evaluate error estimator η_3^m.

Algorithm 4.4: Loops in reduced form of discretized IRGNM

while $\cdots$ (Newton iteration) **do**
  Apply algorithm from [39], i.e.,
  while $\cdots$ (Iteration for β_k) **do**
    Solve linear-quadratic optimization problem (i.e., solve linear PDE).
    Update β and refine eventually.
  Solve nonlinear PDE.

Algorithm 4.5: Loops in inexact Newton method (for nonlinear problems)

while $\cdots$ (Iteration for β) **do**
  Solve nonlinear optimization problem (i.e., solve nonlinear PDE).
  Update β and refine eventually.

In contrast with the nonlinear Tikhonov method

$$\min_{q \in \mathcal{D} \subset Q} \|F(q) - g^\delta\|_G^2 + \frac{1}{\beta} \|q - q_0\|_Q^2$$

investigated in Chapter 3 (cf. Algorithm 3.1 or rather its simplified version Algorithm 4.5), we have one additional loop, but we only have to solve a linear-quadratic optimization problem instead of a nonlinear problem. On the other hand, we still have to solve (at least) one nonlinear PDE in each outer loop. For this reason we doubt whether Algorithm 4.1 pays off with respect to computation time as compared to the method from Chapter 3. Therefore we do not implement this algorithm, but are going to consider more efficient modifications in Section 4.2.

4.1.4. Extension to more general data misfit and regularization terms

Motivated by the increasing use of nonquadratic, non-Hilbert space misfit and regularization terms for modelling, e.g., sparsity of the solution, or non-Gaussian data noise (cf., e.g., [32, 97] for Tikhonov regularization, and [52] for IRGNM), we now extend our results to a more general setting (also see [33, 49]). Inspection of the proofs in Section 4.1.1 gives reason to believe that this can be done in a straight-forward manner.

To this purpose, we consider a more general version of (4.9):

$$\min_{q\in Q} \mathcal{T}_\beta(q) := S(F'(q^k_{\text{old}})(q - q^k_{\text{old}}) + F(q^k_{\text{old}}), g^\delta) + \frac{1}{\beta}R(q) \tag{4.56}$$

with quantities of interest (cf. (4.6))

$$\begin{aligned} I^k_1 &:= S(F'(q^k_{\text{old}})(q^k - q^k_{\text{old}}) + F(q^k_{\text{old}}), g^\delta) + \frac{1}{\beta}R(q^k) \\ I^k_2 &:= S(F'(q^k_{\text{old}})(q^k - q^k_{\text{old}}) + F(q^k_{\text{old}}), g^\delta) \\ I^k_3 &:= S(F(q^k_{\text{old}}), g^\delta) \\ I^k_4 &:= S(F(q^k), g^\delta) \end{aligned} \tag{4.57}$$

and their discrete counterparts

$$\min_{q\in Q_h} S(F'_h(q^k_{\text{old}})(q - q^k_{\text{old}}) + F_h(q^k_{\text{old}}), g^\delta) + \frac{1}{\beta}R(q) \tag{4.58}$$

(cf. (4.12)) with

$$\begin{aligned} I^k_{1,h} &:= S(F'_{h_k}(q^k_{\text{old}})(q^k_{h_k} - q^k_{\text{old}}) + F_{h_k}(q^k_{\text{old}}), g^\delta) + \frac{1}{\beta}R(q^k_{h_k}) \\ I^k_{2,h} &:= S(F'_{h_k}(q^k_{\text{old}})(q^k_{h_k} - q^k_{\text{old}}) + F_{h_k}(q^k_{\text{old}}), g^\delta) \\ I^k_{3,h} &:= S(F_{h_k}(q^k_{\text{old}}), g^\delta) \\ I^k_{4,h} &:= S(F_{h_k}(q^k_{h_k}), g^\delta) \end{aligned} \tag{4.59}$$

(cf. (4.14)).

Let the data misfit and regularization functionals $S\colon G \times G \to \bar{\mathbb{R}}$ and $R\colon Q \to \bar{\mathbb{R}}$ have the following properties.

Assumption 4.6.

1. *The mapping $y \mapsto S(y, g^\delta)$ is convex.*
2. *S is symmetric, i.e., $S(y, \tilde{y}) = S(\tilde{y}, y)$ for all $y, \tilde{y} \in G$.*
3. *S is positive definite, i.e., $S(y, \tilde{y}) \geq 0$ and $S(y, y) = 0$ for all $y, \tilde{y} \in G$.*
4. *There exists a constant c_S such that for all $y, \tilde{y}, \widehat{y} \in G$ there holds $S(y, \tilde{y}) \leq c_S(S(y, \hat{y}) + S(\hat{y}, \tilde{y}))$.*
5. *The regularization functional R is proper (i.e., the domain of R is non-empty) and convex.*

The domain of a functional $\mathcal{R}\colon M \to \bar{\mathbb{R}}$ should be understood as

$$\mathcal{D}(\mathcal{R}) := \{m \in M |\ \mathcal{R}(m) \neq \infty\}\,. \tag{4.60}$$

Remark 4.12. *In fact, in item 3 in Assumption 4.6 it suffices to require $S(y, y) = 0$ only for $y = g$ (i.e., for the exact data), but since item 3 is a more "natural" assumption in terms of general operator properties, we stick with the stronger assumption item 3.*

We refer once more to [52], where convergence and convergence rates for the IRGNM have already been established in an even more general (continuous) setting and mention that we consider a somewhat simpler situation with stronger assumptions on S, R here, since our intention is mainly to demonstrate transferability of the adaptive discretization concept. Moreover, note that we rely on a different choice of the regularization parameter here. The results obtained here will allow us to easily establish convergence rates results for an exact penalty formulation of an all-at-once formulation of the IRGNM in Section 4.2.2.

Although we will, again, restrict ourselves to Hilbert spaces in the next sections, at this point we discuss convergence in a Banach space setting to emphasize the generality of the subsequent results. To this purpose we introduce the *Bregman distance*

$$D_R^{\xi}(q, \overline{q}) := R(q) - R(\overline{q}) - \langle \xi, q - \overline{q} \rangle_{Q^*,Q} \tag{4.61}$$

with some $\xi \in \partial R(q) \subset Q^*$, which coincides with $\frac{1}{2}\|q - \overline{q}\|_Q^2$ for $R(q) = \frac{1}{2}\|q - q_0\|_Q^2$ and $\xi = \overline{q} - q_0$ in a Hilbert space Q. With $\partial R(q)$ we denote the subdifferential of R at q.

Well-definedness (i.e., for every $g^\delta \in G$ and $\beta_k > 0$ there exists a solution $q_{h_k}^k$ to (4.58)) and stable dependence on the data (i.e., for every fixed $\beta_k > 0$ the solution $q_k^{h_k}$ depends continuously on g^δ) can be shown under the following assumptions (cf., e.g., [97, Assumption 1.32], [52, Remark 2.1], [49, Theorem 3.1, Theorem 3.2])

Assumption 4.7.

1. *Q and G are Banach spaces, with which there are associated topologies $\mathcal{T}_Q$ and $\mathcal{T}_G$, which are weaker than the norm topologies.*
2. *The mapping $y \mapsto S(y, g^\delta)$ is sequentially lower semi-continuous with respect to $\mathcal{T}_G$.*
3. *$F'(q_{\text{old}}^k) \colon Q \to G$ is continuous with respect to the topologies $\mathcal{T}_Q$ and $\mathcal{T}_G$.*
4. *$\mathcal{R} \colon Q \to (-\infty, +\infty]$ is proper, convex and $\mathcal{T}_Q$-lower semi continuous.*
5. *$\mathcal{D} := \mathcal{D}(F) \cap \mathcal{D}(R) \neq \emptyset$ is closed with respect to $\mathcal{T}_Q$.*
6. *For every $\rho > 0$ the set*
$$\mathcal{B}_\rho := \{q \in \mathcal{D} |\ R(q) \leq \rho\}$$
is $\mathcal{T}_Q$-sequentially compact in the following sense: every sequence $(q_n)_{n\in\mathbb{N}}$ in this set has a subsequence, which is convergent in Q with respect to the $\mathcal{T}_Q$-topology.

Remark 4.13. *For the setting from Chapter 3, i.e., Hilbert spaces Q and G and the choice $S(y, \tilde{y}) := \frac{1}{2}\|y - \tilde{y}\|_G^2$ and $R(q) := \frac{1}{2}\|q - q_0\|_Q^2$ Assumption 4.6 and 4.7 are obviously fulfilled. As for examples in a real Banach spaces setting, we refer to [48, 71, 97].*

Consistently, the conditions on F, Assumption 3.2 and Assumption 4.4 are generalized to the following two assumptions.

Assumption 4.8. *Let the reduced forward operator F be continuous with respect to $\mathcal{T}_Q$,$\mathcal{T}_G$ and satisfy*

$$(q_n \xrightarrow{\mathcal{T}_Q} q \ \wedge \ S(F(q_n), g) \to 0) \ \Rightarrow \ (q \in \mathcal{D} \ \wedge \ F(q) = g)$$

for all $(q_n)_{n\in\mathbb{N}} \subseteq Q$.

Assumption 4.9. *Let the generalized tangential cone condition*

$$S(F(q), F(\bar{q}) + F'(\bar{q})(q - \bar{q})) \leq c_{tc}^2 S(F(q), F(\bar{q}))$$

hold for all $q, \bar{q} \in Q$ in the level set $\mathcal{B}_\rho \subset \mathcal{D}$ for some $\rho > 0$ and for some $0 < c_{tc} < 1$.

Moreover, the source condition in Assumption 3.5 and Assumption 4.5 are replaced by the following conditions.

Assumption 4.10. *Let the multiplicative variational inequality*

$$|\langle \xi, q - q^\dagger \rangle_{Q^*,Q}| \leq c\sqrt{D_R^\xi(q, q^\dagger)}\kappa \left(\frac{S(F(q), F(q^\dagger))}{D_R^\xi(q, q^\dagger)} \right)$$

hold for all $q \in \mathcal{D}$ (cf. (3.14)*) and some constant $c > 0$, where κ is defined as in Assumption 3.5.*

Assumption 4.11. *Let τ be chosen sufficiently large and $0 < \underline{\tilde{\theta}} < \tilde{\tilde{\theta}}$ sufficiently small (see* (4.18)*,* (4.25)*), such that*

$$c_S \left(c_S c_{tc}^2 + \frac{1 + c_S c_{tc}^2}{\tau^2} \right) < \underline{\tilde{\theta}} \quad \text{and} \quad 0 < \frac{c_S \tilde{\tilde{\theta}} + c_S^2 c_{tc}^2}{1 - c_S^2 c_{tc}^2} < 1\,. \tag{4.62}$$

Finally, let the estimates (4.17) *hold for the discretization error with respect to the quantities of interest* (4.57)*,* (4.59)*, where η_1^k, η_2^k, η_3^k, η_4^k are selected such that*

$$\eta_1^k + c_S^2 c_{tc}^2 \eta_3^k \leq \left(\underline{\tilde{\theta}} - c_S \left(c_S c_{tc}^2 + \frac{1 + c_S c_{tc}^2}{\tau^2} \right) \right) I_{3,h}^k\,, \tag{4.63}$$

$$\eta_3^k \leq c_1 I_{3,h}^k \quad \text{and} \quad \eta_2^k \to 0\,,\ \eta_3^k \to 0\,,\ \eta_4^k \to 0 \ \text{as } k \to \infty\,, \tag{4.64}$$

$$I_{3,h}^k \leq (1 + c_3) I_{4,h}^{k-1} + r^k\,, \quad \text{and} \quad (1 + c_3)\frac{c_S \tilde{\tilde{\theta}} + c_S^2 c_{tc}^2}{1 - c_S^2 c_{tc}^2} \leq c_2 < 1 \tag{4.65}$$

hold for some constants $c_1, c_2, c_3 > 0$, and a sequence $r^k \to 0$ as $k \to \infty$ (where the second condition in (4.65) *is possible due to the right inequality in* (4.62)*).*

Based on this groundwork, we can now formulate a convergence theorem similar to Theorem 4.2:

Theorem 4.4. *Let Assumption 4.6, 4.7, 4.8, 4.9 and 4.11 be satisfied with $0 < c_{tc} < \frac{1}{c_S}$ sufficiently small and let $q^\dagger$ (solution to* (2.1)*) lie in the neighborhood of q_0 where Assumption 4.9 holds, i.e., $q^\dagger \in \mathcal{B}_\rho \subset \mathcal{D}$. Let the starting value be chosen such that $q_{\text{old}}^1 = q_{h_0}^0 \in \mathcal{B}_\rho$ and let*

$$S(g, g^\delta) \leq \delta^2\,. \tag{4.66}$$

Then with β_k and $h = h_k$ fulfilling (4.18) *and k_* selected according to* (4.25)*, there holds:*

(i) the estimate

$$R(q_{h_k}^k) \le R(q^\dagger) \quad \forall 1 \le k < k_*\,, \tag{4.67}$$

(ii) k_* is finite,

(iii) $q_{\text{old}}^{k_*} = q_{\text{old}}^{k_*(\delta)} = q_{\beta_{k_*(\delta)-1},h_{k_*(\delta)-1}}^{k_*(\delta)-1,\delta}$ converges (weakly) subsequentially to a solution of (2.1) *as $\delta \to 0$ in the sense that it has a $\mathcal{T}_Q$-convergent subsequence and each $\mathcal{T}_Q$-convergent subsequence converges to a solution of* (2.1)*. If the solution $q^\dagger$ to* (2.1) *is unique, then $q_{\text{old}}^{k_*(\delta)}$ converges to $q^\dagger$ with respect to $\mathcal{T}_Q$ as $\delta \to 0$.*

Proof. The proof basically follows the lines of the proof of Theorem 4.2, where we have to replace the specific fitting and regularization terms by S and R:

(i): We will prove (4.31) as follows: We will show that, for fixed $k > 0$, provided that $q_{\text{old}}^k = q_{h_{k-1}}^{k-1} \in \mathcal{B}_\rho$, there holds

$$R(q_{h_k}^k) \le R(q^\dagger)$$

which implies $q_{\text{old}}^{k+1} = q_{h_k}^k \in \mathcal{B}_\rho$. Then (4.31) follows recursively, since $q_{\text{old}}^1 = q_{h_0}^0 \in \mathcal{B}_\rho$.

So we assume that $q_{\text{old}}^k \in \mathcal{B}_\rho$ in the following.

According to (4.66), (4.25) and Assumption 4.9 as well as the inequaltity $(a+b)^2 \le 2a^2 + 2b^2$ for arbitrary $a, b \in \mathbb{R}$, we can estimate as follows

$$\begin{aligned}
S(F'(q_{\text{old}}^k)(q^\dagger - q_{\text{old}}^k) + F(q_{\text{old}}^k), g^\delta) &\le c_S\left(S(g, F'(q_{\text{old}}^k)(q^\dagger - q_{\text{old}}^k) + F(q_{\text{old}}^k)) + S(g, g^\delta)\right)\\
&\le c_S\left(S(F(q^\dagger), F'(q_{\text{old}}^k)(q^\dagger - q_{\text{old}}^k) + F(q_{\text{old}}^k)) + \delta^2\right)\\
&\le c_S\left(c_{tc}^2 S(g, F(q_{\text{old}}^k)) + \delta^2\right)\\
&\le c_S\left(c_{tc}^2 c_S\left(S(g^\delta, F(q_{\text{old}}^k)) + \delta^2\right) + \delta^2\right)\\
&= c_S^2 c_{tc}^2\left(I_3^k + \delta^2\right) + c_S\delta^2\\
&= c_S^2 c_{tc}^2 I_3^k + c_S(1 + c_S c_{tc}^2)\delta^2\\
&\le c_S^2 c_{tc}^2 (I_{3,h}^k + \eta_3^k) + c_S(1 + c_S c_{tc}^2)\frac{I_{3,h}^k}{\tau^2}\\
&\le c_S\left(c_S c_{tc}^2 + \frac{1 + c_S c_{tc}^2}{\tau^2}\right) I_{3,h}^k + c_S^2 c_{tc}^2 \eta_3^k\,.
\end{aligned}$$

On the other hand, from (4.18) and the fact that $I_{1,h}^k = I_{2,h}^k + \frac{1}{\beta_k}R(q_h^k)$ (cf. (4.16)) it follows that

$$I_{1,h}^k = I_{2,h}^k + \frac{1}{\beta_k}R(q_h^k) \ge \underline{\tilde\theta} I_{3,h}^k + \frac{1}{\beta_k}R(q_h^k)\,,$$

which together with the previous inequality and (4.63) gives

$$
\begin{aligned}
\underline{\tilde{\theta}} I_{3,h}^k + \frac{1}{\beta_k} R(q_h^k) &\leq I_1^k + \eta_1^k \\
&\leq S(F'(q_{\text{old}}^k)(q^k - q_{\text{old}}^k) + F(q_{\text{old}}^k), g^\delta) + \frac{1}{\beta_k} R(q^k) + \eta_1^k \\
&\leq S(F'(q_{\text{old}}^k)(q^\dagger - q_{\text{old}}^k) + F(q_{\text{old}}^k), g^\delta) + \frac{1}{\beta_k} R(q^\dagger) + \eta_1^k \\
&\leq c_S \left(c_S c_{tc}^2 + \frac{1 + c_S c_{tc}^2}{\tau^2} \right) I_{3,h}^k + c_S^2 c_{tc}^2 \eta_3^k + \frac{1}{\beta_k} R(q^\dagger) + \eta_1^k \\
&\leq \underline{\tilde{\theta}} I_{3,h}^k + \frac{1}{\beta_k} R(q^\dagger) \,,
\end{aligned}
$$

where we have used minimality of q^k for (4.56). This finally implies (4.67).

(ii): Furthermore, for all $k < k_*$ we have by the triangle inequality as well as Assumption 4.9 and (4.18)

$$
\begin{aligned}
I_4^k &= S(F(q^k), g^\delta) \\
&\leq c_S \left(S(F'(q_{\text{old}}^k)(q^k - q_{\text{old}}^k) + F(q_{\text{old}}^k), g^\delta) + S(F'(q_{\text{old}}^k)(q^k - q_{\text{old}}^k) + F(q_{\text{old}}^k), F(q^k)) \right) \\
&\leq c_S \left(I_2^k + c_{tc}^2 S(F(q^k), F(q_{\text{old}}^k)) \right) \\
&\leq c_S \left(I_{2,h}^k + \eta_2^k \right) + c_S^2 c_{tc}^2 \left(S(F(q^k), g^\delta) + S(F(q_{\text{old}}^k), g^\delta) \right) \\
&\leq c_S \left(\tilde{\tilde{\theta}} I_{3,h}^k + \eta_2^k \right) + c_S^2 c_{tc}^2 \left(I_4^k + I_3^k \right) \,,
\end{aligned}
\tag{4.68}
$$

which implies

$$
I_4^k \leq \frac{1}{1 - c_S^2 c_{tc}^2} \left(c_S \tilde{\tilde{\theta}} I_{3,h}^k + c_S \eta_2^k + c_S^2 c_{tc}^2 I_3^k \right) . \tag{4.69}
$$

With (4.17) and (4.65) we can further deduce

$$
\begin{aligned}
I_{4,h}^k &\leq \frac{1}{1 - c_S^2 c_{tc}^2} \left((c_S \tilde{\tilde{\theta}} + c_S^2 c_{tc}^2) I_{3,h}^k + c_S \eta_2^k + c_S^2 c_{tc}^2 \eta_3^k \right) + \eta_4^k \\
&\leq \frac{c_S \tilde{\tilde{\theta}} + c_S^2 c_{tc}^2}{1 - c_S^2 c_{tc}^2} (1 + c_3) I_{4,h}^{k-1} + \frac{1}{1 - c_S^2 c_{tc}^2} \left((c_S \tilde{\tilde{\theta}} + c_S^2 c_{tc}^2) r^k + c_S \eta_2^k + c_S^2 c_{tc}^2 \eta_3^k \right) + \eta_4^k \\
&\leq c_2 I_{4,h}^{k-1} + \frac{1}{1 - c_S^2 c_{tc}^2} \left((c_S \tilde{\tilde{\theta}} + c_S^2 c_{tc}^2) r^k + c_S \eta_2^k + c_S^2 c_{tc}^2 \eta_3^k \right) + \eta_4^k \,.
\end{aligned}
$$

With the notation

$$
a^i := \frac{1}{1 - c_S^2 c_{tc}^2} \left((c_S \tilde{\tilde{\theta}} + c_S^2 c_{tc}^2) r^i + c_S \eta_2^i + c_S^2 c_{tc}^2 \eta_3^i \right) + \eta_4^i \quad \forall i \in \{1, 2, \ldots, k\}
$$

there follows recursively

$$
I_{4,h}^k \leq c_2^k I_{4,h}^0 + \sum_{j=0}^{k-1} c_2^j a^{k-j} \,.
$$

Note that by the second part of (4.64), the second part of (4.65) and the fact that $r^k \to 0$ as $k \to \infty$ (by definition of r^k), we have $c_2^k I_{4,h}^0 + \sum_{j=0}^{k-1} c_2^j a^{k-j} \to 0$ as $k \to \infty$. So, if the discrepancy principle never got active (i.e., $k_* = \infty$), the sequence $(I_{4,h}^k)_{k\in\mathbb{N}}$ and therewith by assumption (4.64) also $(I_{3,h}^k)_{k\in\mathbb{N}}$ would be bounded by a sequence tending to zero as $k \to \infty$, which implies that $I_{3,h}^k$ falls below $\tau^2\delta^2$ for k sufficiently large, thus yielding a contradiction. Hence the stopping index $k_* < \infty$ is well-defined and finite.

(iii): With (2.2), (4.17), (4.64) and definition of k_*, we have

$$\begin{aligned} S(F(q_{\text{old}}^{k_*}), g) &\le c_S\left(S(F(q_{\text{old}}^{k_*}), g^\delta) + \delta^2\right) \le c_S\left(I_{3,h}^{k_*} + \eta_3^k + \delta^2\right) \le c_S\left((1+c_1)I_{3,h}^{k_*} + \delta^2\right) \\ &\le c_S\left((1+c_1)\tau^2 + 1\right)\delta^2 \to 0 \end{aligned} \tag{4.70}$$

as $\delta \to 0$. By (i), (ii) and item 6 in Assumption 4.7 $q_{\text{old}}^{k_*} = q_{h_{k_*-1}}^{k_*}$ has a $\mathcal{T}_Q$-convergent subsequence and due to Assumption 4.8 and (4.70) the limit of every $\mathcal{T}_Q$-convergent subsequence is a solution to $F(q) = g$.

□

As far as convergence rates are concerned, Theorem 3.2 (where R, S are defined by squared Hilbert space norms) can be generalized as follows:

Theorem 4.5. *Let the conditions of Theorem 4.4 be satisfied and let $\tilde{q}$ be a regularized approximation (not necessarily defined by Tikhonov regularization) of $q^\dagger$ such that*

$$R(\tilde{q}) \le R(q^\dagger) \tag{4.71}$$

and

$$S(F(\tilde{q}), g^\delta) \le \hat{\tau}^2\delta^2 \tag{4.72}$$

with some $\hat{\tau}$ independent of δ as well as Assumption 4.10 are fulfilled. Then the rate

$$\sqrt{D_R^\xi(\tilde{q}, q^\dagger)} = \mathcal{O}\left(\frac{\delta}{\sqrt{\psi^{-1}(C\delta)}}\right)$$

with Bregman distance D_R^ξ (cf. (4.61)) holds for some constant $C > 0$ independent of δ.

Proof. The proof can be done similarly to the proof of Theorem 3.2 with the same replacements as in the proof of Theorem 4.4:

If $D_R^\xi(\tilde{q}, q^\dagger)$ vanishes we are trivially done. So we assume $D_R^\xi(\tilde{q}, q^\dagger) \ne 0$ for the rest of the proof.

From (4.71) and Assumption 4.10 there follows

$$\begin{aligned} D_R^\xi(\tilde{q}, q^\dagger) &= R(\tilde{q}) - R(q^\dagger) - \langle \xi, \tilde{q} - q^\dagger\rangle_{Q^*,Q} \le |\langle \xi, \tilde{q} - q^\dagger\rangle_{Q^*,Q}| \\ &\le c\sqrt{D_R^\xi(\tilde{q}, q^\dagger)}\,\kappa\left(\frac{S(F(\tilde{q}), F(q^\dagger))}{D_R^\xi(\tilde{q}, q^\dagger)}\right) . \end{aligned}$$

By monotonicity of κ, (4.72), (4.66) and Assumption 4.6 we further get

$$D_R^\xi(\tilde{q}, q^\dagger) \le c\sqrt{D_R^\xi(\tilde{q}, q^\dagger)}\kappa\left(c_S\frac{S(g^\delta, F(q^\dagger)) + S(F(\tilde{q}), g^\delta)}{D_R^\xi(\tilde{q}, q^\dagger)}\right) \le c\sqrt{D_R^\xi(\tilde{q}, q^\dagger)}\kappa\left(c_S\frac{\delta^2(1+\hat{\tau}^2)}{D_R^\xi(\tilde{q}, q^\dagger)}\right).$$

Since $D_R^\xi(\tilde{q}, q^\dagger) \neq 0$, we can multiply both sides by $\frac{\delta\sqrt{1+\hat{\tau}^2}}{D_R^\xi(\tilde{q},q^\dagger)}$ and obtain

$$\delta\sqrt{1+\hat{\tau}^2} \le c\frac{\delta\sqrt{1+\hat{\tau}^2}}{\sqrt{D_R^\xi(\tilde{q}, q^\dagger)}}\kappa\left(\frac{c_S\delta^2(1+\hat{\tau}^2)}{D_R^\xi(\tilde{q}, q^\dagger)}\right) = \frac{c}{\sqrt{c_S}}\psi\left(\frac{c_S\delta^2(1+\hat{\tau}^2)}{D_R^\xi(\tilde{q}, q^\dagger)}\right).$$

By strict monotonicity of ψ, this yields

$$\psi^{-1}\left(\frac{\sqrt{c_S}}{c}\delta\sqrt{1+\hat{\tau}^2}\right) \le \frac{c_S\delta^2(1+\hat{\tau}^2)}{D_R^\xi(\tilde{q}, q^\dagger)},$$

which implies

$$\sqrt{D_R^\xi(\tilde{q}, q^\dagger)} \le \frac{\delta\sqrt{1+\hat{\tau}^2}}{\psi^{-1}\left(\frac{\sqrt{c_S}}{c}\delta\sqrt{1+\hat{\tau}^2}\right)}.$$

□

Due to Theorem 4.4 these rates particularly hold for $q_h^{k_*}$.

So far we have not commented on well-definedness of the regularization parameter β_k by (4.18) with $I_{2,h}^k$ and $I_{3,h}^k$ according to (4.59). As a matter of fact, this can be shown analogously to [71, Lemma 1, Theorem 3] (where the analysis is restricted to Banach space norms) under some additional assumptions on S and R. For simplicity of exposition we will only do so for the continuous setting. The corresponding proof in the discrete setting can be done in a straight forward manner by combining the proof from the continuous setting (see Lemma 4.6 below) and the proof of Theorem 4.2 (o), where we showed well-definedness in a discrete Hilbert space setting.

Lemma 4.6. *Let Assumption 4.6, 4.7, 4.8 and 4.9 with (lower) semi-continuity replaced by continuity in the item 2 and 4 of Assumption 4.7 and convexity by strict convexity in item 5 of Assumption 4.6 hold. Moreover, let the first inequality in* (4.62) *(from Assumption 4.11) be satisfied and let*

$$S(g, g^\delta) \le \delta^2.$$

For $k < k_*$ we assume for the (then existent and unique) minimizer q_0 of R that

$$S(F'(q_{\text{old}}^k)(q_0 - q_{\text{old}}^k) + F(q_{\text{old}}^k), g^\delta) \ge \tilde{\overline{\theta}} S(F(q_{\text{old}}^k), g^\delta) \tag{4.73}$$

holds. Then there exists β_k such that

$$\underline{\tilde{\theta}} S(F(q_{\text{old}}^k), g^\delta) \le S(F'(q_{\text{old}}^k)(q^k - q_{\text{old}}^k) + F(q_{\text{old}}^k), g^\delta) \le \tilde{\overline{\theta}} S(F(q_{\text{old}}^k), g^\delta).$$

Proof. We define the functional

$$\Psi\colon \mathbb{R}^+ \to \mathbb{R}_0^+ \,, \quad \beta \mapsto S(F'(q_{\text{old}}^k)(q_\beta^\delta - q_{\text{old}}^k) + F(q_{\text{old}}^k), g^\delta)\,,$$

where q_β^δ is the (under the stated conditions existent and unique) minimizer of the Tikhonov functional $\mathcal{T}_\beta(q)$. Since by minimality of q_β^δ we have

$$\begin{aligned}\Psi(\beta) + \frac{1}{\beta}R(q_0) &\le \Psi(\beta) + \frac{1}{\beta}R(q_\beta^\delta) = \mathcal{T}_\beta(q_\beta^\delta) \le \mathcal{T}_\beta(q^\dagger) \\ &= S(F'(q_{\text{old}}^k)(q^\dagger - q_{\text{old}}^k) + F(q_{\text{old}}^k), g^\delta) + \frac{1}{\beta}R(q^\dagger)\,,\end{aligned}$$

we can conclude, using the generalized triangle inequality (item 4 in Assumption 4.6), the generalized tangential cone condition Assumption 4.9 as well as the definition of k_* (3.36)

$$\begin{aligned}\limsup_{\beta\to\infty} \Psi(\beta) &\le S(F'(q_{\text{old}}^k)(q^\dagger - q_{\text{old}}^k) + F(q_{\text{old}}^k), g^\delta) \\ &\le c_S\left(S(F'(q_{\text{old}}^k)(q^\dagger - q_{\text{old}}^k) + F(q_{\text{old}}^k), g) + \delta^2\right) \\ &\le c_S\left(c_{tc}^2 S(F(q_{\text{old}}^k), g) + \delta^2\right) \\ &\le c_S\left(c_{tc}^2 c_S\left[S(F(q_{\text{old}}^k), g^\delta) + \delta^2\right] + \delta^2\right) \\ &\le c_S^2 c_{tc}^2 S(F(q_{\text{old}}^k), g^\delta) + c_S\left(1 + c_S c_{tc}^2\right)\delta^2 \\ &\le c_S^2 c_{tc}^2 S(F(q_{\text{old}}^k), g^\delta) + c_S\left(1 + c_S c_{tc}^2\right)\frac{I_3^k}{\tau^2} \\ &\le c_S\left(c_S c_{tc}^2 + \frac{1 + c_S c_{tc}^2}{\tau^2}\right) S(F(q_{\text{old}}^k), g^\delta) \\ &\le \underline{\tilde{\theta}} S(F(q_{\text{old}}^k), g^\delta)\,.\end{aligned}$$

On the other hand, we have by minimality

$$\begin{aligned}\frac{1}{\beta}R(q_\beta^\delta) &\le \Psi(\beta) + \frac{1}{\beta}R(q_\beta^\delta) = \mathcal{T}_\beta(q_\beta^\delta) \le \mathcal{T}_\beta(q_0) \\ &= S(F'(q_{\text{old}}^k)(q_0 - q_{\text{old}}^k) + F(q_{\text{old}}^k), g^\delta) + \frac{1}{\beta}R(q_0)\,,\end{aligned}$$

hence (by minimality of q_0 for R)

$$0 \le \limsup_{\beta\to 0} R(q_\beta^\delta) - R(q_0) \le \limsup_{\beta\to 0} \beta S(F'(q_{\text{old}}^k)(q_0 - q_{\text{old}}^k) + F(q_{\text{old}}^k), g^\delta) = 0\,,$$

i.e.,

$$R(q_\beta^\delta) \to R(q_0) \quad \text{as } \beta \to 0\,.$$

Hence, by the compactness of the level sets of R (cf. item 6 in Assumption 4.7)), every subsequence of q_β^δ has a $\mathcal{T}_Q$-convergent subsequence whose limit is a minimizer of R due to the lower semi-continuity of R, and hence, by strict convexity of R coincides with q_0. Thus, by a subsequence-subsequence argument q_β^δ converges to q_0 in $\mathcal{T}_Q$ as $\beta \to 0$. By item 2 and 3 in Assumption 4.7 we have

$$\begin{aligned}\liminf_{\beta\to 0} \Psi(\beta) &= \liminf_{\beta\to 0} S(F'(q_{\text{old}}^k)(q_\beta^\delta - q_{\text{old}}^k) + F(q_{\text{old}}^k), g^\delta) \\ &\geq S(F'(q_{\text{old}}^k)(q_0 - q_{\text{old}}^k) + F(q_{\text{old}}^k), g^\delta) \geq \tilde{\bar{\theta}} S(F(q_{\text{old}}^k), g^\delta)\,,\end{aligned}$$

where the latter inequality was just assumed to hold, cf. (4.73). Summarizing, we have

$$\limsup_{\beta\to\infty} \Psi(\beta) \leq \tilde{\underline{\theta}} S(F(q_{\text{old}}^k), g^\delta) < \tilde{\bar{\theta}} S(F(q_{\text{old}}^k), g^\delta) \leq \liminf_{\beta\to 0} \Psi(\beta)\,.$$

If Ψ is continuous, application of the Intermediate Value Theorem yields existence of a β fulfilling

$$\tilde{\underline{\theta}} S(F(q_{\text{old}}^k), g^\delta) \leq \Psi(\beta) \leq \tilde{\bar{\theta}} S(F(q_{\text{old}}^k), g^\delta)\,.$$

Thus, it remains to show the continuity of Ψ, which we will do in the following:

Let $\beta \in (0,\infty)$ be fixed and $(\beta_n)_{n\in\mathbb{N}}$ a sequence converging to β. Then, for sufficiently large n there holds

$$\frac{\beta}{2} \leq \beta_n \leq 2\beta\,.$$

By minimality we have

$$\frac{1}{2\beta} R(q_{\beta_n}^\delta) \leq \frac{1}{\beta_n} R(q_{\beta_n}^\delta) \leq \mathcal{T}_{\beta_n}(q_{\beta_n}^\delta) \leq \mathcal{T}_{\beta_n}(q_\beta^\delta) = \Psi(\beta) + \frac{1}{\beta_n} R(q_\beta^\delta) \leq \Psi(\beta) + \frac{2}{\beta} R(q_\beta^\delta)\,.$$

Hence by item 6 in Assumption 4.7 every subsequence of $q_{\beta_n}^\delta$ has a $\mathcal{T}_Q$-convergent subequence. By item 2., 3., and 4. in Assumption 4.7, and

$$\mathcal{T}_{\beta_n}(q_{\beta_n}^\delta) \leq \mathcal{T}_{\beta_n}(q_\beta^\delta) \to \mathcal{T}_\beta(q_\beta^\delta) \quad \text{as } n \to \infty\,,$$

the limit of every $\mathcal{T}_Q$-convergent subsequence of $q_{\beta_n}^\delta$ is a minimizer of $\mathcal{T}_\beta$. By uniqueness of this minimizer and a subsequence-subsequence argument the whole sequence $q_{\beta_n}^\delta$ converges to q_β^δ in $\mathcal{T}_Q$ as $n \to \infty$. Continuity of $F'(q_{\text{old}}^k)$, $S(\cdot, g^\delta)$, and R implies $\Psi(\beta_n) \to \Psi(\beta)$ as $n \to \infty$. □

In case (4.73) is violated and therefore well-definedness of the choice of β_k cannot be guaranteed, we set the iterate q_h^k to q_0, which formally corresponds to setting $\beta_k = 0$ and implies that the crucial estimate (4.67) trivially holds for such iterates.

4.2. All-at-once formulations

A major drawback of the method discussed in Section 4.1 is the necessity of solving the nonlinear PDE (to a certain precision) in each Newton step in order to evaluate $F(q) = C(S(q))$. As already mentioned in Section 2.5 this can be avoided by considering optimization methods that only work with linearizations of the constraint(s).

To this purpose we simultaneously consider the measurement equation and the PDE:

$$C(u) = g \quad \text{in } G \tag{4.74}$$
$$A(q,u) = f \quad \text{in } W^* \tag{4.75}$$

as a system of operator equations for (q,u), which we will abbreviate by

$$\mathbf{F}(q,u) = \mathbf{g}, \tag{4.76}$$

where

$$\mathbf{F}\colon Q \times V \to G \times W^*\,, \quad \mathbf{F}(q,u) := \begin{pmatrix} C(u) \\ A(q,u) \end{pmatrix}, \qquad \text{and} \qquad \mathbf{g} := \begin{pmatrix} g \\ f \end{pmatrix} \in G \times W^*\,. \tag{4.77}$$

The noisy data for this all-at-once formulation is defined by

$$\mathbf{g}^\delta = \begin{pmatrix} g^\delta \\ f \end{pmatrix} \in G \times W^*\,.$$

Therewith, we will arrive at iterations of the form: Determine (q^k, u^k) as solution to

$$\begin{aligned} \min_{(q,u)\in Q\times V} \varrho \| A_q'(q^{k-1},u^{k-1})(q-q^{k-1}) + A_u'(q^{k-1},u^{k-1})(u-u^{k-1}) + A(q^{k-1},u^{k-1}) - f\|_{W^*}^r \\ + \|C(u^{k-1}) + C'(u^{k-1})(u-u^{k-1}) - g^\delta\|_G^2 + \frac{1}{\beta_k}\left(\|q-q_0\|_Q^2 + \|u-u_0\|_V^2\right) \end{aligned} \tag{4.78}$$

with $\varrho > 0$, $r \in \{1,2\}$.
For $r = 2$, this yields a least squares formulation, see Section 4.2.1.
In case $r = 1$ and ϱ sufficiently large, by exactness of the norm with exponent one as a penalty (cf. Section 2.5.4) this leads to a Generalized Gauss-Newton type [18] form of the IRGNM: Determine (q^k, u^k) as solution to

$$\begin{aligned} &\min_{(q,u)\in Q\times V} \|C(u^{k-1}) + C'(u^{k-1})(u-u^{k-1}) - g^\delta\|_G^2 + \frac{1}{\beta_k}\left(\|q-q_0\|_Q^2 + \|u-u_0\|_V^2\right) \\ &\text{s.t. } A_q'(q^{k-1},u^{k-1})(q-q^{k-1}) + A_u'(q^{k-1},u^{k-1})(u-u^{k-1}) + A(q^{k-1},u^{k-1}) = f \text{ in } W^*\,, \end{aligned}$$

see Section 4.2.2.

Remark 4.14. *Although q^k, u^k obviously depend on δ, i.e., $q^k = q^{k,\delta}, u^k = u^{k,\delta}$, again, we omit the superscript δ for better legibility.*

All-at-once formulations for the solution of inverse problems have also been considered in [22, 23], where they investigate a Levenberg-Marquardt approach. However, we consider a least squares and a Generalized Gauss-Newton formulation of the iteratively regularized Gauss-Newton method, which facilitates the convergence analysis compared to the regularizing Levenberg-Marquardt setting, where optimal convergence rates with a posteriori parameter choice rule have been shown only relatively recently in [43, 47, 56]. Besides we will use the discrepancy principle choice rule as in Chapter 3, whereas Burger and Mühlhuber [22, 23] use the so-called *generalized discrepancy principle* introduced by Goncharsky et al. in [37]. In addition, our focus lies on adaptive discretization using a posteriori error estimators. SQP type formulations are also considered e.g., in [3, 16], but there, the authors put more emphasis on computational aspects and applications than we do here.

For both cases $r = 1$, $r = 2$ in (4.78) we will investigate convergence and convergence rates in the continuous and adaptively discretized setting with discrepancy type choice of β_k and some overall stopping index k_* similar to (3.36), (4.22) and especially (4.25) (with a different choice of I_3). Again, the discretization errors with respect to certain quantities of interest will serve as refinement criteria during the Gauss-Newton iteration, where at the same time, we control the size of the regularization parameter. In Section 4.3, we will provide numerical results and compare the method from this chapter to the one from Chapter 3.

Throughout this section, we assume

Assumption 4.12. *There exists a solution $(q^\dagger, u^\dagger) \in \mathcal{B}_\rho(q_0, u_0) \subset D \times V \subset Q \times V$ to (4.76), where $\mathcal{B}_\rho(q_0, u_0)$ is a ball with some initial guess (q_0, u_0) as center and radius $\rho > 0$, in which local convergence of the Newton type iterations under consideration will be shown. This means*

$$\|q^\dagger - q_0\|_Q^2 + \|u^\dagger - u_0\|_V^2 \leq \rho^2 .$$

For simplicity we assume that the domains of A and C are the whole space, i.e., $\mathcal{D}(A) = Q \times V$ and $\mathcal{D}(C) = V$. Otherwise we would have to restrict our consideration to $(q, u) \in \mathcal{D}(A) \cap (Q \times \mathcal{D}(C))$ in the following.

Remark 4.15. *With the same reasoning as in Remark 4.1 in Section 4.1, we do not minimize over $\mathcal{D}$, but over the whole of Q, since starting with $(q^0, u^0) \in \mathcal{B}_\rho(q_0, u_0)$ (see Assumption 4.12), we will show that all iterates stay in a neighborhood of (q_0, u_0) (cf. Theorem 4.8 and 4.11).*

4.2.1. A least squares formulation

Direct application of the IRGNM to the all-at-once system (4.76) yields the iteration

$$\begin{pmatrix} q^k \\ u^k \end{pmatrix} = \begin{pmatrix} q^{k-1} \\ u^{k-1} \end{pmatrix} - \left(\mathbf{F}'(q^{k-1}, u^{k-1})^* \mathbf{F}'(q^{k-1}, u^{k-1}) + \begin{pmatrix} \alpha_k \,\mathrm{id} & 0 \\ 0 & \mu_k \,\mathrm{id} \end{pmatrix} \right)^{-1} \\ \cdot \left(\mathbf{F}'(q^{k-1}, u^{k-1})^* (\mathbf{F}(q^{k-1}, u^{k-1}) - \mathbf{g}^\delta) + \begin{pmatrix} \alpha_k (q^{k-1} - q_0) \\ \mu_k (u^{k-1} - u_0) \end{pmatrix} \right)$$

with regularization parameters α_k, μ_k for the q and u part of the iterates, respectively.

Although we assumed the unique and stable solvabilty of the state, adjoint, and tangent equations in Section 2.2 already, we state this assumption once more here, in order to point out its importance for this chapter.

Assumption 4.13. *The linear mapping $A'_u(q, u) \colon V \to W^*$ is continuously invertible for all $(q, u) \in Q \times V$.*

We will first of all show that Assumption 4.13 allows us to set the regularization parameter μ_k for the u part to zero. For this purpose, we introduce the abbreviations

$$K\colon V \to W^*\,, \quad K \coloneqq A'_u(q,u) \qquad \text{and} \qquad L\colon Q \to W^*\,, \quad L \coloneqq A'_q(q,u)$$

with Hilbert space adjoints $K^*\colon W^* \to V$ and $L^*\colon W^* \to Q$, i.e.,

$$\begin{aligned}(Lq, w^*)_{W^*} &= (q, L^*w^*)_Q && \forall q \in Q, w^* \in W^*\,,\\ (Kv, w^*)_{W^*} &= (v, K^*w^*)_V && \forall v \in V, w^* \in W^*\,,\end{aligned}$$

where $(\cdot,\cdot)_{W^*}$ and $(\cdot,\cdot)_V$ denote the inner products in W^* and V.

We denote the derivate of $\mathbf{F}$ at a pair (q,u) by $\mathbf{T}$, i.e.,

$$\mathbf{T}\colon Q \times V \to G \times W^*\,, \quad \mathbf{T} = \mathbf{F}'(q,u) = \begin{pmatrix} 0 & C'(u) \\ A'_q(q,u) & A'_u(q,u)\end{pmatrix} = \begin{pmatrix} 0 & C'(u) \\ L & K\end{pmatrix}$$

and define the norm

$$\left\| \begin{pmatrix} q \\ u \end{pmatrix} \right\|^2_{Q\times V} \coloneqq \|q\|_Q^2 + \|u\|_V^2 \quad \text{with operator norm} \quad \|O\|_{Q\times V} \coloneqq \sup_{x\in Q\times V, x\neq 0} \frac{\|Ox\|_{Q\times V}}{\|x\|_{Q\times V}}\,. \tag{4.79}$$

for some $x \in Q \times V$ and some operator $O\colon Q\times V \to Q \times V$.

Further, we define

$$\mathbf{Y}_{\alpha,\mu}\colon Q\times V \to Q\times V\,, \quad \mathbf{Y}_{\alpha,\mu} \coloneqq \mathbf{T}^*\mathbf{T} + \begin{pmatrix} \alpha\,\mathrm{id} & 0 \\ 0 & \mu\,\mathrm{id}\end{pmatrix}$$

for $\alpha > 0$, $\mu \geq 0$.

The following Lemmas yield useful results for the application of the IRGNM (4.2) to the system (4.76), which will be done in the next subsections.

Lemma 4.7. *Under Assumption 4.13 there holds:*

(i) For any $\alpha > 0$, $\mu \geq 0$ the inverse $\mathbf{Y}^{-1}_{\alpha,\mu}$ of $\mathbf{Y}_{\alpha,\mu}$ exists,

(ii)

$$\left\|\mathbf{Y}^{-1}_{\alpha,\mu}\mathbf{T}^*\mathbf{T}\right\|_{Q\times V} \leq 1 + \max\{\alpha,\mu\} \left\|\mathbf{Y}^{-1}_{\alpha,\mu}\right\|_{Q\times V}\,,$$

(iii)

$$\left\|\mathbf{Y}^{-1}_{\alpha,\mu}\right\|_{Q\times V} \leq c_T\left(\frac{1}{\alpha} + 1\right) \tag{4.80}$$

for all $\alpha \in (0,1]$, $\mu \geq 0$ and some $c_T > 0$ independent of α,μ. If the operators K, K^{-1}, and L are bounded uniformly in (q,u), the bound c_T in (4.80) is independent of q and u.

Proof. (i): We introduce the abbreviations

$$P = L^*L + \alpha\,\mathrm{id} \quad \text{and} \quad M = C'(u)^*C'(u) + K^*K + \mu\,\mathrm{id}\,,$$

where $C'(u)^*\colon G \to V$ is the Hilbert space adjoint of $C'(u)\colon V \to G$. We have

$$\mathbf{Y}_{\alpha,\mu} = \begin{pmatrix} P & L^*K \\ K^*L & M \end{pmatrix}.$$

Next, we will show that M is invertible. M is linear and bounded, since

$$\begin{aligned} \|Mv\|_V &\le \|C'(u)^*C'(u)v\|_V + \|K^*Kv\|_V + \mu\|v\|_V \\ &\le \left(\|C'(u)\|^2_{V\to G} + \|K\|^2_{V\to W^*} + \mu\right)\|v\|_V \le c\|v\|_V \quad \forall v \in V \end{aligned}$$

for some constant $c > 0$. M is also positive definite: There holds

$$(Mv, v)_V = \|C'(u)v\|^2_G + \|Kv\|^2_{W^*} + \mu\|v\|^2_V \ge \|Kv\|^2_{W^*} \ge \frac{1}{\|K^{-1}\|^2_{W^*\to V}}\|v\|^2_V \quad \forall v \in V\,, \tag{4.81}$$

since K is invertible due to Assumption 4.13.

Consequently M^{-1} exists (e.g., according to Lax-Milgram, cf., [108, Lemma 2.2.]) and we can define some kind of Schur complement

$$N \coloneqq P - L^*KM^{-1}K^*L = L^*L + \alpha\,\mathrm{id} - L^*KM^{-1}K^*L\,.$$

We will now show that also N is invertible. Using the fact that

$$\begin{aligned} \|M^{-1/2}K^*\|^2_{W^*\to V} &= \|KM^{-1/2}\|^2_{V\to W^*} \\ &= \sup_{v\in V, v\neq 0} \frac{\|KM^{-1/2}v\|^2_{W^*}}{\|v\|^2_V} \\ &= \sup_{v\in V, v\neq 0} \frac{\|Kv\|^2_{W^*}}{\|M^{1/2}v\|^2_V} \\ &= \sup_{v\in V, v\neq 0} \frac{\|Kv\|^2_{W^*}}{\|C'(u)(v)\|^2_G + \|Kv\|^2_{W^*} + \mu\|v\|^2_V} \le 1 \end{aligned}$$

for any $\vartheta \in Q$, we get

$$\begin{aligned} (N\vartheta, \vartheta)_Q &= (L^*L\vartheta + \alpha\vartheta - L^*KM^{-1}K^*L\vartheta, \vartheta)_Q \\ &\ge \|L\vartheta\|^2_{W^*} + \alpha\|\vartheta\|^2_Q - \|M^{-1/2}K^*\|^2_{W^*\to V}\|L\vartheta\|^2_{W^*} \\ &\ge \alpha\|\vartheta\|^2_Q\,, \end{aligned} \tag{4.82}$$

which together with

$$\begin{aligned} \|N\vartheta\|_Q &\le \|L^*L\vartheta\|_Q + \alpha\|\vartheta\|_Q + \|L^*KM^{-1}K^*L\vartheta\|_Q \\ &\le \left(\|L\|^2_{Q\to W^*} + \alpha + \|L\|^2_{Q\to W^*}\|K\|^2_{V\to W^*}\|M^{-1}\|_{V\to V}\right)\|\vartheta\|_Q \\ &\le c\|\vartheta\|_Q \quad \forall \vartheta \in Q \end{aligned}$$

implies the existence of N^{-1}.

For

$$O_{\alpha\mu} := \begin{pmatrix} N^{-1} & -N^{-1}L^*KM^{-1} \\ -M^{-1}K^*LN^{-1} & M^{-1} + M^{-1}K^*LN^{-1}L^*KM^{-1} \end{pmatrix} \tag{4.83}$$

there holds

$$O_{\alpha\mu} \begin{pmatrix} P & L^*K \\ K^*L & M \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

with

$$\begin{aligned}
A &:= N^{-1}\left(P - L^*KM^{-1}K^*L\right) = \mathrm{id} \\
B &:= N^{-1}L^*K - N^{-1}L^*KM^{-1}M = 0 \\
C &:= -M^{-1}K^*LN^{-1}P + \left(M^{-1} + M^{-1}K^*LN^{-1}L^*KM^{-1}\right)K^*L \\
&= -M^{-1}K^*L\left[N^{-1}\left(P - L^*KM^{-1}K^*L\right) - \mathrm{id}\right] = 0 \\
D &:= -M^{-1}K^*LN^{-1}L^*K + \left(M^{-1} + M^{-1}K^*LN^{-1}L^*KM^{-1}\right)M = \mathrm{id}\ .
\end{aligned}$$

Therefore, we have

$$O_{\alpha\mu} = \mathbf{Y}_{\alpha,\mu}^{-1}\,. \tag{4.84}$$

(ii): We immediately get

$$\begin{aligned}
\left\|\mathbf{Y}_{\alpha,\mu}^{-1}\mathbf{T}^*\mathbf{T}\right\|_{Q\times V} &= \left\|\mathbf{Y}_{\alpha,\mu}^{-1}\left(\mathbf{Y}_{\alpha,\mu} - \begin{pmatrix} \alpha\,\mathrm{id} & 0 \\ 0 & \mu\,\mathrm{id} \end{pmatrix}\right)\right\|_{Q\times V} \\
&= \left\|\mathrm{id} - \mathbf{Y}_{\alpha,\mu}^{-1}\begin{pmatrix} \alpha\,\mathrm{id} & 0 \\ 0 & \mu\,\mathrm{id} \end{pmatrix}\right\|_{Q\times V} \\
&\leq 1 + \max\{\alpha,\mu\}\left\|\mathbf{Y}_{\alpha,\mu}^{-1}\right\|_{Q\times V}\,.
\end{aligned}$$

(iii): From (4.81) and (4.82) we can conclude

$$\|N^{-1}\|_{Q\to Q} \leq \frac{1}{\alpha} \quad \text{and} \quad \|M^{-1}\|_{V\to V} \leq \|K^{-1}\|^2_{W^*\to V}\,. \tag{4.85}$$

For $O_{\alpha\mu}$ (cf. (4.83)) this yields

$$
\begin{aligned}
&\|O_{\alpha\mu}\|^2_{Q\times V} \\
&\leq \sup_{(q,u)\in Q\times V,(q,u)\neq 0} \frac{\left\|N^{-1}q - N^{-1}L^*KM^{-1}u\right\|^2_Q}{\|q\|^2_Q + \|u\|^2_V} \\
&\quad + \sup_{(q,u)\in Q\times V,(q,u)\neq 0} \frac{\left\|-M^{-1}K^*LN^{-1}q + \left(M^{-1} + M^{-1}K^*LN^{-1}L^*KM^{-1}\right)u\right\|^2_V}{\|q\|^2_Q + \|u\|^2_V} \\
&\leq \sup_{(q,u)\in Q\times V,(q,u)\neq 0} \frac{2\left\|N^{-1}q\right\|^2_Q}{\|q\|^2_Q + \|u\|^2_V} + \sup_{(q,u)\in Q\times V,(q,u)\neq 0} \frac{2\left\|N^{-1}L^*KM^{-1}u\right\|^2_Q}{\|q\|^2_Q + \|u\|^2_V} \\
&\quad + \sup_{(q,u)\in Q\times V,(q,u)\neq 0} \frac{2\left\|-M^{-1}K^*LN^{-1}q\right\|^2_V}{\|q\|^2_Q + \|u\|^2_V} \\
&\quad + \sup_{(q,u)\in Q\times V,(q,u)\neq 0} \frac{2\left\|\left(M^{-1} + M^{-1}K^*LN^{-1}L^*KM^{-1}\right)u\right\|^2_V}{\|q\|^2_Q + \|u\|^2_V} \\
&\leq \sup_{q\in Q,q\neq 0} \frac{2\left\|N^{-1}q\right\|^2_Q}{\|q\|^2_Q} + \sup_{u\in V,u\neq 0} \frac{2\left\|N^{-1}L^*KM^{-1}u\right\|^2_Q}{\|u\|^2_V} \\
&\quad + \sup_{q\in Q,q\neq 0} \frac{2\left\|-M^{-1}K^*LN^{-1}q\right\|^2_V}{\|q\|^2_Q} + \sup_{u\in V,u0} \frac{2\left\|\left(M^{-1} + M^{-1}K^*LN^{-1}L^*KM^{-1}\right)u\right\|^2_V}{\|u\|^2_V} \\
&\leq 2\Big(\|N^{-1}\|^2_{Q\to Q} + \|N^{-1}L^*KM^{-1}\|^2_{V\to Q} \\
&\quad + \|M^{-1}K^*LN^{-1}\|^2_{Q\to V} + \|M^{-1} + M^{-1}K^*LN^{-1}L^*KM^{-1}\|^2_{V\to V}\Big) \\
&\leq 2\Big(\|N^{-1}\|^2_{Q\to Q} + 2\|N^{-1}\|^2_{Q\to Q}\|L\|^2_{Q\to W^*}\|K\|^2_{V\to W^*}\|M^{-1}\|^2_{V\to V} \\
&\quad + 2\|M^{-1}\|^2_{V\to V} + 2\|N^{-1}\|^2_{Q\to Q}\|L\|^4_{Q\to W^*}\|K\|^4_{V\to W^*}\|M^{-1}\|^4_{V\to V}\Big) \\
&= 2\Big(\|N^{-1}\|^2_{Q\to Q}\Big[1 + 2\|L\|^2_{Q\to W^*}\|K\|^2_{V\to W^*}\|M^{-1}\|^2_{V\to V} \\
&\quad + 2\|L\|^4_{Q\to W^*}\|K\|^4_{V\to W^*}\|M^{-1}\|^4_{V\to V}\Big] + 2\|M^{-1}\|^2_{V\to V}\Big) \\
&\leq 2\|N^{-1}\|^2_{Q\to Q}\Big[1 + 2\|L\|^2_{Q\to W^*}\|K\|^2_{V\to W^*}\|M^{-1}\|^2_{V\to V} \\
&\quad + 2\|L\|^4_{Q\to W^*}\|K\|^4_{V\to W^*}\|K^{-1}\|^8_{W^*\to V}\Big] + 4\|K^{-1}\|^4_{W^*\to V} \\
&\leq \frac{2}{\alpha^2}\Big[1 + 2\|L\|^2_{Q\to W^*}\|K\|^2_{V\to W^*}\|M^{-1}\|^2_{V\to V} \\
&\quad + 2\|L\|^4_{Q\to W^*}\|K\|^4_{V\to W^*}\|K^{-1}\|^8_{W^*\to V}\Big] + 4\|K^{-1}\|^4_{W^*\to V}
\end{aligned}
$$

where we have used that $(a+b)^2 \leq 2a^2 + 2b^2$ for all $a, b \in \mathbb{R}$ twice.

□

Motivated by Lemma 4.7, we set the regularization parameter for the component u to zero (which is justified by (i) in Lemma 4.7), we set $\alpha_k = \frac{1}{\beta_k}$, and define a regularized iteration by

$$\begin{pmatrix} q^k \\ u^k \end{pmatrix} = \begin{pmatrix} q^{k-1} \\ u^{k-1} \end{pmatrix} - \left(\mathbf{F}'(q^{k-1}, u^{k-1})^* \mathbf{F}'(q^{k-1}, u^{k-1}) + \frac{1}{\beta_k} \begin{pmatrix} \mathrm{id} & 0 \\ 0 & 0 \end{pmatrix} \right)^{-1} \cdot \left(\mathbf{F}'(q^{k-1}, u^{k-1})^* (\mathbf{F}(q^{k-1}, u^{k-1}) - \mathbf{g}^\delta) + \frac{1}{\beta_k} \begin{pmatrix} q^{k-1} - q_0 \\ 0 \end{pmatrix} \right)$$

or equivalently $\begin{pmatrix} q^k \\ u^k \end{pmatrix}$ as solution to the unconstrained minimization problem

$$\min_{(q,u)\in Q\times V} \|L_{k-1}(q - q^{k-1}) + K_{k-1}(u - u^{k-1}) + A(q^{k-1}, u^{k-1}) - f\|^2_{W^*} + \|C(u^{k-1}) + C'(u^{k-1})(u - u^{k-1}) - g^\delta\|^2_G + \frac{1}{\beta_k}\|q - q_0\|^2_Q \tag{4.86}$$

with the abbreviations

$$L_{k-1} = A'_q(q^{k-1}, u^{k-1}) \qquad \text{and} \qquad K_{k-1} = A'_u(q^{k-1}, u^{k-1})\,. \tag{4.87}$$

The optimality conditions of first order for (4.86) read

$$\begin{aligned} 0 &= 2(L_{k-1}(q - q^{k-1}) + K_{k-1}(u - u^{k-1}) + A(q^{k-1}, u^{k-1}) - f, L_{k-1}(\delta q))_{W^*} + \frac{2}{\beta_k}(q - q_0, \delta q)_Q\,, \\ 0 &= 2(L_{k-1}(q - q^{k-1}) + K_{k-1}(u - u^{k-1}) + A(q^{k-1}, u^{k-1}) - f, K_{k-1}(\delta u))_{W^*} \\ &\quad + 2(C(u^{k-1}) + C'(u^{k-1})(u - u^{k-1}) - g^\delta, C'(u^{k-1})(\delta u))_G\,. \end{aligned}$$

Like in Section 4.1, in each step k we will replace the infinite dimensional spaces Q, V, W in (4.86) by finite dimensional ones $Q_h = Q_{h_k}, V_h = V_{h_k}, W_h = W_{h_k}$ and get

$$\begin{pmatrix} q_h^k \\ u_h^k \end{pmatrix} = \arg \min_{q\in Q_h, u\in V_h} \|L_{k-1}(q - q_{\mathrm{old}}) + K_{k-1}(u - u_{\mathrm{old}}) + A(q_{\mathrm{old}}, u_{\mathrm{old}}) - f\|^2_{W_h^*} + \|C(u_{\mathrm{old}}) + C'(u_{\mathrm{old}})(u - u_{\mathrm{old}}) - g^\delta\|^2_G + \frac{1}{\beta_k}\|q - q_0\|^2_Q\,.$$

where $(q_{\mathrm{old}}, u_{\mathrm{old}}) = (q^{k-1}, u^{k-1}) = (q_{h_{k-1}}^{k-1}, u_{h_{k-1}}^{k-1})$ is the previous iterate, which itself is discretized by the use of spaces $Q_{h_{k-1}}, V_{h_{k-1}}, W_{h_{k-1}}$. Again (cf. Remark 4.2) the discretization h_k may be different in each Newton step (typically it will get finer for increasing k), but as in Section 4.1 we suppress dependence of h on k in our notation in most of what follows.

Analogously to (4.6) and (4.8), we define quantities of interest via the functionals

$$\begin{aligned} I_1 &: \; Q \times V \times Q \times V \times \mathbb{R} \to \mathbb{R} \\ I_2 &: \; Q \times V \times Q \times V \to \mathbb{R} \\ I_3 &: \; Q \times V \to \mathbb{R} \\ I_4 &: \; Q \times V \to \mathbb{R}\,, \end{aligned}$$

where we insert the previous and current iterates $(q_{\text{old}}, u_{\text{old}})$, (q, u), respectively:

$$
\begin{aligned}
I_1(q_{\text{old}}, u_{\text{old}}, q, u, \beta) &= \left\| \mathbf{F}'(q_{\text{old}}, u_{\text{old}}) \begin{pmatrix} q - q_{\text{old}} \\ u - u_{\text{old}} \end{pmatrix} + \mathbf{F}(q_{\text{old}}, u_{\text{old}}) - \mathbf{g}^\delta \right\|_{G \times W^*}^2 + \frac{1}{\beta} \|q - q_0\|_Q^2 \\
&= \left\| A'_q(q_{\text{old}}, u_{\text{old}})(q - q_{\text{old}}) + A'_u(q_{\text{old}}, u_{\text{old}})(u - u_{\text{old}}) + A(q_{\text{old}}, u_{\text{old}}) - f \right\|_{W^*}^2 \\
&\quad + \left\| C'(u)(u - u_{\text{old}}) + C(u_{\text{old}}) - g^\delta \right\|_G^2 + \frac{1}{\beta} \|q - q_0\|_Q^2 \\
I_2(q_{\text{old}}, u_{\text{old}}, q, u) &= \left\| \mathbf{F}'(q_{\text{old}}, u_{\text{old}}) \begin{pmatrix} q - q_{\text{old}} \\ u - u_{\text{old}} \end{pmatrix} + \mathbf{F}(q_{\text{old}}, u_{\text{old}}) - \mathbf{g}^\delta \right\|_{G \times W^*}^2 \\
&= \left\| A'_q(q_{\text{old}}, u_{\text{old}})(q - q_{\text{old}}) + A'_u(q_{\text{old}}, u_{\text{old}})(u - u_{\text{old}}) + A(q_{\text{old}}, u_{\text{old}}) - f \right\|_{W^*}^2 \\
&\quad + \left\| C'(u)(u - u_{\text{old}}) C(u_{\text{old}}) - g^\delta \right\|_G^2 \\
I_3(q_{\text{old}}, u_{\text{old}}) &= \left\| \mathbf{F}(q_{\text{old}}, u_{\text{old}}) - \mathbf{g}^\delta \right\|_{G \times W^*}^2 \\
&= \|A(q_{\text{old}}, u_{\text{old}}) - f\|_{W^*}^2 + \left\| C(u_{\text{old}}) - g^\delta \right\|_G^2 \\
I_4(q, u) &= \left\| \mathbf{F}(q, u) - \mathbf{g}^\delta \right\|_{G \times W^*}^2 \\
&= \|A(q, u) - f\|_{W^*}^2 + \left\| C(u) - g^\delta \right\|_G^2 .
\end{aligned}
\tag{4.88}
$$

This leads to the quantities of interest

$$
\begin{aligned}
I_1^k &= I_1(q_{\text{old}}^k, u_{\text{old}}^k, q^k, u^k, \beta_k) \\
I_2^k &= I_2(q_{\text{old}}^k, u_{\text{old}}^k, q^k, u^k) \\
I_3^k &= I_3(q_{\text{old}}^k, u_{\text{old}}^k) \\
I_4^k &= I_4(q^k, u^k) .
\end{aligned}
\tag{4.89}
$$

Their discrete analogs (cf. (4.13), (4.14)) are correspondingly defined by

$$
\begin{aligned}
I_{1,h} &\colon\ Q \times V \times Q \times V \times \mathbb{R} \to \mathbb{R} \\
I_{2,h} &\colon\ Q \times V \times Q \times V \to \mathbb{R} \\
I_{3,h} &\colon\ Q \times V \to \mathbb{R} \\
I_{4,h} &\colon\ Q \times V \to \mathbb{R} ,
\end{aligned}
$$

$$
\begin{aligned}
I_{1,h}(q_{\text{old}}, u_{\text{old}}, q, u, \beta) &= \left\| \mathbf{F}'(q_{\text{old}}, u_{\text{old}}) \begin{pmatrix} q - q_{\text{old}} \\ u - u_{\text{old}} \end{pmatrix} + \mathbf{F}(q_{\text{old}}, u_{\text{old}}) - \mathbf{g}^\delta \right\|^2_{G \times W_h^*} + \frac{1}{\beta} \|q - q_0\|_Q^2 \\
&= \|A'_q(q_{\text{old}}, u_{\text{old}})(q - q_{\text{old}}) + A'_u(q_{\text{old}}, u_{\text{old}})(u - u_{\text{old}}) \\
&\quad + A(q_{\text{old}}, u_{\text{old}}) - f\|^2_{W_h^*} \\
&\quad + \|C'(u)(u - u_{\text{old}}) + C(u_{\text{old}}) - g^\delta\|_G^2 + \frac{1}{\beta} \|q - q_0\|_Q^2 \\
I_{2,h}(q_{\text{old}}, u_{\text{old}}, q, u) &= \left\| \mathbf{F}'(q_{\text{old}}, u_{\text{old}}) \begin{pmatrix} q - q_{\text{old}} \\ u - u_{\text{old}} \end{pmatrix} + \mathbf{F}(q_{\text{old}}, u_{\text{old}}) - \mathbf{g}^\delta \right\|^2_{G \times W_h^*} \\
&= \|A'_q(q_{\text{old}}, u_{\text{old}})(q - q_{\text{old}}) + A'_u(q_{\text{old}}, u_{\text{old}})(u - u_{\text{old}}) \\
&\quad + A(q_{\text{old}}, u_{\text{old}}) - f\|^2_{W_h^*} \\
&\quad + \|C'(u)(u - u_{\text{old}})C(u_{\text{old}}) - g^\delta\|_G^2 \\
I_{3,h}(q_{\text{old}}, u_{\text{old}}) &= \left\| \mathbf{F}(q_{\text{old}}, u_{\text{old}}) - \mathbf{g}^\delta \right\|^2_{G \times W_h^*} \\
&= \|A(q_{\text{old}}, u_{\text{old}}) - f\|^2_{W_h^*} + \|C(u_{\text{old}}) - g^\delta\|_G^2 \\
I_{4,h}(q, u) &= \left\| \mathbf{F}(q, u) - \mathbf{g}^\delta \right\|^2_{G \times W_h^*} \\
&= \|A(q, u) - f\|^2_{W_h^*} + \|C(u) - g^\delta\|_G^2 \,,
\end{aligned}
\tag{4.90}
$$

and

$$
\begin{aligned}
I_{1,h}^k &= I_{1,h}(q_{\text{old}}^k, u_{\text{old}}^k, q_h^k, u_h^k, \beta_k) \\
I_{2,h}^k &= I_{2,h}(q_{\text{old}}^k, u_{\text{old}}^k, q_h^k, u_h^k,) \\
I_{3,h}^k &= I_{3,h}(q_{\text{old}}^k, u_{\text{old}}^k) \\
I_{4,h}^k &= I_{4,h}(q_h^k, u_h^k) \,.
\end{aligned}
\tag{4.91}
$$

Just like in Section 4.1 we assume bounds on the discretization errors with respect to these four quantities of interest (see (4.17)), which will serve as refinement criteria and which, at least partly, will be estimated by goal oriented error estimators. At the end of each iteration step we set

$$
q_{\text{old}}^{k+1} = q_h^k \quad \text{and} \quad u_{\text{old}}^{k+1} = u_h^k \,. \tag{4.92}
$$

Remark 4.16. *Note that here, neither q_{old} nor u_{old} are subject to new adaptive discretization in the current step, but they are taken as fixed quantities from the previous step. This is different from Section 4.1, where u_{old} also depends on the current discretization.*

For (4.89) *and* (4.91) *we assume that the norms in G and Q are evaluated exactly cf. Assumption 2.3.*

As in Section 4.1, in our convergence proofs we will compare the quantities of interest $I_{i,h}^k$ with those I_i^k that would be obtained by exact computation on the infinite dimensional spaces, starting from the same $(q_{\text{old}}, u_{\text{old}}) = (q_{\text{old}\,h_{k-1}}, u_{\text{old}\,h_{k-1}})$ as the one underlying $I_{i,h}^k$ (cf. (4.17)*). Thus analogously to Section 4.1 (see Remark 4.2 and Figure 4.1), in our analysis besides the* *actually computed sequence* $(q_h^k, u_h^k) = (q_{h_k}^k, u_{h_k}^k)$ *there appears an auxiliary sequence* (q^k, u^k), *see Figure 4.3.*

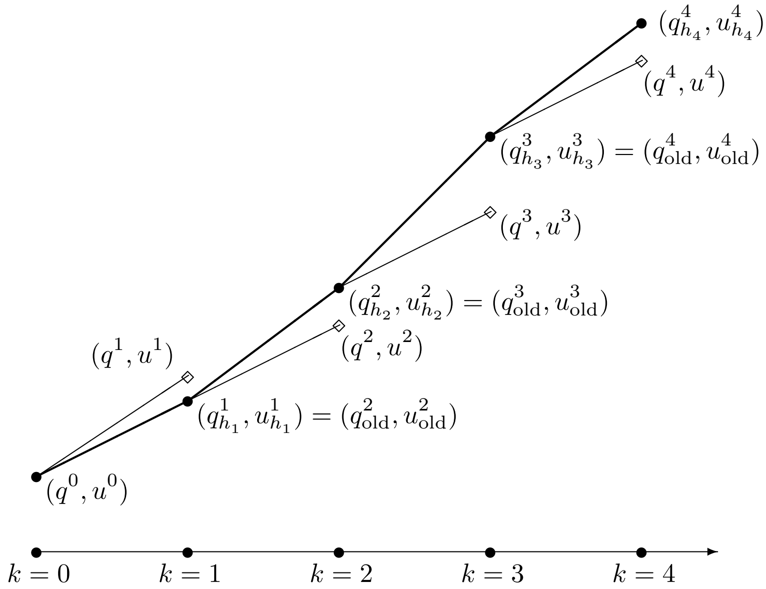

Figure 4.3.: Sequence of discretized iterates and auxiliary sequence of continuous iterates for the all-at-once formulation of IRGNM

We formulate all-at-once versions of Assumption 3.2 and Assumption 4.4, i.e., we assume $\mathbf{F}$ to be weakly sequentially closed and to satisfy a tangential cone condition:

Assumption 4.14. *Let the all-at-once forward operator* $\mathbf{F}$ *be continuous and satisfy*

$$(q_n \rightharpoonup q \wedge u_n \rightharpoonup u \wedge C(u_n) \to g \wedge A(q_n, u_n) \to f) \Rightarrow (q \in \mathcal{D} \wedge C(u) = g \wedge A(q, u) = f)$$

for all sequences $((q_n, u_n))_{n\in\mathbb{N}} \subseteq \mathcal{D} \times V$.

Assumption 4.15. *Let*

$$\begin{aligned}\|C(u) - C(\bar{u}) - C'(u)(u - \bar{u})\|_G + \|A(q, u) - A(\bar{q}, \bar{u}) - A'_q(q, u)(q - \bar{q}) - A'_u(q, u)(u - \bar{u})\|_{W^*} \\ \leq c_{tc} \left(\|C(u) - C(\bar{u})\|_G + \|A(q, u) - A(\bar{q}, \bar{u})\|_{W^*}\right)\end{aligned}$$

hold for all $(q, u), (\bar{q}, \bar{u}) \in \mathcal{B}_\rho(q_0, u_0)$ *and some* $0 < c_{tc} < 1$.

Like in Theorem 4.2 in Section 4.1 we obtain convergence and convergence rates results, which are formulated in the following theorem.

Theorem 4.8. *Let Assumption 4.14 and 4.15 with* c_{tc} *sufficiently small be satisfied. Let the starting value be chosen such that* $(q_{\text{old}}^0, u_{\text{old}}^0) \in \mathcal{B}_\rho(q_0, u_0)$ *and let* $(q^\dagger, u^\dagger) \in \mathcal{B}_\rho(q_0, u_0)$ *be a solution to* (4.76). *For the quantities of interest* (4.89) *and* (4.91), *let further the estimate*

(4.17) *as well as Assumption 4.5 hold. Then with* β_k *according to* (4.18) *and* k_* *selected according to* (4.25), $(q_{\text{old}}^{k_*}, u_{\text{old}}^{k_*}) = (q_{\text{old}}^{k_*(\delta)}, u_{\text{old}}^{k_*(\delta)}) = (q_{\beta_{k_*(\delta)-1},h_{k_*(\delta)-1}}^{k_*(\delta)-1,\delta}, q_{\beta_{k_*(\delta)-1},h_{k_*(\delta)-1}}^{k_*(\delta)-1,\delta})$ *defined by* (4.86) *converges (weakly) subsequentially to a solution* $(q^\dagger, u^\dagger)$ *of* (4.76) *as* $\delta \to 0$ *in the sense that it has a weakly convergent subsequence and each weakly convergent subsequence converges strongly to* $(q^\dagger, u^\dagger)$*. If the solution* $(q^\dagger, u^\dagger)$ *to* (4.76) *is unique, then* $(q_{\text{old}}^{k_*(\delta)}, u_{\text{old}}^{k_*(\delta)})$ *converges strongly to* $(q^\dagger, u^\dagger)$ *as* $\delta \to 0$.

Proof. The claimed convergence (as well as well-definedness of the regularization parameter) follows directly along the lines of the proof of Theorem 4.2 replacing F there by $\mathbf{F}$ according to (4.77). □

For proving rates, as usual (cf., e.g., [5, 27, 51, 69], Assumption 3.5 and 4.10) source conditions are assumed.

Assumption 4.16. *Let*

$$(q^\dagger - q_0, u^\dagger - u_0) \in \mathcal{R}\left(\kappa\left(\mathbf{F}'(q^\dagger, u^\dagger)^*\mathbf{F}'(q^\dagger, u^\dagger)\right)\right)$$

hold with some $\kappa\colon \mathbb{R}^+ \to \mathbb{R}^+$ *such that* $\kappa^2\colon \lambda \mapsto \kappa(\lambda)^2$ *is strictly monotonically increasing on* $(0, \|\mathbf{F}'(q^\dagger, u^\dagger)\|_{Q\times V}^2]$, φ *defined by* $\varphi^{-1}(\lambda) = \kappa^2(\lambda)$ *is convex and* ψ *defined by* $\psi(\lambda) = \kappa(\lambda)\sqrt{\lambda}$ *is strictly monotonically increasing on* $(0, \|\mathbf{F}'(q^\dagger, u^\dagger)\|_{Q\times V}^2]$.

Like in Theorem 3.2, Corollary 4.3, and Theorem 4.5 we get the following convergence rates:

Theorem 4.9. *Let the conditions of Theorem 4.8 and additionally the source condition Assumption 4.16 be fullfiled. Then there exist* $\bar{\delta} > 0$ *and a constant* $C > 0$ *independent of* δ *such that for all* $\delta \in (0, \bar{\delta}]$ *the convergence rates*

$$\|q_{\text{old}}^{k_*} - q^\dagger\|_Q^2 + \|u_{\text{old}}^{k_*} - u^\dagger\|_V^2 = \mathcal{O}\left(\frac{\delta^2}{\psi^{-1}(C\delta)}\right)$$

with $q_{\text{old}}^{k_*} = q_{\text{old}}^{k_*(\delta)} = q_{\beta_{k_*(\delta)-1},h_{k*(\delta)-1}}^{\delta,k_*(\delta)-1}$, $u_{\text{old}}^{k_*} = u_{\text{old}}^{k_*(\delta)} = u_{\beta_{k_*(\delta)-1},h_{k*(\delta)-1}}^{\delta,k_*(\delta)-1}$ *are obtained.*

Proof. The proposition follows directly from Theorem 3.2. □

Remark 4.17. *Comparing the source conditions from Assumption 3.5 for the reduced formulation with Assumption 4.16 for the all-at-once formulation, we consider the case* $\kappa(\lambda) = \sqrt{\lambda}$*. Namely, in that case Assumption 3.5 reads: There exists* $\overline{g} \in G$ *such that*

$$q^\dagger - q_0 = F'(q^\dagger)^*\overline{g} = S'(q^\dagger)^*C'(S(q^\dagger))^*\overline{g}\,. \tag{4.93}$$

On the other hand, Assumption 4.16 with the same κ *reads: There exists* $\tilde{\mathbf{g}} = \begin{pmatrix}\tilde{g}\\ \tilde{f}\end{pmatrix} \in G \times W^*$ *such that*

$$(q^\dagger - q_0, u^\dagger - u_0) = \mathbf{F}'(q^\dagger, u^\dagger)^*\tilde{\mathbf{g}} = \begin{pmatrix} 0 & A_q'(q^\dagger, u^\dagger)^* \\ C'(u^\dagger)^* & A_u'(q^\dagger, u^\dagger)^* \end{pmatrix}\tilde{\mathbf{g}}\,,$$

which is equivalent to

$$\begin{aligned} q^\dagger - q_0 &= A_q'(q^\dagger, u^\dagger)^* \tilde{f} \\ u^\dagger - u_0 &= C'(u^\dagger)^* \tilde{g} + A_u'(q^\dagger, u^\dagger)^* \tilde{f} , \end{aligned}$$

and by elimination of $\tilde{f}$ and use of the identities

$$\begin{aligned} u^\dagger &= S(q^\dagger) \\ S'(q^\dagger) &= -A_u'(q^\dagger, u^\dagger)^{-1} A_q'(q^\dagger, u^\dagger) \end{aligned}$$

we get

$$q^\dagger - q_0 = S'(q^\dagger)^* \Big(C'(S(q^\dagger))^* \tilde{g} + u_0 - u^\dagger \Big) ,$$

which, setting $\bar{g} = \tilde{g} + C'(u^\dagger)^{-*}(u_0 - u^\dagger)$ becomes (4.93)*, provided $u_0 - u^\dagger \in \mathcal{R}(C'(u^\dagger)^*)$. For a further discussion we refer to [66].*

Computation of the error estimators

Theoretically the error estimators for this subsection can be computed similarly to those from Section 4.1. The fact that we consider an unconstrained optimization problem should make things easier, but we get another difficulty in return: For estimating the discretization error with respect to I_1 and I_2 we have to estimate terms like

$$\|E(q,u)\|_{W^*} - \|E(q_h,u_h)\|_{W_h^*} \tag{4.94}$$

for some operator $E\colon Q \times V \to W^*$, which makes things slighty more complicated concerning the use of goal-oriented error estimators. Since we will treat an easier/less costly formulation concerning this matter in Section 4.2.2, here, we won't generate numerical results, but only sketch an idea of how one could compute an error estimator for an error of the form (4.94) for the typical case $V = W = H_0^1(\Omega)$.

We denote the objective functional in (4.86) by $\mathcal{T}$, i.e., $\mathcal{T}\colon Q \times V \to \mathbb{R}$,

$$\begin{aligned} \mathcal{T}(q,u) = \|L_{k-1}(q - q_{\text{old}}) + K_{k-1}(u - u_{\text{old}}) + A(q_{\text{old}}, u_{\text{old}}) - f\|_{V^*}^2 \\ + \|C(u_{\text{old}}) + C'(u_{\text{old}})(u - u_{\text{old}}) - g^\delta\|_G^2 + \frac{1}{\beta_k}\|q - q_0\|_Q^2 \end{aligned}$$

for fixed $\beta_k, q_{\text{old}}, u_{\text{old}}$. We further define $\Psi\colon V \to \mathbb{R}$,

$$\Psi(v) = \|\nabla v\|_{L^2(\Omega)}$$

as well as the auxiliary Lagrangian $M\colon Q \times V \times Q \times V \times V \times V$,

$$M(q,u,\delta q,\delta u,v,w) = \mathcal{T}_q'(q,u)(\delta q) + \mathcal{T}_u'(q,u)(\delta u) + \Psi(v) + \langle E(q,u), w\rangle_{V^*,V} - (\nabla v, \nabla w)_{L^2(\Omega)} .$$

We consider continuous and stationary points $x := (q, u) \in Q \times V$ and $x_h := (q_h, u_h) \in Q_h \times V_h$ of $\mathcal{T}$, i.e., let

$$\mathcal{T}'(x)(\varphi) = 0 \quad \forall \varphi \in Q \times V\,,$$
$$\mathcal{T}'(x_h)(\varphi) = 0 \quad \forall \varphi \in Q_h \times V_h\,.$$

Let $v \in V,\ v_h \in V_h$ solve

$$(\nabla v, \nabla \varphi)_{L^2(\Omega)} = \langle E(q,u), \varphi\rangle_{V^*,V} \quad \forall \varphi \in V\,,$$
$$(\nabla v_h, \nabla \varphi)_{L^2(\Omega)} = \langle E(q_h,u_h), \varphi\rangle_{V^*,V} \quad \forall \varphi \in V_h\,,$$

and w be defined by

$$w = \frac{1}{\|\nabla v\|} v\,, \qquad w_h = \frac{1}{\|\nabla v_h\|} v_h\,.$$

Let further $\delta x := (\delta q, \delta u) \in Q \times V,\ \delta x_h := (\delta q_h, \delta u_h) \in Q_h \times V_h$ solve

$$\mathcal{T}''(x)(\delta x, \varphi) = -\langle E'(x)(\varphi), w\rangle_{V^*,V} \quad \forall \varphi \in Q \times V$$
$$\mathcal{T}''(x_h)(\delta x_h, \varphi) = -\langle E'(x_h)(\varphi), w_h\rangle_{V^*,V} \quad \forall \varphi \in Q_h \times V_h\,.$$

Then it is easily checked that $\xi := (x, \delta x, v, w)$ and $\xi_h := (x_h, \delta x_h, v_h, w_h)$ are continuous and discrete stationary points of M. By the usual DWR reasoning we obtain

$$\begin{aligned}\|E(q,u)\|_{V^*} - \|E(q_h,u_h)\|_{V_h^*} &= \|\nabla v\|_{L^2(\Omega)} - \|\nabla v_h\|_{L^2(\Omega)}\\ &= \Psi(v) - \Psi(v_h)\\ &= \frac{1}{2} M'(\xi_h)(\xi - \tilde{\xi}_h) + R\end{aligned}$$

for arbitrary $\tilde{\xi}_h \in Q_h \times V_h \times Q_h \times V_h \times V_h \times V_h$, where R is some remainder term (see Section 2.4, 3.3 and 4.1.2). The term $\xi - \tilde{\xi}_h$ can be approximated as described in the previous sections about error estimation (Section 2.4, 3.3 and 4.1.2) by high order approximations.

4.2.2. A Generalized Gauss-Newton formulation

A drawback of the unconstrained formulation (4.86) is the fact that a rescaling of the state equation (4.75) changes the solution of the optimization problem. Moreover, depending on the given inverse problem and its application, in some cases, it does not make sense to only minimize the residual of the linearized state equation, instead of setting it to zero. Another disadvantage is the necessity of computing the W^*-norm of the (linearized) residual and especially of computing error estimators for this quantity of interest.

A formulation that is much better tractable is obtained by defining $(q^k, u^k) = (q^{k,\delta}, u^{k,\delta})$ as a solution to the PDE constrained minimization problem

$$\min_{(q,u)\in Q\times V} \mathcal{T}_{\beta_k}(q,u) := \|C(u^{k-1}) + C'(u^{k-1})(u - u^{k-1}) - g^\delta\|_G^2 + \frac{1}{\beta_k}\left(\|q - q_0\|_Q^2 + \|u - u_0\|_V^2\right) \tag{4.95}$$

$$\text{s.t.}\quad L_{k-1}(q - q^{k-1}) + K_{k-1}(u - u^{k-1}) + A(q^{k-1}, u^{k-1}) = f \qquad \text{in } W^* \tag{4.96}$$

(see also [22], [23]) with the abbreviations (4.87).

We consider the Lagrangian $\mathcal{L}\colon Q \times V \times W \to \mathbb{R}$

$$\mathcal{L}(q,u,z) := \mathcal{T}_{\beta_k}(q,u) + \langle f - A(q^{k-1},u^{k-1}) - L_{k-1}(q-q^{k-1}) - K_{k-1}(u-u^{k-1}), z\rangle_{W^*,W}\,,$$

and formulate the optimality conditions of first order for (4.95),(4.96):

$$\mathcal{L}'_z(q,u,z)(\delta z) = \langle f - A(q^{k-1},u^{k-1}) - L_{k-1}(q-q^{k-1}) - K_{k-1}(u-u^{k-1}), \delta z\rangle_{W^*,W} = 0\,, \tag{4.97}$$

$$\begin{aligned}\mathcal{L}'_u(q,u,z)(\delta u) &= 2(C(u^{k-1}) + C'(u^{k-1})(u-u^{k-1}) - g^\delta, C'(u^{k-1})(\delta u))_G \\ &\quad + \tfrac{2}{\beta_k}(u-u_0,\delta u)_V - \langle K_{k-1}\delta u, z\rangle_{W^*,W} = 0\,,\end{aligned} \tag{4.98}$$

$$\mathcal{L}'_q(q,u,z)(\delta q) = \tfrac{2}{\beta_k}(q-q_0,\delta q)_Q - \langle L_{k-1}\delta q, z\rangle_{W^*,W} = 0 \tag{4.99}$$

for all $\delta q \in Q$, $\delta u \in V$, $\delta z \in W$ (cf. Section 2.2).

We assume boundedness of the operators $A(q,u), L_{k-1}, K^*_{k-1}, K^{-1}_{k-1}, C(u)$ and $C'(u)$ in the following sense.

Assumption 4.17. *There holds*

$$\sup_{(q,u)\in\mathcal{B}_\rho(q_0,u_0)} \|A(q,u)\|_{W^*} + \|A'_q(q,u)\|_{Q\to W^*} + \|A'_u(q,u)^*\|_{W^*\to V} + \|A'_u(q,u)^{-1}\|_{W^*\to V} < \infty$$

and

$$\sup_{(q,u)\in\mathcal{B}_\rho(q_0,u_0)} \{\|C(u)\|_G + \|C'(u)\|_{V\to G}\} < \infty\,,$$

where $\|.\|_{X\to Y}$ denotes the operator norm, i.e., for $T\colon X\to Y$, $\|T\|_{X\to Y} := \sup_{\|x\|_X\neq 0} \frac{\|T(x)\|_Y}{\|x\|_X}$.

The following lemma about boundedness of the adjoint will serve as tool for uniformly bounding the penalty parameter ϱ in (4.78).

Lemma 4.10. *Let Assumption 4.17 hold and $(q^{k-1},u^{k-1}) \in \mathcal{B}_\rho(q_0,u_0)$. Provided the sequence of β_k's is bounded away from 0, for a stationary point $(q^k,u^k,z^k) \in Q\times V\times W$ of $\mathcal{L}$ (cf.* (4.97) – (4.99)*) there holds the estimate*

$$\|z^k\|_W \le c_{\mathrm{adj}}\left(\|q^{k-1}-q_0\|_Q + \|u^{k-1}-u_0\|_V + 1\right) \tag{4.100}$$

with a constant c_{adj} independent of k.

Proof. To formulate the optimality system (4.97)–(4.99) in a matrix-vector form, we introduce another dual variable $p \in W^*$ defined by

$$p = J_{W^*} z \in W^* \tag{4.101}$$

via the map J_{W^*}, which maps $z\in W$ to the Riesz representation $J_{W^*}z \in W^*$, such that

$$\mathcal{L}(q,u,z) = \mathcal{T}_{\beta_k}(q,u) + (f - A(q^{k-1},u^{k-1}) - L_{k-1}(q-q^{k-1}) - K_{k-1}(u-u^{k-1}), p)_{W^*}\,.$$

Using the abbreviations (4.87) and

$$C_{k-1} := C'(u^{k-1}), \quad r^f_{k-1} := A(q^{k-1}, u^{k-1}) - f, \quad r^g_{k-1} := C(u^{k-1}) - g^\delta \tag{4.102}$$

the optimality system (4.97)–(4.99) can be written as

$$\mathcal{L}'_z(q^k, u^k, z^k)(\delta z) = \langle -r^f_{k-1} - L_{k-1}(q^k - q^{k-1}) - K_{k-1}(u^k - u^{k-1}), \delta z\rangle_{W^*,W} = 0, \tag{4.103}$$

$$\begin{aligned}\mathcal{L}'_u(q^k, u^k, z^k)(\delta u) &= 2(r^g_{k-1} + C_{k-1}(u^k - u^{k-1}), C_{k-1}\delta u)_G + \tfrac{2}{\beta_k}(u^k - u_0, \delta u)_V \\ &\quad - (K_{k-1}\delta u, p)_{W^*} \\ &= (2C^*_{k-1}[r^g_{k-1} + C_{k-1}(u^k - u^{k-1})] + \tfrac{2}{\beta_k}(u^k - u_0) - K^*_{k-1}p^k, \delta u)_V = 0,\end{aligned} \tag{4.104, 4.105}$$

$$\mathcal{L}'_q(q^k, u^k, z^k)(\delta q) = \tfrac{2}{\beta_k}(q^k - q_0, \delta q)_Q - (L_{k-1}\delta q, p^k)_{W^*} = (\tfrac{2}{\beta_k}(q^k - q_0) - L^*_{k-1}p^k, \delta q)_Q = 0 \tag{4.106}$$

for all $\delta q \in Q$, $\delta u \in V$ and $\delta z \in W$, or equivalently as

$$\begin{aligned}q^k &= q_0 + \frac{\beta_k}{2} L^*_{k-1} p^k \\ u^k &= \left[\frac{2}{\beta_k}\,\mathrm{id} + 2C^*_{k-1}C_{k-1}\right]^{-1}\left(2C^*_{k-1}\left(C_{k-1}u^{k-1} - r^g_{k-1}\right) + \frac{2}{\beta_k}u_0 + K^*_{k-1}p^k\right) \\ u^k &= K^{-1}_{k-1}\left(L_{k-1}q^{k-1} + K_{k-1}u^{k-1} - r^f_{k-1} - L_{k-1}q^k\right).\end{aligned}$$

Eliminating q^k and u^k yields

$$\begin{aligned}&\left[\frac{2}{\beta_k}\,\mathrm{id} + 2C^*_{k-1}C_{k-1}\right]^{-1}\left(2C^*_{k-1}\left(C_{k-1}u^{k-1} - r^g_{k-1}\right) + \frac{2}{\beta_k}u_0 + K^*_{k-1}p^k\right) \\ &= K^{-1}_{k-1}\left(L_{k-1}q^{k-1} + K_{k-1}u^{k-1} - r^f_{k-1} - L_{k-1}\left(q_0 + \frac{\beta_k}{2}L^*_{k-1}p^k\right)\right),\end{aligned}$$

which we reformulate as

$$\begin{aligned}&-\frac{\beta_k}{2}K^{-1}_{k-1}L_{k-1}L^*_{k-1}p^k - \left[\frac{2}{\beta_k}\,\mathrm{id} + 2C^*_{k-1}C_{k-1}\right]^{-1} K^*_{k-1}p^k \\ &= \left[\frac{2}{\beta_k}\,\mathrm{id} + 2C^*_{k-1}C_{k-1}\right]^{-1}\left(2C^*_{k-1}\left(C_{k-1}u^{k-1} - r^g_{k-1}\right) + \frac{2}{\beta_k}u_0\right) \\ &\quad - K^{-1}_{k-1}\left(L_{k-1}(q^{k-1} - q_0) + K_{k-1}u^{k-1} - r^f_{k-1}\right)\end{aligned}$$

and finally

$$\begin{aligned}&-\left[\frac{1}{\beta_k}\,\mathrm{id} + C^*_{k-1}C_{k-1}\right]\beta_k K^{-1}_{k-1}L_{k-1}L^*_{k-1}p^k - K^*_{k-1}p^k \\ &= 2C^*_{k-1}\left(C_{k-1}u^{k-1} - r^g_{k-1}\right) + \frac{2}{\beta_k}u_0 \\ &\quad - 2\left[\frac{1}{\beta_k}\,\mathrm{id} + C^*_{k-1}C_{k-1}\right]K^{-1}_{k-1}\left(L_{k-1}(q^{k-1} - q_0) + K_{k-1}u^{k-1} - r^f_{k-1}\right).\end{aligned}$$

With

$$C_\beta := \left(\frac{1}{\beta_k}\operatorname{id} + C_{k-1}^* C_{k-1}\right)^{\frac{1}{2}}$$

this is equivalent to

$$\begin{aligned}
&-\beta_k C_\beta^2 K_{k-1}^{-1} L_{k-1} L_{k-1}^* p^k - K_{k-1}^* p^k \\
&= 2C_{k-1}^*\left(C_{k-1}u^{k-1} - r_{k-1}^g\right) + \frac{2}{\beta_k}u_0 \\
&\quad - 2C_\beta^2 K_{k-1}^{-1}\left(L_{k-1}(q^{k-1} - q_0) + K_{k-1}u^{k-1} - r_{k-1}^f\right) ,
\end{aligned}$$

which upon premultiplication with C_β^{-1} becomes

$$\begin{aligned}
&-\left(\beta_k C_\beta K_{k-1}^{-1} L_{k-1}(K_{k-1}^{-1}L_{k-1})^* C_\beta + \operatorname{id}\right) C_\beta^{-1} K_{k-1}^* p^k \\
&= -2C_\beta^{-1}C_{k-1}^* r_{k-1}^g + \frac{2}{\beta_k}C_\beta^{-1}(u_0 - u^{k-1}) + 2C_\beta K_{k-1}^{-1}\left(L_{k-1}(q_0 - q^{k-1}) + r_{k-1}^f\right) \\
&\quad + \left[2C_\beta^{-1}C_{k-1}^*C_{k-1} - 2C_\beta + \frac{2}{\beta_k}C_\beta^{-1}\right]u^{k-1} \\
&= -2C_\beta^{-1}C_{k-1}^* r_{k-1}^g + \frac{2}{\beta_k}C_\beta^{-1}(u_0 - u^{k-1}) + 2C_\beta K_{k-1}^{-1}\left(L_{k-1}(q_0 - q^{k-1}) + r_{k-1}^f\right) \\
&\quad + \left[2C_\beta^{-1}\left(\frac{1}{\beta_k}\operatorname{id} + C_{k-1}^*C_{k-1}\right) - 2C_\beta\right]u^{k-1} \\
&= -2C_\beta^{-1}C_{k-1}^* r_{k-1}^g + \frac{2}{\beta_k}C_\beta^{-1}(u_0 - u^{k-1}) + 2C_\beta K_{k-1}^{-1}\left(L_{k-1}(q_0 - q^{k-1}) + r_{k-1}^f\right) .
\end{aligned}$$

Since $\beta_k C_\beta K_{k-1}^{-1} L_{k-1}(K_{k-1}^{-1}L_{k-1})^* C_\beta = \beta_k(C_\beta K_{k-1}^{-1}L_{k-1})(C_\beta K_{k-1}^{-1}L_{k-1})^*$ is positive semidefinite, we can conclude

$$\begin{aligned}
&\left\|C_\beta^{-1}K_{k-1}^* p^k\right\|_V \\
&\leq \quad \left\|-2C_\beta^{-1}C_{k-1}^* r_{k-1}^g + \frac{2}{\beta_k}C_\beta^{-1}(u_0 - u^{k-1}) + 2C_\beta K_{k-1}^{-1}\left(L_{k-1}(q_0 - q^{k-1}) + r_{k-1}^f\right)\right\|_V ,
\end{aligned}$$

and with the estimates

$$\|C_\beta^{-1}C_{k-1}^*\|_{G\to V} \leq 1 , \quad \|C_\beta^{-1}\|_V \leq \beta_k^{\frac{1}{2}} , \qquad \text{and} \quad \|C_\beta\|_V \leq \left(\frac{1}{\beta_k} + \|C_{k-1}\|_{V\to G}^2\right)^{\frac{1}{2}}$$

we have

$$\begin{aligned}
\left\|K_{k-1}^* p^k\right\|_V &\leq \left\|C_\beta C_\beta^{-1} K_{k-1}^* p^k\right\|_V \\
&\leq \left(\frac{1}{\beta_k} + \|C_{k-1}\|_{V\to G}^2\right)^{\frac{1}{2}} \left\|C_\beta^{-1}K_{k-1}^* p^k\right\|_V \\
&\leq 2\left(\frac{1}{\beta_k} + \|C_{k-1}\|_{V\to G}^2\right)^{\frac{1}{2}} \Big\{\|r_{k-1}^g\| + \frac{1}{\sqrt{\beta_k}}\|u_0 - u^{k-1}\|_V \\
&\qquad + \left(\frac{1}{\beta_k} + \|C_{k-1}\|_{V\to G}^2\right)^{\frac{1}{2}} \|K_{k-1}^{-1}\left(L_{k-1}(q_0 - q^{k-1}) + r_{k-1}^f\|\right)\Big\} ,
\end{aligned}$$

which by Assumption 4.17 and $(q^{k-1}, u^{k-1}) \in \mathcal{B}_\rho(q_0, u_0)$ yields (4.100).. □

We will prove inductively that the iterates indeed remain in $\mathcal{B}_\rho(q_0, u_0)$, and even

$$\|q_h^k - q_0\|_Q^2 + \|u_h^k - u_0\|_{V_h}^2 \le \|q^\dagger - q_0\|_Q^2 + \|u^\dagger - u_0\|_V^2 \qquad \forall 1 \le k \le k_* .$$

Thus, due to Lemma 4.10, which remains valid in the discretized setting (4.110), i.e.,

$$\|z_h^k\|_{W_h} \le c_{\mathrm{adj}} \left(\|q_h^k - q_0\|_Q + \|u_h^k - u_0\|_{V_h} + 1 \right) , \tag{4.107}$$

we get uniform boundedness of the dual variables by some sufficiently large ϱ, namely

$$\varrho \ge c_{\mathrm{adj}} \left(\|q^\dagger - q_0\|_Q + \|u^\dagger - u_0\|_V + 1 \right) . \tag{4.108}$$

This allows us to use an exact Penalty formulation

$$\begin{aligned} \min_{(q,u)\in Q\times V} \mathcal{P}(q,u) &:= \|C(u^{k-1}) + C'(u^{k-1})(u - u^{k-1}) - g^\delta\|_G^2 + \frac{1}{\beta_k}(\|q - q_0\|_Q^2 + \|u - u_0\|_V^2) \\ &+ \varrho \|A_q'(q^{k-1}, u^{k-1})(q - q^{k-1}) + A_u'(q^{k-1}, u^{k-1})(u - u^{k-1}) + A(q^{k-1}, u^{k-1}) - f\|_{W^*} , \end{aligned} \tag{4.109}$$

since for ϱ larger than the norm of the dual variable, any solution of (4.95),(4.96) is a solution of (4.109), see Section 2.5.4.

The formulation (4.109) of (4.95),(4.96) will be used in the convergence proofs only. For a practical implementation we will directly discretize (4.95),(4.96).

The discrete version of (4.95),(4.96) reads

$$\begin{aligned} \min_{(q,u)\in Q_{h_k}\times V_{h_k}} \quad & \|C(u_{\mathrm{old}}) + C'(u_{\mathrm{old}})(u - u_{\mathrm{old}}) - g^\delta\|_G^2 + \frac{1}{\beta_k}\left(\|q - q_0\|_Q^2 + \|u - u_0\|_{V_{h_k}}^2 \right) \\ \text{s.t.} \quad & L_{k-1}(q - q_{\mathrm{old}}) + K_{k-1}(u - u_{\mathrm{old}}) + A(q_{\mathrm{old}}, u_{\mathrm{old}}) = f \qquad \text{in } W_{h_k}^* , \end{aligned} \tag{4.110}$$

where $(q_{\mathrm{old}}, u_{\mathrm{old}}) = (q^{k-1}, u^{k-1}) = (q_{h_{k-1}}^{k-1}, u_{h_{k-1}}^{k-1})$ is the previous iterate and we assume again that the norms in G and Q as well as the semilinear form a and the operator C are evaluated exactly (cf. Assumption 2.3).

With ϱ chosen sufficiently large such that (4.108) holds, we define the quantities of interest as follows

$$\begin{aligned} I_1\colon\ & V \times Q \times V \times \mathbb{R} \to \mathbb{R}, && (u_{\mathrm{old}}, q, u, \beta) && \mapsto \left\| C'(u_{\mathrm{old}})(u - u_{\mathrm{old}}) + C(u_{\mathrm{old}}) - g^\delta \right\|_G^2 \\ & && && \quad + \frac{1}{\beta}\left(\|q - q_0\|_Q^2 + \|u - u_0\|_V^2 \right) \\ I_2\colon\ & V \times V \to \mathbb{R}, && (u_{\mathrm{old}}, u) && \mapsto \left\| C'(u_{\mathrm{old}})(u - u_{\mathrm{old}}) + C(u_{\mathrm{old}}) - g^\delta \right\|_G^2 \\ I_3\colon\ & Q \times V \to \mathbb{R}, && (q_{\mathrm{old}}, u_{\mathrm{old}}) && \mapsto \left\| C(u_{\mathrm{old}}) - g^\delta \right\|_G^2 + \varrho \, \|A(q_{\mathrm{old}}, u_{\mathrm{old}}) - f\|_{W^*} \\ I_4\colon\ & Q \times V \to \mathbb{R}, && (q, u) && \mapsto \left\| C(u) - g^\delta \right\|_G^2 + \varrho \, \|A(q, u) - f\|_{W^*} \end{aligned} \tag{4.111}$$

(cf. (4.88)) and

$$\begin{aligned} I_1^k &= I_1(u_{\mathrm{old}}^k, q^k, u^k, \beta_k) \\ I_2^k &= I_2(u_{\mathrm{old}}^k, u^k) \\ I_3^k &= I_3(q_{\mathrm{old}}^k, u_{\mathrm{old}}^k) \\ I_4^k &= I_4(q^k, u^k)\,, \end{aligned} \tag{4.112}$$

(cf. (4.89)), where q_{old}^k, u_{old}^k are fixed from the previous step and q^k, u^k are coupled by the linearized state equation (4.96) for $q^{k-1} = q_{\mathrm{old}}^k$ and $u^{k-1} = u_{\mathrm{old}}^k$.

Consistently, the discrete counterparts to (4.111) and (4.112) are

$$\begin{aligned} I_{1,h}\colon\ & V \times Q \times V \times \mathbb{R} \to \mathbb{R}\,, && (u_{\mathrm{old}}, q, u, \beta) && \mapsto \left\| C'(u_{\mathrm{old}})(u - u_{\mathrm{old}}) + C(u_{\mathrm{old}}) - g^\delta \right\|_G^2 \\ &&&&& \quad + \frac{1}{\beta}\left(\|q - q_0\|_Q^2 + \|u - u_0\|_{V_{h_k}}^2 \right) \\ I_{2,h}\colon\ & V \times V \to \mathbb{R}\,, && (u_{\mathrm{old}}, u) && \mapsto I_2(u_{\mathrm{old}}, u) \\ I_{3,h}\colon\ & Q \times V \to \mathbb{R}\,, && (q_{\mathrm{old}}, u_{\mathrm{old}}) && \mapsto \left\| C(u_{\mathrm{old}}) - g^\delta \right\|_G^2 + \varrho\, \|A(q_{\mathrm{old}}, u_{\mathrm{old}}) - f\|_{W_{h_k}^*} \\ I_{4,h}\colon\ & Q \times V \to \mathbb{R}\,, && (q, u) && \mapsto \left\| C(u) - g^\delta \right\|_G^2 + \varrho\, \|A(q, u) - f\|_{W_{h_k}^*} \end{aligned}$$

and

$$\begin{aligned} I_{1,h}^k &= I_{1,h}(u_{\mathrm{old}}^k, q_{h_k}^k, u_{h_k}^k, \beta_k) \\ I_{2,h}^k &= I_{2,h}(u_{\mathrm{old}}^k, u_{h_k}^k) \\ I_{3,h}^k &= I_{3,h}(q_{\mathrm{old}}^k, u_{\mathrm{old}}^k) \\ I_{4,h}^k &= I_{4,h}(q_{h_k}^k, u_{h_k}^k) \end{aligned} \tag{4.113}$$

(cf. (4.91)), where q_{old}^k, u_{old}^k are fixed from the previous step, since (like in (4.92)) we set $q_{\mathrm{old}}^{k+1} = q_{h_k}^k$ and $u_{\mathrm{old}}^{k+1} = u_{h_k}^k$ at the end of each iteration step.

Remark 4.18. *Here, as compared to* (4.89), *we have removed the W^*-norms in the definition of I_1^k and I_2^k.*

The W^*-norm still appears in I_3^k, but only in connection with the old iterate $(q_{\mathrm{old}}^k, u_{\mathrm{old}}^k)$, such that the only source of error in I_3^k is the evaluation of the W^*-norm. This means, we have replaced the expression

$$\|E(q,u)\|_{W^*} - \|E(q_h, u_h)\|_{W_h^*}$$

from Section 4.2.1 by an expression of the form

$$\|E(q_{\mathrm{old}}, u_{\mathrm{old}})\|_{W^*} - \|E(q_{\mathrm{old}}, u_{\mathrm{old}})\|_{W_h^*}\ . \tag{4.114}$$

For the typical case $W = V = H_0^1(\Omega)$ (see Section 4.3), we will discuss the computation of a goal-oriented error estimator for this kind of error in the subsection "Computation of the error estimators" of this section.

Another way to deal with the discretization error in I_3^k is the following: Tracking the upcoming convergence proof (cf. Theorem 4.11) the reader should realize that the discretization for $I_{3,h}^k$

does not have to be the same as for $I_{1,h}^k$, $I_{2,h}^k$, such that $I_{3,h}^k$ could be evaluated on a very fine separate mesh and η_3^k could be neglected. This alternative is of course, more costly, but since everything else is still done on the adaptively refined (coarser) mesh, the proposed method could still lead to an efficient algorithm.

The W^*-norm also appears in I_4^k, and unfortunately, in combination with the current q and u, which are subject to discretization, such that in principle we face the same situation as in the least squares formulation from Section 4.2.1. Since, however, η_4^k only appears in connection with the very weak assumption $\eta_4^k \to 0$ as $k \to \infty$ (cf. (4.28)*), as in Section 4.1, we save ourselves the computational effort of computing an error estimator for I_4^k.*

Like in Section 4.2.1 we need the weak sequential closedness of $\mathbf{F}$, i.e., Assumption 4.14 and the following tangential cone condition are assumed to hold in this section.

Assumption 4.18. *There holds*

$$\begin{aligned} \|C(u) - C(\bar{u}) - C'(u)(u-\bar{u})\|_G &\leq c_{tc}\|C(u) - C(\bar{u})\|_G \\ \|A(q,u) - A(\bar{q},\bar{u}) - A'_q(q,u)(q-\bar{q}) - A'_u(q,u)(u-\bar{u})\|_{W^*} &\leq 4c_{tc}^2\|A(q,u) - A(\bar{q},\bar{u})\|_{W^*}\,, \end{aligned}$$

for all $(q,u), (\overline{q},\overline{u}) \in \mathcal{B}_\rho(q_0,u_0)$, and some $0 < c_{tc} < 1$.

By means of Lemma 4.10 and Assumption 4.14 and 4.18 we can now formulate a convergence result like in Theorem 4.2, 4.4 and 4.8 for (4.95),(4.96). This can be done similarly to the proof of Theorem 4.4 replacing F there by $\mathbf{F}$ according to (4.77) and setting

$$\begin{aligned} \mathcal{S}\left(\begin{pmatrix} y_C \\ y_A \end{pmatrix}, \begin{pmatrix} \tilde{y}_C \\ \tilde{y}_A \end{pmatrix}\right) &= \|y_C - \tilde{y}_C\|_G^2 + \varrho\|y_A - \tilde{y}_A\|_{W^*}\,, \\ \mathcal{R}\left(\begin{pmatrix} q \\ u \end{pmatrix}\right) &= \|q - q_0\|_Q^2 + \|u - u_0\|_V^2\,, \\ c_S &= 2\,. \end{aligned} \tag{4.115}$$

For clarity of exposition we provide the full convergence proof (Theorem 4.11) without making use of the fact that a solution to the contrained optimization problem (4.95), (4.96) is a solution to the penalty problem (4.109) here (since this implication only holds if the sequence of adjoint states is uniformly bounded, which will be shown at a later point only cf. Corollary 4.12). Only for the convergence rates result Corollary 4.13 we refer to Theorem 4.4 with (4.115) and the relation to (4.109). So in the proof of Theorem 4.11 we will not use minimality with respect to (4.109) but only with respect to the original formulation (4.95),(4.96)

Theorem 4.11. *Let Assumption 4.14 and 4.18 with c_{tc} sufficiently small be satisfied. For the quantities of interest* (4.112) *and* (4.113)*, let further Assumption 4.11 with $c_S = 2$ and* (4.17) *hold. Let the starting value $(q_{\text{old}}^1, u_{\text{old}}^1) = (q_{h_0}^0, u_{h_0}^0)$ be chosen such that*

$$\|q_{\text{old}}^1 - q_0\|_Q^2 + \|u_{\text{old}}^1 - u_0\|_{V_h}^2 \leq \|q^\dagger - q_0\|_Q^2 + \|u^\dagger - u_0\|_V^2$$

(which is obviously satisfied, for instance, by the choice $(q_{\text{old}}^1, u_{\text{old}}^1) = (q_0, u_0)$).

Then with β_k, $h = h_k$ fulfilling (4.18)*, k_* selected according to* (4.25)*, and $(q_{h_k}^k, u_{h_k}^k)$ being a solution to* (4.110)*, for any solution $(q^\dagger, u^\dagger) \in \mathcal{B}_\rho(q_0, u_0)$ of* (4.76) *there holds:*

(i) the estimate

$$\|q^k_{h_k} - q_0\|^2_Q + \|u^k_{h_k} - u_0\|^2_{V_h} \leq \|q^\dagger - q_0\|^2_Q + \|u^\dagger - u_0\|^2_V \qquad \forall 1 \leq k \leq k_*\,, \tag{4.116}$$

which implies that all iterates stay in the ball $\mathcal{B}_\rho(q_0, u_0)$,

(ii) k_* is finite.

(iii) $(q^{k_*}_{\text{old}}, u^{k_*}_{\text{old}}) = (q^{k_*(\delta)}_{\text{old}}, u^{k_*(\delta)}_{\text{old}}) = (q^{k_*(\delta)-1,\delta}_{\beta_{k_*(\delta)-1},h_{k_*(\delta)-1}}, u^{k_*(\delta)-1,\delta}_{\beta_{k_*(\delta)-1},h_{k_*(\delta)-1}})$ converges (weakly) subsequentially to a solution of (4.76) as $\delta \to 0$ in the sense that it has a weakly convergent subsequence and each weakly convergent subsequence converges strongly to a solution of (4.76). If the solution $(q^\dagger, u^\dagger)$ to (4.76) is unique, then $(q^{k_*(\delta)}_{\text{old}}, u^{k_*(\delta)}_{\text{old}})$ converges strongly to $(q^\dagger, u^\dagger)$ as $\delta \to 0$.

Proof. (i): We will prove (4.31) as follows: We will show that, for fixed $k > 0$, provided that

$$\|q^{k-1}_{h_{k-1}} - q_0\|^2_Q + \|u^{k-1}_{h_{k-1}} - u_0\|^2_{V_h} \leq \|q^\dagger - q_0\|^2_Q + \|u^\dagger - u_0\|^2_V$$

(which implies $(q^{k-1}_{h_{k-1}}, u^{k-1}_{h_{k-1}}) \in \mathcal{B}_\rho(q_0, u_0)$), there holds

$$\|q^k_{h_k} - q_0\|^2_Q + \|u^k_{h_k} - u_0\|^2_{V_h} \leq \|q^\dagger - q_0\|^2_Q + \|u^\dagger - u_0\|^2_V\,.$$

Then (4.31) follows recursively, since

$$\|q^1_{\text{old}} - q_0\|^2_Q + \|u^1_{\text{old}} - u_0\|^2_{V_h} \leq \|q^\dagger - q_0\|^2_Q + \|u^\dagger - u_0\|^2_V\,.$$

So we assume

$$\|q^{k-1}_{h_{k-1}} - q_0\|^2_Q + \|u^{k-1}_{h_{k-1}} - u_0\|^2_{V_h} \leq \|q^\dagger - q_0\|^2_Q + \|u^\dagger - u_0\|^2_V \tag{4.117}$$

in the following.

Let

$$C_{h,k-1} := C'(u^{k-1}_{h_{k-1}})\,, \quad L_{h,k-1} := A'_q(q^{k-1}_{h_{k-1}}, u^{k-1}_{h_{k-1}})\,, \quad K_{h,k-1} := A'_u(q^{k-1}_{h_{k-1}}, u^{k-1}_{h_{k-1}}) \tag{4.118}$$

and

$$r^f_{h,k-1} := A(q^{k-1}_{h_{k-1}}, u^{k-1}_{h_{k-1}}) - f\,, \quad r^g_{h,k-1} := C(u^{k-1}_{h_{k-1}}) - g^\delta\,. \tag{4.119}$$

We consider a continuous step emerging from discrete $q^k_{\text{old}} = q^{k-1}_{h_{k-1}}$, $u^k_{\text{old}} = u^{k-1}_{h_{k-1}}$ (cf. Figure 4.3), i.e., let (q^k, u^k) be a solution to (4.95) for $q^{k-1} = q^k_{\text{old}} = q^{k-1}_{h_{k-1}}$. Then the first order optimality condition $\mathcal{L}'(q^k, u^k, z^k) = 0$ (cf. (4.97)–(4.99)) implies

$$\begin{aligned} 0 = (C_{h,k-1}(u^k - u^{k-1}_{h_{k-1}}) + r^g_{h,k-1}, C_{h,k-1}\delta u)_G + \frac{1}{\beta_k}\Big[(q^k - q_0, \delta q)_Q + (u^k - u_0, \delta u)_V\Big] \\ + \frac{1}{2}\langle L_{h,k-1}\delta q + K_{h,k-1}\delta u, z^k\rangle_{W^*,W} \end{aligned}$$

for all $\delta q \in Q$ and $\delta u \in V$, where we have used the same abbreviations as in (4.87) and (4.102) as well as $p^k = J_{W^*} z^k$ (see (4.101)).

Setting $\delta q = q^k - q^\dagger$, $\delta u = u^k - u^\dagger$, this yields

$$
\begin{aligned}
0 = &\|C_{h,k-1}(u^k - u^{k-1}_{h_{k-1}}) + r^g_{h,k-1}\|^2_G \\
&- (C_{h,k-1}(u^k - u^{k-1}_{h_{k-1}}) + r^g_{h,k-1}, C_{h,k-1}(u^\dagger - u^{k-1}_{h_{k-1}}) + r^g_{h,k-1})_G \\
&+ \frac{1}{\beta_k}\|q^k - q_0\|^2_Q - \frac{1}{\beta_k}(q^k - q_0, q^\dagger - q_0)_Q + \frac{1}{\beta_k}\|u^k - u_0\|^2 - \frac{1}{\beta_k}(u^k - u_0, u^\dagger - u_0)_V \\
&- \frac{1}{2}\langle L_{h,k-1}(q^\dagger - q^{k-1}_{h_{k-1}}) + K_{h,k-1}(u^\dagger - u^{k-1}_{h_{k-1}}) + r^f_{h,k-1}, z^k\rangle_{W^*,W}\,,
\end{aligned}
$$

where we have used the fact that (q^k, u^k) satisfies the linearized state equation (4.96), i.e.,

$$
L_{h,k-1}(q^k - q^{k-1}_{h_{k-1}}) + K_{h,k-1}(u^k - u^{k-1}_{h_{k-1}}) + r^f_{h,k-1} = 0\,.
$$

Hence by Cauchy-Schwarz and the fact that $ab \le \frac{1}{2}a^2 + \frac{1}{2}b^2$ for all $a, b \in \mathbb{R}$,

$$
\begin{aligned}
I^k_1 \le\; &\|C_{h,k-1}(u^k - u^{k-1}_{h_{k-1}}) + r^g_{h,k-1}\|_G \|C_{h,k-1}(u^\dagger - u^{k-1}_{h_{k-1}}) + r^g_{h,k-1}\|_G \\
&+ \frac{1}{\beta_k}\|q^k - q_0\|_Q\|q^\dagger - q_0\|_Q + \frac{1}{\beta_k}\|u^k - u_0\|_V\|u^\dagger - u_0\|_V \\
&+ \frac{1}{2}\|L_{h,k-1}(q^\dagger - q^{k-1}_{h_{k-1}}) + K_{h,k-1}(u^\dagger - u^{k-1}_{h_{k-1}}) + r^f_{h,k-1}\|_{W^*}\|z^k\|_W \\
\le\; &\frac{1}{2}\|C_{h,k-1}(u^k - u^{k-1}_{h_{k-1}}) + r^g_{h,k-1}\|^2_G + \frac{1}{2}\|C_{h,k-1}(u^\dagger - u^{k-1}_{h_{k-1}}) + r^g_{h,k-1}\|^2_G \\
&+ \frac{1}{2\beta_k}\|q^k - q_0\|^2_Q + \frac{1}{2\beta_k}\|q^\dagger - q_0\|^2_Q + \frac{1}{2\beta_k}\|u^k - u_0\|^2_V + \frac{1}{2\beta_k}\|u^\dagger - u_0\|^2_V \\
&+ \frac{1}{2}\|L_{h,k-1}(q^\dagger - q^{k-1}_{h_{k-1}}) + K_{h,k-1}(u^\dagger - u^{k-1}_{h_{k-1}}) + r^f_{k-1}\|_{W^*}h\|z^k\|_W \\
=\; &\frac{1}{2}I^k_1 + \frac{1}{2}\|C_{h,k-1}(u^\dagger - u^{k-1}_{h_{k-1}}) + r^g_{h,k-1}\|^2_G \\
&+ \frac{1}{2\beta_k}\|q^\dagger - q_0\|^2_Q + \frac{1}{2\beta_k}\|u^\dagger - u_0\|^2_V \\
&+ \frac{1}{2}\|L_{h,k-1}(q^\dagger - q^{k-1}_{h_{k-1}}) + K_{h,k-1}(u^\dagger - u^{k-1}_{h_{k-1}}) + r^f_{h,k-1}\|_{W^*}\|z^k\|_W\,,
\end{aligned}
$$

which multiplying by 2 and applying Lemma 4.10 with (4.108), and (4.117) leads to

$$
\begin{aligned}
I^k_1 \le\; &\|C_{h,k-1}(u^\dagger - u^{k-1}_{h_{k-1}}) + r^g_{h,k-1}\|^2_G + \frac{1}{\beta_k}\|q^\dagger - q_0\|^2_Q + \frac{1}{\beta_k}\|u^\dagger - u_0\|^2_V \\
&+ \|L_{h,k-1}(q^\dagger - q^{k-1}_{h_{k-1}}) + K_{h,k-1}(u^\dagger - u^{k-1}_{h_{k-1}}) + r^f_{h,k-1}\|_{W^*}\|z^k\|_W \\
\le\; &\|C'(u^{k-1}_{h_{k-1}})(u^\dagger - u^{k-1}_{h_{k-1}}) + C(u^{k-1}_{h_{k-1}}) - g^\delta\|^2_G + \frac{1}{\beta_k}\|q^\dagger - q_0\|^2_Q + \frac{1}{\beta_k}\|u^\dagger - u_0\|^2_V \\
&+ \varrho\|A'_q(q^{k-1}_{h_{k-1}})(q^\dagger - q^{k-1}_{h_{k-1}}) + A'_u(u^{k-1}_{h_{k-1}})(u^\dagger - u^{k-1}_{h_{k-1}}) + A(q^{k-1}_{h_{k-1}}, u^{k-1}_{h_{k-1}}) - A(q^\dagger, u^\dagger)\|_{W^*}
\end{aligned}
\tag{4.120}
$$

for all $1 \le k < k_*$.

The rest of the proof basically follows the lines of the proof of Theorem 4.4 with the choice (4.115), but for convenience of the reader we will follow through the proof anyway.

Using the fact that $(a+b)^2 \le 2a^2 + 2b^2$ for all $a, b \in \mathbb{R}$ and Assumption 4.18, we get from (4.120)

$$\begin{aligned}
I_1^k &\le 2c_{tc}^2 \|C(u_{h_{k-1}}^{k-1}) - C(u^\dagger)\|_G^2 + 2\delta^2 + \frac{1}{\beta_k}\|q^\dagger - q_0\|_Q^2 + \frac{1}{\beta_k}\|u^\dagger - u_0\|_V^2 \\
&\qquad + 4c_{tc}^2 \varrho \|A(q_{h_{k-1}}^{k-1}, u_{h_{k-1}}^{k-1}) - A(q^\dagger, u^\dagger)\|_{W^*} \\
&\le 4c_{tc}^2 \|C(u_{h_{k-1}}^{k-1}) - g^\delta\|_G^2 + 2(1+2c_{tc}^2)\delta^2 + \frac{1}{\beta_k}\|q^\dagger - q_0\|_Q^2 + \frac{1}{\beta_k}\|u^\dagger - u_0\|_V^2 \\
&\qquad + 4c_{tc}^2 \varrho \|A(q_{h_{k-1}}^{k-1}, u_{h_{k-1}}^{k-1}) - f\|_{W^*} \\
&\le 4c_{tc}^2 I_3^k + \frac{2(1+2c_{tc}^2)}{\tau^2} I_{3,h}^k + \frac{1}{\beta_k}\|q^\dagger - q_0\|_Q^2 + \frac{1}{\beta_k}\|u^\dagger - u_0\|_V^2 \\
&\le 2\left(2c_{tc}^2 + \frac{1+2c_{tc}^2}{\tau^2}\right) I_{3,h}^k + 4c_{tc}^2 \eta_3^k + \frac{1}{\beta_k}\left(\|q^\dagger - q_0\|_Q^2 + \|u^\dagger - u_0\|_V^2\right)
\end{aligned}$$

for all $1 \le k < k_*$. This together with (4.18) and the fact that

$$I_{1,h}^k = I_{2,h}^k + \frac{1}{\beta_k}\left(\|q_h^k - q_0\|_Q^2 + \|u_h^k - u_0\|_{V_h}^2\right)$$

yields

$$\begin{aligned}
&\underline{\tilde{\theta}} I_{3,h}^k + \frac{1}{\beta_k}(\|q_{h_k}^k - q_0\|_Q^2 + \|u_{h_k}^k - u_0\|_{V_h}^2) \\
&\le I_{2,h}^k + \frac{1}{\beta_k}(\|q_{h_k}^k - q_0\|_Q^2 + \|u_{h_k}^k - u_0\|_{V_h}^2) \\
&\le I_1^k + \eta_1^k \\
&\le 2\left(2c_{tc}^2 + \frac{(1+2c_{tc})^2}{\tau^2}\right) I_{3,h}^k + \frac{1}{\beta_k}(\|q^\dagger - q_0\|_Q^2 + \|u^\dagger - u_0\|_V^2) + \eta_1^k + 4c_{tc}^2 \eta_3^k\,.
\end{aligned}$$

Hence by (4.63) with $c_S = 2$ we get (4.116), which implies $(q_{\text{old}}^{k+1}, u_{\text{old}}^{k+1}) = (q_{h_k}^k, u_{h_k}^k) \in \mathcal{B}_\rho(q_0, u_0)$ for all $0 \le k \le k_*$.

(ii): By the triangle inequality as well as (4.18), Assumption 4.18 and the fact that (q^k, u^k)

satisfies the linearized state equation (4.96), we have

$$\begin{aligned}
I_4^k &= \|C(u^k) - g^\delta\|_G^2 + \varrho\|A(q^k, u^k) - f\|_{W^*} \\
&\leq 2\|C'(u_{\text{old}}^k)(u^k - u_{\text{old}}^k) + C(u_{\text{old}}^k) - g^\delta\|_G^2 \\
&\quad + 2\|C'(u_{\text{old}}^k)(u^k - u_{\text{old}}^k) + C(u_{\text{old}}^k) - C(u^k)\|_G^2 \\
&\quad + \varrho\|A_q'(q_{\text{old}}^k, u_{\text{old}}^k)(q^k - q_{\text{old}}^k) + A_u'(q_{\text{old}}^k, u_{\text{old}}^k)(u^k - u_{\text{old}}^k) + A(q_{\text{old}}^k, u_{\text{old}}^k) - f\|_{W^*} \\
&\quad + \varrho\|A_q'(q_{\text{old}}^k, u_{\text{old}}^k)(q^k - q_{\text{old}}^k) + A_u'(q_{\text{old}}^k, u_{\text{old}}^k)(u^k - u_{\text{old}}^k) \\
&\qquad + A(q_{\text{old}}^k, u_{\text{old}}^k) - A(q^k, u^k)\|_{W^*} \\
&\leq 2I_2^k + 2c_{tc}^2\|C(u^k) - C(u_{\text{old}}^k)\|_G^2 + 4c_{tc}^2\varrho\|A(q_{\text{old}}^k, u_{\text{old}}^k) - A(q^k, u^k)\|_{W^*} \\
&\leq 2I_2^k + 4c_{tc}^2\left(\|C(u^k) - g^\delta\|_G^2 + \|C(u_{\text{old}}^k) - g^\delta\|_G^2\right) \\
&\quad + 4c_{tc}^2\varrho\left(\|A(q_{\text{old}}^k, u_{\text{old}}^k) - f\|_{W^*} + \|A(q^k, u^k) - f\|_{W^*}\right) \\
&\leq 2\left(\tilde{\bar{\theta}} I_{3,h}^k + \eta_2^k\right) + 4c_{tc}^2(I_4^k + I_3^k)\,,
\end{aligned}$$

which implies

$$I_4^k \leq \frac{1}{1 - 4c_{tc}^2}\left(2\tilde{\bar{\theta}} I_{3,h}^k + 2\eta_2^k + 4c_{tc}^2 I_3^k\right)\,.$$

The rest of the proof of (ii) follows exactly the lines of the proof of Theorem 4.2 after (4.69).

(iii): With (2.2), (4.17), (4.28) and the definition of k_*, we have

$$\begin{aligned}
\|C(u_{\text{old}}^{k_*}) - g\|_G^2 + \varrho\|A(q_{\text{old}}^{k_*}, u_{\text{old}}^{k_*}) - f\|_{W^*} &\leq 2I_3^{k_*} + 2\delta^2 \\
&\leq 2\left(I_{3,h}^{k_*} + \eta_3^k + \delta^2\right) \\
&\leq 2\left((1 + c_1)I_{3,h}^{k_*} + \delta^2\right) \\
&\leq 2\delta^2\left((1 + c_1)\tau^2 + 1\right) \to 0
\end{aligned} \tag{4.121}$$

as $\delta \to 0$. Thus, due to (ii) (4.116) $(q_{\text{old}}^{k_*}, u_{\text{old}}^{k_*}) = (q_{\text{old}}^{k_*(\delta)}, u_{\text{old}}^{k_*(\delta)})$ has a weakly convergent subsequence $\left((q_{\text{old}}^{k_*(\delta_l)}, u_{\text{old}}^{k_*(\delta_l)})\right)_{l\in\mathbb{N}}$ and with Assumption 4.14 and (4.121) the limit of every weakly convergent subsequence is a solution to (4.76). Strong convergence of any weakly convergent subsequence again follows by a standard argument like in (3.9).

□

Well-definedness of the regularization parameter β_k by (4.18) follows along the lines of Lemma 4.6 with the replacements (4.115).

Please note that Theorem 4.11 is a new result also in the continuous case $\eta_i^k = 0$.

Corollary 4.12. *The sequence $(z^k)_{k\in\mathbb{N},k\leq k_*}$ is bounded, i.e.*

$$\bar{\varrho} = \sup_{k\leq k_*}\|z^k\|_W \leq c_{\text{adj}}(\|q^\dagger - q_0\|_Q^2 + \|u^\dagger - u_0\|_V^2 + 1)$$

Proof. The assertion follows directly from Theorem 4.11 (i) and Lemma 4.10. □

The convergences rates from Theorem 3.2, Corollary 4.3, and Theorem 4.5 also hold for the all-at-once formulation (4.76), due to equivalence with (4.109), which we formulate in the following theorem.

Instead of standard source conditions (like in Assumption 3.5 and 4.16) we use variational source conditions (cf., e.g., [32, 48, 52, 97]) here, due to the nonquadratic penalty term in (4.109).

Assumption 4.19. *Let*

$$\begin{aligned}|(q^\dagger - q_0, q - q^\dagger)_Q + (u^\dagger - u_0, u - u^\dagger)_V| \\ \leq c\sqrt{\|q-q^\dagger\|_Q^2 + \|u-u^\dagger\|_V^2}\,\kappa\left(\frac{\|C(u)-C(u^\dagger)\|_G^2 + \varrho\|A(q,u)-A(q^\dagger,u^\dagger)\|_{W^*}}{\|q-q^\dagger\|_Q^2 + \|u-u^\dagger\|_V^2}\right)\end{aligned}$$

for all $(q,u) \in \mathcal{D}\times\mathcal{D}(C)$ *with* ϱ *sufficiently large (cf.* (4.108)*) and independent from* q, u*, hold with* κ *defined as in Assumption 4.16.*

Corollary 4.13. *Let the conditions of Theorem 4.11 and additionally the variational inequality Assumption 4.19 be fulfilled. Then there exist* $\bar{\delta} > 0$ *and a constant* $C > 0$ *independent of* δ *such that for all* $\delta \in (0, \bar{\delta}]$ *the convergence rates*

$$\|q_{\text{old}}^{k_*} - q^\dagger\|_Q^2 + \|u_{\text{old}}^{k_*} - u^\dagger\|_V^2 = \mathcal{O}\left(\frac{\delta^2}{\psi^{-1}(C\delta)}\right)$$

with $q_{\text{old}}^{k_*} = q_{\text{old}}^{k_*(\delta)} = q_{\beta_{k_*(\delta)-1}, h_{k*(\delta)-1}}^{\delta, k_*(\delta)-1}$, $u_{\text{old}}^{k_*} = u_{\text{old}}^{k_*(\delta)} = u_{\beta_{k_*(\delta)-1}, h_{k*(\delta)-1}}^{\delta, k_*(\delta)-1}$ *are obtained.*

Proof. With (4.115) the rates follow directly from Theorem 4.5 due to Theorem 4.11 (especially (4.116) and (4.121)). □

Remark 4.19. *In fact, no regularization of the u part would be needed for stability of the single Gauss-Newton steps, since by Assumption 4.13 the terms*

$$\varrho\|A_q'(q^{k-1}, u^{k-1})(q - q^{k-1}) + A_u'(q^{k-1}, u^{k-1})(u - u^{k-1}) + A(q^{k-1}, u^{k-1}) - f\|_{W^*}$$

and $\frac{1}{\beta_k}\|q - q_0\|_Q^2$ *as regularization term together ensure weak compactness of the level sets of the Tikhonov functional (cf. item 6 in Assumption 4.7).*

However, we require even uniform boundedness of $u_{h_k}^k$ *in order to uniformly bound the dual variable and come up with a penalty parameter* ϱ *that is independent of* k*, cf. the discrete version of Lemma 4.10 (see* (4.107)*). Trying to show this uniform boundedness without regularization term with respect to* u*, we fail at the following point:*

We use the equality constraint (4.96) *(for* $q^{k-1} = q^{k-1}_{h_{k-1}}$*) together with the tangential cone condition Assumption 4.18 for* $c_{tc}^2 < \frac{1}{2}$ *and get*

$$\begin{aligned}
&\|A(q^k, u^k) - f\|_{W^*} \\
&\quad = \|L_{h,k-1}(q^k - q^{k-1}_{h_{k-1}}) + K_{h,k-1}(u^k - u^{k-1}_{h_{k-1}}) + A(q^{k-1}_{h_{k-1}}, u^{k-1}_{h_{k-1}}) - A(q^k, u^k)\|_{W^*} \\
&\quad \leq 4c_{tc}^2 \|A(q^{k-1}_{h_{k-1}}, u^{k-1}_{h_{k-1}}) - A(q^k, u^k)\|_{W^*} \\
&\quad \leq 4c_{tc}^2 \|A(q^{k-1}_{h_{k-1}}, u^{k-1}_{h_{k-1}}) - f\|_{W^*} + 4c_{tc}^2 \|A(q^k, u^k) - f\|_{W^*} \,,
\end{aligned}$$

such that

$$\|A(q^k, u^k) - f\|_{W^*} \leq \frac{4c_{tc}^2}{1 - 4c_{tc}^2} \|A(q^{k-1}_{h_{k-1}}, u^{k-1}_{h_{k-1}}) - f\|_{W^*} \,.$$

However, without error estimators on the difference between $\|A(q^k, u^k) - f\|_{W^*}$ *and* $\|A(q^k_{h_k}, u^k_{h_k}) - f\|_{W^*}$*, this does not give a recursion*

$$\|A(q^k_{h_k}, u^k_{h_k}) - f\|_{W^*} \leq c\|A(q^{k-1}_{h_{k-1}}, u^{k-1}_{h_{k-1}}) - f\|_{W^*}$$

(from which, by uniform boundedness of $q^k_{h_k}$ *and Assumption 4.17 we could conclude uniform boundedness of* $u^k_{h_k}$*).*

Computation of the error estimators

Since – different to Section 4.1 – u_{old} ist not subject to discretization in this section, the computation of the error estimators is "easier" and can be done exactly as in [39] and [65]. Thus, we omit the arguments q_{old} and u_{old} in the quantities of interest in this subsection and (like in Section 4.1.2) we also omit the iteration index k and the explicit dependence on β.

In order to obtain uniform boundedness of $u^k_{h_k}$ we introduced the term $\frac{1}{\beta_k}\|u - u_0\|_V^2$ for theoretical purposes. For our practical computations we will assume that the error by discretization between $\|A(q^k, u^k) - f\|_{W^*}$ and $\|A(q^k_{h_k}, u^k_{h_k}) - f\|_{W^*}$ is small enough, so that the mentioned gap in this argument for uniform boundedness of $u^k_{h_k}$ (see Remark 4.19) can be neglected and the part $\frac{1}{\beta_k}\|u - u_0\|_V^2$ of the regularization term is omitted.

Error estimator for I_1:

We consider

$$I_1(q, u) = \|C'(u_{\text{old}})(u - u_{\text{old}}) + C(u_{\text{old}}) - g^\delta\|_G^2 + \frac{1}{\beta}\|q - q_0\|_Q^2$$

and the Lagrange functional

$$\mathcal{L}(q, u, z) \coloneqq I_1(q, u) + h(z) - B(q, u)(z)$$

with $h \in W^*$ and $B(q, u) \in W^*$ defined as

$$h \coloneqq f - A(q_{\text{old}}, u_{\text{old}}) - A'_q(q_{\text{old}}, u_{\text{old}})(q_{\text{old}}) - A'_u(q_{\text{old}}, u_{\text{old}})(u_{\text{old}})$$

and

$$B(q,u) \coloneqq A'_q(q_{\text{old}},u_{\text{old}})(q) + A'_u(q_{\text{old}},u_{\text{old}})(u)\,.$$

As already mentioned in Section 2.4, 3.3, 4.1.2 and 4.2.1, this approach is based on [11], combining the current quantity of interest with information on the minimization problem. There holds a similar result to Proposition 4.1 (see also [39]), which allows to estimate the difference $I_1(q,u) - I_1(q_h,u_h)$ by computing a discrete stationary point $x_h = (q_h,u_h,z_h) \in X_h = Q_h \times V_h \times W_h$ of $\mathcal{L}$. This is done by solving the equations

$$\begin{aligned}
z_h \in W_h\colon &\quad A'_u(q_{\text{old}},u_{\text{old}})(du)(z_h) = I'_{1,u}(q_h,u_h)(du) && \forall du \in V_h && (4.122)\\
u_h \in V_h\colon &\quad A'_q(q_{\text{old}},u_{\text{old}})(q_h)(dz) + A'_u(q_{\text{old}},u_{\text{old}})(u_h)(dz) = h(dz) && \forall dz \in W_h && (4.123)\\
q_h \in Q_h\colon &\quad I'_q(q_h,u_h)(dq) = A'_q(q_{\text{old}},u_{\text{old}})(dq)(z_h) && \forall dq \in Q_h && (4.124)
\end{aligned}$$

Then the error estimator η_1 for I_1 can be computed as

$$I_1 - I_{1,h} = I_1(q,u) - I_2(q_h,u_h) \approx \frac{1}{2}\mathcal{L}'(x_h)(\pi_h x_h - x_h) = \eta_1 \tag{4.125}$$

(cf. Section 4.1.2 and [39]).

Remark 4.20. *Please note that the equations* (4.122)–(4.124) *are solved anyway in the process of solving the optimization problem* (4.95),(4.96).

Error estimator for I_2:

The computation of the error estimator for I_2 can be done similarly to the computation of η_2 in Section 4.1.2 (or η^I from [39]) by means of the Lagrange functional $\mathcal{L}$. We consider

$$I_2(u) \coloneqq \|C'(u_{\text{old}})(u-u_{\text{old}}) + C(u_{\text{old}}) - g^\delta\|_G^2$$

and compute a discrete stationary point $y_h \coloneqq (x_h, x_{1,h}) \in X_h \times X_h$ of the auxiliary Lagrange functional

$$\mathcal{M}(y) \coloneqq I'_2(u) + \mathcal{L}''(x)(x_1)$$

by solving the equations

$$\begin{aligned}
x_h \in X_h\colon &\quad \mathcal{L}'(x_h)(dx_1) = 0 && \forall dx_1 \in X_h\\
x_{1,h} \in X_h\colon &\quad \mathcal{L}''(x_h)(x_{1,h},dx) = -I'_2(u_h)(du) && \forall dx \in X_h\,.
\end{aligned}$$

Then we compute the error estimator η_2 for I_2 by

$$I_2 - I_{2,h} = I_2(u) - I_2(u_h) \approx \frac{1}{2}\mathcal{M}'(y_h)(\pi_h y_h - y_h) = \eta_2\,.$$

Error estimator for I_3:

In Remark 4.18 we already mentioned that the W^*-norm in I_3 can be evaluated on a separate very fine mesh, so that we will neglect the difference between $\|A(q_{\text{old}},u_{\text{old}}) - f\|_{W^*}$ and

$\|A(q_{\text{old}},u_{\text{old}})-f\|_{W_h^*}$. This implies that we don't need to compute the error estimator η_3, since then we assume $I_3 = I_{3,h}$, such that (4.29) holds, and the first part of (4.28) is trivially fullfilled.

Nevertheless, for completness, we will give a rough idea how one could estimate an error of the form

$$\|E(q_{\text{old}},u_{\text{old}})\|_{W^*} - \|E(q_{\text{old}},u_{\text{old}})\|_{W_h^*}$$

for the typical case $V = W = H_0^1(\Omega)$. We exclude the trivial case $E(q_{\text{old}},u_{\text{old}}) = 0$ and define the functional $\Psi\colon V \to \mathbb{R}$,

$$\Psi(v) = \|\nabla v\|_{L^2(\Omega)}$$

and the auxiliary Lagrangian $M\colon V \times V$,

$$M(v,w) = \Psi(v) + \langle E(q_{\text{old}},u_{\text{old}}),w\rangle_{V^*,V} - (\nabla v,\nabla w)_{L^2(\Omega)}\,.$$

Let $v \in V$, $v_h \in V_h$ solve the equations

$$\begin{aligned}(\nabla v,\nabla\varphi)_{L^2(\Omega)} &= \langle E(q_{\text{old}},u_{\text{old}}),\varphi\rangle_{V^*,V} \quad \forall\varphi\in V\,,\\ (\nabla v_h,\nabla\varphi)_{L^2(\Omega)} &= \langle E(q_{\text{old}},u_{\text{old}}),\varphi\rangle_{V^*,V} \quad \forall\varphi\in V_h\end{aligned}$$

and let

$$w = \frac{1}{\|\nabla v\|}v\,, \qquad w_h = \frac{1}{\|\nabla v_h\|}v_h\,. \tag{4.126}$$

Then it is easily checked that $(v,w) \in V \times V$ and $(v_h,w_h) \in V_h \times V_h$ are continuous and discrete stationary points of M, i.e.,

$$\begin{aligned}M_v'(v,w)(\varphi) &= (\|\nabla v\|_{L^2(\Omega)})^{-1}(\nabla v,\nabla\varphi)_{L^2(\Omega)} - (\nabla\varphi,\nabla w)_{L^2(\Omega)} = 0 \qquad \forall\varphi\in V\,, &(4.127)\\ M_w'(v,w)(\varphi) &= \langle E(q_{\text{old}},u_{\text{old}}),\varphi\rangle_{V^*,V} - (\nabla v,\nabla\varphi)_{L^2(\Omega)} = 0 \qquad \forall\varphi\in V\,,\\ M_v'(v_h,w_h)(\varphi) &= (\|\nabla v_h\|_{L^2(\Omega)})^{-1}(\nabla v_h,\nabla\varphi)_{L^2(\Omega)} - (\nabla\varphi,\nabla w_h)_{L^2(\Omega)} = 0 \quad \forall\varphi\in V_h\,, &(4.128)\\ M_w'(v_h,w_h)(\varphi) &= \langle E(q_{\text{old}},u_{\text{old}}),\varphi\rangle_{V^*,V} - (\nabla v_h,\nabla\varphi)_{L^2(\Omega)} = 0 \qquad \forall\varphi\in V_h\,,\end{aligned}$$

and there holds

$$\begin{aligned}&\|E(q_{\text{old}},u_{\text{old}})\|_{V^*} - \|E(q_{\text{old}},u_{\text{old}})\|_{V_h^*}\\ &\quad= \|\nabla v\|_{L^2(\Omega)} - \|\nabla v_h\|_{L^2(\Omega)}\\ &\quad= \Psi(v) - \Psi(v_h)\\ &\quad= \tfrac{1}{2}M'(v_h,w_h)(v-\widehat{v_h},w-\widehat{w_h}) + R\\ &\quad= \tfrac{1}{2}\left[\langle E(q_{\text{old}},u_{\text{old}}),w-\widehat{w_h}\rangle_{V^*,V} - (\nabla v_h,\nabla(w-\widehat{w_h}))_{L^2(\Omega)}\right] + R\end{aligned}$$

for arbitrary $\widehat{v_h},\widehat{w_h} \in V_h$ with some remainder term R (cf. Section 2.4, 3.3 and 4.1.2). Please note that due to the relation (4.126) the dual variable w_h is obtained without solving an additional system of equations; different to Section 4.2.1, where we have an additional variable δx, for which this is not the case.

Error estimator for I_4:

We already mentioned in Remark 4.18 that we will not compute η_4 (with the same reasoning as in Section 4.1.2), as the error $|I_4 - I_{4,h}|$ needs to be controlled only through the very weak assumption $\eta_4^k \to 0$ as $k \to \infty$ (cf. (4.28)).

Remark 4.21. *We wish to mention once more that these error estimators are not reliable, see Remark 3.10.*

Algorithm

Since we only know about the existence of an upper bound $\overline{\varrho}$ on $\|z^k\|_W$ and $\|z^k_{h_k}\|_{W_{h_k}}$ (cf. Corollary 4.12), but not its value, we choose ϱ (cf. (4.108)) heuristically, i.e., in each iteration step we set $\varrho = \varrho_k = \max\{\varrho_{k-1}\,, \|z^k_{h_k}\|_{W_{h_k}}\}$ for the discrete counterpart $z^k_{h_k}$ of z^k.

As already mentioned in the discussion about error estimation in this section, we neglect the part $\frac{1}{\beta_k}\|u - u_0\|^2_V$ of the regularization term.

As motivated in the previous subsection "Computation of the error estimators", we assume $\eta_3^k = 0$ for all k, such that we neither compute η_3 nor η_4, although, for simplicity, we evaluate $I^k_{3,h}$ on the current mesh instead of a very fine mesh. Thus, we only check for the condition

$$\eta_1^k \leq \left(\underline{\tilde{\theta}} - 2\left(2c_{tc}^2 + \frac{(1+2c_{tc})^2}{\tau^2}\right)\right) I^k_{3,h}$$

on η_1^k in Assumption 4.11.

For computing β_k, $h = h_k$ fulfilling (4.18), we can resort to the Algorithm from [39] (cf. Algorithm 4.2, Theorem 4.1).

The presented Generalized Gauss-Newton formulation can be implemented according to the following Algorithm 4.6. The algorithm is illustrated by a flowchart in Figure 4.4.

The structure of the loops is the same as in Algorithm 4.1 from Section 4.1, but here, we only have to solve linear PDEs (i.e., step 6 in Algorithm 4.4 is replaced by "Solve linear PDE"), which justifies the drawback of an additional loop in comparison to [65] (see also Algorithm 4.5).

We mention once again that the Penalty problem is only considered for the evaluation of $I_{3,h}$. When it comes to optimization, we always solve the constrained problem (4.110), and not the penalty problem (4.78).

Algorithm 4.6: Generalized Gauss-Newton Method

Choose τ, τ_β, $\tilde{\tau}_\beta$, $\underline{\tilde{\theta}}$, $\tilde{\tilde{\theta}}$ such that $0 < \underline{\tilde{\theta}} \leq \tilde{\tilde{\theta}} < 1$ and (4.62) holds with $c_S = 2$. $\tilde{\theta} = (\underline{\tilde{\theta}} + \tilde{\tilde{\theta}})/2$ and $\max\{1\,, \tilde{\tau}_\beta\} < \tau_\beta \leq \tau$, and choose c_1, c_2 and c_3, such that the second part of (4.29) is fulfilled.
Choose a discretization $h = h_0$ and starting value $q_h^0 (= q_{h_0}^0)$ and set $q_{\text{old}}^0 = q_h^0$.
Choose a starting value $u_h^0 (= u_{h_0}^0)$ (e.g., by solving the nonlinear PDE $A(q_{\text{old}}^0, u_{\text{old}}^0) = f$ in $W_{h_0}^*$) and set $u_{\text{old}}^0 = u_h^0$.
Compute the adjoint state $z_h^0 (= z_{h_0}^0)$ (see (4.98)), evaluate $\|z_h^0\|_{W_h}$, set $\varrho_0 = \|z_h^0\|_{W_h}$. and evaluate $I_{3,h}^0$ (cf. (4.91)).
Set $k = 0$ and $h = h_0^1 = h_0$.
while $I_{3,h}^k > \tau^2 \delta^2$ or $k = 0$ **do**
  Set $h = h_k^1$.
  Solve the optimization problem (4.55).
  Set $h_k^2 = h_k^1$ and $\delta_\beta = \sqrt{\tilde{\theta} I_{3,h}^k}$.
  while $I_{2,h}^k > \left(\tau_\beta^2 + \frac{\tilde{\tau}_\beta^2}{2} \right) \delta_\beta^2$ **do**
    With q_{old}^k, u_{old}^k fixed, apply Algorithm 4.2 (with quantity of interest I_2^k and noise level δ_β) starting with the current mesh $h (= h_k^1)$ to obtain a regularization parameter β_k and a possibly different discretization h_k^2 such that (4.18) holds. Therewith, also the corresponding $v_h^k = v_{h_k^2}^k$, $q_h^k = q_{h_k^2}^k$ according to (4.55) are computed.
  Set $h = h_k^2$.
  Evaluate the error estimator η_1^k (cf. (4.89), (4.91)).
  Set $h_k^3 = h_k^2$.
  if (4.63) with $c_S = 2$ and $\eta_3^k = 0$ is violated **then**
    Refine grid with respect to η_1^k such that we obtain a finer discretization h_k^3.
    Solve the optimization problem (4.55) and evaluate η_1^k.
  else
    Set $q_{\text{old}}^{k+1} = q_h^k$, $u_{\text{old}}^{k+1} = u_{\text{old}}^k + v_h^k$.
  Set $h = h_k^3$.
  Compute the adjoint state $z_h^{k+1} (= z_{h_k^3}^{k+1})$ (see (4.98)), evaluate $\|z_h^{k+1}\|_{W_h}$, set $\varrho_k = \max\{\varrho_{k-1}\,, \|z_h^{k+1}\|_{W_h}\}$ and evaluate $I_{3,h}^{k+1}$.
  Set $h_{k+1}^1 = h_k^3$ (i.e., use the current mesh as a starting mesh for the next iteration).
  Set $k = k + 1$.

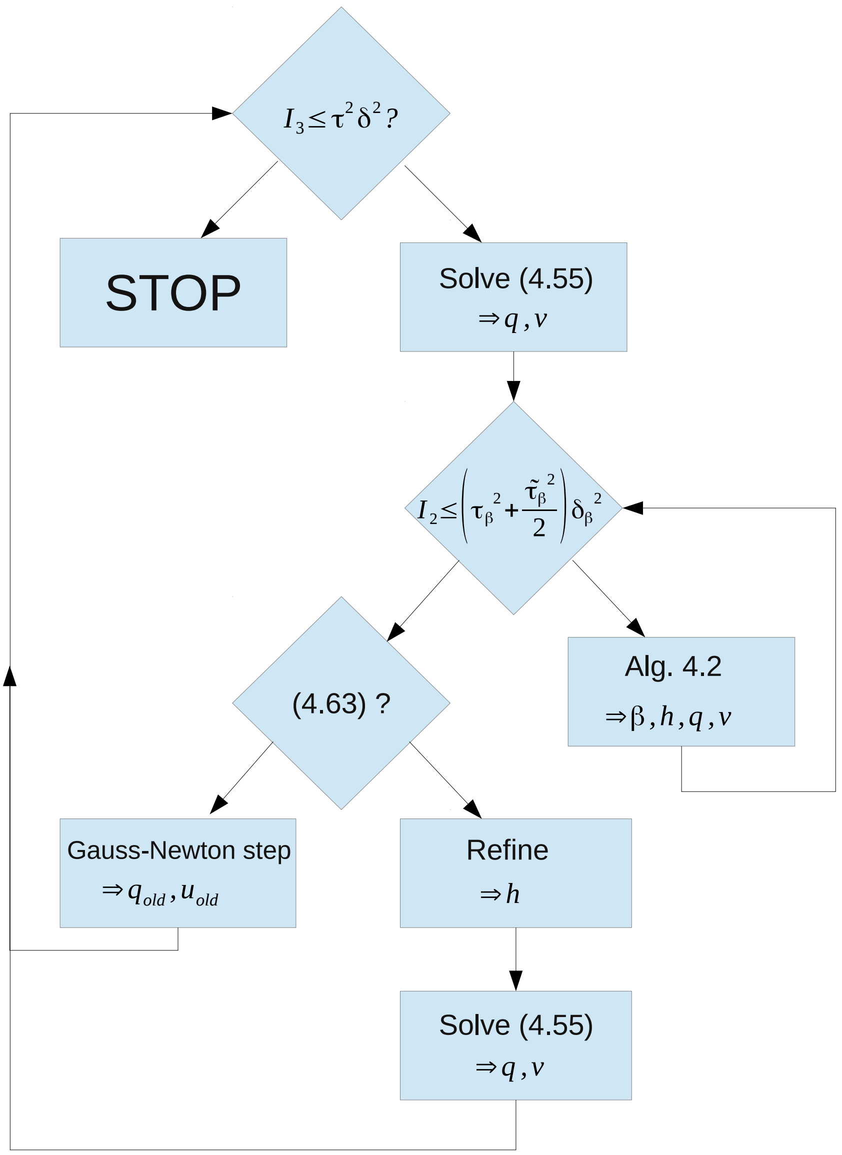

Figure 4.4.: Illustration of Algorithm 4.6. The arrow within the blocks point to variables that are computed in this step.

4.3. Numerical results

As mentioned and justified in Section 4.1.3, we will not implement the reduced form of IRGNM from Section 4.1 (cf. Algorithm 4.1), but concentrate on the all-at-once formulation from Section 4.2. We already discussed the implementational aspect in the subsection “Algorithm” in Section 4.2 and we presented a possible implementation in Algorithm 4.6.

We consider again the test examples from Section 3.4 with the same sources (a), (b), and (c) and observation operators (i), (ii) as in Section 3.6.

The concrete choice of the parameters for the numerical tests is as follows: $c_{tc} = 10^{-7}$, $\tilde{\tilde{\theta}} = 0.4999$, $\underset{\sim}{\theta} = 0.2$, $\tau = 5$, $\tau_\beta = 1.66$, $\tilde{\tau}_\beta = 1$, ($c_2 = 0.9999$, $c_3 = 0.0001$). The coarsest (starting mesh) consists of 25 nodes and 16 equally sized squares, the inital values for the control and the state are $q^0 = 0$ and $u^0 = 0$ and we start with a regularization parameter $\beta = 10$.

Considering the numerical tests, we are mainly interested in saving computation time compared to Algorithm 3.1 (NT) from Chapter 3, where the inexact Newton method for the determination of the regularization parameter β is applied directly to the nonlinear problem, instead of the linearized subproblems (4.55). The choice of the parameters from Section 3.6 implies that both algorithms (NT) and the Generalized Gauss-Newton Algorithm 4.6 (GGN) are stopped, if the concerning quantities of interest fall below the same bound ($\bar{\bar{\tau}}^2\delta^2$ for (NT) and $\tau^2\delta^2$ for (GGN)).

4.3.1. Example 1

First we consider Example 3.1 from Section 3.4, for which we have already seen the numerical results using Algorithm 3.1 (NT):

For $q \in L^2(\Omega)$ find $u \in H_0^1(\Omega)$ such that

$$\begin{cases} -\Delta u + \zeta u^3 = q & \text{in } \Omega \\ \qquad\qquad u = 0 & \text{on } \partial\Omega \,. \end{cases}$$

For the same configuration as in Section 3.6.1, i.e., source (a), point measurements (i), $\zeta = 1$ and 1% noise, we present the reconstructions obtained by Algorithm 4.6 (GGN). Despite the linearizations, the algorithm detects the location of the source very well and refines the mesh accordingly.

Taking a look at Figure 4.6 the reader can track the behavior of Algorithm 4.6 (GGN). The algorithm goes from right to left in Figure 4.6, where the quantities of interest I_2 and I_3 (or rather their discrete counterparts $I_{3,h}$ and $I_{2,h}$) are rather large. The noise level for the inner iteration $\tilde{\theta} I_{3,h}$ is about 0.68 in the beginning. For this noise level the stopping criterion for the β-algorithm (step 10 and 11 in Algorithm 4.6) is already fulfilled, such that only one Gauss-Newton step is made without refining or updating β. This decreases the noise level $\tilde{\theta} I_{3,h}$ to about 0.41. Then the β-algorithm comes into play, with one refinement step and three β-steps, which in total reduces $I_{2,h}$ from 1.340 to 0.48, with which the β-algorithm

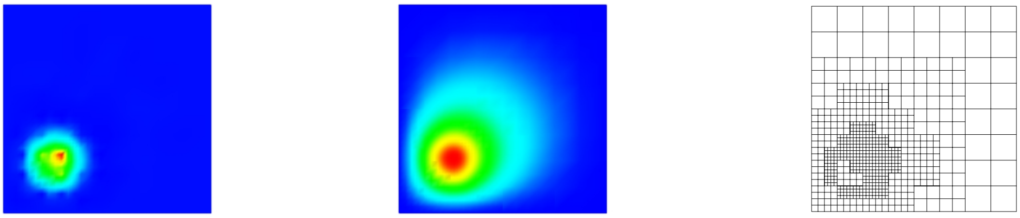

Figure 4.5.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 1 (a)(i), $\zeta = 1$, 1% noise using (GGN)

terminates. The subsequent run of the β-algorithm (with a smaller noise level of 0.25) consists of one refinement step and two β enlargement steps and so on. Finally, after 7 Gauss-Newton iterations, both quantites of interest $I_{2,h}$ and $I_{3,h}$ fulfill the required smallness conditions such that the whole Gauss-Newton Algorithm terminates. (The fact that it looks like the line for I_2 touches the one for I_3 is just a coincidence, see the corresponding graph for another test configuration Figure 4.7.)

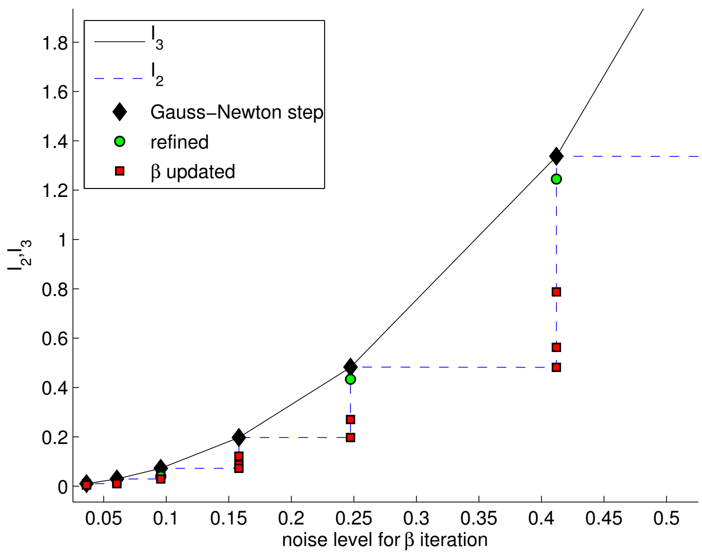

Figure 4.6.: GGN algorithm for Example 1 (a)(i) with $\zeta = 1$ and 1% noise using (GGN)

In Table 4.1 we present the respective results (i.e., Example 1 (a) (i), 1% noise) for different choices of ζ (first column). In order to better be able to compare the results from Algorithm 3.1 with the ones from Algorithm 4.6, we show here once more the values from Table 3.2. In the second and fifth column in Table 4.1 one can see the relative control error (3.92), in the third and sixth column is the number of nodes in the adaptively refined mesh, and in the forth and seventh column one can see the regularization parameter obtained by (GGN) and

(NT) respectively. The eighth column shows the gain with respect to computation time using (GGN) instead of (NT). We can see that we have the same situation as in Section 3.6: The reduction of computation time is about the same for all ζ, where the computation times in general are higher (thus more representative) for larger ζ. The same holds for the number of nodes and the relative control error. Only the obtained regularization parameter is larger, the larger ζ is. At the same time, it yields a smaller control error, which at first glance seems to be somewhat surprising, since (GGN) only works with the linearized state equation. One possible explanation could be the less accurate evaluation of the norms in (3.92), since we evaluate these norms on the final mesh of the particular algorithm, i.e., for (GGN) we evaluated on a much coarser mesh than for (NT). But on the other hand, a comparison between the reconstructions from Figure 4.5 and Figure 3.4 confirms the better results by (GGN). So it seems that (GGN) finds a better regularization parameter, which causes the better reconstruction.

Table 4.1.: Algorithm 4.6 (GGN) versus Algorithm 3.1 (NT) for (a)(i) for different choices of ζ with 1% noise. CTR: Computation time reduction using (GGN) instead of (NT)

ζ	GN			NT			CTR
	error	β	# nodes	error	β	# nodes	
1	0.395	4899	781	0.456	3085	38665	98%
10	0.395	5155	769	0.452	3194	38377	99%
100	0.393	7579	761	0.443	4999	31967	95%
1000	0.421	17445	677	0.475	11463	44413	99%

Due to the mentioned larger and consequently more significant computation times, we only consider $\zeta = 1000$ or $\zeta = 100$ in the following.

To see that the algorithm behaves similarly when considering an L^2 tracking type functional, i.e., configuration (ii) (see the beginning of Section 3.6), we invite the reader to take a look at Figure 4.7, where the Gauss-Newton steps, the refinement steps, and the β updating steps are presented for Example 1 (a) (ii) with $\zeta = 1000$ and 1% noise.

Algorithm 4.2 starts with noise level $\tilde{\theta} I_{3,h} = 2.8 \cdot 10^{-2}$, decides to increase β three times, until $I_{2,h}$ is small enough (decreased from $6.2 \cdot 10^{-3}$ to $2.9 \cdot 10^{-3}$). There follows a Gauss-Newton step, which slightly increases $I_{2,h}$, but at this point more importantly decreases $I_{3,h}$. The next run of Algorithm 4.2 consists of a refinement step and three β steps. After six Gauss-Newton iterations, $I_{2,h}$ is decreased to $1.2 \cdot 10^{-5}$ and $I_{3,h}$ to $3.1 \cdot 10^{-5}$ and Algorithm 4.6 terminates.

In Figure 4.8 we present the reconstruction and the adaptively refined mesh produced by Algorithm 4.6 (GGN). The mesh is much coarser than with Algorithm 3.1 (NT) (only 137 instead of 1305 nodes), which causes a saving of computation time of 99%. The regularization parameter is slightly smaller and the relative control error (3.92) is a bit larger, see Table 4.2.

The results for a different exact source distribution (b) can be found in Figure 4.9 for point measurements (i) and in Figure 4.10 for the L^2-projection alternative (ii).

In both cases (GGN) produces much coarser meshes than (NT), especially for (i), which leads to a remarkable reduction of computation time, which can be read out of Table 4.2. The

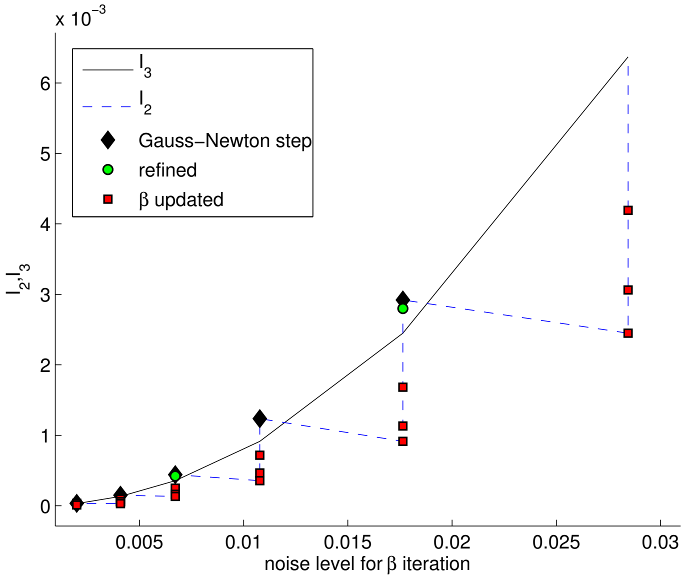

Figure 4.7.: GGN algorithm for Example 1 (a)(ii) with $\zeta = 1000$ and 1% noise using (GGN)

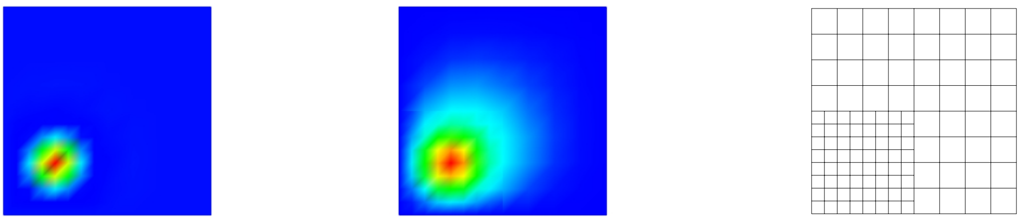

Figure 4.8.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 1 (a)(ii), $\zeta = 1000$, 1% noise using (GGN)

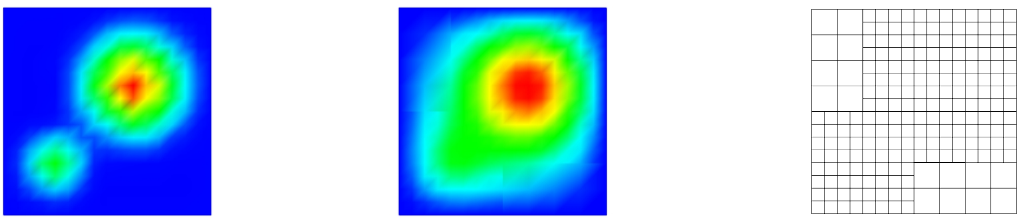

Figure 4.9.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 1 (b)(i), $\zeta = 1000$, 1% noise using (GGN)

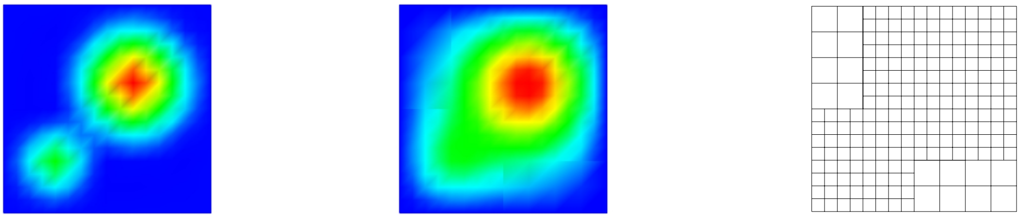

Figure 4.10.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 1 (b)(ii), $\zeta = 1000$, 1% noise using (GGN)

obtained regularization parameter is larger than for (NT), which explains the smaller control error.

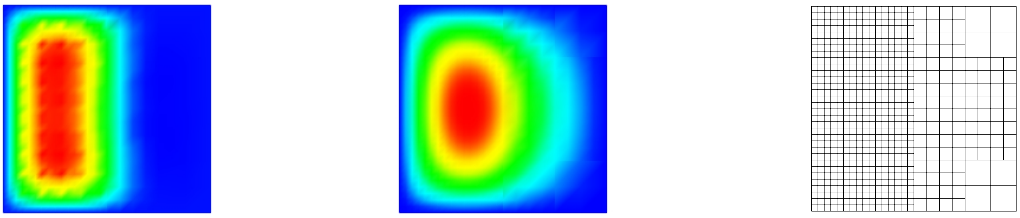

Figure 4.11.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 1 (c)(i), $\zeta = 1000$, 1% noise using (GGN)

The same observations can be made for the reconstruction of the source (c), see Figure 4.11 and Figure 4.12.

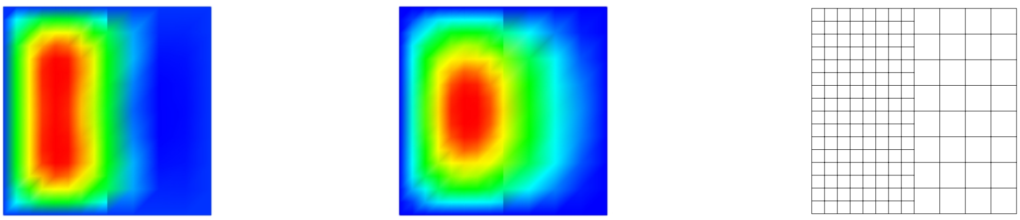

Figure 4.12.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 1 (c)(ii), $\zeta = 1000$, 1% noise using (GGN)

The key results of the previous test configuration for Example 1 for (GGN) as well as (NT) are collected in Table 4.2.

We come back to the example source (a) and examine the results for different noise levels.

Table 4.2.: Algorithm 4.6 (GGN) versus Algorithm 3.1 (NT) for Example 1 with 1% noise. CTR: Computation time reduction using (GGN) instead of (NT)

Example	GN			NT			CTR
	error	β	# nodes	error	β	# nodes	
(a) (i)	0.421	17445	677	0.475	11463	44413	99%
(a)(ii)	0.318	3449334	137	0.302	4368044	1305	53%
(b) (i)	0.225	1764	241	0.300	947	20231	98%
(b)(ii)	0.247	243290	241	0.276	151048	761	33%
(c) (i)	0.431	396	673	0.450	261	11785	91%
(c)(ii)	0.467	36831	189	0.461	42541	651	28%

In Figure 4.13 and Figure 4.14 we display the reconstructions of the control and the state obtained by (GGN) for configuration (i) with $\zeta = 100$ for the noise levels 1%, 2%, and 4%.

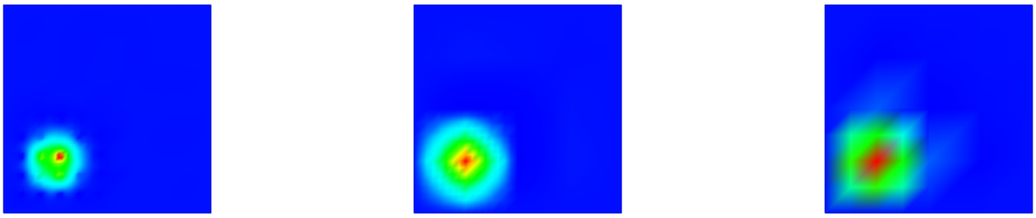

Figure 4.13.: Reconstructed control for Example 1 (a)(i), $\zeta = 100$ for different noise levels. FLTR: 1%, 2%, 4% noise using (GGN)

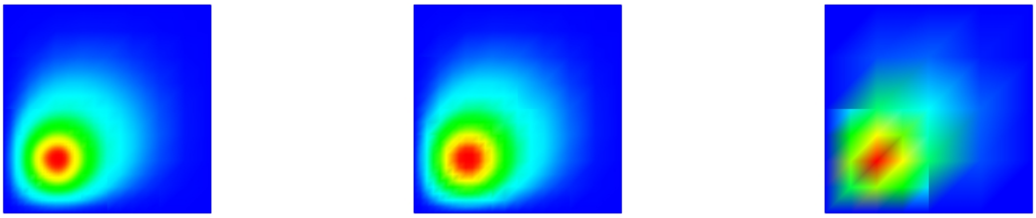

Figure 4.14.: Reconstructed state for Example 1 (a)(i), $\zeta = 100$ for different noise levels. FLTR: 1%, 2%, 4% noise using (GGN)

The corresponding refined meshes can be found in Figure 4.15.

In Table 4.3 we list the relative control error (3.92), the computed final regularization parameter β, and the number of nodes in the obtained mesh for $p = 0.5\%$, $p = 1\%$, $p = 2\%$, $p = 4\%$, and $p = 8\%$ and come to the same conclusion as for Algorithm 3.1 (NT) in Section 3.6: The more noise, the worse the reconstruction, the stronger the regularization, the coarser the mesh.

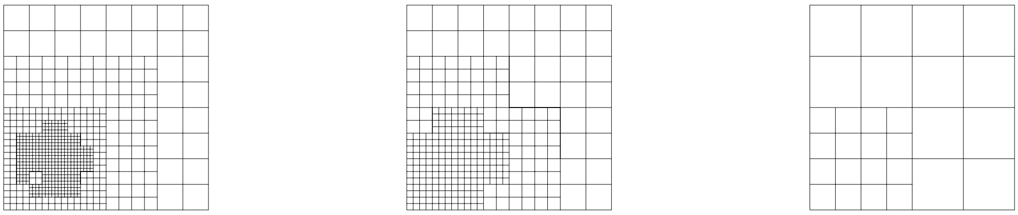

Figure 4.15.: Adaptively refined mesh for Example 1 (a)(i), $\zeta = 100$ for different noise levels. FLTR: 1%, 2%, 4% noise using (GGN)

Table 4.3.: Different noise levels for Example 1 (a)(i), $\zeta = 100$ using (GGN)

noise	error	β	# nodes
0.5%	0.360	13689	761
1%	0.393	7579	761
2%	0.549	1748	361
4%	0.770	817	41
8%	1.03	53	41

4.3.2. Example 2

In the following we consider Example 3.2 (also see Section 3.6.2):

For $q \in L^2(\Omega)$, $q \geq 0$ almost everywhere in Ω find $u \in H_0^1(\Omega)$ such that

$$\begin{cases} -\Delta u + qu = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \,. \end{cases}$$

Like in Section 3.6.2, we choose the initial guess as well as the starting value $q_0 = q^0 = 0 \in L^2(\Omega)$ for q and neglect the constraint $q \geq 0$ in our computations.

In Section 3.6.2 we have seen the results of Algorithm 3.1 (NT). Here, we will present how the Generalized Gauss-Newton Algorithm 4.6 (GGN) performs on this particular example.

Figure 4.16 shows the reconstructions of the control and the state as well as the adaptively refined mesh produced by (GGN) for the exact source distribution (a) (see the beginning of Section 3.6) and 1% noise.

For the source distribution (b) the corresponding reconstructions and the obtained mesh are displayed in Figure 4.17. The key results for both configurations are collected in Table 4.4, where we present the results of (NT) from Table 3.11 once again for comparison.

For both, (a) and (b), (GGN) is faster than (NT), mainly because it refines a lot less. The control errors are within the same range. For (a) (NT) yields a better reconstruction – probably due to a larger regularization parameter – while for (b) it is the other way around. For (a) the error ist strikingly large for both algorithms. In Section 3.6.2 we already tried to give an explanation for this behavior of (NT), namely the negligence of the control error. This

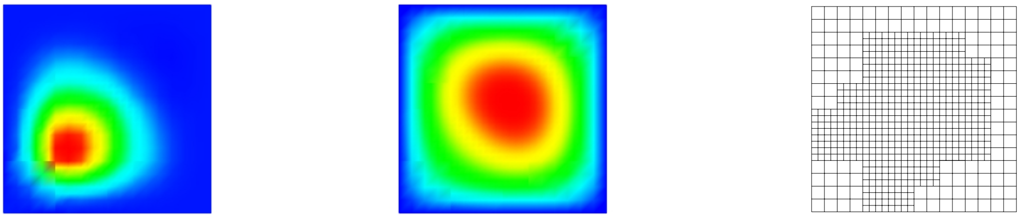

Figure 4.16.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 2 (a)(i), 1% noise using (GGN)

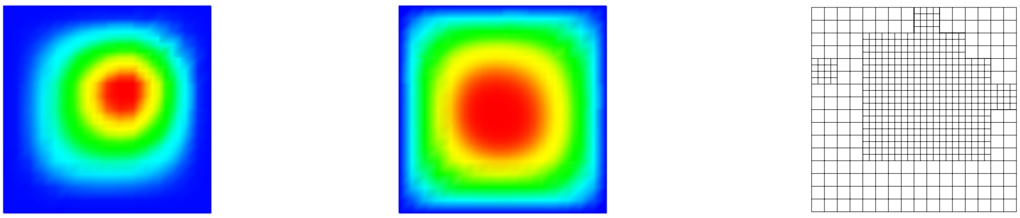

Figure 4.17.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 2 (b)(i), 1% noise using (GGN)

Table 4.4.: Algorithm 4.6 (GGN) versus Algorithm 3.1 (NT) for Example 2 with 1% noise. CTR: Computation time reduction using (GGN) instead of (NT)

Example	GN			NT			CTR
	error	β	# nodes	error	β	# nodes	
(a)	0.946	30168	725	0.930	43504	2033	10%
(b)	0.590	11732	641	0.669	6027	7663	35%

argumentation, however, cannot hold for (GGN), since there, the control enters the right-hand side linearly:

$$\begin{cases} -\Delta u + q_{\text{old}} u = f + q_{\text{old}} u_{\text{old}} - q u_{\text{old}} & \text{in } \Omega \\ \qquad\qquad u = 0 & \text{on } \partial\Omega \end{cases},$$

such that the error estimator for the control discretization vanishes (see the explanation in Section 3.6.2). So the large error seems actually to be down to the bilinearity of the PDE (also see Section 3.4). In order to still obtain a better reconstruction one could for instance reduce the size of τ ($\overline{\overline{\tau}}$ for (NT)).

4.3.3. Example 3

Finally, we consider Example 3.3 and Example 3.4, where (as in Section 3.6) we replaced q by $(1 + q)$:

For $q \in L^2(\Omega)$, $q \geq 0$ almost everywhere in Ω find $u \in H^1(\Omega)$ such that

$$-\Delta u - (1 + q)u = f \quad \text{in } \Omega$$

under Dirichlet zero boundary conditions

$$u = 0 \quad \text{on } \partial\Omega\,,$$

or Neumann zero boundary conditions

$$\partial_n u = 0 \quad \text{on } \partial\Omega\,.$$

We consider two possible right-hand sides:

$$\begin{aligned} f_1 &= -1 \\ f_2 &= 4\pi^2 \left[\sin(2\pi x - \tfrac{\pi}{2})(\sin(2\pi y - \tfrac{\pi}{2}) + 1) + \sin(2\pi y - \tfrac{\pi}{2})(\sin(2\pi x - \tfrac{\pi}{2}) + 1)\right] \\ &\quad - (\sin(2\pi x - \tfrac{\pi}{2}) + 1)(\sin(2\pi y - \tfrac{\pi}{2}) + 1)\,. \end{aligned}$$

Like for the previous example (Example 2), we choose the initial guess and the starting value $q_0 = q^0 = 0 \in L^2(\Omega)$ for q and neglect the constraint $q \geq 0$ for simplicity.

We restrict ourselves to the test configuration (a), where, in order to compare the results to the ones from Section 3.6 (NT), we choose $c = 1$ and 0.01% of noise.

The reconstructions of the control and the state as well as the adaptively refined mesh produced by (GGN) for the right-hand side $f = f_1$ and Dirichlet zero boundary conditions can be seen in Figure 4.18.

The corresponding results for $f = f_2$ are shown in Figure 4.19 and the key results for both right-hand sides are collected in Table 4.5.

The observations from the previous examples are confirmed once again. (GGN) refines a lot less than (NT), where the resulting control error and the regularization parameter only differ slightly. Overall (GGN) is again much faster than (NT).

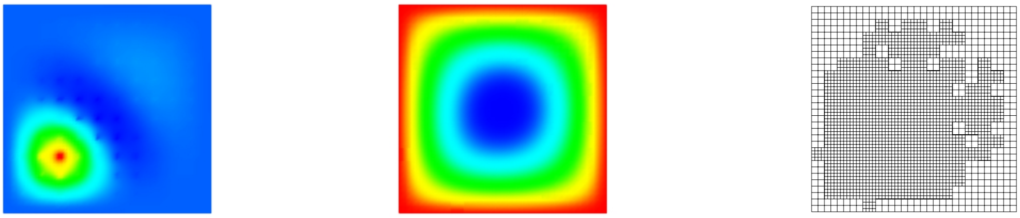

Figure 4.18.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 3 (a)(i), 1% noise, $f = f_1$, Dirichlet zero boundary conditions using (GGN)

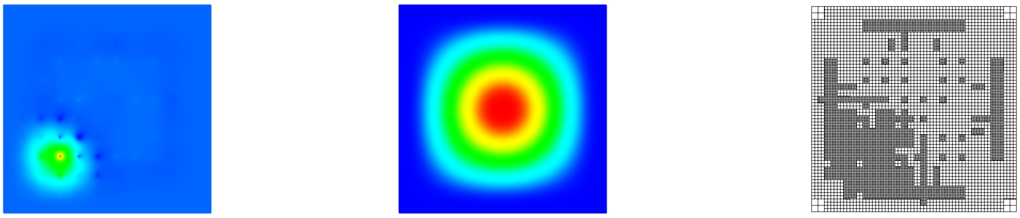

Figure 4.19.: FLTR: reconstructed control, reconstructed state, adaptively refined mesh for Example 3 (a)(i), 1% noise, $f = f_2$, Dirichlet zero boundary conditions using (GGN)

Table 4.5.: Algorithm 4.6 (GGN) versus Algorithm 3.1 (NT) for Example 3 (a)(i) with 1% noise. CTR: Computation time reduction using (GGN) instead of (NT)

right-hand side	NT			GN			CTR
	error	β	# nodes	error	β	# nodes	
f_1	0.649	682138	3001	0.655	715479	6789	45%
f_2	0.419	2859	8397	0.415	3008	65623	82%

To summarize the results from Section 4.3, the proposed Generalized Gauss-Newton method (GGN) seems to avoid overrefinement, which leads to a remarkable reduction of computation time with respect to (NT). At the same time it usually finds a larger regularization parameter, which yields good reconstructions despite the coarse mesh. Unfortunately the replacement of the nonlinear state equation by a linear one as such was not the crucial factor for the shorter computation time, but rather the fact that this enabled the use of Algorithm 4.2, which exploits the advantages of a linear-quadratic optimal control problem.

For this example, we also tested another way of replacing the nonlinear PDE in (NT) by a linearized one. Since this, however, does not fit into the framework of this chapter, we refer the reader to Appendix D, where we consider the Born approximation for the Helmholtz equation (cf., e.g., [15]).

5. Conclusion and perspectives

In this thesis, we considered strategies for the efficient solution of nonlinear inverse problems combining the computation of a regularization parameter with adaptive discretization methods.

In the first part (Chapter 3), we presented an extension of the adaptive discretization strategy for parameter identification from [39] to the case of nonlinear ill-posed problems. For this purpose we used the discrepancy principle and applied an inexact Newton method to the resulting quantity of interest. By imposing conditions on the quantity of interest and the reduced forward operator F, we provided a general rates result based on Jensen's inequality, which is not restricted to the context of this thesis. In order to guarantee convergence of the resulting sequence of regularization parameters, we derived refinement criteria, which we combined with goal-oriented error estimators [10, 11]. Main difference of our approach to [39] consisted in compensating the lack of convexity of the cost funtional and the lack of knowledge about lower and upper bounds on the first, second and third derivative of the quantity of interest with respect to the regularization parameter. We considered quadratic convergence (as in [39]), but also linear convergence conditions, where we extended the idea from [98] regarding bounds on the second derivative. In addition to the original stationary formulation, we also considered parabolic problems in Section 3.7, where the proposed algorithm decides whether to refine in space or in time. The strength of the presented method lies in the ability to save a considerable amount of degrees of freedom in the beginning of the Newton type iteration for determining the regularization parameter. Indeed, our numerical experiments showed a remarkable gain in computation time with our strategy as compared to uniform refinement.

In the second part (Chapter 4), we combined this idea with iteratively regularized Gauss-Newton methods; first, in a reduced form using the parameter-to-solution map (see Section 4.1), and second in an all-at-once formulation (see Section 4.2), where the measurement equation and the PDE are treated simultaneously. We considered two types of all-at-once approaches: a least squares form and a Generalized Gauss-Newton formulation. In both cases, i.e., for reduced and all-at-once form, we showed convergence and proved convergence rates which we carried over from the continuous to the discretized setting by controlling precision only in four real valued quantities per Newton step. The choice of the regularization parameter in each Newton step and of the overall stopping index have been done a posteriori, via a discrepancy type principle. We only provided numerical experiments for the all-at-once formulation, more specifically for the Generalized Gauss-Newton approach. The reason for this is that this formulation allows us to consider only the linearized PDE as a constraint in each Newton step, which seemed to be the more promising alternative to save computational effort. From the numerical tests we have seen that the presented method yields reasonable reconstructions and can even lead to a large reduction of computation time compared to the non-iterative method from Chapter 3.

An idea to possibly even further reduce the computational effort lies in the combination with the Hierarchical Trust region method proposed in [72]. Instead of using Newton's method for the solution of the optimal control problem (in each step of the inexact Newton method for the computation of the regularization parameter), we could use this modification of the usual trust region method (cf., e.g., [26, 95]), where second order information is computed on a coarser mesh than all other computations in order to save computational costs. The choice of this mesh (in terms of resolution) and the trust region radius are controlled simultaneously, which finally leads to a respectable reduction of computation time (see [72]) compared to a usual trust region method.

We expect that the adaptive strategy using goal oriented error estimators can also be used in the context of different parameter choice rules such as the balancing principle [80], which is of particular interest in case $\nu > \frac{1}{2}$ in the Hölder type source condition (cf. Assumption 3.5, Remark 3.5), where the discrepancy principle fails to provide optimal rates. Moreover, we hope that the use of the balancing principle with a sufficiently strong source condition will allow us to avoid conditions on F like the tangential cone condition (see Assumption 3.3), such that we can also take strongly nonlinear problems into consideration. Another (discrepancy type) choice rule is presented in [114], which enables the use of Poisson instead of Gaussian noise.

In our considerations for the parabolic setting, we assumed that the control is constant in time and that the spatial discretization is fixed for all time intervals. However, the proposed method could be extended to the case of time-dependent control and dynamic meshes, i.e., meshes that change in each time step. A corresponding formulation for the discretization of the optimal control problem including adaptivity is given in [104].

Another interesting question is if and how the derived results can be extended to the case of control bounds. In this context, approaches combining the choice of a regularization parameter with adaptive discretizations can be found in [110, 111].

Moreover, a possible combination of adaptivity with regularization by projection onto finite dimensional spaces in place of, or in addition to, Tikhonov regularization could be interesting. Instead of the regularization parameter, we would have to choose the discretization level. Using dual-weighted-residual error estimators, this could be done in an efficient manner during a multi-level method, where the discrepancy principle serves as a stopping criterion. We distinguish two approaches: discretization in the image space and discretization in the preimage space. The former, however, seems to require a large number of quantities of interest, which may diminish the efficiency of the adaptive approach. Theoretical results concerning regularization by discretization in the preimage space without adaptivity are provided in [70].

Acknowledgments

I gratefully acknowledge the funding received by the German Research Foundation ("Deutsche Forschungsgesellschaft", DFG) within the frame of the project "Adaptive Discretization methods for the Regularization of Inverse Problems" (KA 1778/5-1 and VE 368/2-1) for most of my time at the Technische Universität München. I also thank the Federal Ministry of Education and Research ("Bundesministerium für Bildung und Forschung", BMBF) for their financial support for two months. My research stay at the University of Texas at Austin was financially supported by the entrepreneurial program of the TUM program "Women for Math Science" and by the TUM Graduate School, for which I am very grateful.

Above all, I would like to thank my supervisor Prof. Dr. Boris Vexler and my mentor and reviewer Prof. Dr. Barbara Kaltenbacher for proposing the interesting DFG project and subject of my thesis, and especially for their help, advice, patience as well as their friendly ear and their empathy. I am also very grateful that Prof. Dr. Boris Vexler gave me the opportunity of attending international conferences to broaden my horizon on a scientific level and for letting me concentrate on my research as well as enable me to gain more experience in teaching. I especially appreciate the close collaboration with Prof. Dr. Barbara Kaltenbacher, since this can not be taken for granted living in two different countries.

Further, I thank Prof. Dr. Arnd Rösch for reviewing this thesis and for the interesting discussions we had on conferences/workshops.

Also thanks to Prof. Dr. George Biros for inviting me for a research visit at the University of Texas at Austin and for giving me an insight in his work.

All the numerical experiments in this thesis have been done with the optimization toolkit RoDoBo [101] and the finite element library Gascoigne [35]. So a special credit is due to all the people who have contributed to the development of these software packages.

Next, I would like to thank my colleagues at the university for the enjoyable time we spent together inside and outside the workplace, for the exchange of ideas, and for their open doors when it comes to organizational, as well as scientific questions. Especially I thank Martin, Moritz, Konstantin, and Andreas, the colleagues I shared my office with, for treating all my questions with patience and for the nice atmosphere they created. Thanks to Dominik, Konstantin, and Andreas for helping me with implementational and technical matters and thanks to Moritz, Florian K., Florian L., Ira, Dominik, Christian, and Andreas for proof-reading parts of the manuscript.

Last but not least I thank my family and friends for the support they gave me over the years in every kind of way and for always believing in me.

A. State, dual, and tangent equations

"State": Find $u \in V$ such that

$$\mathcal{L}'_z(q,u,z)(\varphi) = 0$$

for all $\varphi \in W$.

"Dual": Find $z \in W$ such that

$$\mathcal{L}'_u(q,u,z)(\varphi) = 0$$

for all $\varphi \in V$.

"Gradient": Find $q \in Q$ such that

$$\mathcal{L}'_q(q,u,z)(\varphi) = 0$$

for all $\varphi \in Q$.

"Tangent": Find $\delta u \in V$ such that

$$\mathcal{L}''_{uz}(q,u,z)(\delta u,\varphi) = -\mathcal{L}''_{qz}(q,u,z)(\delta q,\varphi)$$

for all $\varphi \in W$.

"DualForHessian": Find $\delta z \in W$ such that

$$\mathcal{L}''_{zu}(q,u,z)(\delta z,\varphi) = -\mathcal{L}''_{qu}(q,u,z)(\delta q,\varphi) - \mathcal{L}''_{uu}(q,u,z)(\delta u,\varphi)$$

for all $\varphi \in V$.

"DualQI": Find $z_1^{(0)} \in W$ such that

$$\mathcal{L}''_{zu}(x)(z_1^{(0)},\varphi) = -I'_u(u)(\varphi)$$

for all $\varphi \in V$.

"QIEq": Find $q_1 \in Q$ such that

$$j''(q)(q_1,\varphi) = -\mathcal{L}''_{zq}(x)(z_1^{(0)},\varphi)$$

for all $\varphi \in Q$.

"K1Eq": Find $q_2 \in Q$ such that

$$j''(q)(q_2,\varphi) = -K'_{q_1}(q,q_1)(\varphi)$$

for all $\varphi \in Q$.

“TangentK2For3rdDerivative”: Find $u_3^{(0)} \in V$ such that

$$\begin{aligned}\mathcal{L}''_{uq}(x)(u_3^{(0)}, \varphi) = &-\mathcal{L}'''_{quz}(x)(q_1, u_2, \varphi) - \mathcal{L}'''_{quz}(x)(q_2, u_1, \varphi) \\ &- \mathcal{L}'''_{uuz}(x)(u_1, u_2, \varphi) + \mathcal{L}'''_{qqz}(x)(q_1, q_2, \varphi)\end{aligned}$$

for all $\varphi \in W$.

“DualK2”: Find $z_3^{(0)} \in W$ such that

$$\begin{aligned}\mathcal{L}''_{zu}(x)(z_3^{(0)}, \varphi) = &-I''_{uu}(x)(u_2, \varphi) - \mathcal{L}''_{uu}(x)(u_3^{(0)}, \varphi) - \mathcal{L}''_{quu}(x)(q_1, u_2, \varphi) \\ &- \mathcal{L}'''_{qzu}(x)(q_1, z_2, \varphi) - \mathcal{L}'''_{quu}(x)(q_2, u_1, \varphi) \\ &- \mathcal{L}'''_{qzu}(x)(q_2, z_1, \varphi) - \mathcal{L}'''_{zuu}(x)(z_1, u_2, \varphi) \\ &- \mathcal{L}'''_{uuu}(x)(u_1, u_2, \varphi) - \mathcal{L}'''_{zuu}(x)(z_2, u_1, \varphi) \\ &- \mathcal{L}'''_{qqu}(x)(q_1, q_2, \varphi)\end{aligned}$$

for all $\varphi \in V$.

“K2Eq”: Find $q_3 \in Q$ such that

$$\begin{aligned}j''(q)(q_3, \varphi) = &-\mathcal{L}''_{uq}(x)(u_3^{(0)}, \varphi) - \mathcal{L}''_{zq}(x)(z_3^{(0)}, \varphi) - K'_q(q, q_1)(\varphi) \\ &- \mathcal{L}'''_{uuq}(x)(u_1, u_2, \varphi) - \mathcal{L}'''_{uzq}(x)(u_1, z_2, \varphi) \\ &- \mathcal{L}'''_{uzq}(x)(u_1, z_1, \varphi) - \mathcal{L}'''_{qqq}(x)(q_1, q_2, \varphi) \\ &- \mathcal{L}'''_{qzq}(x)(q_1, z_2, \varphi) - \mathcal{L}'''_{qzq}(q_2, z_1, \varphi) \\ &- \mathcal{L}'''_{quq}(x)(q_1, u_2, \varphi) - \mathcal{L}'''_{quq}(q_2, u_1, \varphi)\end{aligned}$$

for all $\varphi \in Q$.

B. Dual and tangent equations for the parabolic setting

"Adjoint" (2.26): Find $z \in \tilde{V}$ such that

$$-(\partial_t z, \varphi)_I + \tilde{a}'_u(q,u)(z,\varphi) + (z(T), \varphi(T)) = \int_I J'_1(u(t))(\varphi(t))\,dt + J'_2(u(T))(\varphi(T)) \tag{B.1}$$

for all $\varphi \in \tilde{V}$.

"Tangent" (2.29): Find $\delta u \in \tilde{V}$ such that

$$(\partial_t \delta u, \varphi_I) + \tilde{a}'_u(q,u)(\delta u, \varphi) + (\delta u(0), \varphi(0)) = -\tilde{a}'_q(q,u)(\delta q, \varphi) + (u'_0(q)(\delta q), \varphi(0))$$

for all $\varphi \in \tilde{V}$.

"Additional Adjoint" (2.30): Find $\delta z \in \tilde{V}$ such that

$$\begin{aligned}-(\partial_t \delta z, \varphi)_I + \tilde{a}'_u(q,u)(\varphi, \delta z) + (\delta z(T), \varphi(T)) = &-\tilde{a}''_{uu}(q,u)(\delta u, \varphi, z) - \tilde{a}''_{qu}(q,u)(\delta q, \varphi, z)\\ &+ \int_I J''_1(u(t))(\delta u(t), \varphi(t))\,dt\\ &+ J''_2(u(T))(\delta u(T), \varphi(T))\end{aligned}$$

for all $\varphi \in \tilde{V}$.

"DualQI": Find $z_1^{(0)} \in \tilde{V}$ such that

$$-(\partial_t z_1^{(0)}, \varphi)_I + \tilde{a}'_u(q,u)(\varphi, z_1^{(0)}) + (z_1^{(0)}(0), \varphi(0)) = 2\int_I J'_1(u(t))(\varphi(t))\,dt + 2J'_2(u(T))(\varphi(T))$$

for all $\varphi \in \tilde{V}$.

"TangentK2For3rdDerivative": Find $u_3^{(0)} \in \tilde{V}$ such that

$$\begin{aligned}&(\partial_t u_3^{(0)}, \varphi)_I + \tilde{a}'_u(q,u)(u_3^{(0)}, \varphi) + (u_3^{(0)}(0), \varphi(0)) - (u''_0(q)(q_1, q_2), \varphi(0))\\ &\quad = -\tilde{a}''_{qu}(q,u)(q_1, u_2, \varphi) - \tilde{a}''_{qu}(q,u)(q_2, u_1, \varphi) - \tilde{a}''_{uu}(q,u)(u_1, u_2, \varphi) - \tilde{a}''_{qq}(q,u)(q_1, q_2, \varphi)\end{aligned}$$

for all $\varphi \in \tilde{V}$.

"DualK2": Find $z_3^{(0)} \in \tilde{V}$ such that

$$
\begin{aligned}
&-(\partial_t z_3^{(0)}, \varphi)_I + \tilde{\tilde{a}}'_u(q,u)(\varphi, z_3^{(0)}) + (z_3^{(0)}(T), \varphi(T)) \\
&= \int_I J_1''(u(t))(u_2(t), \varphi(t))\,dt + J_2''(u(T))(u_2(T), \varphi(T)) \\
&\quad + \int_I J_1''(u(t))(u_3(t), \varphi(t))\,dt + J_2''(u(T))(u_3(T), \varphi(T)) \\
&\quad - \tilde{a}''_{uu}(q,u)(u_3, \varphi, z) - \tilde{a}''_{qu}(q,u)(q_1, \varphi, z_2) - \tilde{a}''_{qu}(q,u)(q_2, \varphi, z_1) - \tilde{a}''_{uu}(q,u)(u_2, \varphi, z_1) \\
&\quad - \tilde{a}''_{uu}(q,u)(u_1, \varphi, z_2) - \tilde{a}'''_{quu}(q,u)(q_1, u_2, \varphi, z) - \tilde{a}'''_{quu}(q,u)(q_2, u_1, \varphi, z) \\
&\quad - \tilde{a}'''_{uuu}(q,u)(u_1, u_2, \varphi, z) - \tilde{a}'''_{qqu}(q,u)(q_1, q_2, \varphi, z) \\
&\quad + \int_I J_1'''(u(t))(u_1(t), u_2(t), \varphi(t))\,dt + J_2'''(u(T))(u_1(T), u_2(T), \varphi(T))
\end{aligned}
$$

for all $\varphi \in \tilde{V}$.

C. Implicit Euler scheme

Using the abbreviations

$$
\begin{aligned}
Q_m &\coloneqq q^-_{h,m}\,, & U_m &\coloneqq u^-_{h,m}\,, & Z_m &\coloneqq z^-_{h,m} \\
\Delta Q_m &\coloneqq \delta q^-_{h,m}\,, & \Delta U_m &\coloneqq \delta u^-_{h,m}\,, & \Delta Z_m &\coloneqq \delta z^-_{h,m} \\
Q^i_m &\coloneqq q^{(0),-}_{i,h,m}\,, & U^i_m &\coloneqq \delta u^{(0),-}_{i,h,m}\,, & Z^i_m &\coloneqq z^{(0),-}_{i,h,m} \qquad \text{for } i = 1,2,3
\end{aligned}
$$

the implicit Euler scheme for the State, Adjoint, Additional Adjoint, DualQI, TangentK2For3rdDerivative, and DualK2 equations read for all $\Psi \in W_h$:

State:

$$
\begin{aligned}
m = 0: &\qquad (U_0, \Psi) = (u_0(Q_0), \Psi) \\
m = 1, \dots, M: &\qquad (U_m, \Psi) + k_m \tilde{a}(Q_m, U_m)(\Psi) = (U_{m-1}, \Psi) + k_m (f(t_m), \Psi)
\end{aligned}
$$

Adjoint:

$$
\begin{aligned}
m = M: &\qquad (\Psi, Z_M) + k_M \tilde{a}'_u(Q_M, U_M)(\Psi, Z_M) = -k_M J'_1(U_M)(\Psi) + J'_2(U_M)(\Psi) \\
m = M-1, \dots, 1: &\qquad (\Psi, Z_m) + k_m \tilde{a}'_u(Q_m, U_m)(\Psi, Z_m) = (\Psi, Z_{m+1}) + k_m J'_1(U_m)(\Psi) \\
m = 0: &\qquad (\Psi, Z_0) = (\Psi, Z_1)
\end{aligned}
$$

Tangent:

$$
\begin{aligned}
m = 0: &\qquad (\Delta U_0, \Psi) = (u'_0(Q_0)(\Delta Q_0), \Psi) \\
m = 1, \dots, M: &\qquad (\Delta U_m, \Psi) + k_m \tilde{a}'_u(Q_m, U_m)(\Delta U_m, \Psi) = (\Delta U_{m-1}, \Psi) - k_m \tilde{a}'_q(Q_m, U_m)(\Delta Q_m, \Psi)
\end{aligned}
$$

Additional Adjoint:

$$
\begin{aligned}
m = M: \qquad & (\Psi, \Delta Z_M) + k_M \tilde{a}'_u(Q_M, U_M)(\Psi, \Delta Z_M) \\
&= -k_M \tilde{a}''_{uu}(Q_M, U_M)(\Delta U_M, \Psi, Z_M) - k_M \tilde{a}''_{qu}(Q_M, U_M)(\Delta Q_M, \Psi, \Delta Z_M) \\
&\quad + k_M J''_1(U_M)(\Delta U_M, \Psi) + J''_2(U_M)(\Delta U_M, \Psi)
\end{aligned}
$$

$$
\begin{aligned}
m = M-1, \dots, 1: \qquad & (\Psi, \Delta Z_m) + k_m \tilde{a}'_u(Q_m, U_m)(\Psi, \Delta Z_m) \\
&= (\Psi, \Delta Z_{m+1}) - k_m \tilde{a}''_{uu}(Q_m, U_m)(\Delta U_m, \Psi, \Delta Z_m) \\
&\quad - k_m \tilde{a}''_{qu}(Q_m, U_m)(\Delta Q_m, \Psi, Z_m) + k_m J''_1(U_m)(\Delta U_m)(\Psi)
\end{aligned}
$$

$$
m = 0: \qquad (\Psi, \Delta Z_0) = (\Psi, \Delta Z_1)
$$

(cf. [9]). Similarly we can formulate the implicit Euler scheme for

DualQI:

$$
\begin{aligned}
m = M: \quad & (\Psi, Z_M^1) + k_M \tilde{a}'_u(Q_M, U_M)(\Psi, Z_M^1) = 2k_M J'_1(U_M)(\Psi) + 2J'_2(U_M)(\Psi) \\
m = M-1,\dots,1: \quad & (\Psi, Z_m^1) + k_m \tilde{a}'_u(Q_m, U_m)(\Psi, Z_m^1) = (\Psi, Z_{m+1}^1) + 2k_m J'_1(U_m)(\Psi) \\
m = 0: \quad & (\Psi, Z_0^1) = (\Psi, Z_1^1)
\end{aligned}
$$

TangentK2For3rdDerivative:

$$
m = 0: \quad (U_0^3, \Psi) = (u_0''(Q_0)(Q_1, Q_2), \Psi)
$$

$$
\begin{aligned}
m = 1,\dots,M: \quad & (U_m^3, \Psi) + k_m \tilde{a}'_u(Q_m, U_m)(U_m^3, \Psi) \\
& = (U_{m-1}^3, \Psi) - k_m \tilde{a}''_{qu}(Q_m, U_m)(Q_m^1, U_m^2, \Psi) - k_m \tilde{a}''_{qu}(Q_m, U_m)(Q_m^2, U_m^1, \Psi) \\
& \quad - k_m \tilde{a}''_{uu}(Q_m, U_m)(U_m^1, U_m^2, \Psi) - k_m \tilde{a}''_{qq}(Q_m, U_m)(Q_m^1, Q_m^2, \Psi)
\end{aligned}
$$

DualK2:

$$
\begin{aligned}
m = M: \quad & (\Psi, Z_M^3) + k_M \tilde{a}'_u(Q_M, U_M)(\Psi, Z_M^3) \\
& = k_M J''_1(U_M)(U_M^3, \Psi) + J''_2(U_M)(U_M^3, \Psi) + k_M J''_1(U_M)(U_M^2, \Psi) + J''_2(U_M)(U_M^2, \Psi) \\
& \quad - k_M \tilde{a}''_{uu}(Q_M, U_M)(U_M^3, \Psi, Z_M) - k_M \tilde{a}''_{qu}(Q_M, U_M)(Q_M^1, \Psi, Z_M^2) \\
& \quad - k_M \tilde{a}''_{qu}(Q_M, U_M)(Q_M^2, \Psi, Z_M^1) - k_M \tilde{a}''_{uu}(Q_M, U_M)(U_M^2, \Psi, Z_M^1) \\
& \quad - k_M \tilde{a}''_{uu}(Q_M, U_M)(U_M^1, \Psi, Z_M^2) - k_M \tilde{a}'''_{quu}(Q_M, U_M)(Q_M^1, U_M^2, \Psi, Z_M) \\
& \quad - k_M \tilde{a}'''_{quu}(Q_M, U_M)(Q_M^2, U_M^1, \Psi, Z_M) - k_M \tilde{a}'''_{uuu}(Q_M, U_M)(U_M^1, U_M^2, \Psi, Z_M) \\
& \quad - k_M \tilde{a}'''_{qqu}(Q_M, U_M)(Q_M^1, Q_M^2, \Psi, Z_M) + k_M J'''_1(U_M)(U_M^1, U_M^2, \Psi) \\
& \quad + J'''_2(U_M)(U_M^1, U_M^2, \Psi)
\end{aligned}
$$

$$
\begin{aligned}
m = M-1,\dots,1: \quad & (\Psi, Z_m^3) + k_m \tilde{a}'_u(Q_m, U_m)(\Psi, Z_m^3) \\
& = (\Psi, Z_{m+1}^3) - k_m \tilde{a}''_{uu}(Q_m, U_m)(U_m^3, \Psi, Z_m) \\
& \quad + k_m J''_1(U_m)(U_m^3, \Psi) + k_m J''_1(U_m)(U_m^2, \Psi) \\
& \quad - k_m \tilde{a}'''_{quu}(Q_m, U_m)(Q_m^1, U_m^2, \Psi, Z_m) - k_m \tilde{a}''_{qu}(Q_m, U_m)(Q_m^1, \Psi, Z_m^2) \\
& \quad - k_m \tilde{a}'''_{quu}(Q_m, U_m)(Q_m^2, U_m^1, \Psi, Z_m) - k_m \tilde{a}''_{qu}(Q_m, U_m)(Q_m^2, \Psi, Z_m^1) \\
& \quad - k_m \tilde{a}''_{uu}(Q_m, U_m)(U_m^2, \Psi, Z_m^1) - k_m \tilde{a}''_{uu}(Q_m, U_m)(U_m^1, \Psi, Z_m^2) \\
& \quad - k_m \tilde{a}'''_{qqu}(Q_m, U_m)(Q_m^1, Q_m^2, \Psi, Z_m) \\
& \quad - k_m \tilde{a}'''_{uuu}(Q_m, U_m)(U_m^1, U_m^2, \Psi, Z_m) + k_m J'''_1(U_m)(U_m^1, U_m^2, \Psi)
\end{aligned}
$$

$$
m = 0: \quad (\Psi, Z_0^3) = (\Psi, Z_1^3)
$$

D. Numerical results for the Helmholtz equation via Born approximation

Instead of considering a sequence of linearized problems and iteratively updating the iterates q^k like we did in Chapter 4 , we treat here only the linearization at one initial guess q_0 by Example 3. We will not provide any theoretical convergence results, but give a short motivation of the basic idea and present numerical results for this approach.

For a motivation let us first consider the inverse medium problem in scalar scattering (cf. [15]), i.e., the time-harmonic scalar wave equation

$$-\Delta\tilde{u} - \tilde{k}^2\tilde{u} = f \tag{D.1}$$

(plus some boundary conditions) with wavenumber $\tilde{k}$, which we divided into a parameter k_0 characterizing the background, and a parameter k denoting the unkown perturbation of the background medium, i.e., $\tilde{k}^2 = k_0^2 + k^2$. The total scattered field is given as $\tilde{u} = u_0 + u$ with incident field u_0. Therewith, (D.1) becomes

$$-\Delta(u_0+u) - (k_0^2+k^2)(u_0+u) = -\Delta u_0 - k_0^2 u_0 - k^2 u_0 - \Delta u - k_0^2 u - k^2 u = f$$

Via Born approximation, i.e., considering the incident field instead of the total field, we neglect the term $k^2 u$, such that we get

$$-\Delta u - k_0^2 u = k^2 u_0$$

with u_0 solving

$$-\Delta u_0 - k_0^2 u_0 = f\,.$$

Transfering this reasoning to our setting

$$-\Delta\tilde{u} - (q_0+q)\tilde{u} = f \tag{D.2}$$

we arrive at the following optimal control problem

$$\begin{aligned} &\min_{(q,u)\in Q\times V} \|C(u+u_0) - \tilde{g}^\delta\|_G^2 + \frac{1}{\beta}\|q\|_Q^2 \\ &\text{s.t.}\ \ -\Delta u - q_0 u = q u_0\,, \end{aligned} \tag{D.3}$$

where u_0 is given as solution to

$$-\Delta u_0 - q_0 u_0 = f\,, \tag{D.4}$$

and noisy data $\tilde{g}^\delta$ created as described in Section 3.6 (i),(ii) via exact data $\tilde{u}^\dagger$ simulated as solution to the bilinear PDE (D.2) on a very fine mesh. For our tests, we solved the linear PDE (D.4) analytically.

Since (D.3) is a linear-quadratic optimization problem, we can apply Algorithm 4.2. The obtained reconstructions of the control and the state as well as the adaptively refined mesh for the exact source distribution (a) with point measurements (i) (see the beginning of Section 3.6) with $q_0 = 1$, $f = f_1$ (see Section 3.6.3) and Neumann zero boundary conditions are displayed in Figure D.1.

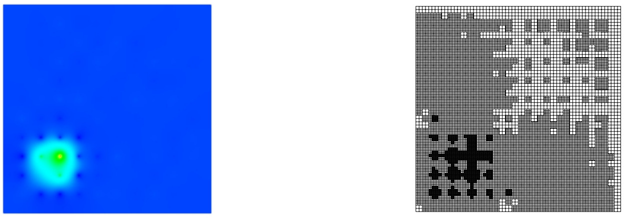

Figure D.1.: Reconstructed control and adaptively refined mesh for the linearized version of Example 3 (a)(i), 1% noise, $f = f_1$, Neumann zero boundary conditions

The corresponding results for $f = f_2$ with Neumann and also with Dirichlet zero boundary conditions can be found in Figure D.2 and Figure D.3.

Figure D.2.: Reconstructed control and adaptively refined mesh for the linearized version of Example 3 (a)(i), 1% noise, $f = f_2$, Neumann zero boundary conditions

The resulting relative control error (3.92), the number of mesh nodes, and the computed regularization parameter are shown in Table D.1.

Table D.1.: Algorithm 4.2 for the linearized version of Example 3 (a) (i), 1%

right-hand side and boundary conditions	error	β	# nodes
f_1, Neumann	0,38	15669	26048
f_2, Neumann	0,427	8371	9942
f_2, Diichlet	0,363	9877	12417

For $f = f_1$ with Neumann boundary conditions we save 47% of computation time with respect

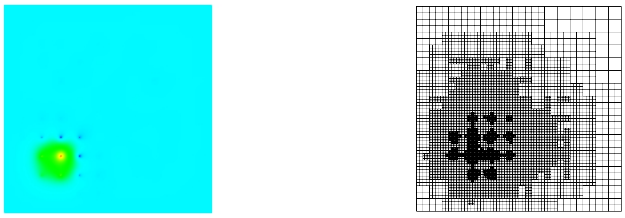

Figure D.3.: Reconstructed control and adaptively refined mesh for the linearized version of Example 3 (a)(i), 1% noise, $f = f_2$, Dirichlet zero boundary conditions

to (NT) and for $f = f_2$ with Neumann boundary, we even save 95% of computation time. For $f = f_2$ and Dirichlet boundary conditions, we save 84% of computation time compared to Algorithm 3.1 (NT) and 11% compared to Algorithm 4.6 (GN). To put this in perspective, we mention once more that, although the proposed approach seems to work well for the tested configurations, we did not give any theoretical results.

Nomenclature

- F: operator in inverse problem equation (cf. (2.1))
- q: searched-for control/parameter (cf. (2.1), (2.15))
- Q: control space (Hilbert space) (cf. (2.1), (2.15))
- $(.,.)_Q$: inner product in Q (cf. beginning of Section 2.2) (analogously for $L^2(\Omega)$)
- $\mathcal{D}$: subdomain of Q (e.g., domain of F) (cf. (2.1))
- g: exact data (cf. (2.1))
- $q^\dagger$: exact solution to $F(q) = g$ (cf. Assumption 3.1,Assumption 4.1)
- G: observation space (Hilbert space) (cf. (2.1))
- g^δ: noisy data (cf. (2.2))
- δ: noise (cf. (2.2))
- q_0: initial guess for q
- $\mathcal{R}$: range (cf. (2.6),Assumption 3.5)
- $\|.\|_{X\to Y}$: operator norm (cf. Assumption 3.5)
- β: regularization parameter (cf. (2.4))
- $i(\beta)$: misfit term of the reduced cost functional j_β for a solution of (2.4) in terms of β.
- τ: given parameter $\tau \geq 1$ in the discrepancy principle (cf. (2.8))
- β_*: "optimal" regularization parameter according to the discrepancy principle (cf. (2.8))
- C: observation operator (cf. (2.13))
- S: control-to-state operator (cf. (2.14))
- u: state (cf. (2.15))
- f: right-hand side of the state equation (cf. (2.15))
- A: state equation operator (cf. (2.15))
- V: state space (Hilbert space) (cf. (2.15))
- W: test space (Hilbert space) (cf. (2.15))
- W^* dual spaces of W (cf. beginning of Section 2.2)(analogously for V^*)

- $\langle\cdot,\cdot\rangle_{W^*,W}$: duality pairing between W and W^* (cf. beginning of Section 2.2) (analogously for V, V^*)
- a: state equation operator (cf. (2.16))
- $J_\beta(q,u)$ cost functional (non-reduced) (cf. (2.18))
- $I(u)$ misfit term of $J_\beta(q,u)$ (cf. (2.18))
- $j_\beta(q)$: reduced cost functional (cf. (2.20))
- $j'_\beta(q)(\delta)$: directional derivative (cf. (2.23))
- $\iota(q)$: misfit term of the reduced cost functional j_β
- $i(\beta)$: misfit term of the reduced cost functional j_β for a solution of (2.4) in terms of β (cf. (2.10))
- $\mathcal{L}$: Lagrangian (cf. (2.24), if not defined otherwise locally)
- q_β^δ solution/stationary point to (2.4) for fixed β and δ. (cf. Remark 2.4, Chapter 2, Assumption 3.4)
- $V_h, Q_h, W_h, S_h, F_h, j_{\beta,h}(= j_h), \iota_h, i_h, q_{\beta,h}^\delta$ discretized versions of the quantities above (cf. Section 2.3)
- id: identity map (cf., e.g., (2.42))
- $\nabla j_\beta(q)$: gradient of j at q, i.e., $j'_\beta(q)(\delta q) = (\nabla j_\beta(q), \delta q)_Q$ (cf. Section 2.5.1)
- $\nabla^2 j_\beta(q)$: Hessian of j at q, i.e., $j''_\beta(q)(\delta q, \tau q) = (\delta q, \nabla^2 j_\beta(q)\tau q)_Q$ (cf. Section 2.5.1)
- c_{tc}: constant in tangential cone condition (cf. Assumption 3.3, 4.4, 4.9 and 4.18)
- ψ: function used in the source condition Assumption 3.5
- $\underline{\beta}, \overline{\beta}$: lower and upper bound on β (cf. beginning of Section 3.2)
- $\underline{\tau}, \overline{\tau}, \tilde{\tau}, \underline{\underline{\tau}}, \overline{\overline{\tau}}$: bounds on τ (cf. Theorem 3.1, Theorem 3.4)
- η: constant in Assumption 3.6
- $\underline{\gamma}, \overline{\gamma}$: bounds on $i''(\beta)$ (cf. (3.26))
- k_* stopping index for inexact Newton method for β (cf. (3.36) in Theorem 3.7)
- $i_h^k, {i'}_h^k$: approximations to $i(\beta^k)$,$i'(\beta^k)$ (cf. Theorem 3.7)
- $\tilde{V}$: “parabolic version” of V (cf. beginning of Section 3.7)
- $(\cdot,\cdot)$: L^2 scalar product (cf. Section 3.4 below Example 3.4)
- $(\cdot,\cdot)_{\mathbb{R}^d}$: euclidian scalar product (cf. Section 3.4 below Example 3.4)
- c_s: constant used in Sobolev’s embedding inequality (cf. (3.79))
- c_p: Poincaré’s constant (cf. (3.80))

- $c_{sp} = c_s\sqrt{1+c_p^2}$ (cf. Example 3.1 (iii))
- (NT): "Nonlinear Tikhonov", standing for the method proposed in Chapter 3 (also see Algorithm 3.1)
- FLTR: "From Left To Right"
- $L^2(I,H)$: space containing functions living in $L^2(\Omega)$ in time with values in H (cf. Section 3.7)
- $(\cdot,\cdot)_I$: $L^2(\Omega)$ inner product integrated over time (cf. (3.94))
- $\tilde{a}$: "parabolic version" of a (cf. (3.95))
- $J(q,u), I(u)$ in Section 3.7: "parabolic versions" of J, I listed above (cf. (3.97))
- $\tilde{\mathcal{L}}$: "parabolic version" of $\mathcal{L}$ (cf. (3.98))
- (IRGNM): "Iteratively Regularized Gauss-Newton Method", the method proposed in Chapter 4 (also see Algorithm 4.6)
- I_1, I_2, I_3, I_4: quantities of interest for IRGNM in reduced form (cf. (4.6), (4.88))
- $I_1^k, I_2^k, I_3^k, I_4^k$ quantities of interest for IRGNM in step k in reduced form (cf. (4.8), (4.57), (4.89))
- $I_{1,h}, I_{2,h}, I_{3,h}, I_{4,h}$: quantities of interest for IRGNM in reduced and discrete form (cf. (4.13), (4.90))
- $I_{1,h}^k, I_{2,h}^k, I_{3,h}^k, I_{4,h}^k$: quantities of interest for IRGNM in step k in reduced and discrete form (cf. (4.14), (4.59), (4.91))
- $q_{\text{old}}^k, u_{\text{old}}^k, u^k, v^k, u_{\text{old},h}, u_h, v_h, u_{\text{old},h}^k, u_h^k, w_h^k$ in Section 4.1: cf. beginning of Section 4.1
- τ_β, δ_β: cf. Theorem 4.1
- $q^k = q^{\delta,k} = q_\beta^{\delta,k}$ and $q_h^k = q_{h_k}^k = q_{h_k}^{\delta,k} = q_{\beta,h_k}^{\delta,k}$ (cf. Assumption 4.2)
- c_1, c_2, c_3: constants used in Theorem 4.2
- c_4, c_5: constants used in (4.50), (4.51)
- S, R general misfit and regularization terms (cf. beginning of Section 4.1.4)
- $\mathcal{D}(R)$: domain of R (cf. (4.60))
- $D_R^\xi(q,\overline{q})$: Bregman distance (cf. (4.61))
- $c_{\mathcal{S}}$: constant used in Theorem 4.4
- $\mathcal{T}_Q, \mathcal{T}_V$: topologies (cf. Assumption 4.7)
- $\mathbf{F}, \mathbf{g}, \mathbf{g}^\delta, u^\dagger, \mathbf{T}, L, K, P, M, \mathbf{Y}_{\alpha,\mu}$: cf. beginning of Section 4.2
- $\|.\|_{Q\times V}$: cf. (4.79)
- ϱ: penalty parameter, cf. (4.78)

- L_{k-1}, K_{k-1}: cf (4.87)
- $C_{k-1}, r^g_{k-1}, r^f_{k-1}$: cf. (4.102)
- $C_{h,k-1}, L_{h,k-1}, K_{h,k-1}, r^g_{h,k-1}, r^f_{h,k-1}$: cf. (4.118), (4.119)
- $\mathcal{T}_{\beta_k}(q, u)$ objective functional for Generalized Gauss-Newton formulation (cf. (4.95)).
- $\|.\|_{X\to Y}$: operator norm (cf. Assumption 4.17)
- (GGN): "Generalized Gauss-Newton", the method proposed in Section 4.2.2 (see also Algorithm 4.6)

List of Figures

List of Tables

List of Algorithms

Bibliography

[1] H. Ben Ameur, G. Chavent, and J. Jaffré. Refinement and coarsening indicators for adaptive parametrization to the estimation of hydraulic transmissivities. *Inverse Problems*, 18(3):775–794, 2002.

[2] H. Ben Ameur and B. Kaltenbacher. Regularization of parameter estimation by adaptive discretization using refinement and coarsening indicators. *J. Inv. Ill-Posed Probl.*, 10(6):561–583, 2002.

[3] U. Ascher and E. Haber. A multigrid method for distributed parameter estimation problems. *ETNA*, 15:1–12, 2003.

[4] A. B. Bakushinskii. The Problem of the convergence of the iteratively regularized Gauss-Newton method. *Comput. Math. Math. Phys*, 32:1353–1359, 1992.

[5] A. B. Bakushinsky and M. Kokurin. *Iterative Methods for Approximate solution of Inverse Problems*. Springer, 2005.

[6] J. Baumeister. *Stable Solution of Inverse Problems*. Vieweg Advanced Lectures in Mathematics, Braunschweig/Wiesbaden, 1987.

[7] Ph. Becker. Approximation by finite element functions using local regularization. *Revue Franc. Automat. Inform. Rech. Operat.*, 9(2), 1975.

[8] R. Becker and M. Braack. Multigrid techniques for finite elements on locally refined meshes. *Numer. Linear Algebra Appl.*, 7(6):363–379, 2000.

[9] R. Becker, D. Meidner, and B. Vexler. Efficient numerical solution of parabolic optimization problems by finite element methods. *Optim. Methods Softw.*, 22(5):813–833, 2007.

[10] R. Becker and R. Rannacher. An optimal control approach to a posteriori error estimation. *Acta Numer.*, pages 1–102, 2011.

[11] R. Becker and B. Vexler. A posteriori error estimation for finite element discretizations of parameter identification problems. *Numer. Math.*, 96(3):435–459, 2004.

[12] R. Becker and B. Vexler. Mesh refinement and numerical sensitivity analysis for parameter calibration of partial differential equations. *J. Comput. Phys.*, 206(1):95–110, 2005.

[13] L. Beilina and V. Klibanov. *Approximate Global Convergence and Adaptivity for Coefficient Inverse Problems*. Springer, 2012.

[14] O. Benedix and B. Vexler. A posteriori error estimation and adaptivity for elliptic optimal control problems with state constraints. *Comput. Optim. Appl.*, 2008.

[15] G. Biros and S. Chaillat. FaIMS: A fast algorithm for the inverse medium problem with multiple frequencies and multiple sources for the scalar Helmholtz equation. *Journal of Computational Physics*, 321(12):4403–4421, 2012.

[16] G. Biros and O. Ghattas. Parallel Lagrange-Newton-Krylov-Schur Methods for PDE-Constrained Optimization. Part I: The Krylov-Schur Solver. *SIAM J. Sci. Comput.*, 27:687–713, 2005.

[17] B. Blaschke(-Kaltenbacher), A. Neubauer, and O. Scherzer. On convergence rates for the iteratively regularized Gauss-Newton method. *IMA J. Numer. Anal.*, 17:421–436, 1997.

[18] H.-G. Bock. Randwertproblemmethoden zur Parameteridentifizierung in Systemen nichtlinearer Differentialgleichungen. *Bonner Mathematische Schriften*, 183, 1987.

[19] J. F. Bonnans, J. C. Gilbert, C. Lemarchal, and C. A. Sagastizbal. *Numerical Optimization – Theoretical and Practical Aspects.* Springer, 2003.

[20] D. Braess. *Finite Elemente-Theorie, schnelle Löser und Anwendungen in der Elastizitätstheorie.* Springer, 2007.

[21] A. L. Bukhgeim. *Introduction to the Theory of Inverse Problems.* VSP, 2000.

[22] M. Burger and W. Mühlhuber. Iterative regularization of parameter identification problems by sequential quadratic programming methods. *Inverse Problems*, 18:943–969, 2002.

[23] M. Burger and W. Mühlhuber. Numerical approximation of an SQP-type method for parameter identification. *SIAM J. Numer. Anal.*, 40:1775–1797, 2002.

[24] G. Chavent and R. Bissell. Indicators for the refinement of parametrization. In *Proceedings of the third International Symposium on Inverse Problems in Engineering Mechanics*, Nagano, Japan, 1998. Elsevier.

[25] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*, volume 40 of *Classics Appl. Math.* SIAM, Philadelphia, 2002.

[26] A. R. Conn, N. I. M. Gould, and Ph. L. Toint. *Trust-Region Methods*, volume 1 of *MPS-SIAM Series on Optimization.* SIAM, Philadelphia, 2000.

[27] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems.* Kluwer Academic Publishers, 1996.

[28] H. .W. Engl and A. Neubauer. Convergence rates for Tikhonov regularization in finite-dimensional subspaces of Hilbert scales. *Proc. Amer. Math. Soc.*, 102:587–592, 1988.

[29] K. Eriksson, D. Estep, P. Hansbo, and C. Johnson. *Computational differential equations.* Cambridge University Press, 1996.

[30] M. Ern and A. Ern. A posteriori control of modeling errors and discretization errors. *Multiscale Model. Simul.*, 1(2):221–238, 2003.

[31] L. C. Evans. *Partial Differential Equations.* Springer, 1998.

[32] J. Flemming. Theory and examples of variational regularisation with non-metric fitting functionals. *J. Inverse Ill-Posed Prob.*, 18(6), 2010.

[33] J. Flemming. *General Tikhonov regularization – Basic theory and comprehensive results on convergence rates.* PhD thesis, Technische Universität Chemnitz, 2011.

[34] A. Fursikov. *Optimal Control of Distributed Systems: Theory and Applications*, volume 187 of *Transl. Math. Monogr.* AMS, Providence, 1999.

[35] Gascoigne. *Documentation.* Available from: "`http://www.gascoigne.de`".

[36] C. Geiger and C. Kanzow. *Numerische Verfahren zur Lösung unrestringierter Optimierungsaufgaben.* Springer, 1999.

[37] A. V. Goncharsky, A. S. Leonov, and A. G. Yagola. A generalized discrepancy principle for the case of operators specified with an error. *DAN SSSR*, 203(6):1238–12, 1972.

[38] H. P. W. Gottlieb. Eigenvalues of the Laplacian with Neumann boundary conditions. *The ANZIAM Journal*, 12:293–309, 1984.

[39] A. Griesbaum, B. Kaltenbacher, and B. Vexler. Efficient computation of the Tikhonov regularization parameter by goal-oriented adaptive discretization. *Inverse Problems*, 24, 2008.

[40] P. Grisvard. *Elliptic problems in nonsmooth domains*, volume 24 of *Monographs and Studies in Mathematics.* Pitman Advanced Publishing Program, 1985.

[41] J. Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, pages 49–52, 1902.

[42] M. Hanke. A regularization Levenberg-Marquardt scheme, with applications to inverse groundwater filtration problems. *Inverse Problems*, 24:79–95, 1997.

[43] M. Hanke. The regularizing Levenberg-Marquardt scheme is of optimal order. *J. Integral Equations Appl.*, 22(2):259–283, 2010.

[44] M. Hanke, A. Neubauer, and O. Scherzer. A convergence analysis of the Landweber iteration for nonlinear ill-posed problems. *Numer. Math.*, 72:21–37, 1995.

[45] P. C. Hansen. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Rev.*, 34(4):561–580, 1992.

[46] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE Constraints.* Springer, 2009.

[47] M. Hochbruck and M. Hönig. On the convergence of a regularizing Levenberg-Marquardt scheme for nonlinear ill-posed problems. *Numerical Mathematics*, 115:71–79, 2010.

[48] B. Hofmann and B. Kaltenbacher. Convergence rates for the Iteratively Regularized Gauss-Newton method in Banach spaces. *Inverse Problems*, 26, 2010.

[49] B. Hofmann, B. Kaltenbacher, C. Pöschl, and O. Scherzer. A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators. *Inverse Problems*, 23(3), 2007.

[50] T. Hohage. Logarithmic convergence rates of the iteratively regularized Gauß-Newton method for an inverse potential and an inverse scattering problem. *Inverse Problems*, 13:1279–1299, 1997.

[51] T. Hohage. Iterative Methods in Inverse Obstacle Scattering: Regularization Theory of Linear and Nonlinear Exponentially Ill-Posed Problems. *PhD thesis, University of Linz*, 1999.

[52] T. Hohage and F. Werner. Iteratively regularized Newton methods with general data misfit functionals and applications to Poisson data. *Numer. Math.*, 123(4), 2013.

[53] V. Isakov. *Inverse Problems for Partial Differential Equations*, volume 127 of *Applied Mathematical Sciences*. Springer, New York, 2006.

[54] Q. Jin. On the iteratively regularized Gauss-Newton method for solving nonlinear ill-posed problems. *Math. Comp.*, 69(232):1603–1623, 2000.

[55] Q. Jin. A convergence analysis of the iteratively regularized Gauss-Newton method under the Lipschitz condition. *Inverse Problems*, 24(4), 2008.

[56] Q. Jin. On a regularized Levenberg-Marquardt method for solving nonlinear inverse problems. *Numer. Math.*, 115:229–259, 2010.

[57] Q. Jin and Z.-Y. Hou. On the choice of the regularization parameter for ordinary and iterated Tikhonov regularization of nonlinear ill-posed problems. *Inverse Problems*, 13(3):815, 1997.

[58] Q. Jin and U. Tautenhahn. Tikhonov regularization and a posteriori rules for solving nonlinear ill-posed problems. *Inverse Problems*, 19:1–21, 2003.

[59] Q. Jin and U. Tautenhahn. On the discrepancy principle for some Newton type methods for solving nonlinear inverse problems. *Numer. Math.*, 111(4), 2009.

[60] B. Kaltenbacher. Some Newton type methods for the regularization of nonlinear ill-posed problems. *Inverse Problems*, 13(729), 1997.

[61] B. Kaltenbacher. Determination of paramteres in nonlinear hyperbolic PDEs via a multiharmonic formulation, used in piezoelectric material characterization. *Math. Meth. Mod. Appl. Sci.*, 16:1–27, 2006.

[62] B. Kaltenbacher. A note on logarithmic convergence rates for nonlinear Tikhonov regularization. *J. Inverse Ill-Posed Prob.*, 16:79–88, 2008.

[63] B. Kaltenbacher, M. Kaltenbacher, and S. Reitzinger. Identification of nonlinear B-H curves based on magnetic field computations and multigrid methods for ill-posed problems. *Eur. J. Appl. Math.*, 14:13–38, 2003.

[64] B. Kaltenbacher, A. Kirchner, and S. Veljovic. Goal oriented adaptivity in the IRGNM for parameter identification in PDEs I: reduced formulation. *Inverse Problems*, 30, 2014.

[65] B. Kaltenbacher, A. Kirchner, and B. Vexler. Adaptive discretizations for the choice of a Tikhonov regularization parameter in nonlinear inverse problems. *Inverse Problems*, 27, 2011.

[66] B. Kaltenbacher, A. Kirchner, and B. Vexler. Goal oriented adaptivity in the IRGNM for parameter identification in PDEs II: all-at-once formulations. *Inverse Problems*, 30, 2014.

[67] B. Kaltenbacher and A. Neubauer. Convergence of projected iterative regularization methods for nonlinear problems with smooth solutions. *Inverse Problems*, 22:1105–1119, 2006.

[68] B. Kaltenbacher, A. Neubauer, and A. G. Ramm. Convergence rates of the continuous regularized Gauss-Newton method. *Inverse Ill-Posed Probl.*, 10:261–280, 2002.

[69] B. Kaltenbacher, A. Neubauer, and O. Scherzer. *Iterative Regularization Methods for Nonlinear Ill-Posed Problems.* De Gruyter, 2008.

[70] B. Kaltenbacher and J. Offtermatt. A convergence analysis of regularization by discretization in preimage space. Technical Report 2010-44, University of Stuttgart, 2010.

[71] B. Kaltenbacher, F. Schöpfer, and T. Schuster. Iterative methods for nonlinear ill-posed problems in Banach spaces: Convergence and applications to parameter identification problems. *Inverse Problems*, 2009.

[72] A. Kirchner, D. Meidner, and B. Vexler. Trust region methods with hierarchical finite element models for PDE constrained optimization. *Control Cybernet.*, 40(4):1019–1042, 2011.

[73] P. Kügler and H. W. Engl. Identification of a temperature dependent heat conductivity by Tikhonov regularization. *Inverse Ill-Posed Problems*, 10:67–90, 2002.

[74] K. Kunisch and F. Colonius. Stability for parameter estimation in two point boundary value problems. *Journal für die reine und angewandte Mathematik*, 370:1–29, 1986.

[75] J. Lang. *Adaptive Multilevel Solution of Nonlinear Parabolic PDE Systems. Theory, Algorithm, and Applications.* Springer, 16 edition, 1999.

[76] S. Larsson and V. Thomée. *Partial Differential Equations with Numerical Methods.* Springer, 2003.

[77] R. Lazarov, S. Lu, and S. V. Pereverzev. On the balancing principle for some problems of Numerical Analysis. *Numer. Math.*, 106:659–689, 2007.

[78] J.-L. Lions. *Optimal Control of Systems Governed by Partial Differential Equations*, volume 170 of *Grundlehren Math. Wiss.* Springer-Verlag, Berlin, 1971.

[79] C. Liu. The Helmholtz equation on Lipschitz domains. *IMA preprint*, 1356, 1995.

[80] S. Lu, S. V. Pereverzev, and R. Ramlau. An Analysis of Tikhonov regularization for nonlinear ill-posed problems under a general smoothness assumption. *Inverse Problems*, 23:217–230, 2007.

[81] P. Mahale and T. Nair. A simplified generalized Gauss-Newton method for nonlinear ill-posed problems. *Math. Comp.*, 78(265), 2009.

[82] The MathWorks, Inc., Natick, Massachusetts, United States. *MATLAB R2012a.*

[83] V. Maz'ya and J. Rossmann. *Elliptic Equations in Polyhedral Domains*, volume 162. Math. Surv. Monogr., 2010.

[84] D. Meidner. *Adaptive Space-Time Finite Element Methods for Optimization Problems Governed by Nonlinear Parabolic Systems.* PhD thesis, Ruprecht-Karls-Universität, Heidelberg, 2008.

[85] D. Meidner and B. Vexler. Adaptive space-time finite element methods for parabolic optimization problems. *SIAM J. Control Optim.*, 46(1):116–142, 2007.

[86] D. Meidner and B. Vexler. A priori error estimates for space-time finite element approximation of parabolic optimal control problems. Part I: Problems with control constraints. *SIAM J. Control Optim.*, 47(3):1301–1329, 2008.

[87] D. Meidner and B. Vexler. A priori error estimates for space-time finite element approximation of parabolic optimal control problems. Part I: Problems without control constraints. *SIAM J. Control Optim.*, 47(3):1150–1177, 2008.

[88] A. Moiola and E. A. Spence. Is the Helmholtz equation really sign-indefinite? Technical report, University of Reading, 2012.

[89] V. A. Morozov. *On the solution of functional equations by the method of regularization.* Soviet Math. Dokl. American Mathematical Society, 1966.

[90] V. A. Morozov. *Regularization methods for ill-posed problems.* CRC Press, 1993.

[91] M. T. Nair, E. Schock, and U. Tautenhahn. Morozov's discrepancy principle under general source conditions. *Anal. Anw.*, 22:199–214, 2003.

[92] F. Natterer. The finite element method for ill-posed problems. *RAIRO Anal. Numér.*, pages 271–278, 1977.

[93] P. Neittaanmäki and D. Tiba. *Optimal Control of Nonlinear Parabolic Systems*, volume 179. Monographs and Textbooks in Pure and Applied Mathematics, 1994.

[94] A. Neubauer. Solution of ill-posed problems via adaptive grid regularization: Convergence analysis. *Numer. Funct. Anal. Optim.*, 28(3–4):405–423, 2007.

[95] J. Nocedal and S. J. Wright. *Numerical Optimization.* Springer Series in Operations Research. Springer, New York, 1999.

[96] S. V. Pereverzev, S. G. Solodky, and E. A. Volynets. The balancing principle in solving semi-discrete inverse problems in Sobolev scales by Tikhonov method. *Appl. Anal.*, 91(2):435–446, 2012.

[97] C. Poeschl. Tikhonov regularization with general residual term. *PhD thesis, Leopold Franzens Universität*, 2008.

[98] F. Potra. On a monotone Newton-like method. *Computing*, 39:233–246, 1987.

[99] R. Rannacher and J. Vihharev. Adaptive finite element analysis of nonlinear problems balancing of discretization and iteration errors. *J. Numer. Math.*, 21(1):23–62, 2012.

[100] A. Rieder. On the regularization of nonlinear ill-posed problems via inexact Newton iterations. *Inverse Problems*, 15:309–327, 1999.

[101] RoDoBo. *Documentation.* Available from: "`http://www.rodobo.org`".

[102] O. Scherzer. Convergence criteria of iterative methods based on Landweber iteration for nonlinear problems. *J. Math. Anal. Appl.*, 194:911–933, 1995.

[103] O. Scherzer, H. W. Engl, and K. Kunisch. Optimal a posteriori parameter choice for Tikhonov regularization for solving nonlinear ill-posed problems. *SIAM J. Numer. Anal.*, 30(6):1796–1838, 1993.

[104] M. Schmich and B. Vexler. Adaptivity with dynamic meshes for space-time finite element discretizations of parabolic equations. *SIAM J. Sci. Comput.*, 30(1), 2008.

[105] T. I. Seidman and C. R. Vogel. Well-posedness and convergence of some regularization methods for nonlinear ill-posed problems. *Inverse Problems*, 5:227–238, 1989.

[106] V. Thomée. *Galerkin Finite Element Methods for Parabolic Problems*, volume 25 of *Springer Series Computational Mathematics.* Springer, Berlin, 2006.

[107] A. N. Tikhonov and V. Y. Arsenin. *Solutions of Ill-Posed Problems.* Scripta series in mathematics. Winston, New York, 1977.

[108] F. Tröltzsch. *Optimal control of partial differential equations – Theory, methods and applications*, volume 112. American Mathematical Society, Rhode Island, graduate studies in mathematics edition, 2010.

[109] VisuSimple. *Documentation.* Available from: "`http://www.visusimple.uni-hd.de`".

[110] D. Wachsmuth. Adaptive regularization and discretization of bang-bang optimal control problems. *Electron. Trans. Numer. Anal.*, 40:249–267, 2013.

[111] D. Wachsmuth and G. Wachsmuth. Regularization error estimates and discrepancy principle for optimal control problems with inequality constraints. *Control Cybernet.*, 40:1125–1158, 2011.

[112] L. Yang, J.-N. Yu, and Z.-C. Deng. An inverse problem of identifying the coefficient of parabolic equations. *Applied Mathematical Modelling*, 32(10):1984–1995, 2008.

[113] S. S. Yang and G. Biros. Multigrid algorithm for inverse problems with linear parabolic PDE constraints. *Siam Journal of Scientific Computing*, 31(1):369–397, 2008.

[114] R. Zanella, P. Boccacci, L. Zanni, and M. Bertero. Efficient gradient projection methods for edge-preserving removal of Poisson noise. *Inverse Problems*, 25, 2009.