Oversmoothing Tikhonov regularization for ill-posed inverse problems

DISSERTATION

zur Erlangung des Grades eines Doktors
der Naturwissenschaften

vorgelegt von

Chantal Klinkhammer

eingereicht bei der Naturwissenschaftlich-Technischen Fakultät
der Universität Siegen
Siegen 2024

Betreuer und erster Gutachter

apl. Prof. Dr. Robert Plato

Universität Siegen

Zweiter Gutachter

Prof. Dr. Bernd Hofmann

Technische Universität Chemnitz

Tag der mündlichen Prüfung

24. September 2024

Abstract

Inverse problems arise when causes cannot be measured directly but must be concluded from observed effects. The inaccuracies arising from measurements of the effects can lead to significant deviations in determining the causes due to the typically inherent ill-posedness in inverse problems. Regularization methods overcome this ill-posedness by finding an approximation of the solution that is stable with respect to the measured data. The regularization parameter should be chosen optimally to achieve a balance between stability and approximation, minimizing the deviation of the regularized solution from the actual solution. This thesis examines Tikhonov regularization for solving nonlinear ill-posed inverse problems. The considered Tikhonov functional has an oversmoothing penalty term, such that minimization of the Tikhonov functional determines regularized solutions that are, in a certain sense, smoother than the actual solution of the inverse problem. Research on oversmoothing Tikhonov regularization has rapidly advanced, focusing on convergence rates under various conditions. Extensions to nonlinear operator equations and exploration of different source conditions and parameter choice strategies have enriched this field. This work contributes by generalizing results to a mixed source condition and providing convergence rates for a priori strategies and for the discrepancy principle as methods to select the regularization parameter. Another focus is on oversmoothing Tikhonov regularization in the finite-dimensional setting, where discretization is achieved through projection methods. This is an area that has yet to be thoroughly explored in this context.

Zusammenfassung

Inverse Probleme entstehen, wenn Ursachen nicht direkt gemessen werden können, sondern aus beobachteten Effekten geschlossen werden. Die bei der Messung der Effekte auftretenden Ungenauigkeiten können aufgrund der üblicherweise inhärenten Schlechtgestelltheit bei inversen Problemen zu erheblichen Abweichungen in den zu ermittelnden Ursachen führen. Regularisierungsmethoden überwinden diese Schlechtgestelltheit, indem sie eine Approximation der Lösung ermitteln, welche stabil von den gemessenen Daten abhängt. Der Regularisierungsparameter ist optimal zu wählen, sodass ein Kompromiss zwischen Stabilität und Approximation entsteht und dadurch die Abweichung der regularisierten zu der tatsächlichen Lösung minimal gehalten wird. Diese Arbeit untersucht die Tikhonov-Regularisierung zur Lösung nichtlinearer schlechtgestellter inverser Probleme. Das betrachtete Tikhonov-Funktional enthält einen überglättenden Strafterm, sodass die Minimierung des Funktionals regularisierte Lösungen bestimmt, die in gewissem Sinne glatter sind als die tatsächliche Lösung des inversen Problems. Die Forschung zur überglättenden Tikhonov-Regularisierung hat sich fortlaufend entwickelt und setzt den Fokus insbesondere auf Konvergenzraten unter verschiedenen Bedingungen. Erweiterungen auf nichtlineare Operatorgleichungen sowie die Erforschung verschiedener Quellbedingungen und Parameterwahlstrategien haben dieses Feld bereichert. Diese Arbeit trägt dazu bei, indem sie Ergebnisse auf eine gemischte Quellbedingung verallgemeinert und Konvergenzraten für a priori Strategien und das Diskrepanzprinzip als Methoden zur Auswahl des Regularisierungsparameters liefert. Ein weiterer Schwerpunkt liegt auf der überglättenden Tikhonov-Regularisierung im Rahmen endlichdimensionaler Räume, wobei die Diskretisierung durch Projektionsverfahren erfolgt. Dieser in diesem Zusammenhang noch nicht eingehend erforschte Bereich eröffnet neue Forschungsmöglichkeiten.

Acknowledgments

I wish to express my sincere gratitude to Prof. Robert Plato, whose insightful ideas and steadfast guidance during the course of my doctoral studies had a profound impact on the development of this thesis. In acknowledgment of his support and confidence in my abilities, I diligently pursued the completion of the project.

Furthermore, I extend my deepest appreciation to Prof. Bernd Hofmann for his instrumental role in initiating this project and for his dedicated efforts and enthusiasm throughout its progression. Finally, I am indebted to the DFG (Deutsche Forschungsgemeinschaft) for their generous financial support of our project, "Oversmoothing regularization models in light of local ill-posedness phenomena," supported under grant PL 182/8-1.

Contents

1 Introduction and outline

Introduction

Direct problems involve the computation of effects from a given cause using a mathematical model typically represented by a set of equations, such as ordinary or partial differential equations, or integral equations. In contrast, indirect or *inverse* problems aim to deduce a cause from an observed effect, relying on the model's description. Those problems occur in practice whenever relevant information is not accessible directly. In many cases, a model is known or can be constructed that establishes a connection between measurable data and the desired information. A classic example is computed tomography, in which the density of body tissue is sought, but only the intensity of x-rays can be measured. Utilizing the known properties of x-rays, a mathematical model describes the relationship between these quantities.

Inverse problems encompass a broad spectrum of practical challenges across various research domains. Monographs such as those authored by Hofmann [18] and Groetsch [14] provide valuable insights into the diverse nature of inverse problems. Inverse problems tend to be *ill-posed*. The presence of ill-posedness leads to high sensitivity of the model to perturbations within the observed data. In practical contexts, the presence of perturbations, typically in the form of measurement errors, remains unavoidable.

Regularization methods contribute to alleviating the effect of ill-posedness and help to reconstruct an adequate solution. Distinct regularization methods evolved in the recent years. Broadly, these methods can be categorized as variational methods, iterative techniques, and the method of approximate inverse [54]. The most common variational regularization method is the *Tikhonov regularization*. It has been studied rigorously in recent years. This thesis contributes to existing research by analyzing Tikhonov regularization with an *oversmoothing* penalty term. When the operator equation solution does not satisfy certain properties, which are specified by the penalty term, an application of the Tikhonov regularization leads to regularized solutions that, according to that penalty term, are smoother than the actual solution. We call the Tikhonov regularization oversmoothing in such situations. Mathematically, the absence of a certain degree of smoothness, as prescribed by the penalty term, is characterized by the fact that the searched-for solution of the operator equation does not belong to the domain of definition of the penalty term. Consequently, the Tikhonov functional fails to attain a finite value for this solution. The regularized solutions, obtained as minimizers of the Tikhonov functional with an adequate regularization parameter, however, lie in the domain of definition of the penalty term. They exhibit a degree of smoothness surpassing that of the actual solution. The oversmoothing case may occur intentionally, for example, when a smoother regularized solution is desired to facilitate numerical implementation. Conversely, oversmoothing can also arise unintentionally, for instance, if the smoothness of the searched-for solution is overestimated.

Over the past few years, research on oversmoothing regularization has developed rapidly. For the period prior to publication of this thesis, we briefly summarize the state of research. In his seminal article [41], Natterer addressed the oversmoothing Tikhonov regularization for linear operator equations. His research revealed that, under certain conditions, optimal convergence rates remain unaffected by oversmoothing penalty terms. One critical condition is a two-sided Lipschitz-type inequality. Its formulation incorporates the concept of *Hilbert scales*, a concept that Natterer also utilized in his main Theorem's proof. Some decades later, this approach has been extended to oversmoothing Tikhonov regularization for nonlinear operator equations. In particular, in the article [23], Hofmann and Mathé introduced and utilized so-called *auxiliary elements* to establish convergence rates for oversmoothing Tikhonov regularization for ill-posed nonlinear operator equations.

As explained in [8, p. 57], convergence rates can depend on a priori knowledge on the solution of the operator equation. This knowledge is generally provided by *source conditions*, containing information on the solution smoothness. Classical source conditions are of Hölder type. Accordingly, the first results on convergence rates for the oversmoothing Tikhonov regularization were based on Hölder-type source conditions. Along with a specification of a source condition, a specification of a parameter choice strategy is significant. In the mentioned paper [23], the authors provide convergence rates for an *a posteriori* parameter choice strategy. The same authors [21], and additional authors [12] completed these results by convergence rates results for an *a priori parameter* choice. In [27], a different type of source condition, specifically a low-order source condition of logarithmic type, has been examined for a non-oversmoothing setting. Such types of source conditions are of particular interest due to their less restrictive nature. Hofmann and Plato [24] incorporated this source condition and the situation of missing smoothness assumptions in their study on oversmoothing Tikhonov regularization. Their analysis thus covers three cases: the absence of smoothness assumptions, Hölder-type smoothness assumptions, and low-order smoothness assumptions. For all three cases, convergence rates for an a priori parameter choice are provided. For the cases of a missing smoothness assumption and a Hölder-type smoothness assumption, [24] also presents convergence assertions for a discrepancy principle. However, low-order source conditions under the discrepancy principle have not been investigated. One main topic of this thesis is to address this research opportunity. The result was published in [33]. In this thesis, we give a further generalization of the result in [33] by generalizing the source condition to a *mixed sourced condition* and involving both, an a priori and an a posteriori parameter choice strategy in our analysis.

We now acknowledge research that either explores more general source conditions, without addressing the oversmoothing situation or generalizes our findings to the Banach space space setting, without considering the more general mixed source condition. In the classical, non-oversmoothing setting, Tikhonov regularization under a general source condition by means of index functions has been investigated in [55], [38], and [22] for linear inverse problems. In addition, there has been a recent increase in interest in variational source conditions. Flemming [10] extensively explored variational source conditions and existing studies. Ongoing research has successfully extended results on oversmoothing Tikhonov regularization within the Hilbert space setting to the Banach space setting. A first step for the analysis of oversmoothing Tikhonov regularization towards the Banach space setting was made in [5]. Establishing low-order source conditions in the Banach space setting poses challenges due to the involvement of the logarithm of operators, which are not as easily defined as in the Hilbert space setting. The authors in [49] effectively addressed this difficulty.

Another topic of this thesis is *discretization* within the oversmoothing Tikhonov regularization. In practical applications, discretization of continuous problems is inevitable, because numerical implementation requires a finite framework. We focus on discretization by *projection* methods. Our literature research indicates that oversmoothing in the context of discretization within regularization has not been investigated yet. We postpone a general literature review concerning discretization in regularization theory to the introductory remarks of Part II of this thesis.

Outline

In Chapter 2, we provide the mathematical framework and the theoretical background essential for understanding the methodology used in this thesis. This chapter begins with a brief introduction to ill-posed inverse problems and Section 2.2 continues to introduce the oversmoothing Tikhonov regularization, the method utilized in this thesis to address these problems. Section 2.3 presents basic results on spectral calculus for certain operators, laying the groundwork for understanding the mathematical tools utilized in the subsequent analysis. In the oversmoothing setting, we perform convergence analysis within so-called pre-Hilbert scales. These are introduced in Section 2.4.

After introducing the shared mathematical framework, the thesis divides into two parts: The first part deals with the oversmoothing Tikhonov regularization in the infinite-dimensional setting. The second part examines a discretized version of the oversmoothing Tikhonov regularization. Both parts begin with chapters on fundamental requirements. These chapters, namely Chapter 3 and Chapter 7, establish well-posedness of the oversmoothing Tikhonov regularization and incorporate additional assumptions, including the source condition, necessary for proving convergence rates. The assumptions in the second part involve the projection operator as well. Therefore, with Section 7.4 an additional section is included in Part II to provide an example that confirms these assumptions. Chapters 3 and 7 also introduce auxiliary elements that serve as fundamental tools in the convergence analysis. Both parts proceed with a chapter on convergence analysis. While Chapter 4 of the first part involves both an a priori and an a posteriori parameter choice strategy, Chapter 8 of the second part focuses solely on an a priori parameter choice strategy. We confirm the established results numerically in Chapters 6 and 9. To validate the appropriateness of the considered examples, Chapter 5 establishes the related Fourier series. Specifically, the Fourier coefficients are used to confirm that the examples are appropriate for the oversmoothing setting and that they satisfy the source conditions.

Although the structure of both parts is nearly identical, we emphasize the main differences within their subject matter: The first part analyzes a more general mixed source condition and considers an a priori and an a posteriori parameter choice strategy. The second part analyzes the classical source condition of Hölder-type under an a priori parameter choice for the regularization and discretization parameters. In sake of variety, each part presents a different approach to constructing auxiliary elements.

This thesis concludes with a “Conclusion and outlook” chapter, summarizing our main findings, and proposing future research directions and opportunities for further exploration and development of the findings.

2 Theoretical background

Beginning with a mathematical formulation of nonlinear inverse and ill-posed problems in Section 2.1, this chapter introduces the central objective of this thesis. Subsequently, in Section 2.2, it presents the oversmoothing Tikhonov regularization as the method we employ to address these problems. To establish convergence rates for this method under the considered source condition, we require the understanding of functions of bounded and selfadjoint operators. Spectral calculus provides the techniques to define functions of such operators. The main statements are summarized in Section 2.3. Afterwards, in Section 2.4, pre-Hilbert scales are introduced, which serve as the framework for our analysis. These pre-Hilbert scales provide the necessary framework to prove convergence rates. Unless otherwise specified, this chapter is based on the monographs [18, Chapters 2.2, 3.2, and 4.3], [8, Chapter 10], and the article [24].

2.1 Ill-posed inverse problems

In this section, we introduce the setting considered throughout this thesis. Let X and Y be real Hilbert spaces with inner products $\langle \cdot, \cdot \rangle$ and norms $\| \cdot \|$, respectively. We omit indices within the norms and inner products that indicate the corresponding space whenever they can be identified from the context. We consider a nonlinear operator $F : X \supset \mathcal{D}(F) \to Y$. Our objective is to recover a solution $u^\dagger$, presumed to exist within the domain $\mathcal{D}(F)$ of F, of the nonlinear operator equation

$$F(u) = f^\dagger, \tag{1}$$

where the right-hand side $f^\dagger \in Y$ is only given by noisy observations $f^\delta \in Y$. These observations f^δ are assumed to satisfy the deterministic noise model

$$\| f^\dagger - f^\delta \| \leq \delta \tag{2}$$

for a known noise level $\delta > 0$.

Linear inverse problems are typically ill-posed according to Hadamard [16], meaning that they violate one of the following conditions of well-posedness:

- The operator F is surjective, that is for every $f^\dagger \in Y$ there exists a solution of the operator equation (1).
- The operator F is injective, that is for every $f^\dagger \in Y$ the solution of (1) is unique.

- The solution of (1) depends continuously on the data.

This concept of well-posedness is global for linear operators F. The local behavior of nonlinear operators might vary along their domain of definition, thus we should focus on a local analysis in the case of nonlinear operators. Therefore, we assume that the operator equation (1) is at least at the solution $u^\dagger$ locally ill-posed in the sense of the following definition, see [26], [25], or [18, Definition 2.7].

Definition 2.1. We call an operator equation (1) *locally well-posed* at $u^\dagger$ if there is a closed ball $\mathcal{B}_r(u^\dagger) = \{u \in X : \| u - u^\dagger \| \leq r\}$ with radius $r > 0$ and center $u^\dagger$, such that for every sequence $(u_n)_{n\in\mathbb{N}} \subset \mathcal{B}_r(u^\dagger) \cap \mathcal{D}(F)$, the convergence of the images $\lim_{n\to\infty} \| F(u_n) - F(u) \| = 0$ implies the convergence of the preimages $\lim_{n\to\infty} \| u_n - u \| = 0$. Otherwise, the operator equation is called *locally ill-posed* at $u^\dagger$.

This definition suggests the local analysis of an operator equation focusing on the stability of a solution. Since we presuppose that only noisy data f^δ, satisfying (2), is available, the absence of stability can lead to significant errors in reconstructed solutions. Regularization methods improve the stability of an inverse problem. The regularization method that we use, is the oversmoothing Tikhonov regularization specified in the next section.

2.2 Oversmoothing Tikhonov regularization

Regularization methods are used to reconstruct a solution of an ill-posed operator equation from non-exact data. Those methods attempt to overcome the ill-posedness of an operator equation by finding a stable solution.

To obtain a regularized solution for $u^\dagger$, classical Tikhonov regularization requires solving the minimization problem

$$\min_{u\in\mathcal{D}(F)} \left\{ \| F(u) - f^\delta \|^2 + \alpha \| u - \overline{u} \|^2 \right\}$$

for an adequate regularization parameter $\alpha > 0$. The expression to be minimized is known as the *Tikhonov functional*. The term $\| u - \overline{u} \|^2$ is called *penalty term*. The element $\overline{u}$ denotes a reference element and can be interpreted as an initial guess of the actual solution $u^\dagger$.

Nevertheless, enhanced stability is accompanied by a decrease in the accuracy of approximation. The regularization parameter α controls the interplay between stability and approximation of the minimizer. Hence, one objective is to determine a suitable value for the regularization parameter that effectively balances this trade-off. There exist a priori parameter choice strategies, depending on the noise level δ only, as well as a posteriori parameter choice strategies, depending on δ and f^δ. In both parts of this thesis, we examine an a priori parameter choice strategy. In the first part, we additionally examine the discrepancy principle as an a posteriori parameter choice rule. Heuristic parameter choice rules, depending only on f^δ, are disregarded here.

In the oversmoothing setting, we consider a Tikhonov functional which distinguishes in the penalty term. Specifically, the norm $\|\cdot\|$ in the penalty term is replaced by a stronger norm $\|\cdot\|_1$ equipping a densely defined subspace $X_1 \subset X$. The condition that $\|\cdot\|_1$ is stronger than $\|\cdot\|$ means that

$$\|u\| \leq K\|u\|_1 \qquad \text{for all} \quad u \in X_1$$

and for some positive finite constant K. We postpone a more precise definition of the space $X_1 \subset X$ and its norm $\|\cdot\|_1$ to Section 2.4.

The Tikhonov functional that we analyze is thus given by

$$T_\alpha^\delta(u) := \|F(u) - f^\delta\|^2 + \alpha\|u - \overline{u}\|_1^2 \qquad \text{for} \quad u \in X \quad \text{and} \quad \alpha > 0. \tag{3}$$

The oversmoothing situation is evoked by the condition

$$u^\dagger \notin X_1.$$

In this situation, the Tikhonov functional (3) fails to have a finite value at the searched-for solution $u^\dagger$ of the operator equation (1). Minimization of the Tikhonov functional, however, results in a solution

$$u_\alpha^\delta := \underset{u \in \mathcal{D}(F) \cap X_1}{\arg\min}\ T_\alpha^\delta(u) \tag{4}$$

that lies in X_1. For this reason, we call the Tikhonov regularization oversmoothing, because it generates a solution that is in some sense smother than the actual solution $u^\dagger$.

The question arises if u_α^δ is still a good approximation for $u^\dagger$. Convergence rates, given by the limiting behavior of the expression

$$\|u_\alpha^\delta - u^\dagger\|$$

as the noise level δ approaches zero, determine the precision of approximation.

The main goal of our study is to answer this question by providing convergence rates. We also address this question in the finite-dimensional setting in Part II of this thesis, where we minimize the Tikhonov functional T_α^δ over a finite-dimensional subspace $V_h \subset X_1$. The next two sections provide fundamental background knowledge required for the analysis in both parts.

2.3 Basic results on spectral calculus for linear selfadjoint operators

Spectral calculus provides a method for analyzing functions of certain operators, extending classical concepts of calculus to the realm of operators. Based on the presentations in [17, Chapter VI], we summarize the essential fundamentals of spectral calculus for operators which are assumed to be linear, selfadjoint, and bounded. The theory can be expanded to cover unbounded operators, but the theory for bounded operators is sufficient in our setting.

Central to spectral calculus is the spectral decomposition given below in Theorem 2.3. One main ingredient for the decomposition is the spectrum of an operator. The *spectrum* $\sigma(A)$ of a bounded, linear, and selfadjoint operator $A : X \supset \mathcal{D}(A) \to X$ is given by the set of all numbers $\lambda \in \mathbb{R}$ for which the operator

$$A - \lambda I$$

is not bijective. As usual, the operator $I : X \to X$ denotes the identity operator.

In this section, we use the following notation:

$$\sigma_{\min} := \min\{\lambda : \lambda \in \sigma(A)\} \quad \text{and} \quad \sigma_{\max} := \max\{\lambda : \lambda \in \sigma(A)\}.$$

It holds that

$$\sigma_{\min} = \inf\{\langle Au, u\rangle : \| u \| = 1\} \text{ and } \sigma_{\max} = \sup\{\langle Au, u\rangle : \| u \| = 1\} = \| A \|,$$

such that

$$\sigma(A) \subset [\sigma_{\min}, \| A \|].$$

Along with the spectrum, the notion of a spectral family plays a crucial role. It provides a systematic way of decomposing an operator into orthogonal projections associated with its spectrum. In the definition below, we formalize the concept of a spectral family and outline its key properties.

Definition 2.2 (Spectral family)**.** Let A be a bounded linear and selfadjoint operator in X. We call a family $\{P(\lambda)\}_{\lambda\in\mathbb{R}}$ of orthogonal projections in X a *spectral family* of A if it satisfies the following properties for a decreasing sequence $(\theta_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ of positive numbers converging to 0 as $n \to \infty$:

- $P(\lambda_1) \leq P(\lambda_2)$, for $\lambda_1 \leq \lambda_2$,
- $\lim_{n\to\infty} P(\sigma_{\min} - \theta_n)u = 0$ for all $u \in X$ and $P(\sigma_{\max}) = I$,
- $\lim_{n\to\infty} P(\lambda + \theta_n)u = P(\lambda)u$ for all $u \in X$ and all $\lambda \in \mathbb{R}$,
- $A = \int_{\sigma_{\min}-\theta}^{\sigma_{\max}} \lambda \, \mathrm{d}P(\lambda)$, for $\theta > 0$ arbitrary but fixed.

The integral is of Riemann-Sieltjes type, which means that for every $\varepsilon > 0$ there exists some $\tau > 0$ such that

$$\| A - \sum_{n=1}^{N} \lambda_n^{'} (P(\lambda_n) - P(\lambda_{n-1})) \| \le \varepsilon,$$

whenever

$$\left.\begin{array}{lll} \lambda_0 < \sigma_{\min} = \lambda_1 < \cdots < \lambda_{N-1} < \lambda_N = \sigma_{\max}, & & \\ \lambda_n - \lambda_{n-1} \le \tau & \text{for} & 1 \le n \le N, \\ \lambda_{n-1} \le \lambda_n^{'} \le \lambda_n & \text{for} & 1 \le n \le N. \end{array}\right\} \tag{5}$$

We can now formulate the spectral Theorem, which is a central result in the theory of spectral calculus. This theorem provides a powerful framework for understanding continuous functions when applied to operators by connecting the spectral decomposition of an operator to its functional calculus.

Theorem 2.3 (Spectral Theorem). *Let A be a bounded linear and selfadjoint operator on X. Then there exists a unique spectral family $\{P(\lambda)\}_{\lambda \in \mathbb{R}}$ such that for every continuous function $\psi : \mathbb{R} \to \mathbb{C}$ and $\theta > 0$ fixed, we have*

$$\begin{aligned} \psi(A) &= \int_{\sigma_{\min}-\theta}^{\sigma_{\max}} \psi(\lambda) \, \mathrm{d}P(\lambda), \\ \psi(A)u &= \int_{\sigma_{\min}-\theta}^{\sigma_{\max}} \psi(\lambda) \, \mathrm{d}P(\lambda)u, \quad \textit{and} \\ \langle \psi(A)u, v \rangle &= \int_{\sigma_{\min}-\theta}^{\sigma_{\max}} \psi(\lambda) \, \mathrm{d}\langle P(\lambda)u, v \rangle \end{aligned}$$

for all $u \in X$ and all $v \in X$. As above, the integrals are of Riemann-Stieljtes type. That means that for every $\varepsilon > 0$ there exists some $\tau > 0$, such that

$$\begin{aligned} \| \psi(A) - \sum_{n=1}^{N} \psi(\lambda_n^{'})(P(\lambda_n) - P(\lambda_{n-1})) \| &\le \varepsilon \\ \| \psi(A)u - \sum_{n=1}^{N} \psi(\lambda_n^{'})(P(\lambda_n)u - P(\lambda_{n-1})u) \| &\le \varepsilon \| u \|, \quad \textit{and} \\ |\langle \psi(A)u, v \rangle - \sum_{n=1}^{N} \psi(\lambda_n^{'})(\langle P(\lambda_n)u, v \rangle - \langle P(\lambda_{n-1})u), v \rangle| &\le \varepsilon \| u \| \| v \| \end{aligned}$$

hold for all $u, v \in X$, whenever (5) hold.

Without a proof, we collect useful properties concerning a continuous function of an operator. The detailed proofs are available under Corollary 2.1, Theorem 2, and Corollary 3.2 in [17, § 32].

Corollary 2.4. *Let A be a linear, bounded, and selfadjoint operator on X. Moreover let $\psi : \mathbb{R} \to \mathbb{R}$ be continuous. Then the following statements hold:*

- *$\psi(A)$ is selfadjoint as well,*
- *if $AB = BA$ for a linear and bounded operator B, then $\psi(A)B = B\psi(A)$, and*
- *$\| \psi(A) \| = \sup\{|\psi(\lambda)| : \lambda \in \sigma(A)\}$.*

2.4 The concept of pre-Hilbert scales

The estimate $T_\alpha^\delta(u_\alpha^\delta) \le T_\alpha^\delta(u^\dagger)$, which is typically used to perform convergence analysis in Tikhonov regularization, contains no information in the oversmoothing situation. However, within the framework of Hilbert scales, or more specifically pre-Hilbert scales, we can still establish convergence rates effectively.

The concept is built based on an operator B that plays an important role in our analysis. In addition to establishing the framework of pre-Hilbert scales, the operator B specifies the penalty term of the Tikhonov functional. Its inverse is used to set up smoothness assumptions on the operator F in Section 3.1 and to define the source condition for the solution $u^\dagger$ in Sections 3.2 and 7.2. In Section 3.3 and Section 7.3, we use its inverse to define auxiliary elements, which are essential for proving convergence rates in the oversmoothing setting.

Let

$$B : \mathcal{D}(B) \to X \tag{6}$$

be a selfadjoint, unbounded linear operator with dense domain $\mathcal{D}(B) \subset X$. Further, we assume that there exists a constant $k > 0$ such that

$$\|Bu\| \ge k\|u\| \qquad \text{for all } \; u \in \mathcal{D}(B). \tag{7}$$

Since the adjoint of a densely defined operator is closed, it follows that the operator B is closed. By (7), we know that B is injective, and we can define its inverse

$$B^{-1} : \mathcal{R}(B) \to X,$$

with $\mathcal{R}(B)$ denoting the range of B. It follows, for example, by Theorem 4.2-C in [56], that B^{-1} is closed as well. Using Theorem 4.2-D of the referenced monograph, we can deduce that $\mathcal{D}(B^{-1}) = \mathcal{R}(B)$ is closed. Consequently, X can be decomposed as the orthogonal sum $\mathcal{R}(B) \oplus \mathcal{R}(B)^{\perp}$ of $\mathcal{R}(B)$ and its orthogonal complement $\mathcal{R}(B)^{\perp}$. The latter is trivial because $\mathcal{R}(B)^{\perp} = \mathcal{N}(B) = 0$. From $\mathcal{R}(B) = X$, we can thus deduce that B is a bijection from $\mathcal{D}(B)$ onto X. This means that the inverse $B^{-1} : X \to X$ is well-defined on the entire space X. Condition (7) moreover implies that B^{-1} is bounded.

Based on the inverse B^{-1}, which is bounded and selfadjoint, we use spectral calculus to define for the fractional powers

$$(B^{-1})^{\tau} \qquad \text{for } \ \tau \in \mathbb{R}.$$

These are injective allowing as to express B^{τ} as

$$B^{\tau} = ((B^{-1})^{\tau})^{-1} \qquad \text{for } \ \tau \in \mathbb{R}.$$

Now we come to the definition of pre-Hilbert scales. We set

$$X_{\tau} := \begin{cases} \mathcal{D}(B^{\tau}) & \text{for } \tau > 0, \\ X & \text{for } \tau \leq 0 \end{cases}$$

with norms

$$\|u\|_{\tau} = \|B^{\tau} u\|.$$

We call the system of spaces $(X_{\tau})_{\tau \in \mathbb{R}}$ *pre-Hilbert scale* in this thesis. This system of spaces is closely related to Hilbert scales, which are introduced, for example, in [37, p. 92] or [8, Section 10.4]. A comprehensive introduction to Hilbert scales for compact operators is given in [3, Section 5.1]. In the case of negative τ, the definition of Hilbert scales requires a stricter definition of the underlying spaces, namely that each space X_{τ} is defined as the completion with respect to the corresponding norm $\| \cdot \|_{\tau}$ of the set of elements for which each power of the operator B is defined. Since for negative values of τ, the corresponding spaces $X_{\tau} = X$ with norms $\| \cdot \|_{\tau}$ are pre-Hilbert spaces in our setting, we call the introduced system of spaces *pre*-Hilbert scale. This terminology is in accordance to that in [36, p. 36].

Pre-Hilbert scales maintain useful properties of Hilbert scales. The powerful interpolation inequality, stated next, is one of them.

Lemma 2.5 (Interpolation inequality)**.** *For all $p \geq q \geq s$, $p \neq s$, and $u \in X_p$, we have*

$$\|u\|_q \leq \|u\|_p^{1-\nu} \|u\|_s^{\nu}, \tag{8}$$

where $\nu = (p-q)/(p-s)$.

The interpolation inequality can be proven using spectral calculus and the Hölder inequality, see [34] or [8, p. 213].

Part I
Oversmoothing Tikhonov regularization – the infinite-dimensional setting

We begin this part by ensuring that, under specific assumptions, minimization of the Tikhonov functional T_α^δ in (3) is well-posed. The notion of well-posedness that we use, ensures that the minimizers exist and that they are stable with respect to perturbations in the data f. Subsequently, we formulate additional assumptions in Section 3.2, as well as auxiliary elements in Section 3.3. Based on these assumptions and auxiliary elements, we can proceed with convergence analysis for the oversmoothing Tikhonov regularization in Chapter 4. This chapter involves the examination of an a priori and an a posteriori parameter choice for the regularization parameter α. To numerically confirm the convergence rates for an example in Chapter 6, Chapter 5 serves as a preparatory chapter to justify the appropriateness of that example.

Passages of Chapter 3, Chapter 4, and Chapter 6 build upon the article [33], expanding its findings to encompass a mixed source condition.

3 Fundamental requirements

3.1 Well-posedness

To ensure the effectiveness and reliability of the oversmoothing Tikhonov regularization, we validate certain regularization properties of the extremal problem

$$\min_{u \in \mathcal{D}(F) \cap X_1} T_\alpha^\delta(u). \tag{9}$$

The regularization properties include the well-posedness of the extremal problem (9), the stability of minimizers (4), and convergence of regularized solutions. We focus on well-posedness and stability assertions in this section. Convergence rates will be discussed in Chapter 4. For pure convergence assertions, without any smoothness assumption, we refer to [24].

In our approach, we require the following assumptions, which are identical to those in [24].

Assumption 3.1. (a) The operator $F : \mathcal{D}(F) \to Y$ is sequentially continuous on its domain $\mathcal{D}(F) \subset X$ with respect to the weak topologies of the Hilbert spaces X and Y. That means for a sequence $(u_n)_{n\in\mathbb{N}} \subset \mathcal{D}(F)$ converging with respect to the weak topology of X to some $u \in \mathcal{D}(F)$ as $n \to \infty$ implies the convergence of the image sequence $(F(u_n))_{n\in\mathbb{N}}$ to $F(u)$ with respect to the weak topology of Y as $n \to \infty$.

(b) The domain $\mathcal{D}(F)$ is a closed and convex subset of X.

(c) $\mathcal{D} := \mathcal{D}(F) \cap X_1 \neq \emptyset$ and $\overline{u} \in \mathcal{D}$.

The following definition, based on [58, p. 59], which appears to originate from the Russian monograph [59], provides the criterion for well-posedness of the extremal problem (9). This definition inherits the existence of minimizers. Moreover, we will observe in the progression of this section that stability of minimizers is an implication of this definition, shedding light on its significance.

Definition 3.2 (Well-posedness (Vainikko))**.** We call the extremal problem (9) *well-posed* if each corresponding minimizing sequence $(u_n)_{n\in\mathbb{N}} \subset \mathcal{D}$, that means each sequence $(u_n)_{n\in\mathbb{N}} \subset \mathcal{D}$ satisfying

$$\lim_{n\to\infty} T_\alpha^\delta(u_n) = \inf_{u\in\mathcal{D}} T_\alpha^\delta(u),$$

has a subsequence that converges with respect to a given topology of X to a minimizer of the extremal problem.

The next theorem shows, that the oversmoothing Tikhonov regularization is well-posed in the sense of Definition 3.2. It establishes the well-posedness condition under the stronger norm $\|\cdot\|_1$, and inspects the limiting behavior of the corresponding misfit and penalty functional sequences. A key aspect utilized in its proof and throughout this section is the fact that the space X_1, defined as the domain of the operator B and equipped with the norm $\|\cdot\|_1$, is a Hilbert space.

Building upon this knowledge, we identify two important properties:

- The set $\mathcal{D}$ is closed in X_1.
- The operator F when considered as $F : X_1 \supset \mathcal{D} \to Y$ is sequentially continuous with respect to the weak topologies of X_1 and Y.

The first property follows, because the embedding operator

$$i : X_1 \to X, \qquad u \mapsto u$$

is continuous, and $\mathcal{D}(F)$ is closed in X by assumption. More precisely, the continuity of the embedding operator implies that its inverse i^{-1} maps closed sets to closed sets. Therefore $i^{-1}(\mathcal{D}(F)) = X_1 \cap \mathcal{D}(F) = \mathcal{D}$ is closed in X_1. The closedness of $\mathcal{D}$ in X_1 together with its convexity imply that $\mathcal{D}$ is weakly closed, and consequently, weakly sequentially closed in X_1. This latter characteristic is the one we specifically utilize in our proof of well-posedness.

The second property follows because, by Assumption 3.1, the operator F is sequentially continuous with respect to the weak topologies of X and Y and because convergence of a sequence with respect to the weak topology of X_1 implies its convergence with respect to the weak topology of X.

Theorem 3.3 (Well-posedness). *Under Assumption 3.1, for each $\alpha > 0$ and each $f^\delta \in Y$ the extremal problem $\min_{u\in\mathcal{D}} T_\alpha^\delta(u)$ is well-posed with respect to the norm topology of X_1 in the sense of Definition 3.2. Moreover, the convergence of any minimizing subsequence $(u_{n_k})_{k\in\mathbb{N}}$ with respect to the weak topology of X_1 implies its convergence with respect to the norm topology of X_1 to a minimizer $u_\alpha^\delta := \arg\min_{u\in\mathcal{D}} T_\alpha^\delta(u)$ as $k \to \infty$, and each such subsequence satisfies*

$$\lim_{k\to\infty} \| F(u_{n_k}) - f^\delta \| = \| F(u_\alpha^\delta) - f^\delta \| \quad \text{and} \quad \lim_{k\to\infty} \| u_{n_k} - \overline{u} \|_1 = \| u_\alpha^\delta - \overline{u} \|_1. \tag{10}$$

Proof. The proof is inspired by the proof of Propositions 4.1 and 4.2 in [54]. Since $\mathcal{D} \neq \emptyset$ by Assumption 3.1 (c), the infimum

$$T_* := \inf\{T_\alpha^\delta(u) : u \in \mathcal{D}\} \geq 0$$

exists. This implies the existence of a minimizing sequence $(u_n)_{n\in\mathbb{N}} \subset \mathcal{D}$ such that

$$\lim_{n\to\infty} T_\alpha^\delta(u_n) = T_*.$$

Then $(T_\alpha^\delta(u_n))_{n\in\mathbb{N}}$ and consequently $(\| u_n - \overline{u} \|_1^2)_{n\in\mathbb{N}}$ are bounded sequences in $\mathbb{R}$, implying that the sequence $(u_n)_{n\in\mathbb{N}}$ is bounded in X_1. Thus, there exists a subsequence $(u_{n_k})_{k\in\mathbb{N}} \subset \mathcal{D}$ converging with respect to the weak topology of X_1 to some element u_* as $k \to \infty$. Since $\mathcal{D}$ is weakly sequentially closed in X_1, the limit u_* lies in $\mathcal{D}$ as well. Now let $(u_{n_k})_{k\in\mathbb{N}} \subset \mathcal{D}$ be any minimizing subsequence converging with respect to the weak topology of X_1 to some element $u_* \in \mathcal{D}$ as $k \to \infty$. Then, $(F(u_{n_k}))_{k\in\mathbb{N}}$ converges with respect to the weak topology of Y to $F(u_*)$ as $k \to \infty$, because the operator F, when considered as $F : X_1 \supset \mathcal{D} \to Y$, is sequentially continuous with respect to the weak topologies of X_1 and Y. The weak sequential lower semi-continuity of the norms $\| \cdot \|$ and $\| \cdot \|_1$ implies

$$\begin{aligned} T_\alpha^\delta(u_*) &= \| F(u_*) - f^\delta \|^2 + \alpha \| u_* - \overline{u} \|_1^2 \\ &\leq \liminf_{k\to\infty} \| F(u_{n_k}) - f^\delta \|^2 + \alpha \liminf_{k\to\infty} \| u_{n_k} - \overline{u} \|_1^2 \end{aligned}$$

$$\leq \liminf_{k\to\infty}(\| F(u_{n_k}) - f^\delta \|^2 + \alpha\| u_{n_k} - \overline{u} \|_1^2) = T_*.$$

Hence, u_* minimizes T_α^δ, justifying the representation $u_* = u_\alpha^\delta$, which we adopt henceforth. From the calculation above, we can conclude that

$$\lim_{k\to\infty} T_\alpha^\delta(u_{n_k}) = T_\alpha^\delta(u_*) = T_\alpha^\delta(u_\alpha^\delta). \tag{11}$$

Now we verify the convergence assertions in (10), beginning with the first. On the one hand,

$$\begin{aligned}
\| F(u_\alpha^\delta) - f^\delta \|^2 &= T_\alpha^\delta(u_\alpha^\delta) - \alpha\| u_\alpha^\delta - \overline{u} \|_1^2 \\
&\geq \lim_{k\to\infty} T_\alpha^\delta(u_{n_k}) - \alpha \liminf_{k\to\infty} \| u_{n_k} - \overline{u} \|_1^2 \\
&= \lim_{k\to\infty} T_\alpha^\delta(u_{n_k}) + \limsup_{k\to\infty}(-\alpha\| u_{n_k} - \overline{u} \|_1^2) \\
&= \limsup_{k\to\infty}(T_\alpha^\delta(u_{n_k}) - \alpha\| u_{n_k} - \overline{u} \|_1^2) = \limsup_{k\to\infty} \| F(u_{n_k}) - f^\delta \|^2.
\end{aligned}$$

On the other hand, the weak sequential lower semi-continuity of the norm $\| \cdot \|$ implies

$$\| F(u_\alpha^\delta) - f^\delta \|^2 \leq \liminf_{k\to\infty} \| F(u_{n_k}) - f^\delta \|^2.$$

In combination, these estimates yield that

$$\lim_{k\to\infty} \| F(u_{n_k}) - f^\delta \|^2 = \| F(u_\alpha^\delta) - f^\delta \|^2.$$

We use this convergence and the convergence of the Tikhonov functional (11) to verify the second statement of (10):

$$\begin{aligned}
\lim_{k\to\infty} \alpha\| u_{n_k} - \overline{u} \|_1^2 &= \lim_{k\to\infty} (T_\alpha^\delta(u_{n_k}) - \| F(u_{n_k}) - f^\delta \|^2) \\
&= T_\alpha^\delta(u_\alpha^\delta) - \| F(u_\alpha^\delta) - f^\delta \|^2 = \alpha\| u_\alpha^\delta - \overline{u} \|_1^2.
\end{aligned}$$

This convergence of the penalty functional, together with the convergence of $(u_{n_k})_{k\in\mathbb{N}}$ with respect to the weak topology of X_1 and the Hilbert space property of $(X_1, \| \cdot \|_1)$, yields that $(u_{n_k})_{k\in\mathbb{N}}$ converges to u_α^δ as $k \to \infty$ with respect to the norm topology of X_1:

$$\begin{aligned}
\lim_{k\to\infty} \| u_{n_k} - u_\alpha^\delta \|_1^2 &= \lim_{k\to\infty} \| u_{n_k} - \overline{u} + \overline{u} - u_\alpha^\delta \|_1^2 \\
&= \lim_{k\to\infty} (\| u_{n_k} - \overline{u} \|_1^2 + \| \overline{u} - u_\alpha^\delta \|_1^2 + 2\langle u_{n_k} - \overline{u}, \overline{u} - u_\alpha^\delta \rangle_1) \\
&= \| u_\alpha^\delta - \overline{u} \|_1^2 + \| \overline{u} - u_\alpha^\delta \|_1^2 - 2\langle u_\alpha^\delta - \overline{u}, u_\alpha^\delta - \overline{u} \rangle_1 = 0.
\end{aligned}$$

This confirms the well-posedness of the extremal problem with respect to the norm topology of X_1. □

Having established the well-posedness of the extremal problem (9), we now utilize it to demonstrate the stability of regularized solutions in Theorem 3.5 below. Stability is a fundamental aspect that ensures the resilience of regularized solutions against perturbations in the data.

In this regard, we turn to the concept of stability as defined by Schuster et al. [54, Proposition 4.2] in Definition 3.4 below. This definition guarantees that minor changes in the data do not induce significant deviations in the regularized solution.

Definition 3.4 (Stability (Schuster et al.)). Let $\alpha > 0$. For a data sequence $(f_n)_{n\in\mathbb{N}} \subset Y$, with $\lim_{n\to\infty} \| f_n - f^\delta \| = 0$, we introduce the functionals

$$T_\alpha^n(u) := \| F(u) - f_n \|^2 + \alpha \| u - \overline{u} \|_1^2 \qquad \text{for all} \;\; n \in \mathbb{N}. \tag{12}$$

By $(u_n^*)_{n\in\mathbb{N}}$ we denote the corresponding sequence of minimizers of $T_\alpha^n(u)$ over $\mathcal{D}$, that means

$$u_n^* := \arg\min_{u\in\mathcal{D}} T_\alpha^n(u) \qquad \text{for all} \;\; n \in \mathbb{N}.$$

We say that the minimizers u_α^δ of (9) are stable with respect to the data f^δ, if every sequence $(u_n^*)_{n\in\mathbb{N}}$ of minimizers has a subsequence $(u_{n_k}^*)_{k\in\mathbb{N}}$, which converges with respect to the weak topology of X, and the weak limit $u^* \in X$ of each such subsequence is a minimizer u_α^δ of (9).

In addition to addressing a stronger formulation of the stability definition of regularized solutions, the following theorem includes an assertion concerning the convergence of the misfit and penalty functionals.

Theorem 3.5 (Stability of regularized solutions). *Under Assumption 3.1, for each $\alpha > 0$, the minimizers of (9) are stable with respect to small perturbations in the data f^δ in the sense of Definition 3.4. Moreover, each subsequence $(u_{n_k}^*)_{k\in\mathbb{N}}$ of minimizers of T_α^n that converges with respect to the weak topology of X_1 as $k \to \infty$, converges with respect to the norm topology of X_1 to a minimizer u_α^δ of the Tikhonov functional T_α^δ as $k \to \infty$. Additionally, we can conclude that*

$$\lim_{k\to\infty} \| F(u_{n_k}^*) - f_{n_k} \| = \lim_{k\to\infty} \| F(u_{n_k}^*) - f^\delta \| = \| F(u_\alpha^\delta) - f^\delta \| \quad \textit{and}$$
$$\lim_{k\to\infty} \| u_{n_k}^* - \overline{u} \|_1 = \| u_\alpha^\delta - \overline{u} \|_1.$$

Proof. For $\alpha > 0$ and a data sequence $(f_n)_{n\in\mathbb{N}} \subset Y$ with $\lim_{n\to\infty} \| f_n - f^\delta \| = 0$ let $(u_n^*)_{n\in\mathbb{N}}$ be a sequence of minimizers of the functional (12). That is

$$u_n^* := \arg\min_{u\in\mathcal{D}} \left\{ \| F(u) - f_n \|^2 + \alpha \| u - \overline{u} \|_1^2 \right\} \qquad \text{for all} \;\; n \in \mathbb{N}.$$

According to Theorem 3.3, these minimizers u_n^* exist for all $n \in \mathbb{N}$. We show that $(u_n^*)_{n\in\mathbb{N}}$ minimizes $T_\alpha^\delta(u) = \| F(u) - f^\delta \|^2 + \alpha \| u - \overline{u} \|_1^2$ as $n \to \infty$. One relevant tool is the following estimate:

$$|T_\alpha^\delta(u) - T_\alpha^n(u)| \leq \| f^\delta - f_n \|^2 + 2\| F(u) - f_n \| \| f_n - f^\delta \|, \; u \in \mathcal{D}, \; n \in \mathbb{N}. \tag{13}$$

The right-hand side vanishes as $n \to \infty$ if $(\| F(u) - f_n \|)_{n\in\mathbb{N}}$ is a bounded sequence. We proceed with the usual notation $u_\alpha^\delta := \arg\min_{u\in\mathcal{D}} T_\alpha^\delta(u)$, and apply the estimate (13) twice to obtain

$$\begin{aligned} T_\alpha^\delta(u_\alpha^\delta) &\leq T_\alpha^\delta(u_n^*) \leq |T_\alpha^\delta(u_n^*) - T_\alpha^n(u_n^*)| + T_\alpha^n(u_n^*) \\ &\leq \| f^\delta - f_n \|^2 + 2\| F(u_n^*) - f_n \| \| f_n - f^\delta \| + T_\alpha^n(u_n^*) \\ &\leq \| f^\delta - f_n \|^2 + 2\| F(u_n^*) - f_n \| \| f_n - f^\delta \| + T_\alpha^n(u_\alpha^\delta) \\ &\leq 2(\| f^\delta - f_n \|^2 + 2\| F(u_n^*) - f_n \| \| f_n - f^\delta \|) + T_\alpha^\delta(u_\alpha^\delta). \end{aligned} \tag{14}$$

Since $\| F(u_n^*) - f_n \| \leq T_\alpha^n(u_n^*) \leq T_\alpha^n(u)$ for all $u \in \mathcal{D}$ and $n \in \mathbb{N}$, it follows that $\limsup_{n\to\infty} \| F(u_n^*) - f_n \| \leq \limsup_{n\to\infty} T_\alpha^n(u) = T_\alpha^\delta(u) < \infty$ for all $u \in \mathcal{D}$. Therefore, the sequence $(\| F(u_n^*) - f_n \|)_{n\in\mathbb{N}}$ is bounded, and the right-hand side of estimate (14) converges to $T_\alpha^\delta(u_\alpha^\delta)$ as $n \to \infty$.

Thus, $(u_n^*)_{n\in\mathbb{N}}$ is a minimizing sequence for T_α^δ. According to Theorem 3.3, there exists a subsequence $(u_{n_k}^*)_{k\in\mathbb{N}}$ which converges with respect to the norm topology of X_1 to a minimizer of T_α^δ as $k \to \infty$. Furthermore, each cluster point of the sequence $(u_n^*)_{n\in\mathbb{N}}$ is a minimizer of T_α^δ. The assertions within Theorem 3.3 addressing the convergence of the misfit and penalty functional directly contribute to the proof's conclusion. □

Notably, convergence in the norm topology of X_1 implies convergence with respect to the norm and weak topology of X, aligning with the classical stability condition according to Definition 3.4. Note, moreover, that the article [49] presents a proof of a more general result, specifically the exponents in the Tikhonov functional are replaced by arbitrary values.

Lastly, we mention that a minimizer of the Tikhonov functional may be not unique because the misfit functional $u \mapsto \|F(u) - f^\delta\|^2$ and hence the Tikhonov functional T_α^δ may be non-convex.

3.2 Additional assumptions and the mixed source condition

It is possible to enhance the precision of the regularized solutions by presuming information in a model. Such information can, for example, address the underlying spaces X and Y, the operator F, the operators domain of definition $\mathcal{D}(F)$, or the solution $u^\dagger$. In this section, we introduce assumptions essential for establishing convergence rates of the oversmoothing Tikhonov regularization. One of these assumptions concerns a smoothness condition on the solution $u^\dagger$, specified through a mixed source condition. As mentioned above, we cite [24] as the seminal work that presents pure convergence assertions, which do not rely on any smoothness assumption on the solution $u^\dagger$.

To specify the source condition, we recall the operator B, introduced in (6) in Section 2.4. By means of B, we define the linear operator G by

$$G : X \to X, \qquad G := B^{-(2a+2)}, \tag{15}$$

where $a > 0$ is determined through item (f) of Assumption 3.7 below. The operator G is bounded, injective, selfadjoint, and positive semidefinite. Throughout Part I of this thesis, we assume that the solution $u^\dagger$ obeys the mixed source condition defined as follows:

Definition 3.6 (Mixed source condition)**.** We define the function

$$\varphi : (0, \|G\|] \to (0, \infty), \qquad t \mapsto (-\ln ct)^{-\kappa}, \tag{16}$$

for $\kappa > 0$ and with

$$0 < c < \| G \|^{-1}. \tag{17}$$

If the relation

$$u^\dagger - \overline{u} = G^{\frac{p}{2a+2}} \varphi(G) w \tag{18}$$

holds for some $0 \leq p < 1$ and an element $w \in X$ with $\|w\| \leq \rho$, where $\rho > 0$ is a constant, then we say that $u^\dagger$ satisfies a *mixed source condition.*

If $p = 0$, the source condition (18) is of logarithmic type and hence of low-order. If $\kappa = 0$, the source condition is of Hölder-type. The results established in this part hold for both of these situations as well. The requirement in (17) ensures that φ does not attain singularities. Throughout this part of the thesis, the constant c denotes the constant specified for the function φ in (16).

Note that many studies and monographs involve the operator F, or its Fréchet, Gâteaux, or directional derivative instead of B to define source conditions; see, for example, [54, p. 65], [8, p. 247], or [26].

The following assumption summarizes conditions that are used to prove convergence rates in Chapter 4. Since the assumptions supplement those in Assumption 3.1, they commence with the item label (d).

Assumption 3.7. (d) The solution $u^\dagger \in \mathcal{D}(F)$ is an interior point of $\mathcal{D}(F)$.
(e) The observations f^δ satisfy $\| f^\delta - f^\dagger \| \leq \delta$, for $\delta > 0$.

(f) Let $a > 0$, and let there exist positive finite constants $c_a \le C_a$ and c_0, c_1 such that

$$\| F(u) - f^\dagger \| \le C_a \| u - u^\dagger \|_{-a} \quad \text{for each } u \in \mathcal{D} \text{ with } \| u - u^\dagger \|_{-a} \le c_0 \tag{19}$$

and

$$c_a \| u - u^\dagger \|_{-a} \le \| F(u) - f^\dagger \| \quad \text{for each } u \in \mathcal{D} \text{ with } \| F(u) - f^\dagger \| \le c_1. \tag{20}$$

(g) Source condition (18) applies.

3.3 Auxiliary elements

In this section, we define auxiliary elements and establish some related results, which are essential to determine convergence rates. As auxiliary elements, also referred to as smooth approximations (see e.g. [39]), we consider the minimizers $\hat{u}_\beta$ of the artificial Tikhonov functional

$$T_{a,\beta}(u) := \|u - u^\dagger\|_{-a}^2 + \beta \|u - \overline{u}\|_1^2 \qquad \text{for } \beta > 0,$$

over the whole space X. These are uniquely defined, and it can be shown that they admit the representation

$$\hat{u}_\beta := \overline{u} + G(G + \beta I)^{-1}(u^\dagger - \overline{u}) = u^\dagger - \beta(G + \beta I)^{-1}(u^\dagger - \overline{u}) \tag{21}$$

for $\beta > 0$, and G as defined in (15). This specific form of auxiliary elements was introduced in [23]. A similar approach, involving the Fréchet derivative of F at $u^\dagger$ instead of G, can be found in the proof of Theorem 10.7 in [8]. Together with this representation, the following lemma, which corresponds to Corollary 5.2 in [24], allows us to make the statements in Lemma 3.10.

Lemma 3.8. *Let $f : (0, \|G\|] \to (0, \infty)$ be a continuous, monotonically non-decreasing function with*

$$\lim_{t \downarrow 0} f(t) = 0.$$

Moreover, suppose that for each exponent η and sufficiently small $t > 0$, the quotient function $t \mapsto t^\eta / f(t)$ is strictly increasing. Then, for each $0 \le \theta < 1$, there exist finite constants $\overline{\beta} > 0$ and $\overline{C} > 0$ such that

$$\sup_{0 < \lambda \le \|G\|} \frac{\beta \lambda^\theta f(\lambda)}{\lambda + \beta} \le \overline{C} \beta^\theta f(\beta)$$

holds for all $0 < \beta \le \overline{\beta}$.

Proof. For the proof, we refer the interested reader to the aforementioned paper [24]. □

Remark 3.9. The function φ, defined as in (16), fulfills the conditions of Lemma 3.8:

Obviously, φ is a continuous, monotonically non-decreasing function and satisfies $\lim_{t\downarrow 0}\varphi(t) = 0$. The last property, requiring quotient function $t \mapsto t^\eta/\varphi(t)$ to be strictly increasing for all $\eta > 0$ and sufficiently small $t > 0$, is confirmed for $0 < t < c^{-1}\,\mathrm{e}^{-\kappa/\eta}$ through the following calculation:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{t^\eta}{(-\ln ct)^{-\kappa}}\right) &= \eta t^{\eta-1}(-\ln ct)^\kappa - t^{\eta-1}\kappa(-\ln ct)^{\kappa-1}\\ &= t^{\eta-1}(-\ln ct)^{\kappa-1}\left[\eta(-\ln ct) - \kappa\right] > 0.\end{aligned}$$

The next result in Lemma 3.10 provides bounds for norms involving the auxiliary elements. Throughout this thesis, we make use of these bounds on several occasions. We begin by applying them in the proof of the two forthcoming lemmas, Lemma 3.11 and Lemma 4.1, and they continue to be instrumental for the proofs of two main theorems of this part, namely Theorem 4.4 and Theorem 4.14.

Lemma 3.10. *Let source condition (18) be satisfied for some* $0 \le p < 1$ *and* $a > 0$*. Then, there exist positive constants* C_i*, for* $i = 1, 2, 3,$ *and* $\beta_0 \le \|G\|$ *such that the following estimates hold for all* $0 < \beta \le \beta_0$*:*

(i) $\|\hat{u}_\beta - u^\dagger\| \le C_1\beta^{\frac{p}{2a+2}}\varphi(\beta)$,

(ii) $\|\hat{u}_\beta - u^\dagger\|_{-a} \le C_2\beta^{\frac{a+p}{2a+2}}\varphi(\beta)$,

(iii) $\|\hat{u}_\beta - \overline{u}\|_1 \le C_3\beta^{\frac{p-1}{2a+2}}\varphi(\beta)$.

Proof. To deduce the items of the lemma, we use representation (21), source condition (18), as well as results on spectral theory along with Lemma 3.8. According to Lemma 3.8, there exist positive constants $\overline{C}_i$ and $\overline{\beta}_i$ for $i = 1, 2, 3,$ such that

$$\begin{aligned}\|\hat{u}_\beta - u^\dagger\| &= \|-\beta(G+\beta I)^{-1}(u^\dagger - \overline{u})\| = \|\beta(G+\beta I)^{-1}G^{\frac{p}{2a+2}}\varphi(G)w\|\\ &\le \overline{C}_1\beta^{\frac{p}{2a+2}}\varphi(\beta)\rho = C_1\beta^{\frac{p}{2a+2}}\varphi(\beta), \qquad 0 < \beta \le \overline{\beta}_1,\end{aligned}$$

and

$$\begin{aligned}\|\hat{u}_\beta - u^\dagger\|_{-a} &= \|B^{-a}(\hat{u}_\beta - u^\dagger)\| = \|G^{\frac{a}{2a+2}}(\hat{u}_\beta - u^\dagger)\|\\ &= \|\beta G^{\frac{a}{2a+2}}(G+\beta I)^{-1}G^{\frac{p}{2a+2}}\varphi(G)w\|\\ &\le \overline{C}_2\beta^{\frac{a+p}{2a+2}}\varphi(\beta)\rho = C_2\beta^{\frac{a+p}{2a+2}}\varphi(\beta), \qquad 0 < \beta \le \overline{\beta}_2,\end{aligned}$$

as well as

$$\begin{aligned}\|\hat{u}_\beta - \overline{u}\|_1 &= \|B\,(\hat{u}_\beta - \overline{u})\| = \|\tfrac{1}{\beta}\beta G^{-\frac{1}{2a+2}}G(G+\beta I)^{-1}(u^\dagger - \overline{u})\| \\ &= \|\tfrac{1}{\beta}\beta G^{-\frac{1}{2a+2}}G(G+\beta I)^{-1}G^{\frac{p}{2a+2}}\varphi(G)w\| \\ &\leq \tfrac{1}{\beta}\overline{C}_3\rho\beta^{1+\frac{p-1}{2a+2}}\varphi(\beta) = C_3\beta^{\frac{p-1}{2a+2}}\varphi(\beta), \qquad 0 < \beta \leq \overline{\beta}_3,\end{aligned}$$

with $C_i = \overline{C}_i\rho > 0$, for $i = 1, 2, 3$, and ρ defined as the upper bound for the source element w in (18). Setting $\beta_0 = \min\left\{\overline{\beta}_1, \overline{\beta}_2, \overline{\beta}_3\right\}$ yields the assertion. Note that Lemma 3.8 can be applied because the exponents of G within the norms $\|\cdot\|$ sum up to values that are smaller than 1, because $0 \leq p < 1$. $\square$

Based on Lemma 3.10, we establish the following lemma, which gives upper bounds for the same norms as in Lemma 3.10 when the parameter β is chosen a priori in a specific manner. The corresponding bounds are expressed through functions depending on the noise level δ. Although this lemma is required only for proving convergence rates for the discrepancy principle in Section 4.2, we include it in this current section, as the statement concerns the auxiliary elements.

Lemma 3.11. *Let source condition (18) hold with $a > 0$ and $0 \leq p < 1$. By $\hat{u}_{\beta_*}$ we denote the auxiliary elements defined in (21), with regularization parameter $\beta = \beta_*$ chosen a priori as*

$$\beta_* := \beta_*(\delta) = \delta^r\varphi(\delta)^{-r},$$

where

$$r := \frac{2a+2}{a+p}.$$

For this parameter choice, the following bounds hold as $\delta \downarrow 0$:

(i) $\|\hat{u}_{\beta_*} - u^\dagger\| = \mathcal{O}\big(\delta^{\frac{p}{a+p}}\varphi(\delta)^{\frac{a}{a+p}}\big)$,

(ii) $\|\hat{u}_{\beta_*} - u^\dagger\|_{-a} = \mathcal{O}(\delta)$,

(iii) $\|\hat{u}_{\beta_*} - \overline{u}\|_1 = \mathcal{O}\big(\delta^{\frac{p-1}{a+p}}\varphi(\delta)^{\frac{a+1}{a+p}}\big)$.

Proof. Without loss of generality, we can assume that $\beta_* \leq \beta_0$, where β_0 is the constant from Lemma 3.10. According to that lemma, it is sufficient to establish the estimates

(A) $\beta_*^{\frac{p}{2a+2}}\varphi(\beta_*) = \mathcal{O}\big(\delta^{\frac{p}{a+p}}\varphi(\delta)^{\frac{a}{a+p}}\big)$,

(B) $\beta_*^{\frac{a+p}{2a+2}}\varphi(\beta_*) = \mathcal{O}(\delta)$, and

(C) $\beta_*^{\frac{p-1}{2a+2}}\varphi(\beta_*) = \mathcal{O}\big(\delta^{\frac{p-1}{a+p}}\varphi(\delta)^{\frac{a+1}{a+p}}\big)$

as $\delta \downarrow 0$. As a preliminary step, we verify that β_* satisfies $\varphi(\beta_*) = \mathcal{O}(\varphi(\delta))$ as $\delta \downarrow 0$:

$$\begin{aligned}\lim_{\delta\downarrow 0} \frac{\varphi(\beta_*)}{\varphi(\delta)} &= \lim_{\delta\downarrow 0} \frac{\left(-\ln\left(c\delta^r\left(-\ln c\delta\right)^{\kappa r}\right)\right)^{-\kappa}}{(-\ln c\delta)^{-\kappa}} \\ &= \lim_{\delta\downarrow 0}\Big(\frac{r\ln c\delta + \ln c^{1-r} + \kappa r\ln(-\ln c\delta)}{\ln c\delta}\Big)^{-\kappa} \\ &= \lim_{\delta\downarrow 0}\Big(r\Big(1+\frac{r^{-1}\ln(c^{1-r}) + \kappa\ln(-\ln c\delta)}{\ln c\delta}\Big)\Big)^{-\kappa} = r^{-\kappa}.\end{aligned}$$

For item (A), we thus calculate that

$$\begin{aligned}\beta_*^{\frac{p}{2a+2}}\varphi(\beta_*) &= (\delta^{\frac{2a+2}{a+p}}\varphi(\delta)^{-\frac{2a+2}{a+p}})^{\frac{p}{2a+2}}\mathcal{O}(\varphi(\delta)) = \mathcal{O}(\delta^{\frac{p}{a+p}}\varphi(\delta)^{1-\frac{p}{a+p}}) \\ &= \mathcal{O}(\delta^{\frac{p}{a+p}}\varphi(\delta)^{\frac{a}{a+p}}) \qquad \text{as} \quad \delta\downarrow 0.\end{aligned}$$

Similarly, simple calculations verify (B) and (C):

$$\beta_*^{\frac{a+p}{2a+2}}\varphi(\beta_*) = \big(\delta^{\frac{2a+2}{a+p}}\varphi(\delta)^{-\frac{2a+2}{a+p}}\big)^{\frac{a+p}{2a+2}}\mathcal{O}(\varphi(\delta)) = \mathcal{O}(\delta) \qquad \text{as} \quad \delta\downarrow 0,$$

and

$$\begin{aligned}\beta_*^{\frac{p-1}{2a+2}}\varphi(\beta_*) &= (\delta^{\frac{2a+2}{a+p}}\varphi(\delta)^{-\frac{2a+2}{a+p}})^{\frac{p-1}{2a+2}}\mathcal{O}(\varphi(\delta)) \\ &= \mathcal{O}(\delta^{\frac{p-1}{a+p}}\varphi(\delta)^{\frac{a+1}{a+p}}) \qquad \text{as} \quad \delta\downarrow 0.\end{aligned}$$

□

4 Convergence analysis

This chapter divides into two sections. In Section 4.1, the first of these two, we establish convergence rates for an a priori parameter choice of α. The analysis is similar to that in [21], in which the authors consider a Hölder-type source condition. In the second section, Section 4.2, after introducing the discrepancy principle and related results, we establish convergence rates when the regularization parameter α is chosen according to that discrepancy principle. Our analysis relies on the auxiliary elements presented in Section 3.3.

In a first step, we establish an upper bound for the components of the Tikhonov functional when applied to a minimizer. This upper bound is of relevance in both sections.

Lemma 4.1. *Under Assumptions 3.1 and 3.7, there exists a positive and finite constant C_4, such that the following estimate is satisfied for a minimizer u_α^δ of the Tikhonov functional (3) for all $0 < \alpha \le \|G\|$ and $\delta > 0$:*

$$\max\left\{\|F(u_\alpha^\delta) - f^\delta\|, \sqrt{\alpha}\|u_\alpha^\delta - \overline{u}\|_1\right\} \le C_4 \alpha^{\frac{a+p}{2a+2}} \varphi(\alpha) + \delta.$$

Proof. The proof follows the proof of Lemma 3.2 in [24]. Let $\beta_0 \le \|G\|$ be the constant from Lemma 3.10. Further, we set $\beta = \alpha$ as the regularization parameter for the auxiliary elements $\hat{u}_\beta$ defined in (21) and denote these auxiliary elements by $\hat{u}_\alpha$. From item (i) of Lemma 3.10, we can deduce that $\hat{u}_\alpha$ belongs to $\mathcal{D}$ for sufficiently small $\alpha > 0$, say $\alpha \le \alpha_0$ for some $\alpha_0 \le \beta_0$. This conclusion stems from Assumption 3.7 (d) that $u^\dagger$ is an interior point of $\mathcal{D}(F)$. Based on this information, the following inequality chain holds for $\alpha \le \alpha_0$:

$$\begin{aligned}\left(\|F(u_\alpha^\delta) - f^\delta\|^2 + \alpha\|u_\alpha^\delta - \overline{u}\|_1^2\right)^{\frac{1}{2}} &\le \left(\|F(\hat{u}_\alpha) - f^\delta\|^2 + \alpha\|\hat{u}_\alpha - \overline{u}\|_1^2\right)^{\frac{1}{2}} \\ &\le \|F(\hat{u}_\alpha) - f^\delta\| + \sqrt{\alpha}\|\hat{u}_\alpha - \overline{u}\|_1 \\ &\le \|F(\hat{u}_\alpha) - f^\dagger\| + \sqrt{\alpha}\|\hat{u}_\alpha - \overline{u}\|_1 + \delta. \quad (22)\end{aligned}$$

According to item (ii) of Lemma 3.10, we have

$$\|\hat{u}_\alpha - u^\dagger\|_{-a} \le C_2 \alpha^{\frac{a+p}{2a+2}} \varphi(\alpha) \qquad \text{for } \alpha \le \alpha_0.$$

The function $\alpha \mapsto \alpha^{\frac{a+p}{2a+2}} \varphi(\alpha)$ is monotonically increasing on $(0, \|G\|]$. Hence

$$\|\hat{u}_\alpha - u^\dagger\|_{-a} \le C_2 \alpha_0^{\frac{a+p}{2a+2}} \varphi(\alpha_0) =: c_0 \qquad \text{for } \alpha \le \alpha_0.$$

Consequently, estimate (19) applies for $\alpha \le \alpha_0$, which allows us to determine an upper bound for first summand of (22):

$$\|F(\hat{u}_\alpha) - f^\dagger\| \le C_a \|\hat{u}_\alpha - u^\dagger\|_{-a} \le C_a C_2 \alpha^{\frac{a+p}{2a+2}} \varphi(\alpha) \qquad \text{for } \alpha \le \alpha_0.$$

Using item (iii) of Lemma 3.10, we find for the second summand of (22) that

$$\sqrt{\alpha}\|\hat{u}_\alpha - \overline{u}\|_1 \le C_3\alpha^{\frac{1}{2}+\frac{p-1}{2a+2}}\varphi(\alpha) \qquad \text{for} \;\; \alpha \le \alpha_0.$$

The claim thus holds for $\alpha \le \alpha_0$ with $C_4 = C_a C_2 + C_3$.

Now we consider $\alpha_0 < \alpha \le \|G\|$. In this case

$$\frac{\alpha^{\frac{a+p}{2a+2}}\varphi(\alpha)}{\alpha_0^{\frac{a+p}{2a+2}}\varphi(\alpha_0)} \ge 1$$

such that

$$\begin{aligned} T_\alpha^\delta(u_\alpha^\delta)^{\frac{1}{2}} &\le \|F(\overline{u}) - f^\delta\| \le \|F(\overline{u}) - f^\dagger\| + \delta \\ &\le \frac{\alpha^{\frac{a+p}{2a+2}}\varphi(\alpha)}{\alpha_0^{\frac{a+p}{2a+2}}\varphi(\alpha_0)}\|F(\overline{u}) - f^\dagger\| + \delta \qquad \text{for} \;\; \alpha_0 < \alpha \le \|G\|. \end{aligned}$$

Setting

$$C_4 = \frac{\|F(\overline{u}) - f^\dagger\|}{\alpha_0^{\frac{a+p}{2a+2}}\varphi(\alpha_0)}$$

yields the assertion for $\alpha_0 < \alpha \le \|G\|$. □

4.1 A priori parameter choice

Throughout this section, we assume that Assumptions 3.1 and 3.7 are satisfied. As a preparation for the proof of this section's main theorem, we collect some estimates for norms involving minimizers u_α^δ of (9). An immediate consequence of Lemma 4.1 is the subsequent corollary.

Corollary 4.2. *Let C_4 be the constant from Lemma 4.1. Then for each $0 < \alpha \le \|G\|$ and $\delta > 0$, we have*

$$\|u_\alpha^\delta - \overline{u}\|_1 \le \frac{1}{\sqrt{\alpha}}(C_4\alpha^{\frac{a+p}{2a+2}}\varphi(\alpha) + \delta).$$

Lemma 4.3. *Let C_4 be the constant from Lemma 4.1. For each $0 < \alpha \le \|G\|$ and $0 < \delta < \infty$, we have*

$$\|u_\alpha^\delta - u^\dagger\|_{-a} \le \frac{2}{c_a}\big(\frac{C_4}{2}\alpha^{\frac{a+p}{2a+2}}\varphi(\alpha) + \delta\big).$$

Proof. To obtain the result, we use the estimate (20), the triangle inequality, and Lemma 4.1 in the given order:

$$\begin{aligned} c_a \| u_\alpha^\delta - u^\dagger \|_{-a} \le \| F(u_\alpha^\delta) - f^\dagger \| &\le \| F(u_\alpha^\delta) - f^\delta \| + \delta \\ &\le C_4 \alpha^{\frac{a+p}{2a+2}} \varphi(\alpha) + 2\delta \end{aligned} \tag{23}$$

for $0 < \alpha \le \| G \|$. Since $\delta \le \delta_0$, for some $\delta_0 < \infty$, and because the mapping $\alpha \mapsto \alpha^{(a+p)/(2a+2)} \varphi(\alpha)$ is monotonically increasing on the interval $(0, \| G \|]$, the right-hand side of (23) is bounded from above by $C_4 \| G \|^{(a+p)/(2a+2)} \varphi(\| G \|) + 2\delta_0 =: c_1$ for $\alpha \le \| G \|$, validating the applicability of estimate (20). Dividing the estimate in (23) by c_a yields the assertion. □

We have now everything at hand to prove the next theorem, which gives an upper bound for the error term $\| u_\alpha^\delta - u^\dagger \|$. This upper bound consists of two components depending on α. The first component increases in α, while the second component decreases in α. This highlights the importance of choosing α appropriately.

Theorem 4.4. *Let $\beta_0 \le \| G \|$ be the constant from Lemma 3.10. There exists a positive and finite constant K_1 such that the estimate*

$$\| u_\alpha^\delta - u^\dagger \| \le \Big(\frac{2}{c_a}\Big)^{\frac{1}{a+1}} \big(K_1 \alpha^{\frac{p}{2a+2}} \varphi(\alpha) + \delta \alpha^{-\frac{a}{2a+2}}\big)$$

holds for each $0 < \alpha \le \beta_0$ and $0 < \delta < \infty$.

Proof. As in the proof of Lemma 4.1, we consider the auxiliary elements $\hat{u}_\beta$ defined in (21) and set $\beta = \alpha$ as their regularization parameter. Accordingly, we denote them by $\hat{u}_\beta = \hat{u}_\alpha$. Using the triangle and interpolation inequality (8) gives

$$\begin{aligned} \| u_\alpha^\delta - u^\dagger \| &\le \| u_\alpha^\delta - \hat{u}_\alpha \| + \| \hat{u}_\alpha - u^\dagger \| \\ &\le \underbrace{\| u_\alpha^\delta - \hat{u}_\alpha \|_{-a}}_{=:I}^{\frac{1}{a+1}} \underbrace{\| u_\alpha^\delta - \hat{u}_\alpha \|_1}_{=:II}^{\frac{a}{a+1}} + \underbrace{\| \hat{u}_\alpha - u^\dagger \|}_{=:III}. \end{aligned} \tag{24}$$

For the term in I, another application of the triangle inequality along with Lemma 4.3 and Lemma 3.10 yields that

$$\begin{aligned} I &\le \| u_\alpha^\delta - u^\dagger \|_{-a} + \| u^\dagger - \hat{u}_\alpha \|_{-a} \le \frac{2}{c_a} \Big(\frac{C_4}{2} \alpha^{\frac{a+p}{2a+2}} \varphi(\alpha) + \delta\Big) + C_2 \alpha^{\frac{a+p}{2a+2}} \varphi(\alpha) \\ &= \frac{2}{c_a} \Big(\frac{(C_4 + C_2 c_a)}{2} \alpha^{\frac{a+p}{2a+2}} \varphi(\alpha) + \delta\Big), \end{aligned}$$

for $0 < \alpha \le \beta_0$ and $0 < \delta < \infty$. We determine a bound for the term in II by making use of the triangle inequality, Corollary 4.2, and Lemma 3.10:

$$II \le \| u_\alpha^\delta - \overline{u} \|_1 + \| \overline{u} - \hat{u}_\alpha \|_1 \le \frac{1}{\sqrt{\alpha}} (C_4 \alpha^{\frac{a+p}{2a+2}} \varphi(\alpha) + \delta) + C_3 \alpha^{\frac{p-1}{2a+2}} \varphi(\alpha)$$
$$= \frac{1}{\sqrt{\alpha}} ((C_3 + C_4) \alpha^{\frac{a+p}{2a+2}} \varphi(\alpha) + \delta)$$

for $0 < \alpha \le \beta_0$ and $0 < \delta < \infty$. Item (i) of Lemma 3.10 gives an upper bound for the term in III:

$$III \le C_1 \alpha^{\frac{p}{2a+2}} \varphi(\alpha) \qquad \text{for each } \ 0 < \alpha \le \beta_0.$$

By inserting I, II, and III into (24), we obtain for $\widetilde{C} := \max\{\frac{C_4 + C_2 c_a}{2}, C_3 + C_4\}$ that

$$\|u_\alpha^\delta - u^\dagger\| \le \big(\frac{2}{c_a} (\widetilde{C} \alpha^{\frac{a+p}{2a+2}} \varphi(\alpha) + \delta)\big)^{\frac{1}{a+1}} \big(\frac{1}{\sqrt{\alpha}} (\widetilde{C} \alpha^{\frac{a+p}{2a+2}} \varphi(\alpha) + \delta)\big)^{\frac{a}{a+1}}$$
$$+ C_1 \alpha^{\frac{p}{2a+2}} \varphi(\alpha)$$
$$= \big(\frac{2}{c_a}\big)^{\frac{1}{a+1}} \alpha^{-\frac{a}{2a+2}} (\widetilde{C} \alpha^{\frac{a+p}{2a+2}} \varphi(\alpha) + \delta) + C_1 \alpha^{\frac{p}{2a+2}} \varphi(\alpha)$$
$$= \big(\frac{2}{c_a}\big)^{\frac{1}{a+1}} (K_1 \alpha^{\frac{p}{2a+2}} \varphi(\alpha) + \alpha^{-\frac{a}{2a+2}} \delta)$$

for each $0 < \alpha \le \beta_0$ and $0 < \delta < \infty$, and with $K_1 := \widetilde{C} + (2/c_a)^{-1/(a+1)} C_1$. □

Theorem 4.5 (Convergence rates under a mixed source condition)**.** *For the a priori parameter choice*

$$\alpha_* := \alpha_*(\delta) = \delta^r \varphi(\delta)^{-r},$$

where

$$r := \frac{2a+2}{a+p},$$

we have

$$\| u_{\alpha_*}^\delta - u^\dagger \| = \mathcal{O}\big(\delta^{\frac{p}{a+p}} \varphi(\delta)^{\frac{a}{a+p}}\big) \qquad \text{as } \ \delta \downarrow 0.$$

Proof. Noticing that α_* coincides with β_* of Lemma 3.11, we can conclude as in the preparatory step of the proof of that lemma that $\varphi(\alpha_*) = \mathcal{O}(\varphi(\delta))$ as $\delta \downarrow 0$. Without loss of generality, we can assume $\alpha_* \le \beta_0$, with β_0 denoting the constant from Lemma 3.10, and apply Theorem 4.4 to the a priori parameter choice α_*:

$$\| u_{\alpha_*}^\delta - u^\dagger \| \le \big(\frac{2}{c_a}\big)^{\frac{1}{a+1}} (K_1 \alpha_*^{\frac{p}{2a+2}} \varphi(\alpha_*) + \delta \alpha_*^{-\frac{a}{2a+2}})$$
$$= \big(\frac{2}{c_a}\big)^{\frac{1}{a+1}} (K_1 \delta^{\frac{p}{a+p}} \varphi(\delta)^{-\frac{p}{a+p}} \varphi(\alpha_*) + \delta \delta^{-\frac{a}{a+p}} \varphi(\delta)^{\frac{a}{a+p}})$$
$$= \mathcal{O}\big(\delta^{\frac{p}{a+p}} \varphi(\delta)^{\frac{a}{a+p}}\big) \qquad \text{as } \ \delta \downarrow 0.$$

□

4.2 A posteriori parameter choice: Discrepancy principle

The classical a posteriori parameter choice is a discrepancy principle according to Mozorov [40, p. 53], where the objective is to find a parameter $\alpha := \alpha(f^\delta, \delta)$ such that the misfit functional $\|F(u_\alpha^\delta) - f^\delta\|$ approximately behaves like δ. That is

$$k_1\delta \leq \|F(u_\alpha^\delta) - f^\delta\| \leq k_2\delta$$

for constants $1 \leq k_1 \leq k_2 < \infty$. In this thesis, following the research of Hofmann and Plato [24], we employ a version of the discrepancy principle that, particularly in the presence of potential discontinuities within the function $\alpha \mapsto \|F(u_\alpha^\delta) - f^\delta\|$, holds greater significance than the classical discrepancy principle. Moreover, it allows a sequential implementation, as demonstrated by Algorithm 1.

Before detailing the procedure of the discrepancy principle, we present relevant background information. Lemma 4.1 already suggests the potential monotonic behavior of both, the misfit and penalty functionals of the Tikhonov functional (3) when considered as functions of α. Now, we examine this monotonic behavior in more detail. In combination with findings for the limiting behavior of the misfit functional and the minimizers of (9), we can substantiate the validity of the discrepancy principle introduced thereafter. The next two lemmas cite the outcome of Proposition 4.5 from the work of [24]. We have separated the assertions from that proposition and incorporated an auxiliary result of its proof into two distinct lemmas. Lemma 4.7, the first of them, provides insights into the monotonic behavior of the misfit and the penalty functional when considered as functions $\alpha \mapsto \|F(u_\alpha^\delta) - f^\delta\|$, and $\alpha \mapsto \|u_\alpha^\delta - \overline{u}\|_1^2$ respectively, for fixed $\delta > 0$.

The second lemma, Lemma 4.6, provides an overview of the limiting behavior of the misfit functional when considered as a function $\alpha \mapsto \|F(u_\alpha^\delta) - f^\delta\|$ for fixed $\delta > 0$. Additionally, it provides a convergence assertion for the minimizers u_α^δ for fixed $\delta > 0$.

The proofs follow the proof of Proposition 4.5 in [24]. For convenience, we include them here.

Lemma 4.6. *Under Assumption 3.1, for fixed $\delta > 0$, the function*

$$\alpha \mapsto \|F(u_\alpha^\delta) - f^\delta\|$$

is non-decreasing and the function

$$\alpha \mapsto \|u_\alpha^\delta - \overline{u}\|_1$$

is non-increasing.

Proof. We begin by showing that the function $\alpha \mapsto \| u_\alpha^\delta - \overline{u} \|_1$ is non-increasing.

Let $0 < \alpha \leq \beta$. Then

$$\begin{aligned} T_\beta^\delta(u_\beta^\delta) &\leq T_\beta^\delta(u_\alpha^\delta) = T_\alpha^\delta(u_\alpha^\delta) + (\beta - \alpha)\| u_\alpha^\delta - \overline{u} \|_1^2 \\ &\leq T_\alpha^\delta(u_\beta^\delta) + (\beta - \alpha)\| u_\alpha^\delta - \overline{u} \|_1^2 \\ &= T_\beta^\delta(u_\beta^\delta) + (\beta - \alpha)(\| u_\alpha^\delta - \overline{u} \|_1^2 - \| u_\beta^\delta - \overline{u} \|_1^2). \end{aligned}$$

Therefore, we have $\| u_\alpha^\delta - \overline{u} \|_1^2 \geq \| u_\beta^\delta - \overline{u} \|_1^2$, which confirms the first statement. This allows us to compute

$$\begin{aligned} \| F(u_\alpha^\delta) - f^\delta \|^2 + \alpha \| u_\alpha^\delta - \overline{u} \|_1^2 &= T_\alpha^\delta(u_\alpha^\delta) \\ &\leq T_\alpha^\delta(u_\beta^\delta) = \| F(u_\beta^\delta) - f^\delta \|^2 + \alpha \| u_\beta^\delta - \overline{u} \|_1^2 \\ &\leq \| F(u_\beta^\delta) - f^\delta \|^2 + \alpha \| u_\alpha^\delta - \overline{u} \|_1^2. \end{aligned}$$

It follows that $\| F(u_\alpha^\delta) - f^\delta \|^2 \leq \| F(u_\beta^\delta) - f^\delta \|^2$, confirming the monotonic behavior of $\alpha \mapsto \|F(u_\alpha^\delta) - f^\delta\|$. □

Lemma 4.7. *Under Assumptions 3.1 and 3.7, for fixed* $\delta > 0$, *the function* $\alpha \mapsto \|F(u_\alpha^\delta) - f^\delta\|$ *satisfies*

$$\lim_{\alpha \downarrow 0} \|F(u_\alpha^\delta) - f^\delta\| \leq \delta \quad \textit{and} \quad \lim_{\alpha \to \infty} \|F(u_\alpha^\delta) - f^\delta\| = \| F(\overline{u}) - f^\delta \|. \tag{25}$$

Furthermore,

$$\lim_{\alpha \to \infty} \| u_\alpha^\delta - \overline{u} \| = 0.$$

Proof. We begin by showing the last assertion, $\lim_{\alpha \to \infty} \|u_\alpha^\delta - \overline{u}\| = 0$. The estimate

$$\|F(u_\alpha^\delta) - f^\delta\|^2 + \alpha \|u_\alpha^\delta - \overline{u}\|_1^2 = T_\alpha^\delta(u_\alpha^\delta) \leq T_\alpha^\delta(\overline{u}) = \|F(\overline{u}) - f^\delta\|^2 \tag{26}$$

yields $\|u_\alpha^\delta - \overline{u}\|_1 = \mathcal{O}(\alpha^{-1/2})$ as $\alpha \to \infty$. This, together with condition (7) of the operator B, implies that $\|u_\alpha^\delta - \overline{u}\| = \mathcal{O}(\alpha^{-1/2})$ as $\alpha \to \infty$.

The first statement of (25) follows from Lemma 4.1.

It remains to verify the second statement of (25). On the one hand, taking the limit as $\alpha \to \infty$ in (26) gives

$$\lim_{\alpha \to \infty} \| F(u_\alpha^\delta) - f^\delta \| \leq \| F(\overline{u}) - f^\delta \|.$$

On the other hand, the convergence $\lim_{\alpha \to \infty} \| u_\alpha^\delta - \overline{u} \| = 0$, together with the sequential continuity of F with respect to the weak topologies of X and Y and the weak lower semi-continuity of the norm $\| \cdot \|$ in Y, yields $\lim_{\alpha \to \infty} \| F(u_\alpha^\delta) - f^\delta \| \geq \| F(\overline{u}) - f^\delta \|$. This shows $\lim_{\alpha \to \infty} \| F(u_\alpha^\delta) - f^\delta \| = \| F(\overline{u}) - f^\delta \|$. □

Remark 4.8. • The statements of Lemma 4.6 and Lemma 4.7 concerning non-oversmoothing Tikhonov regularization can be found in [58, Section 3.2].

• Note that Assumption 3.7 in Lemma 4.7 is only required to show that $\lim_{\alpha\downarrow 0}\|F(u_\alpha^\delta) - f^\delta\| \leq \delta$ through Lemma 4.1.

The statements of Lemma 4.6 and Lemma 4.7 justify the discrepancy principle, defined in Definition 4.9 below, as a method for selecting the regularization parameter α. Just like Lemma 4.6 and Lemma 4.7, the definition for the sequential discrepancy principle originates from [24] and [58, p. 70].

Definition 4.9. (Discrepancy principle) For constants $k, l \in (1, \infty)$ proceed as follows:

(a) If $\|F(\overline{u}) - f^\delta\| \leq k\delta$ holds, choose $\alpha_{\text{dis}} = \infty$, which means $u_{\alpha_{\text{dis}}}^\delta := \overline{u} \in \mathcal{D}$.

(b) If $\|F(\overline{u}) - f^\delta\| > k\delta$, determine $\alpha =: \alpha_{\text{dis}} \in (0, \infty)$ such that

$$\|F(u_{\alpha_{\text{dis}}}^\delta) - f^\delta\| \leq k\delta \leq \|F(u_{\gamma_*}^\delta) - f^\delta\|, \tag{27}$$

for some $\gamma_* \in [\alpha_{\text{dis}}, l\alpha_{\text{dis}}]$.

We can now proceed to establish relevant bounds for norms involving the regularized solution $u_{\alpha_{\text{dis}}}^\delta$.

Corollary 4.10. *Let Assumptions 3.1 and 3.7 be satisfied. Further, let $k > 1$ be the constant specified in the discrepancy principle 4.9. Then, for $\alpha = \alpha_{\text{dis}}$ determined through the discrepancy principle, we have*

$$\|u_{\alpha_{\text{dis}}}^\delta - u^\dagger\|_{-a} \leq \frac{k+1}{c_a}\delta \qquad \textit{for each} \;\; 0 < \delta < \infty.$$

Proof. An application of estimate (20) and the triangle inequality yields

$$c_a\| u_{\alpha_{\text{dis}}}^\delta - u^\dagger \|_{-a} \leq \| F(u_{\alpha_{\text{dis}}}^\delta) - f^\dagger \| \leq \| F(u_{\alpha_{\text{dis}}}^\delta) - f^\delta \| + \delta \leq (k+1)\delta. \tag{28}$$

The right-hand side of (28) is bounded from above by $(k+1)\delta_0$ for $\delta \leq \delta_0$, validating the applicability of the estimate (20). Division of (28) by c_a completes the proof. □

As in the proof of the main theorem of the previous section, we require a result similar to that of Corollary 4.2, tailored to the situation where $\alpha = \alpha_{\text{dis}}$ is selected by the discrepancy principle. To establish such a result, we seek a lower bound for α_{dis}. The following result in Lemma 4.11 gives a lower bound for the inverse of a specific function, which is instrumental in determining this lower bound for α_{dis}.

Lemma 4.11. *Let $b, d > 0$ be finite constants. We define the function $\chi_{b,d}$ by*

$$\chi_{b,d} : (0, \|G\|] \to \mathbb{R}, \qquad t \mapsto t^{\frac{1}{b}} \left(-\ln ct\right)^{-d}. \tag{29}$$

Its inverse $\chi_{b,d}^{-1}$ satisfies

$$\chi_{b,d}^{-1}(t) \geq C_5 t^b \left(-\ln ct\right)^{bd} \qquad \text{for } \ 0 < t \leq \| G \|$$

and for some constant $C_5 > 0$.

Proof. The basic idea of this proof stems from the proof of Lemma 3.3 in [55]. Noticing that $\chi_{b,d}$ is continuous and strictly monotonically increasing with $\lim_{t\downarrow 0} \chi_{b,d}(t) = 0$, we set $\chi_{b,d}^{-1}(t) = \lambda$. By multiplying $t^b(-\ln ct)^{bd}/(t^b(-\ln ct)^{bd})$, we artificially expand this latter equation to

$$\chi_{b,d}^{-1}(t) = t^b(-\ln ct)^{bd} \frac{\lambda}{t^b(-\ln ct)^{bd}}. \tag{30}$$

We consider the fraction in (30) separately and substitute $t = \chi_{b,d}(\lambda)$:

$$\begin{aligned}
\frac{\lambda}{t^b(-\ln ct)^{bd}} &= \frac{\lambda}{\left(\lambda^{\frac{1}{b}} \left(-\ln c\lambda\right)^{-d}\right)^b \left(-\ln\left(c\lambda^{\frac{1}{b}} \left(-\ln c\lambda\right)^{-d}\right)\right)^{bd}} \\
&= \left(\frac{-\ln\left(c\lambda^{\frac{1}{b}} \left(-\ln c\lambda\right)^{-d}\right)}{-\ln c\lambda}\right)^{-bd} \\
&= \left(\frac{\frac{1}{b}\ln c\lambda + \ln(c^{1-\frac{1}{b}}) - d\ln\left(-\ln c\lambda\right)}{\ln c\lambda}\right)^{-bd} \\
&= \left(\frac{1}{b} + \frac{\ln(c^{1-\frac{1}{b}}) - d\ln\left(-\ln c\lambda\right)}{\ln(c\lambda)}\right)^{-bd}.
\end{aligned}$$

Now as $\lambda \downarrow 0$, or equivalently $t \downarrow 0$, we have

$$\lim_{\lambda\downarrow 0}\left(\frac{1}{b} + \frac{\ln c^{1-\frac{1}{b}} - d\ln\left(-\ln c\lambda\right)}{\ln c\lambda}\right)^{-bd} = b^{bd}.$$

It follows

$$\chi_{b,d}^{-1}(t) = t^b(-\ln ct)^{bd}(b^{bd} + \mathrm{o}(1)) \qquad \text{as } \ t \downarrow 0.$$

Arguing that $\chi_{b,d}^{-1}$ is continuous on the compact interval $[\varepsilon, \|G\|]$, for $\varepsilon > 0$ small, yields the existence of a constant $C_5 > 0$ such that

$$\chi_{b,d}^{-1}(t) \geq C_5 t^b(-\ln ct)^{bd} \qquad \text{for } \ 0 < t \leq \|G\|.$$

□

Lemma 4.12 (Lower bound for α_{dis})**.** *Assume that case (b) of Definition 4.9 applies and that Assumptions 3.1 and 3.7 are satisfied. Then there exist positive and finite constants C and δ_0 such that the regularization parameter α_{dis}, chosen according to Definition 4.9, satisfies*

$$\alpha_{\text{dis}} \geq C\delta^r \left(-\ln c\delta\right)^{\kappa r} = C\delta^r \varphi(\delta)^{-r} \qquad \textit{for} \ \ 0 < \delta \leq \delta_0,$$

with r given in Lemma 3.11.

In the following proof and throughout this thesis, the symbol $\sim$ denotes asymptotic equivalence in the usual sense. This means we consider two functions or sequences to be asymptotically equivalent if their behavior becomes increasingly similar as a certain parameter approaches infinity or some other limit. Specifically, $f(x) \sim g(x)$ as $x \to x_0$, for some functions f and g and some value x_0, means that $\lim_{x \to x_0} f(x)/g(x) = 1$.

Proof. Let C_4 and C_5 be the constants from Lemma 4.1 and Lemma 4.11, respectively. Further, let k, l, and γ_* be the constants from Definition 4.9. We consider the case $0 < l\alpha_{\text{dis}} \leq \|G\|$ first. Condition (27) in case (b) of Definition 4.9, the monotonic behavior of $\alpha \mapsto \|F(u_\alpha^\delta) - f^\delta\|$, as described in Lemma 4.6, and Lemma 4.1 yield

$$\begin{aligned} k\delta \leq \|F(u_{\gamma_*}^\delta) - f^\delta\| \leq \|F(u_{l\alpha_{\text{dis}}}^\delta) - f^\delta\| &\leq C_4 (l\alpha_{\text{dis}})^{\frac{a+p}{2a+2}} \varphi(l\alpha_{\text{dis}}) + \delta \\ &= C_4 \chi_{r,\kappa}(l\alpha_{\text{dis}}) + \delta, \end{aligned} \tag{31}$$

where $\chi_{r,\kappa}$ is defined as in (29). The function $\chi_{r,\kappa}$ can be applied because $l\alpha_{\text{dis}} \leq \|G\|$. We rearrange the estimate in (31) and write

$$C_4^{-1}(k-1)\delta \leq \chi_{r,\kappa}(l\alpha_{\text{dis}}).$$

The left-hand side is positive, because $k > 1$. Further, we can assume that $C_4^{-1}(k-1)\delta \leq \|G\|$, for $\delta \leq \delta_0$ small enough, which allows an application of the estimate for the inverse of $\chi_{r,\kappa}$ as given in Lemma 4.11:

$$l\alpha_{\text{dis}} \geq C_5 (C_4^{-1}(k-1)\delta)^r (-\ln c C_4^{-1}(k-1)\delta)^{\kappa r} \qquad \text{for} \ \ \delta \leq \delta_0.$$

We use the asymptotic equivalence

$$\ln c\delta \sim \ln c C_4^{-1}(k-1)\delta \qquad \text{as} \ \ \delta \downarrow 0$$

to deduce

$$\alpha_{\text{dis}} \geq C_6 \delta^r (-\ln c\delta)^{\kappa r} \qquad \text{for} \ \ \delta \leq \delta_0,$$

for some positive and finite constant C_6 and provided that δ_0 is small enough. In the case $\|G\| < l\alpha_{\text{dis}} < \infty$ the assertion still follows, because

$$\begin{aligned} & l\alpha_{\text{dis}} > \|G\| \\ \Leftrightarrow \quad & \alpha_{\text{dis}} > \frac{\|G\|}{l} \geq \frac{\|G\|}{l}\delta^r \left(-\ln c\delta\right)^{\kappa r} \qquad \text{for} \quad \delta \leq \delta_0, \end{aligned}$$

with δ_0 sufficiently small. □

The lower bound of Lemma 4.12 allows us to verify the next estimate in Lemma 4.13. This estimate is essential for proving this section's main result, stated in Theorem 4.14.

Lemma 4.13. *Let Assumptions 3.1 and 3.7 be satisfied and suppose that α_{dis} is determined through the discrepancy principle in Definition 4.9. Then, we can conclude that*

$$\|u^\delta_{\alpha_{\text{dis}}} - \overline{u}\|_1 = \mathcal{O}\big(\delta^{\frac{p-1}{a+p}}\varphi\left(\delta\right)^{\frac{a+1}{a+p}}\big) \qquad \textit{as} \quad \delta \downarrow 0.$$

Proof. To prove the statement, we consider three different scenarios for possible values of α_{dis}:

- $0 < \alpha_{\text{dis}} \leq \|G\|$,
- $\|G\| < \alpha_{\text{dis}} < \infty$, and
- $\alpha_{\text{dis}} = \infty$.

We start with the examination of the first case. Corollary 4.2 yields

$$\|u^\delta_{\alpha_{\text{dis}}} - \overline{u}\|_1 \leq C_4 \alpha_{\text{dis}}^{\frac{p-1}{2a+2}} \varphi(\alpha_{\text{dis}}) + \frac{\delta}{\sqrt{\alpha_{\text{dis}}}} \qquad \text{for} \quad 0 < \alpha_{\text{dis}} \leq \|G\|.$$

The term on the right-hand side is monotonically non-increasing for $0 < \alpha_{\text{dis}} \leq \|G\|$. Therefore, we insert the lower bound for α_{dis} given in Lemma 4.12 to obtain

$$\|u^\delta_{\alpha_{\text{dis}}} - \overline{u}\|_1 \leq C_4 (C\delta^r \varphi(\delta)^{-r})^{\frac{p-1}{2a+2}} \varphi(C\delta^r\varphi(\delta)^{-r}) + \frac{\delta}{\sqrt{C\delta^r\varphi(\delta)^{-r}}}, \tag{32}$$

with $0 < \delta \leq \delta_0$ small enough. Noticing that the lower bound $C\delta^r\varphi(\delta)^{-r}$ of α_{dis} up to the constant C coincides with the a priori parameter choice β_* for the auxiliary elements in Lemma 3.11, we can show in the same manner as in the proof of that lemma that

$$\varphi(C\delta^r\varphi(\delta)^{-r}) = \mathcal{O}(\varphi(\delta)) \qquad \text{as} \quad \delta \downarrow 0.$$

We use this asymptotic behavior together with

$$C_4(C\delta^r\varphi(\delta)^{-r})^{\frac{p-1}{2a+2}} = \widetilde{C}(\delta^{\frac{2a+2}{a+p}}\varphi(\delta)^{-\frac{2a+2}{a+p}})^{\frac{p-1}{2a+2}} = \widetilde{C}\delta^{\frac{p-1}{a+p}}\varphi(\delta)^{\frac{1-p}{a+p}},$$

where $\widetilde{C} := C_4 C^{\frac{p-1}{2a+2}}$, to deduce the following asymptotic behavior for the first summand of (32):

$$\begin{aligned} C_4(C\delta^r\varphi(\delta)^{-r})^{\frac{p-1}{2a+2}}\varphi(C\delta^r\varphi(\delta)^{-r}) &= \widetilde{C}\delta^{\frac{p-1}{a+p}}\varphi(\delta)^{-\frac{p-1}{a+p}}\mathcal{O}(\varphi(\delta)) \\ &= \mathcal{O}(\delta^{\frac{p-1}{a+p}}\varphi(\delta)^{\frac{a+1}{a+p}}) \qquad \text{as} \quad \delta \downarrow 0. \end{aligned}$$

For the second summand in (32), we calculate

$$\frac{\delta}{\sqrt{C\delta^r\varphi(\delta)^{-r}}} = \mathcal{O}(\delta^{1-\frac{a+1}{a+p}}\varphi(\delta)^{\frac{a+1}{a+p}}) = \mathcal{O}(\delta^{\frac{p-1}{a+p}}\varphi(\delta)^{\frac{a+1}{a+p}}) \qquad \text{as} \quad \delta \downarrow 0.$$

Hence, we conclude

$$\|u^\delta_{\alpha_{\text{dis}}} - \overline{u}\|_1 = \mathcal{O}(\delta^{\frac{p-1}{a+p}}\varphi(\delta)^{\frac{a+1}{a+p}}) \qquad \text{as} \quad \delta \downarrow 0.$$

Next, let $\|G\| < \alpha_{\text{dis}} < \infty$. Since

$$\|F(u^\delta_{\alpha_{\text{dis}}}) - f^\delta\|^2 + \alpha_{\text{dis}}\|u^\delta_{\alpha_{\text{dis}}} - \overline{u}\|_1^2 = T^\delta_{\alpha_{\text{dis}}}(u^\delta_{\alpha_{\text{dis}}}) \le T^\delta_{\alpha_{\text{dis}}}(\overline{u}) = \|F(\overline{u}) - f^\delta\|^2,$$

we have

$$\|u^\delta_{\alpha_{\text{dis}}} - \overline{u}\|_1 \le \sqrt{\frac{1}{\alpha_{\text{dis}}}}\|F(\overline{u}) - f^\delta\| = \mathcal{O}(1) \qquad \text{as} \quad \delta \downarrow 0.$$

Thus, especially

$$\|u^\delta_{\alpha_{\text{dis}}} - \overline{u}\|_1 = \mathcal{O}(\delta^{\frac{p-1}{a+p}}\varphi(\delta)^{\frac{a+1}{a+p}}) \qquad \text{as} \quad \delta \downarrow 0,$$

because $\lim_{\delta\downarrow 0}\delta^{\frac{p-1}{a+p}}\varphi(\delta)^{\frac{a+1}{a+p}} = \infty$.

The third case $\alpha_{\text{dis}} = \infty$ coincides with case (a) of Definition 4.9, respectively with the choice of $u^\delta_{\alpha_{\text{dis}}} = \overline{u}$. Thereupon, the claim of the lemma is an immediate consequence. □

The previous results provide all the necessary components to prove this section's main result, which expresses the convergence rate of the discrepancy principle in the given setting. This convergence rate is asymptotically optimal according to [28]. It is the same convergence rate as for the a priori parameter choice established in Section 4.1.

Theorem 4.14 (Convergence rates for the discrepancy principle)**.** *Under Assumptions 3.1 and 3.7, the convergence rate*

$$\|u^\delta_{\alpha_{\rm dis}} - u^\dagger\| = \mathcal{O}(\delta^{\frac{p}{a+p}} \varphi(\delta)^{\frac{a}{a+p}}) \qquad \textit{as} \;\; \delta \downarrow 0,$$

is attained for $\alpha_{\rm dis}$ *selected by the discrepancy principle, specified in Definition 4.9.*

Proof. We follow the decomposition procedure of $\|u^\delta_{\alpha_{\rm dis}} - u^\dagger\|$ as in the proof of Theorem 4.4. For the auxiliary elements $\hat{u}_\beta$, we use the regularization parameter β_* as in Lemma 3.11 and write $\hat{u}_{\beta_*}$. The triangle and the interpolation inequality (8) yield

$$\begin{aligned}\|u^\delta_{\alpha_{\rm dis}} - u^\dagger\| &\le \|u^\delta_{\alpha_{\rm dis}} - \hat{u}_{\beta_*}\| + \|\hat{u}_{\beta_*} - u^\dagger\| \\ &\le \underbrace{\|u^\delta_{\alpha_{\rm dis}} - \hat{u}_{\beta_*}\|_{-a}}_{=:I}{}^{\frac{1}{a+1}} \underbrace{\|u^\delta_{\alpha_{\rm dis}} - \hat{u}_{\beta_*}\|_1}_{=:II}{}^{\frac{a}{a+1}} + \underbrace{\|\hat{u}_{\beta_*} - u^\dagger\|}_{=:III}. \end{aligned} \tag{33}$$

First, we establish a bound for the expression in *I*. The triangle inequality yields

$$I = \|u^\delta_{\alpha_{\rm dis}} - \hat{u}_{\beta_*}\|_{-a} \le \|u^\delta_{\alpha_{\rm dis}} - u^\dagger\|_{-a} + \|u^\dagger - \hat{u}_{\beta_*}\|_{-a}. \tag{34}$$

According to Corollary 4.10, we have

$$\|u^\delta_{\alpha_{\rm dis}} - u^\dagger\|_{-a} = \mathcal{O}(\delta) \qquad \text{as} \;\; \delta \downarrow 0.$$

For the second summand in (34), item (ii) of Lemma 3.11 yields

$$\|u^\dagger - \hat{u}_{\beta_*}\|_{-a} = \mathcal{O}(\delta) \qquad \text{as} \;\; \delta \downarrow 0.$$

It follows

$$I = \mathcal{O}(\delta) \qquad \text{as} \;\; \delta \downarrow 0.$$

Next, we estimate the expression in *II*. The triangle inequality yields

$$II = \|u^\delta_{\alpha_{\rm dis}} - \hat{u}_{\beta_*}\|_1 \le \|u^\delta_{\alpha_{\rm dis}} - \overline{u}\|_1 + \|\overline{u} - \hat{u}_{\beta_*}\|_1.$$

It follows from Lemma 4.13 and item (iii) of Lemma 3.11 that

$$II = \mathcal{O}(\delta^{\frac{p-1}{a+p}} \varphi(\delta)^{\frac{a+1}{a+p}}) \qquad \text{as} \;\; \delta \downarrow 0.$$

An estimate for the term in *III* is given by item (i) of Lemma 3.11:

$$III = \|\hat{u}_{\beta_*} - u^\dagger\| = \mathcal{O}(\delta^{\frac{p}{a+p}} \varphi(\delta)^{\frac{a}{a+p}}) \qquad \text{as} \;\; \delta \downarrow 0.$$

Inserting the bounds for the components I, II, and III into estimate (33), we can confirm the assertion of the theorem:

$$\begin{aligned}\|u^\delta_{\alpha_{\rm dis}} - u^\dagger\| &\le \mathcal{O}(\delta)^{\frac{1}{a+1}} \mathcal{O}(\delta^{\frac{p-1}{a+p}} \varphi(\delta)^{\frac{a+1}{a+p}})^{\frac{a}{a+1}} + \mathcal{O}(\delta^{\frac{p}{a+p}} \varphi(\delta)^{\frac{a}{a+p}}) \\ &= \mathcal{O}(\delta^{\frac{p}{a+p}} \varphi(\delta)^{\frac{a}{a+p}}) \qquad \text{as} \;\; \delta \downarrow 0.\end{aligned}$$

□

5 Fourier series of a specific function

In this chapter, we seek a cosine series of the form

$$u_*(t) = \sum_{n=1}^{\infty} a_n \sqrt{2}\cos(n - \tfrac{1}{2})\pi t \tag{35}$$

for u_* given by

$$u_* : (0,1] \to [0,\infty), \qquad t \mapsto t^{-\mu}(-\ln \theta t)^{-\nu}, \qquad 0 < \mu, \theta < 1,\ \nu > 0. \tag{36}$$

The motivation for seeking such a representation becomes evident in Chapter 6. Representation (35) allows us to verify the smoothness conditions on u_* for the example considered in Chapter 6. Specifically, in Chapter 6, we use representation (35) to confirm that the considered solution $u^\dagger$ is smooth enough to meet the mixed source condition (18) but not sufficiently smooth to lie in X_1. Consequently, it is an appropriate example in the oversmoothing analysis.

We show first that the system $\{\sqrt{2}\cos(n-1/2)\pi t \mid 0 \le t \le 1\}_{n\in\mathbb{N}}$ indeed forms a complete orthonormal system, that is an orthonormal basis, of $L^2(0,1)$. This ensures that every element $u \in L^2(0,1)$ admits the representation

$$u(t) = \sum_{n=1}^{\infty} a_n \sqrt{2}\cos(n - \tfrac{1}{2})\pi t, \qquad 0 \le t \le 1,$$

with coefficients a_n given by

$$a_n = \sqrt{2}\int_0^1 u(t)\cos(n - \tfrac{1}{2})\pi t \,\mathrm{d}t, \qquad n \in \mathbb{N}.$$

The orthonormality can be shown straightforwardly, so we omit these steps here. To verify completeness of the system, we construct this representation from the original Fourier series representation. According to Fourier analysis, every square-integrable and 2π-periodic function

$$\widetilde{f} : [-\pi,\pi] \to \mathbb{C}$$

has the representation

$$\widetilde{f}(x) = \sum_{n\in\mathbb{Z}} \frac{1}{2\pi}\int_{-\pi}^{\pi} \widetilde{f}(t)\,\mathrm{e}^{-int}\,\mathrm{d}t\,\mathrm{e}^{inx}, \qquad -\pi \le x \le \pi.$$

For $u \in L^2(0,1)$ arbitrary, we define its even extension $\widetilde{u}$ by

$$\widetilde{u}(x) = \begin{cases} u(x), & 0 \le x \le 1, \\ u(-x), & -1 \le x < 0, \end{cases}$$

and consider

$$\widetilde{f}(x) = \widetilde{u}(\tfrac{x}{\pi}) \, \mathrm{e}^{\frac{ix}{2}}.$$

This function has the series representation

$$\widetilde{u}(\tfrac{x}{\pi}) \, \mathrm{e}^{\frac{ix}{2}} = \sum_{n \in \mathbb{Z}} \frac{1}{2\pi} \int_{-\pi}^{\pi} \widetilde{u}(\tfrac{t}{\pi}) \, \mathrm{e}^{\frac{it}{2}} \, \mathrm{e}^{-int} \, \mathrm{d}t \, \mathrm{e}^{inx}, \qquad -\pi \le x \le \pi.$$

Substitution of $x = x/\pi$ and $t = t/\pi$ yields

$$\widetilde{u}(x) \, \mathrm{e}^{\frac{i\pi x}{2}} = \sum_{n \in \mathbb{Z}} \frac{1}{2} \int_{-1}^{1} \widetilde{u}(t) \, \mathrm{e}^{-i(n-\frac{1}{2})\pi t} \, \mathrm{d}t \, \mathrm{e}^{i\pi n x}, \qquad -1 \le x \le 1,$$

which is equivalent to

$$\widetilde{u}(x) = \sum_{n \in \mathbb{Z}} \frac{1}{2} \int_{-1}^{1} \widetilde{u}(t) \, \mathrm{e}^{-i(n-\frac{1}{2})\pi t} \, \mathrm{d}t \, \mathrm{e}^{i(n-\frac{1}{2})\pi x}, \qquad -1 \le x \le 1.$$

Using

$$\mathrm{e}^{-i(n-\frac{1}{2})\pi t} = \cos(n - \tfrac{1}{2})\pi t - i \sin(n - \tfrac{1}{2})\pi t$$

and the fact that $\widetilde{u}$ is even, we can conclude that

$$\widetilde{u}(x) = \sum_{n \in \mathbb{Z}} \int_0^1 u(t) \cos(n - \tfrac{1}{2})\pi t \, \mathrm{d}t \cos(n - \tfrac{1}{2})\pi x, \qquad -1 \le x \le 1.$$

We split the sum into two sums, shift the index in the first sum, and use the property that cosine is an even function to obtain

$$\begin{aligned} \widetilde{u}(x) &= \sum_{n \in \mathbb{N}_0} \int_0^1 u(t) \cos(-n - \tfrac{1}{2})\pi t \, \mathrm{d}t \cos(-n - \tfrac{1}{2})\pi x \\ &\quad + \sum_{n \in \mathbb{N}} \int_0^1 u(t) \cos(n - \tfrac{1}{2})\pi t \, \mathrm{d}t \cos(n - \tfrac{1}{2})\pi x \\ &= 2 \sum_{n \in \mathbb{N}} \int_0^1 u(t) \cos(n - \tfrac{1}{2})\pi t \, \mathrm{d}t \cos(n - \tfrac{1}{2})\pi x, \qquad -1 \le x \le 1. \end{aligned}$$

Restricting the domain of $\widetilde{u}$ to the interval $[0,1]$ yields the desired representation:

$$u(x) = \sum_{n\in\mathbb{N}} \sqrt{2} \int_0^1 u(t)\cos(n-\tfrac{1}{2})\pi t \,\mathrm{d}t \sqrt{2}\cos(n-\tfrac{1}{2})\pi x, \qquad 0 \le x \le 1.$$

The coefficients for the function u_* in (36) cannot be determined explicitly, because the integral needed to calculate them is too complex to evaluate. However, the following result gives the asymptotic behavior of the coefficients for large $n \in \mathbb{N}$. At this point, we reiterate that the symbol $\sim$ denotes asymptotic equivalence in the usual manner, as mentioned prior to the proof of Lemma 4.12.

Theorem 5.1. *The coefficients a_n, $n \in \mathbb{N}$ of the series representation (35) for u_* given in (36) have the asymptotic behavior*

$$a_n \sim \sqrt{2}\pi^{\mu-1}(n-\tfrac{1}{2})^{\mu-1}(\ln n)^{-\nu}\Gamma(1-\mu)\sin\frac{\pi\mu}{2} \qquad \text{as} \quad n \to \infty.$$

The Greek letter Γ represents the Gamma function defined as

$$\Gamma(z) = \int_0^\infty t^{z-1}\,\mathrm{e}^{-t}\,\mathrm{d}t, \qquad z \in \mathbb{C}, \quad \mathrm{Re}(z) > 0,$$

where $\mathrm{Re}(z)$ denotes the real part of z.

Remark 5.2. The result of Theorem 5.1 can be generalized by replacing the expression $(-\ln\theta t)^{-\nu}$ in (36) with a function $L : (0,K] \to (0,\infty)$, $x \mapsto L(x)$, $K > 0$, that is slowly varying as $x \to 0$ and of bounded variation in every interval (ε, K) for $\varepsilon > 0$. In this context, we call a function slowly varying as $x \to 0$, if for all $\delta > 0$, the function $x^{-\delta}L(x)$ is decreasing and $x^{\delta}L(x)$ is increasing in some neighborhood of $x = 0$. Such a generalized result is given by Theorem 2.24 in [60, Chapter V]. The asymptotic behavior of the series coefficients there are given based on classical orthonormal basis functions $e_n(t) = \mathrm{e}^{int}$ for $n \in \mathbb{N}_0$ and $t \in (0,\pi]$, whereas we consider here the functions $\sqrt{2}\cos((n-1/2)\pi t)$ with $n \in \mathbb{N}$ and $t \in (0,1]$ as orthonormal basis functions. A sketch of a proof of that theorem is provided, primarily referring to the proof of Theorem 2.6 in the same chapter of [60]. This referenced theorem provides a converse result: for a given Fourier series of a function, it determines the asymptotic behavior of the explicit representation of that function.

To prepare the proof of Theorem 5.1, we present some required results. Firstly, we state the following mean value Theorem for integrals. This theorem is well known and it can be found, for example, in [11, p. 134].

Theorem 5.3 (Mean value Theorem for definite integrals)**.** *Let* $f : [a, b] \to \mathbb{R}$ *be a continuous function and let* $g : [a, b] \to [0, \infty)$ *be integrable, then there exists a number* $\xi \in [a, b]$ *such that*

$$\int_a^b f(x)g(x)\,\mathrm{d}x = f(\xi)\int_a^b g(x)\,\mathrm{d}x.$$

Proof. See for example [11]. □

A second result that we use to prove Theorem 5.1 is a conclusion of the following representation of the Gamma function. This representation can be found, for example, in [51, Theorem 2.12] or under equation (13·10) in [60, Chapter II].

Lemma 5.4. *Let* $z \in \mathbb{C}$ *with* $0 < \mathrm{Re}(z) < 1$*, then*

$$\mathrm{e}^{-\frac{\pi i z}{2}}\Gamma(z) = \int_0^\infty t^{z-1}\,\mathrm{e}^{-it}\,\mathrm{d}t. \tag{37}$$

Proof. The proof requires knowledge on complex analysis and line integrals, so we refer to the textbook [51] for a detailed proof. In [60], the procedure to deduce the result from the classical integral representation of the gamma function is explained, as well. □

Based on Lemma 5.4, we obtain the following result, which is essential for proving Theorem 5.1. It is available, for example, as equation (37) in [2, Section 1.6].

Corollary 5.5. *For* $x \in \mathbb{R}$*, with* $0 < x < 1$*, we have*

$$\sin\frac{\pi x}{2}\Gamma(1-x) = \int_0^\infty t^{-x}\cos t\,\mathrm{d}t.$$

Proof. By taking the real part of (37) with z replaced by $1 - x$, we deduce

$$\begin{aligned} &\mathrm{Re}\big(\mathrm{e}^{-\frac{\pi i(1-x)}{2}}\Gamma(1-x)\big) = \mathrm{Re}\Big(\int_0^\infty t^{-x}\,\mathrm{e}^{-it}\,\mathrm{d}t\Big), \qquad 0 < x < 1 \\ \Leftrightarrow\ &\cos\frac{\pi(1-x)}{2}\Gamma(1-x) = \int_0^\infty t^{-x}\cos t\,\mathrm{d}t, \qquad 0 < x < 1. \end{aligned}$$

Because

$$\cos\frac{\pi(1-x)}{2} = \sin\frac{\pi x}{2},$$

the statement of the corollary follows. □

Now we can proceed to the proof of Theorem 5.1.

Proof of Theorem 5.1. The proof is inspired by the approach in [60], where, as mentioned in Remark 5.2, a converse result is proven. To keep the notation simple, we use the notation ξ in each application of the mean value Theorem 5.3 for definite integrals, even though its value may differ depending on individual applications.

To give an approximate behavior of a_n as $n \to \infty$, we split the integral

$$a_n = \sqrt{2} \int_0^1 u_*(t) \cos(n - \tfrac{1}{2})\pi t \, \mathrm{d}t, \qquad n \in \mathbb{N},$$

into three parts:

$$a_n = \sqrt{2}(T_1 + T_2 + T_3),$$

with

$$\begin{aligned} T_1 &:= \int_0^{\frac{\omega}{(n-1/2)\pi}} t^{-\mu}(-\ln \theta t)^{-\nu} \cos(n - \tfrac{1}{2})\pi t \, \mathrm{d}t, \\ T_2 &:= \int_{\frac{\omega}{(n-1/2)\pi}}^{\frac{\Omega}{(n-1/2)\pi}} t^{-\mu}(-\ln \theta t)^{-\nu} \cos(n - \tfrac{1}{2})\pi t \, \mathrm{d}t, \quad \text{and} \\ T_3 &:= \int_{\frac{\Omega}{(n-1/2)\pi}}^{1} t^{-\mu}(-\ln \theta t)^{-\nu} \cos(n - \tfrac{1}{2})\pi t \, \mathrm{d}t, \end{aligned}$$

where $\omega > 0$ is sufficiently small and $\Omega > \omega$ is sufficiently big. Let $\varepsilon > 0$. Moreover, let $n_0 \geq \Omega/\pi + 1/2$ represent a sufficiently large natural number.

An application of Theorem 5.3 yields for some $\xi \in [0, \frac{\omega}{(n-1/2)\pi}]$ that

$$\begin{aligned} |T_1| &\leq \int_0^{\frac{\omega}{(n-1/2)\pi}} \left| t^{-\mu}(-\ln \theta t)^{-\nu} \cos(n - \tfrac{1}{2})\pi t \right| \mathrm{d}t \\ &\leq \int_0^{\frac{\omega}{(n-1/2)\pi}} t^{-\mu}(-\ln \theta t)^{-\nu} \, \mathrm{d}t = (-\ln \theta \xi)^{-\nu} \int_0^{\frac{\omega}{(n-1/2)\pi}} t^{-\mu} \, \mathrm{d}t \\ &\leq \Big(-\ln \theta \frac{\omega}{(n - \frac{1}{2})\pi}\Big)^{-\nu} \frac{1}{1-\mu} \Big(\frac{\omega}{(n - \frac{1}{2})\pi}\Big)^{1-\mu} \\ &\leq \varepsilon (n - \tfrac{1}{2})^{\mu - 1} \pi^{\mu - 1} (\ln n)^{-\nu}, \end{aligned}$$

for ω small and $n \geq n_0$. We continue with the examination of T_2:

$$T_2 = \int_{\frac{\omega}{(n-1/2)\pi}}^{\frac{\Omega}{(n-1/2)\pi}} t^{-\mu}(-\ln \theta t)^{-\nu} \cos(n - \tfrac{1}{2})\pi t \, \mathrm{d}t$$

$$
\begin{aligned}
&= \int_{\frac{\omega}{(n-1/2)\pi}}^{\frac{\Omega}{(n-1/2)\pi}} t^{-\mu}(\ln n)^{-\nu}\cos(n-\tfrac{1}{2})\pi t \,\mathrm{d}t \\
&\quad + \int_{\frac{\omega}{(n-1/2)\pi}}^{\frac{\Omega}{(n-1/2)\pi}} t^{-\mu}((-\ln\theta t)^{-\nu}-(\ln n)^{-\nu})\cos(n-\tfrac{1}{2})\pi t \,\mathrm{d}t \\
&= T_2^{'} + T_2^{''},
\end{aligned}
$$

with

$$
\begin{aligned}
T_2^{'} &:= \int_{\frac{\omega}{(n-1/2)\pi}}^{\frac{\Omega}{(n-1/2)\pi}} t^{-\mu}(\ln n)^{-\nu}\cos(n-\tfrac{1}{2})\pi t \,\mathrm{d}t \quad \text{and} \\
T_2^{''} &:= \int_{\frac{\omega}{(n-1/2)\pi}}^{\frac{\Omega}{(n-1/2)\pi}} t^{-\mu}((-\ln\theta t)^{-\nu}-(\ln n)^{-\nu})\cos(n-\tfrac{1}{2})\pi t \,\mathrm{d}t.
\end{aligned}
$$

Substituting $x = (n-1/2)\pi t$ in $T_2^{'}$ gives

$$
T_2^{'} = (\ln n)^{-\nu}(n-\tfrac{1}{2})^{\mu-1}\pi^{\mu-1}\int_{\omega}^{\Omega} x^{-\mu}\cos x \,\mathrm{d}x.
$$

For the integral, we have according to Corollary 5.5 that

$$
\int_{\omega}^{\Omega} x^{-\mu}\cos x \,\mathrm{d}x \in (\Gamma(1-\mu)\sin(\tfrac{1}{2}\pi\mu) \pm \varepsilon),
$$

for ω small and Ω big. Hence

$$
T_2^{'} \in (\gamma \pm \varepsilon)(\ln n)^{-\nu}(n-\tfrac{1}{2})^{\mu-1}\pi^{\mu-1},
$$

with

$$
\gamma := \Gamma(1-\mu)\sin\tfrac{1}{2}\pi\mu.
$$

An application of Theorem 5.3 to $T_2^{''}$, for $\xi \in \left[\frac{\omega}{(n-1/2)\pi}, \frac{\Omega}{(n-1/2)\pi}\right]$, results in

$$
\begin{aligned}
|T_2^{''}| &\le \int_{\frac{\omega}{(n-1/2)\pi}}^{\frac{\Omega}{(n-1/2)\pi}} |t^{-\mu}((-\ln\theta t)^{-\nu}-(\ln n)^{-\nu})\cos(n-\tfrac{1}{2})\pi t| \,\mathrm{d}t \\
&\le \int_{\frac{\omega}{(n-1/2)\pi}}^{\frac{\Omega}{(n-1/2)\pi}} \left|t^{-\mu}((-\ln\theta t)^{-\nu}-(\ln n)^{-\nu})\right| \mathrm{d}t \\
&= \left|(-\ln\theta\xi)^{-\nu}-(\ln n)^{-\nu}\right| \int_{\frac{\omega}{(n-1/2)\pi}}^{\frac{\Omega}{(n-1/2)\pi}} t^{-\mu}\,\mathrm{d}t \\
&= \frac{|(-\ln\theta\xi)^{-\nu}-(\ln n)^{-\nu}|}{1-\mu}\left(\left(\frac{\Omega}{(n-\frac{1}{2})\pi}\right)^{1-\mu} - \left(\frac{\omega}{(n-\frac{1}{2})\pi}\right)^{1-\mu}\right)
\end{aligned}
$$

$$< \varepsilon(\ln n)^{-\nu}(n-\tfrac{1}{2})^{\mu-1}\pi^{\mu-1} \qquad \text{for} \quad n \geq n_0.$$

The last inequality follows because the asymptotic equivalence

$$(\ln \tfrac{K}{\theta}(n-\tfrac{1}{2}))^{-\nu} \sim (\ln n)^{-\nu} \qquad \text{as} \quad n \to \infty$$

holds uniformly for $K \in \big[\pi/\Omega, \pi/\omega\big]$.

To give an estimate for T_3, we first note that the function u_*, given as in (36), is monotonically non-increasing on $(0, \theta^{-1}\,\mathrm{e}^{-\nu/\mu}]$ and monotonically increasing on $(\theta^{-1}\,\mathrm{e}^{-\nu/\mu}, \infty)$. Thus, depending on whether $t_* := \theta^{-1}\,\mathrm{e}^{-\nu/\mu}$ lies underneath, within, or above the interval

$$I := \big[\frac{\Omega}{(n-1/2)\pi}, 1\big],$$

we distinguish three cases that determine the piecewise monotonic behavior of u_* on I:

(a) If $t_* = \theta^{-1}\,\mathrm{e}^{-\nu/\mu} < \frac{\Omega}{(n-1/2)\pi}$, then u_* is monotonically increasing on I.

(b) If $\frac{\Omega}{(n-1/2)\pi} \leq t_* < 1$, then u_* is monotonically non-increasing on $[\frac{\Omega}{(n-1/2)\pi}, t_*]$ and monotonically increasing on $(t_*, 1]$.

(c) If $1 \leq t_*$, then u_* is monotonically non-increasing on I.

By partial integration, with the notation $u_*'(t) := \mathrm{d}/\,\mathrm{d}t(u_*(t))$, it follows

$$\begin{aligned} T_3 &= \int_{\frac{\Omega}{(n-1/2)\pi}}^{1} u_*(t)\cos(n-\tfrac{1}{2})\pi t\,\mathrm{d}t \\ &= \frac{1}{(n-\frac{1}{2})\pi}\Big(\sin((n-\tfrac{1}{2})\pi t)u_*(1) - \sin(\Omega)u_*\Big(\frac{\Omega}{(n-\frac{1}{2})\pi}\Big)\Big) \\ &\quad - \frac{1}{(n-\frac{1}{2})\pi}\int_{\frac{\Omega}{(n-1/2)\pi}}^{1} u_*'(t)\sin(n-\tfrac{1}{2})\pi t\,\mathrm{d}t. \end{aligned}$$

Thus

$$\begin{aligned} |T_3| &\leq \frac{1}{(n-\frac{1}{2})\pi}\Big(\Big|\sin((n-\tfrac{1}{2})\pi t)u_*(1) - \sin(\Omega)u_*\Big(\frac{\Omega}{(n-\frac{1}{2})\pi}\Big)\Big| \\ &\quad + \Big|\int_{\frac{\Omega}{(n-1/2)\pi}}^{1} u_*'(t)\sin(n-\tfrac{1}{2})\pi t\,\mathrm{d}t\Big|\Big) \\ &\leq \frac{1}{(n-\frac{1}{2})\pi}\Big(u_*(1) + u_*\Big(\frac{\Omega}{(n-\frac{1}{2})\pi}\Big) + \Big|\int_{\frac{\Omega}{(n-1/2)\pi}}^{1} u_*'(t)\sin(n-\tfrac{1}{2})\pi t\,\mathrm{d}t\Big|\Big) \\ &\leq \frac{1}{(n-\frac{1}{2})\pi}\Big(u_*(1) + u_*\Big(\frac{\Omega}{(n-\frac{1}{2})\pi}\Big) + \int_{\frac{\Omega}{(n-1/2)\pi}}^{1} |u_*'(t)|\,\mathrm{d}t\Big). \end{aligned} \tag{38}$$

Based on this latter estimate (38), we proceed to examine the three cases (a), (b), and (c), which determine the sign of the integral in (38). We examine case (a) first. In this case $u_*'(t) > 0$, hence

$$
\begin{aligned}
|T_3| &\le \frac{1}{(n-\frac{1}{2})\pi}\Big(u_*(1) + u_*\Big(\frac{\Omega}{(n-\frac{1}{2})\pi}\Big) + \int_{\frac{\Omega}{(n-1/2)\pi}}^{1} u_*'(t)\,\mathrm{d}t\Big) \\
&= \frac{1}{(n-\frac{1}{2})\pi}\Big(u_*(1) + u_*\Big(\frac{\Omega}{(n-\frac{1}{2})\pi}\Big) + u_*(1) - u_*\Big(\frac{\Omega}{(n-\frac{1}{2})\pi}\Big)\Big) \\
&= \frac{2}{(n-\frac{1}{2})\pi} u_*(1) \\
&= o((\ln n)^{-\nu}(n-\tfrac{1}{2})^{\mu-1}\pi^{\mu-1}) \qquad \text{as} \ \ n \to \infty.
\end{aligned}
$$

The last equation follows, because

$$
\lim_{n\to\infty} \frac{\frac{2}{(n-\frac{1}{2})\pi}u_*(1)}{(\ln n)^{-\nu}\big((n-\frac{1}{2})\pi\big)^{\mu-1}} = 2(-\ln\theta)^{-\nu} \lim_{n\to\infty} \frac{(\ln n)^{\nu}}{\big((n-\frac{1}{2})\pi\big)^{\mu}} = 0.
$$

Now we examine case (c). In this case $u_*'(t) \le 0$, hence

$$
\begin{aligned}
|T_3| &\le \frac{1}{(n-\frac{1}{2})\pi}\Big(u^{\dagger}(1) + u_*\Big(\frac{\Omega}{(n-\frac{1}{2})\pi}\Big) - \int_{\frac{\Omega}{(n-1/2)\pi}}^{1} u_*'(t)\,\mathrm{d}t\Big) \\
&= \frac{1}{(n-\frac{1}{2})\pi}\Big(u_*(1) + u_*\Big(\frac{\Omega}{(n-\frac{1}{2})\pi}\Big) - u_*(1) + u_*\Big(\frac{\Omega}{(n-\frac{1}{2})\pi}\Big)\Big) \\
&= \frac{2}{(n-\frac{1}{2})\pi} u_*\Big(\frac{\Omega}{(n-\frac{1}{2})\pi}\Big) \\
&= 2((n-\tfrac{1}{2})\pi)^{\mu-1}\Omega^{-\mu}\Big(\ln\frac{(n-\frac{1}{2})\pi}{\theta\Omega}\Big)^{-\nu} \\
&\le \varepsilon(\ln n)^{-\nu}(n-\tfrac{1}{2})^{\mu-1}\pi^{\mu-1},
\end{aligned}
$$

for Ω big and $n \ge n_0$. For case (b) we split the integral in (38) into two integrals :

$$
\begin{aligned}
&\int_{\frac{\Omega}{(n-1/2)\pi}}^{t_*} |u_*'(t)|\,\mathrm{d}t + \int_{t_*}^{1} |u_*'(t)|\,\mathrm{d}t \\
\le& -u_*(t_*) + u_*\Big(\frac{\Omega}{(n-\frac{1}{2})\pi}\Big) + u_*(1) - u_*(t_*) \\
\le& u_*\Big(\frac{\Omega}{(n-\frac{1}{2})\pi}\Big) + u_*(1).
\end{aligned}
$$

An upper bound of $|T_3|$ is thus given by

$$
|T_3| \le \frac{2}{(n-\frac{1}{2})\pi}\Big(u_*\Big(\frac{\Omega}{(n-\frac{1}{2})\pi}\Big) + u_*(1)\Big)
$$

$$\leq \varepsilon(\ln n)^{-\nu}(n-\tfrac{1}{2})^{\mu-1}\pi^{\mu-1},$$

for Ω big and $n \geq n_0$. Collecting the estimates for T_1, T_2, and T_3, it follows

$$\sqrt{2}(\gamma-4\varepsilon)(n-\tfrac{1}{2})^{\mu-1}\pi^{\mu-1}(\ln n)^{-\nu} < a_n < \sqrt{2}(\gamma+4\varepsilon)(n-\tfrac{1}{2})^{\mu-1}\pi^{\mu-1}(\ln n)^{-\nu}$$

for $n \geq n_0$. As $\varepsilon > 0$ approaches 0, the assertion of the theorem follows. $\square$

6 Numerical results

In this chapter, we present the numerical results, both for α chosen a priori and according to the discrepancy principle. The discrepancy principle is realized by Algorithm 1 that is based on the procedure described in [24].

As an example, we consider the exponential growth model. In the oversmoothing setting this example is also considered in the appendix of [23], as well as in [19], and in [20] in a Banach space setting. Initially, it was presented in [14, Section 3.1]. In what follows, we denote by $u' \in L^1(0,1)$ the weak derivative of a function $u \in L^1(0,1)$, that means

$$\int_0^1 u(t) f'(t)\,\mathrm{d}t = -\int_0^1 u'(t) f(t)\,\mathrm{d}t,$$

for all infinitely differentiable functions f with $f(0) = f(1) = 0$. The exponential growth model is described by the initial value problem

$$f'(t) = u(t) f(t), \qquad 0 \le t \le T, \quad f(0) = f_0, \tag{39}$$

for some positive constants T and f_0. The function $f : [0,T] \to [0,\infty)$ is interpreted as the time dependent size of a population. It is assumed that the population size commences with an initial population size f_0 and that it can be measured throughout the observation period $[0,T]$. The function $u : [0,T] \to \mathbb{R}$ is interpreted as the time dependent growth rate. Defining the forward operator $F : L^2(0,T) \to L^2(0,T)$ by

$$[F(u)](t) = f_0 \exp\left(\int_0^t u(z)\mathrm{d}z\right), \qquad 0 \le t \le T,$$

we can rewrite the initial value problem (39) as the operator equation

$$F(u) = f,$$

which is ill-posed on the whole domain $\mathcal{D}(F) := L^2(0,T)$. The aim is to determine the growth rate

$$u(t), \qquad 0 \le t \le T,$$

from noisy observations $f^\delta \in L^2(0,T)$. To define an appropriate Hilbert scale, we introduce the integration operator

$$J : X \to X, \qquad [Ju](t) := \int_0^t u(z)\mathrm{d}z, \qquad 0 \le t \le T.$$

By means of this operator, we define the linear, selfadjoint, and unbounded operator

$$B = (J^*J)^{-\frac{1}{2}}.$$

Moreover, we set

$$T = 1 \quad \text{and} \quad f_0 = 1.$$

The scale generated by B is then given by

$$X_p = \mathcal{D}(B^p) = \mathcal{R}((J^*J)^{\frac{p}{2}}) = \mathcal{R}((J^*)^p), \qquad 0 < p \leq 1,$$

where $(J^*)^p = (J^p)^*$, with $0 < p \leq 1$, is the adjoint of a Riemann-Liouville fractional integral operator J^p. Depending on the value of p, the spaces X_p exhibit the following relation to Sobolev or fractional Sobolev spaces:

$$X_p = \begin{cases} H^p(0,1) & \text{for } 0 < p < 1/2 \\ \{u \in H^{1/2}(0,1) : \int_0^1 |u(x)|^2/(1-x)\,\mathrm{d}x < \infty\} & \text{for } p = 1/2 \\ \{u \in H^p(0,1) : u(1) = 0\} & \text{for } 1/2 < p < 3/2 \\ \{u \in H^2(0,1) : u'(0) = 0,\ u(1) = 0\} & \text{for } p = 2 \end{cases},$$

see [12], or the references therein, that is [13, Lemma 8] and [36, Example 2.1.5]. The Sobolev spaces $H^1(0,1)$ and $H^2(0,1)$ consist of all square integrable functions whose first and second weak derivatives, respectively, exist and lie in $L^2(0,1)$ as well. For non-integer values $p > 0$, the Sobolev spaces $H^p(0,1)$ can be defined by an interpolation argument, and the corresponding norms can be defined through Fourier analysis. See [12] or [1] for more details. The appendix of [23] demonstrates that this example satisfies Assumption 3.7 (f) with $a = 1$. In the next steps, we verify that the solution

$$u^\dagger(t) = (1-\mu)t^{-\mu}(-\ln\theta t)^{-\nu} + \nu t^{-\mu}(-\ln\theta t)^{-\nu-1}, \qquad 0 < t \leq 1,$$

with positive constants

$$\mu < 1,\ \nu,\ \text{and}\ \theta < 1,$$

having

$$f^\dagger(t) = \exp(t^{1-\mu}(-\ln\theta t)^{-\nu})$$

as an image under the operator F, is an appropriate example for the numerical analysis of the oversmoothing Tikhonov regularization. Specifically, we show that $u^\dagger \notin X_1$, or equivalently $u^\dagger \notin \mathcal{R}(B^{-1})$, but $u^\dagger \in \mathcal{R}(G^p\varphi(G))$ if the constants μ, ν, κ, and p are chosen properly. As presented in [18, p. 101], the singular value decomposition for the integration operator J is given by

$$Ju = \sum_{i=1}^{\infty} \sigma_i \langle u, u_i\rangle v_i,$$

with singular values

$$\sigma_i = \frac{1}{(i-\frac{1}{2})\pi}, \qquad i \in \mathbb{N},$$

and eigenfunctions

$$u_i(t) = \sqrt{2}\cos(i-\tfrac{1}{2})\pi t, \quad v_i(t) = \sqrt{2}\sin(i-\tfrac{1}{2})\pi t, \qquad 0 \le t \le 1, \quad i \in \mathbb{N}.$$

We have seen in the preliminary steps of Chapter 5 that the system $\{u_i(t)\}_{i\in\mathbb{N}}$ forms an orthonormal basis of the space $L^2(0,1)$. Analogously, it can be shown that the system $\{v_i(t)\}_{i\in\mathbb{N}}$ also forms an orthonormal basis of that space. The singular value decomposition of J allows us to formulate the operators G and $G^p\varphi(G)$ as

$$Gu = \sum_{i=1}^{\infty} \sigma_i^4 \langle u, u_i\rangle u_i \quad \text{and} \quad G^p\varphi(G)u = \sum_{i=1}^{\infty} \sigma_i^{4p}\varphi(\sigma_i^4)\langle u, u_i\rangle u_i,$$

respectively. According to Theorem 5.1, the solution $u^\dagger$ has the representation

$$u^\dagger = \sum_{i=1}^{\infty} |\langle u^\dagger, u_i\rangle| u_i = \sum_{i=1}^{\infty} (a_i + b_i) u_i$$

with

$$\begin{aligned} a_i &\sim \sqrt{2}\gamma(1-\mu)\pi^{\mu-1}(i-\tfrac{1}{2})^{\mu-1}(\ln \tfrac{\pi}{c^{1/4}}(i-\tfrac{1}{2}))^{-\nu} \quad \text{and} \\ b_i &\sim \sqrt{2}\gamma\nu\pi^{\mu-1}(i-\tfrac{1}{2})^{\mu-1}(\ln \tfrac{\pi}{c^{1/4}}(i-\tfrac{1}{2}))^{-\nu-1} \end{aligned} \tag{40}$$

as $i \to \infty$, where $\mu < 1$, $\nu > 0$, $c \le (\pi/2)^4$, and $\gamma = \Gamma(1-\mu)\sin(\pi\mu/2)$. Without loss of generality, we have adjusted the expression in the logarithm. The Picard criterion, which is given for example in [8, Theorem 2.8], states that for the compact operator J with singular system $\{\sigma_i, u_i, v_i\}_{i\in\mathbb{N}}$, an element $u \in \overline{\mathcal{R}(J)}$ belongs to $\mathcal{R}(J)$ if and only if the series $\sum_{i=1}^{\infty} \sigma_i^{-2}|\langle u, v_i\rangle|^2$ converges. We apply the Picard criterion to the compact operator $G^{p/4}$ to determine which values for μ and ν yield $u^\dagger \notin X_q = \mathcal{D}(B^q) = \mathcal{R}(G^{q/4})$. In a first step, we realize that because of $a_i \ge b_i$ for all sufficiently large $i \in \mathbb{N}$, the following estimate holds:

$$C + \sum_{i=i_0}^{\infty} \frac{a_i^2}{\sigma_i^{2q}} \le \sum_{i=1}^{\infty} \frac{|\langle u^\dagger, u_i\rangle|^2}{\sigma_i^{2q}} = \sum_{i=1}^{\infty} \frac{(a_i+b_i)^2}{\sigma_i^{2q}} \le C + 4\sum_{i=i_0}^{\infty} \frac{a_i^2}{\sigma_i^{2q}}, \tag{41}$$

where C is a positive and finite constant representing the value of the partial sum $\sum_{i=1}^{i_0}(a_i+b_i)^2/\sigma_i^{2q}$, and $i_0 \in \mathbb{N}$ is assumed to be large enough such that the asymptotic behavior of a_i as $i \to \infty$ applies. Therefore, it is sufficient to examine the series $\sum_{i=i_0}^{\infty} a_i^2/\sigma_i^{2q}$ with a_i replaced by the sequence in (40) describing its asymptotic behavior:

$$2(\gamma(1-\mu))^2 \sum_{i=i_0}^{\infty} \frac{(\pi(i-\frac{1}{2}))^{2(q+\mu-1)}}{(\ln \frac{\pi}{c^{1/4}}(i-\frac{1}{2}))^{2\nu}}. \tag{42}$$

If $q+\mu=1$, the series in (42) dominates the harmonic series, which diverges. If $q+\mu>1$, the sequence within the series is not a null sequence and hence the series diverges. If $q+\mu<1$, the sequence is non-increasing and positive, so we can apply the integral test and analyze the behavior of

$$2(\gamma(1-\mu))^2 \int_{i_0}^{\infty} \frac{(\pi(x-\frac{1}{2}))^{2(q+\mu-1)}}{(\ln \frac{\pi}{c^{1/4}}(x-\frac{1}{2}))^{2\nu}}\, \mathrm{d}x.$$

Substitution of $\mathrm{e}^{\,t} = \frac{\pi}{c^{1/4}}(x-\frac{1}{2})$ leads to

$$2(\gamma(1-\mu))^2 \frac{c^{\frac{1}{2}(p+\mu-\frac{1}{2})}}{\pi} \int_{\ln(c^{-1/4}\pi(i_0-1/2))}^{\infty} \mathrm{e}^{\,2t(q+\mu-1)+t} t^{-2\nu}\, \mathrm{d}t,$$

which converges if and only if either $q+\mu<1/2$, or $q+\mu=1/2$ and $\nu>1/2$. Otherwise, the series diverges. Thus $u^\dagger \in X_q$, if $q<1/2-\mu$, or if $q=1/2-\mu$ and $\nu>1/2$. Conversely, $u^\dagger \notin X_q$ if $q>1/2-\mu$ or if $q=1/2-\mu$ and $\nu \le 1/2$. Therefore, the choice $\mu=1/2$ yields $u^\dagger \notin X_q$ for any $0<q\le 1$.

Now we demonstrate that with $\overline{u}=0$, the solution $u^\dagger$ satisfies the source condition (18). To do this, we apply the Picard criterion to $G^p\varphi(G)$. Following the same reasoning as in (41), it is sufficient to examine the behavior of the series beginning at an index i_0 and with the asymptotic behavior of a_i^2 replacing $|\langle u^\dagger, u_i\rangle|^2$. Specifically, we examine the behavior of

$$\begin{aligned}\sum_{i=i_0}^{\infty} \frac{a_i^2}{\sigma_i^{8p}\varphi(\sigma_i^4)^2} &\sim \sum_{i=i_0}^{\infty} 2(\gamma(1-\mu))^2 \frac{(\pi(i-\frac{1}{2}))^{2(\mu-1+4p)}}{(\ln \frac{\pi}{c^{1/4}}(i-\frac{1}{2}))^{2\nu}(-\ln(c\sigma_i^4))^{-2\kappa}} \\ &= 2(\gamma(1-\mu))^2 4^{2\kappa} \sum_{i=i_0}^{\infty} \frac{(\pi(i-\frac{1}{2}))^{2(\mu-1+4p)}}{(\ln \frac{\pi}{c^{1/4}}(i-\frac{1}{2}))^{2(\nu-\kappa)}}.\end{aligned}$$

A first necessary condition for the sequence within the series to be a null sequence is $\mu-1+4p<0$. So we assume that $\mu-1+4p<0$. The sequence within the series is positive and non-increasing for

$$i \ge \max\Big\{1, \frac{c^{\frac{1}{4}}}{\pi} \exp\Big(\frac{\nu-\kappa}{\mu-1+4p}\Big) + \tfrac{1}{2}\Big\} =: i^*.$$

Without loss of generality, we can assume that $i_0 \geq i^*$ and apply the integral test followed by substitution of $\mathrm{e}^{t} = (x - 1/2)\pi/c^{1/4}$:

$$\overline{\gamma} \int_{i_0}^{\infty} \frac{(\pi(x-\frac{1}{2}))^{2(\mu-1+4p)}}{(\ln \frac{\pi}{c^{1/4}}(x-\frac{1}{2}))^{2(\nu-\kappa)}} \,\mathrm{d}x = \overline{\gamma} \int_{t_0}^{\infty} c^{\frac{1}{2}(\mu-\frac{1}{2}+4p)} \pi^{-1} \,\mathrm{e}^{2t(\mu-\frac{1}{2}+4p)} t^{2(\kappa-\nu)} \,\mathrm{d}t,$$

where $t_0 := \ln(i_0 - 1/2)\pi/c^{1/4}$ and $\overline{\gamma} := 2(\gamma(1-\mu))^2 4^{2\kappa}$. The integral converges if and only if either $\mu < 1/2 - 4p$, or $\mu = 1/2 - 4p$ and $\kappa < \nu - 1/2$. Hence, according to the Picard criterion, $u^\dagger \in \mathcal{R}(G^p \varphi(G))$ if and only if $\mu < 1/2 - 4p$, or $\mu = 1/2 - 4p$ and $\kappa < \nu - 1/2$.

We are interested in the case $u^\dagger \notin X_q$ for any $0 < q \leq 1$, in which the choice $\mu = 1/2$ is necessary. In this case, the exponent p in the source condition (18) has to be zero and the condition $\kappa < \nu - 1/2$ has to hold. If alternatively $\mu < 1/2$, then source condition (18) applies for $p < 1/4(1/2 - \mu)$. In this case, we have to take into account that the Hölder-type source condition $u^\dagger \in X_q$ applies for $0 \leq q < 1/2 - \mu$, which might lead to faster convergence rates of the form $\delta^{q/(a+q)}$.

The numerical computations were carried out in RStudio [52] under the details summarized below.

- The minimization step utilized the "`fmincon`" command provided by the "`pracma`" package [4].

- We discretized the observation period $[0.001, 1]$ by $N = 100$ grid points. The value 0.001 was chosen instead of 0 as a left endpoint because of the singularity of $u^\dagger$ at 0.

- The data $f^\delta \in \mathbb{R}^N$ were perturbed by a normalized random vector $\Lambda \in [-1, 1]^N$ with entries following a standard normal distribution: $f^\delta = f^\dagger + \Lambda$.

- For $u^\dagger$, we chose the exponents $\mu = 0.5$ and $\nu = 1.25$, and the constant $\theta = 0.6$. This choice satisfies the source condition for $p = 0$ and $\kappa < 0.75$.

- For the function φ, we chose the constants $\kappa = 0.7$ and $c = 0.6$.

- The initial constants for the discrepancy principle in Algorithm 1 were chosen as $\alpha^{(0)} = 0.001$, $\vartheta = 2$, and $k = 5$.

- The integrals were discretized using the trapezoidal rule and the derivatives required to compute the norm $\|\cdot\|_1$ were obtained by finite difference approximations.

Algorithm 1: Sequential discrepancy principle

Result: Parameter α_{dis} selected according to the discrepancy principle in Definition 4.9.

Choose an initial guess $\alpha^{(0)}$ and constants $\vartheta, k > 1$;
if $\|F(u^\delta_{\alpha^{(0)}}) - f^\delta\| \geq k\delta$ **then**
  repeat
    $\alpha^{(i)} = \vartheta^{-i}\alpha^{(0)}$, $i = 1, 2, \dots,$
    determine $u^\delta_{\alpha^{(i)}}$ and $\|Fu^\delta_{\alpha^{(i)}} - f^\delta\|$
  until $\|F(u^\delta_{\alpha^{(i)}}) - f^\delta\| \leq k\delta \leq \|F(u^\delta_{\alpha^{(i-1)}}) - f^\delta\|$ for the first time;
set $\alpha_{\text{dis}} = \alpha^{(i)}$;
if $\|F(u^\delta_{\alpha^{(0)}}) - f^\delta\| \leq k\delta$ **then**
  repeat
    $\alpha^{(i)} = \vartheta^{i}\alpha^{(0)}$, $i = 1, 2, \dots,$
    determine $u^\delta_{\alpha^{(i)}}$ and $\|F(u^\delta_{\alpha^{(i)}}) - f^\delta\|$
  until $\|F(u^\delta_{\alpha^{(i-1)}}) - f^\delta\| \leq k\delta \leq \|F(u^\delta_{\alpha^{(i)}}) - f^\delta\|$ for the first time;
set $\alpha_{\text{dis}} = \alpha^{(i-1)}$;
return α_{dis}

δ	α_*	$\|u^\delta_{\alpha_*} - u^\dagger\|$	$\frac{\|u^\delta_{\alpha_*} - u^\dagger\|}{\|u^\dagger\|}$	$\delta^{\frac{p}{a+p}}\varphi(\delta)^{\frac{a}{a+p}}$
0.1000	$1.81 \cdot 10^{-3}$	2.1595	0.7640	0.4848
0.0500	$2.10 \cdot 10^{-4}$	1.7272	0.6110	0.4155
0.0250	$2.17 \cdot 10^{-5}$	1.3521	0.4784	0.3662
0.0125	$2.08 \cdot 10^{-6}$	1.0516	0.3720	0.3291
0.0062	$1.89 \cdot 10^{-7}$	0.8459	0.2993	0.2999
0.0031	$1.64 \cdot 10^{-8}$	0.7316	0.2588	0.2764

Table 1: A priori parameter choices for $\alpha = \alpha_*$ and errors of the regularized solutions for different values of δ.

Table 1 and Table 2 present the numerical results for the oversmoothing Tikhonov regularization. Table 1 shows the results for $\alpha = \alpha_*$ chosen a priori as in Theorem 4.5. Table 2 presents the results for the discrepancy principle with $\alpha = \alpha_{\text{dis}}$ determined sequentially through Algorithm 1. In each table, the second column illustrates values of the regularization parameter α for the different noise levels in column 1. The third and the fourth column show the error and the relative error of the regularized solution, respectively. The fifth column includes the values for established converge rate, which is given as $\varphi(\delta)$. The values for α are rounded to three significant digits in scientific notation. All other values are rounded to four decimal places.

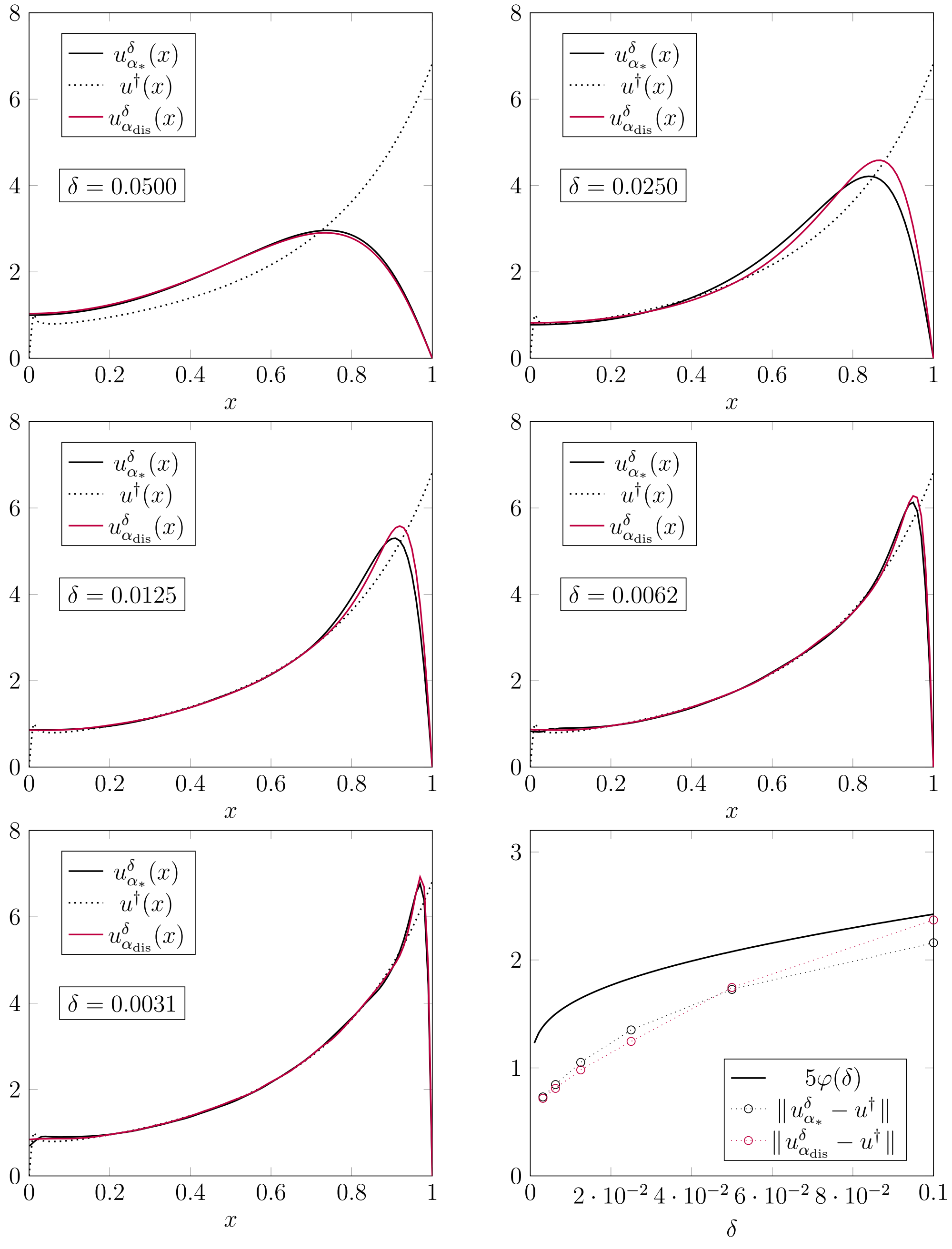

Figure 1: The upper graphs and the graphs on the bottom left show the regularized solutions for different noise levels δ, compared to the actual solution $u^\dagger$. The bottom-right graph shows a comparison of the errors with the established convergence rate.

δ	α_{dis}	$\| u^\delta_{\alpha_{\mathrm{dis}}} - u^\dagger \|$	$\frac{\| u^\delta_{\alpha_{\mathrm{dis}}} - u^\dagger \|}{\| u^\dagger \|}$	$\delta^{\frac{p}{a+p}} \varphi(\delta)^{\frac{a}{a+p}}$
0.1000	$2.00 \cdot 10^{-3}$	2.3705	0.8386	0.4848
0.0500	$2.50 \cdot 10^{-4}$	1.7453	0.6175	0.4155
0.0250	$1.56 \cdot 10^{-5}$	1.2457	0.4407	0.3662
0.0125	$1.95 \cdot 10^{-6}$	0.9821	0.3475	0.3291
0.0063	$1.22 \cdot 10^{-7}$	0.8110	0.2869	0.2999
0.0031	$3.05 \cdot 10^{-8}$	0.7187	0.2543	0.2764

Table 2: Parameters $\alpha = \alpha_{\mathrm{dis}}$ determined through the discrepancy principle for different values of δ and the corresponding errors of the regularized solutions.

Figure 1 illustrates the regularized solutions for $\alpha = \alpha_*$ chosen a priori and $\alpha = \alpha_{\mathrm{dis}}$ determined by the discrepancy principle. The figure consists of six graphs: the upper four and the bottom-left graphs show the regularized solutions as solid lines for the different indicated noise levels in comparison to the actual solution $u^\dagger$, which is displayed as a dotted line. Regularized solutions for $\alpha = \alpha_*$ chosen a priori are colored black, while those resulting for the discrepancy principle are colored purple. In the bottom-right graph, the errors for different noise levels, as detailed in Tables 1 and 2, are depicted by circles. Consistent with the coloring of the regularized solutions, black-colored circles correspond to errors for the a priori parameter choice, and purple circles correspond to errors for the discrepancy principle. For comparison, the function $\varphi(\delta)$, representing the convergence rate, is included as a solid line with $\kappa = 0.7$ and $c = 0.6$. Note that it is scaled by a factor of 5. The figure illustrates how the approximations for $u^\dagger$ improve for decreasing noise levels δ. The solution $u^\dagger$ is approximated quite precisely except near the right endpoint, where the boundary condition $u^\delta_\alpha(1) = 0$ causes the regularized solutions to deviate from the actual solution $u^\dagger$. The tables and the graphs confirm our theoretical findings and reveal that both considered parameter choice strategies lead to very similar solutions.

Part II
Oversmoothing Tikhonov regularization – the finite-dimensional setting

Discretization is a fundamental step in practically applying regularization methods to solve ill-posed problems. In many real-world scenarios, the problems we encounter are continuous in nature, but for computational purposes, we need to transform them into a discrete form that can be handled by computers. Discretization involves approximating continuous functions, operators, or equations by a finite set of discrete elements or values.

We focus on the discretization with projection methods.

For linear ill-posed problems, projection methods have been studied thoroughly:

While projection methods, such as the collocation or Galerkin methods, are a general tool for solving linear operator equations [35, Chapter 13], Natterer [42] demonstrated that projection methods can serve as regularization methods by projecting onto a suitable subspace. The corresponding regularization parameter is governed by the discretization level. Examples of such methods include the least squares method, the Ritz method, and the collocation method. Another classical example is the truncated singular value decomposition for compact operators.

Fine discretization, however, leaves opportunities for incorporating regularization methods. The order in which discretization and regularization are applied affects the analysis. Engl et al [8, p. 127] provide an overview of research concerning various approaches on combining regularization and discretization by projection. We briefly recapitulate this overview considering the linear operator equation

$$Au = f$$

for a linear operator $A : X \to Y$, and orthogonal projection operators $P_h : X \to V_h$ and $Q_h : Y \to Y_h$, where $V_h \subset X$ and $Y_h \subset Y$ are suitable (finite-dimensional) subspaces.

The approach

$$A_h = AP_h$$

was analyzed in [15, Chapter 4]. For Tikhonov regularization this approach is equivalent to minimizing the Tikhonov functional over the subspace V_h. Another approach

$$A_h = Q_hA$$

corresponds to the case, where A is only known approximately. This was studied by King and Neubauer [32] and Vainikko [58, Chapter 5]. Throughout Section 5.2 in [8], the advantages of this approach over the strategy $A_h = AP_h$ are remarked.

Yet another approach

$$A_h = Q_h A P_h$$

was studied by Plato and Vainikko [46].

Now we turn to research within discretization by projection methods for *nonlinear* ill-posed and inverse problems. Kaltenbacher [31] showed that for nonlinear ill-posed problems, projection methods can be considered as regularization methods as well. She considered a priori as well as a posteriori choice strategies for the discretization level. Another projection-based approach is to linearize the nonlinear problem using Newton methods and regularize the linearized equation within each Newton step by projection. See, for example, [30] and [29]. When it comes to combination of regularization and projection methods, especially iterative multilevel discretization methods have been studied. These methods combine regularization and discretization across various discretization levels. More precisely, a discretization level is initially fixed and a regularization method is applied. The obtained regularized solution is used as an initial guess for a finer or coarser discretization. Scherzer [53] and Ramlau [50] analyzed Landweber's regularization in this context. In [7], an iteratively regularized Gauss–Newton method turned out to be advantageous compared to Landweber's method.

Neubauer [43], [45] examined Tikhonov regularization in Hilbert scales in combination with finite-dimensional approximation. Additionally, in his work [44], he explored an approach closely related to ours, involving the minimization of the Tikhonov functional over a finite-dimensional subspace.

The oversmoothing situation has not been investigated. Our objective is to study the oversmoothing Tikhonov regularization in a finite-dimensional setting. In this setting, we prove convergence rates for a priori parameter choices for α and for the discretization level under a Hölder-type source condition. As in Part I, this analysis relies on the use of auxiliary elements. In Section 7.3, we propose an alternative approach to the auxiliary elements in Section 3.3. The structure of this part is analogous to the structure of Part I: We begin with a chapter on fundamental requirements. After that, Chapter 8 establishes convergence rates, which are confirmed in Chapter 9 by numerical results.

Now, we proceed to specify our framework: For $h > 0$, which in general describes the step size of the discretization, let

$$V_h \subset \mathcal{D} = \mathcal{D}(F) \cap X_1$$

be a finite-dimensional subspace of X_1, where $X_1 \subset X$, equipped with the norm $\| \cdot \|_1$, is specified in Section 2.4. We consider the orthogonal projection

$$P_h : X \to V_h.$$

Given an initial guess $\overline{u} \in \mathcal{D}(F) \cap X_1$ and $\alpha > 0$, we minimize the Tikhonov functional

$$T_\alpha^\delta(u) = \| F(u) - f^\delta \|^2 + \alpha \| u - \overline{u} \|_1^2,$$

with respect to $u \in V_h$, to obtain

$$u_{h,\alpha}^\delta = \arg\min_{u \in V_h} T_\alpha^\delta(u).$$

For notational convenience, we omit the indices h for elements $u \in V_h$.

In line with the structure of the first part of this thesis, we begin with the establishment of a well-posedness assertion in Chapter 7. This chapter includes required assumptions, auxiliary elements, and presents an example that assures the practicality of the required conditions on the projection operator P_h.

7 Fundamental requirements

In this chapter, we establish the groundwork towards proving convergence rates for the setting explained above. While maintaining the same structure as Chapter 3, this chapter introduces additional assumptions concerning the projection operator P_h. Section 7.4 presents an example that validates these new assumptions.

7.1 Well-posedness

As in the first part of this thesis, we establish regularization properties of the extremal problem

$$\min_{u \in V_h} T_\alpha^\delta(u). \tag{43}$$

This is an essential step for validating the utility of our approach. Specifically, we establish well-posedness of the extremal problem, as well as stability of minimizers.

Assumption 7.1 below summarizes the required assumptions.

Assumption 7.1. (a) The operator $F : X \supset \mathcal{D}(F) \to Y$ is sequentially continuous with respect to the norm topology of X and the weak topology of Y.

(b) $V_h \subset \mathcal{D} = \mathcal{D}(F) \cap X_1$.

(c) $\overline{u} \in \mathcal{D}$.

We highlight the differences between this assumption and Assumption 3.1 from Part I. First, the operator F is now assumed to be sequentially continuous with respect to the norm topology of X and the weak topology of Y instead of the weak topologies of X and Y. Second, Assumption 7.1 (b) not only replaces Assumption 3.1 (b), which required $\mathcal{D}(F)$ to be a closed and convex subset of X, but also replaces Assumption 3.7 (d), which required $u^\dagger$ to be an interior point of $\mathcal{D}(F)$.

Based on these assumptions we can now verify regularization properties equivalent to those in Section 3.1. The assertions hold for any norm in V_h because the norms are equivalent in V_h.

Theorem 7.2 (Well-posedness)**.** *Under Assumption 7.1, for each $\alpha > 0$ and each $f^\delta \in Y$, the extremal problem (43) is well-posed in the sense of Definition 3.2. Moreover, the convergent minimizing subsequences $(u_{n_k})_{k\in\mathbb{N}} \subset V_h$ of (43) have a minimizer $u_{h,\alpha}^\delta$ of (43) as a limit and satisfy*

$$\lim_{k\to\infty} \| F(u_{n_k}) - f^\delta \| = \| F(u_{h,\alpha}^\delta) - f^\delta \| \quad \textit{and} \tag{44}$$

$$\lim_{k\to\infty} \| u_{n_k} - \overline{u} \|_1 = \| u_{h,\alpha}^\delta - \overline{u} \|_1. \tag{45}$$

Proof. Since $V_h \neq \emptyset$, the infimum

$$T_{h,*} := \inf\{T_\alpha^\delta(u) : u \in V_h\} \geq 0$$

exists. Let $(u_n)_{n\in\mathbb{N}} \subset V_h$ be a miminizing sequence such that

$$\lim_{n\to\infty} T_\alpha^\delta(u_n) = T_{h,*}.$$

Then $(T_\alpha^\delta(u_n))_{n\in\mathbb{N}}$ and consequently $(\| u_n - \overline{u}_h \|_1)_{n\in\mathbb{N}}$ are bounded sequences in $\mathbb{R}$. This implies that the sequence $(u_n)_{n\in\mathbb{N}} \subset V_h$ is bounded in the finite-dimensional subspace $(V_h, \| \cdot \|_1)$. Since in finite-dimensional spaces, every bounded sequence has a convergent subsequence, there exists a subsequence $(u_{n_k})_{k\in\mathbb{N}} \subset V_h$ converging to some element $u_{h,*}$ as $k \to \infty$. The limit $u_{h,*}$ lies in V_h as well, because V_h is closed.

Now let $(u_{n_k})_{k\in\mathbb{N}}$ be any subsequence of the minimizing sequence that converges to some limit $u_{h,*}$ as $k \to \infty$. We show that this limit $u_{h,*}$ indeed minimizes the Tikhonov functional T_α^δ. From the sequential continuity of F with respect to the norm topology of X and the weak topology of Y, it follows that $F(u_{n_k})$ converges to $F(u_{h,*})$ with respect to the weak topology of Y. The weakly lower semi-continuity of the norm $\| \cdot \|$ and the continuity of the norm $\| \cdot \|_1$ imply

$$\begin{aligned} T_\alpha^\delta(u_{h,*}) &= \| F(u_{h,*}) - f^\delta \|^2 + \alpha \| u_{h,*} - \overline{u}_h \|_1^2 \\ &\leq \liminf_{k\to\infty} \| F(u_{n_k}) - f^\delta \|^2 + \alpha \lim_{k\to\infty} \| u_{n_k} - \overline{u}_h \|_1^2 \\ &\leq \liminf_{k\to\infty} (\| F(u_{n_k}) - f^\delta \|^2 + \alpha \| u_{n_k} - \overline{u}_h \|_1^2) = T_{h,*}. \end{aligned}$$

Hence $u_{h,*}$ minimizes T_α^δ, and we write $u_{h,\alpha}^\delta := u_{h,*}$. The convergence of the penalty functional (45) is an immediate consequence of the continuity of the norm $\| \cdot \|_1$. The convergence of the misfit functional (44) can then be deduced easily:

$$\begin{aligned} \lim_{k\to\infty} \| F(u_{n_k}) - f^\delta \|^2 &= \lim_{k\to\infty} (T_\alpha^\delta(u_{n_k}) - \alpha \| u_{n_k} - \overline{u} \|_1^2) \\ &= T_\alpha^\delta(u_{h,\alpha}^\delta) - \alpha \| u_{h,\alpha}^\delta - \overline{u} \|_1^2 = \| F(u_{h,\alpha}^\delta) - f^\delta \|^2. \end{aligned}$$

□

Theorem 7.3 (Stability of regularized solutions)**.** *Under Assumption 7.1, for each $\alpha > 0$, minimizers of (43) are stable with respect to small perturbations in the data f^δ in the sense of Definition 3.4. Specifically, for a data sequence*

$(f_n)_{n\in\mathbb{N}} \subset Y$ with $\lim_{n\to\infty} \| f_n - f^\delta \| = 0$, let $(u_n^*)_{n\in\mathbb{N}} \subset V_h$ be a sequence of minimizers of the Tikhonov functional over V_h with f^δ replaced by f_n. That is

$$u_n^* := \arg\min_{u\in V_h} T_\alpha^n(u), \qquad n \in \mathbb{N},$$

where T_α^n is given in (12). Then there exists a convergent subsequence $(u_{n_k}^*)_{k\in\mathbb{N}}$ and each cluster point of $(u_n^*)_{n\in\mathbb{N}}$ is a minimizer of T_α^δ. Furthermore, each such convergent subsequence $(u_{n_k}^*)_{k\in\mathbb{N}}$ satisfies

$$\lim_{k\to\infty} \| F(u_{n_k}^*) - f_{n_k} \| = \| F(u_{h,\alpha}^\delta) - f^\delta \| \quad \text{and}$$
$$\lim_{k\to\infty} \| u_{n_k}^* - \overline{u} \|_1 = \| u_{h,\alpha}^\delta - \overline{u} \|_1.$$

Proof. According to Theorem 7.2, the minimizers u_n^* exist for all $n \in \mathbb{N}$. Analogously to the proof of Theorem 3.5, it can be shown that $(u_n^*)_{n\in\mathbb{N}}$ minimizes $T_\alpha^\delta(u) = \| F(u) - f^\delta \|^2 + \alpha \| u - \overline{u} \|_1^2$ over V_h as $n \to \infty$.

Thus, according to Theorem 7.2, there exists a subsequence $(u_{n_k}^*)_{k\in\mathbb{N}} \subset V_h$ converging to $u_{h,\alpha}^\delta := \arg\min_{u\in V_h} T_\alpha^\delta(u)$ as $k \to \infty$. The limit of each such subsequence is a minimizer of T_α^δ over V_h. Moreover, Theorem 7.2 implies the convergence of the corresponding misfit and penalty functionals. □

7.2 Additional assumptions

In this section, we summarize additional assumptions. These encompass the source condition, assumptions on the projection operator P_h, and conditions facilitating convergence rates similar to those in Assumption 3.7.

To define the source condition, we recall the operator B introduced in Section 2.4. Instead of $G = B^{-(2a+2)}$ as in (15) in Section 3.2, we set

$$G := B^{-1}.$$

The underlying *Hölder-type* source condition, which accompanies our study in this part, is provided by

$$u^\dagger - \overline{u} \in X_p,$$

or equivalently

$$u^\dagger - \overline{u} = G^p w, \tag{46}$$

for some $0 < p \le 1$ and a source element $w \in X$, with $\| w \| \le \rho$, $\rho > 0$. The case $p < 1$ leads to the oversmoothing situation that we are particularly interested in. However, we include $p = 1$ in our calculations since the established results remain valid in that case as well.

Along with this source condition, we assume that there is a constant

$$p_0 \geq 2, \tag{47}$$

such that

$$\|(I - P_h)u\| \leq k_0 h^{p_0} \|u\|_{p_0} \qquad \text{for all } \ u \in X_{p_0} \ \text{ and} \tag{48}$$

$$\|P_h u\|_1 \leq k_* \|u\|_1 \qquad \text{for all } \ u \in X_1 \tag{49}$$

hold throughout this part, where k_0 and k_* are positive and finite constants. The value 2 in (47) is chosen arbitrarily; a higher value is desirable, because as we shall see, a higher value of p_0 permits a bigger scope for the value of a that is determined through item (e) of Assumption 7.5.

Since G is selfadjoint, for any $q \geq 0$, the operators G^q and G^{-1} are selfadjoint, as well. Moreover, the operators $(I - P_h)$, P_h are selfadjoint, thus for any $q \geq 0$

$$\|G^q(I - P_h)\| = \|(I - P_h)G^q\| \quad \text{and} \tag{50}$$

$$\|G^{-1}P_h\| = \|P_h G^{-1}\|. \tag{51}$$

We are going to use these properties frequently throughout this chapter. A first application allows us to formulate assumption (48) as

$$\|G^{p_0}(I - P_h)w\| \leq k_0 h^{p_0} \|w\| \qquad \text{for all } \ w \in X, \tag{52}$$

where we used that any $u \in X_{p_0}$ can be written as $u = G^{p_0} w$ for some $w \in X$. Based on (52) the next lemma extends the estimate of (48) to cover any $0 \leq p \leq p_0$.

Lemma 7.4. *Assumption (48) implies*

$$\|(I - P_h)u\| \leq k_p h^p \|u\|_p, \qquad u \in X_p, \quad 0 \leq p \leq p_0, \tag{53}$$

or equivalently

$$\|G^p(I - P_h)w\| \leq k_p h^p \|w\|, \qquad w \in X, \quad 0 \leq p \leq p_0, \tag{54}$$

where $k_p = k_0^{p/p_0}$ and k_0 is the constant from assumption (48). Obviously, k_p is bounded when considered as a function of p on $[0, p_0]$.

Proof. For $p = 0$, the estimates in (53) and (54) hold trivially. For $0 < p \le 1$, the elements $u \in X_p$ can be written as $u = G^p w$ for some $w \in X$. Therefore, by substitution and using (50), the estimates in (53) and (54) are equivalent. We focus on verifying (54). Using the interpolation inequality (8) and assumption (52), we show that for $0 < p \le p_0$, with $p_0 \ge 2$, the following holds:

$$\begin{aligned} \| G^p (I - P_h) w \| &\le \| G^{p_0} (I - P_h) w \|^{\frac{p}{p_0}} \| (I - P_h) w \|^{\frac{p_0 - p}{p_0}} \\ &\le k_0^{\frac{p}{p_0}} h^p \| I - P_h \|^{\frac{p_0 - p}{p_0}} \| w \|, \qquad w \in X. \end{aligned}$$

Since $\| I - P_h \| \le 1$, the assertion of the lemma is confirmed. □

The subsequent assumption concludes our collection of assumptions, mirroring Assumption 3.7 (e)-(g) in the preceding part of this thesis, albeit with the source condition replaced.

Assumption 7.5. (d) The observations f^δ satisfy $\| f^\delta - f^\dagger \| \le \delta$ for $\delta > 0$.

(e) Let $a > 0$, and let there exist positive finite constants $c_a \le C_a$ and c_0, c_1 such that

$$\| F(u) - f^\dagger \| \le C_a \| u - u^\dagger \|_{-a} \quad \text{for each } u \in \mathcal{D} \quad \text{with } \| u - u^\dagger \|_{-a} \le c_0 \tag{55}$$

and

$$c_a \| u - u^\dagger \|_{-a} \le \| F(u) - f^\dagger \| \quad \text{for each } u \in \mathcal{D} \quad \text{with } \| F(u) - f^\dagger \| \le c_1. \tag{56}$$

(f) Source condition (46) applies.

7.3 Discretized auxiliary elements and their properties

In this section, we provide another approach to auxiliary elements. This approach was used for the oversmoothing Tikhonov regularization in the Banach space setting, for example, in [49] or in [20], and the idea traces back to the articles [48] and [46]. In this section, it will become clear why the exponent $2a + 2$, with a determined through item (e) of Assumption 7.5, is chosen for the inverse of B in Part I. To introduce the auxiliary elements, let

$$R_\beta : X \to X \qquad \text{for } \beta > 0$$

be an arbitrary family of regularizing operators. For $G = B^{-1}$ set

$$S_\beta := I - R_\beta G \qquad \text{for } \beta > 0.$$

We make the following assumptions on these operators:

$$\| R_\beta \| \le \frac{c_*}{\beta}, \tag{57}$$

$$\| S_\beta G^p \| \le c_p \beta^p, \qquad 0 \le p \le \overline{p}, \tag{58}$$

$$R_\beta G = G R_\beta, \tag{59}$$

for $\beta > 0$, finite constants c_* and c_p, and a finite constant $\overline{p}$ that represents the saturation of R_β. Further, we assume that c_p is bounded as a function of p.

Remark 7.6. For a linear operator equation $Au = f$, where $A : X \to Y$ is a bounded operator mapping between Hilbert spaces X and Y, regularization methods can rely on the approach of approximating the Moore-Penrose inverse of the linear operator. Such approximations are constructed by a family $\{g_\beta\}_{\beta>0}$ of Borel-measurable functions

$$g_\beta : [0, \| A \|] \to \mathbb{R}, \qquad \beta > 0$$

satisfying

$$\sup_{0 \le \lambda \le \| A \|} |g_\beta(\lambda)| \le c_* \beta^{-1}, \qquad \beta > 0 \quad \text{and} \tag{60}$$

$$\sup_{0 \le \lambda \le \| A \|} \lambda^p |1 - \lambda g_\beta(\lambda)| \le c_p \beta^p, \qquad \beta > 0, \quad 0 \le p \le \overline{p}. \tag{61}$$

The functions g_β can be traced back to [57] and they are called generator functions [54, p. 58], or filter functions in the case of compact operators [36, p. 55].

Therefore, in our setting, the regularization operators R_β, satisfying (57) and (58), can be defined by $R_\beta := g_\beta(G)$ for g_β satisfying (60) and (61) for $A = G = B^{-1}$.

For compact operators, the truncated singular value decomposition is a regularization method that satisfies the conditions (57)–(59) for all $\overline{p} > 0$, see for example [36, Theorem 4.1.2]. For operators G, which are not necessarily compact, an example for a family of regularization operators $\{R_\beta\}_{\beta>0}$ that satisfies the assumptions in (57)-(59) with an integer $\overline{p} = m \ge 1$, is given by Lavrientiev's m-times iterated method, see [20]. In the following example, we briefly outline this method.

Example 7.7 (Lavrientiev's m-times iterated method)**.** Let $m \in \mathbb{N}$. For $f \in X$ and $\overline{u} = 0$ the element $R_\beta f$ is given by

$$(G + \beta I) u_n = \beta u_{n-1} \qquad \text{for} \;\; n = 1, 2, \ldots, m, \qquad R_\beta f := u_m.$$

The operator R_β can be written as

$$R_\beta = \beta^{-1} \sum_{j=1}^{m} \beta^j (G + \beta I)^{-j} = G^{-1}(I - \beta^m(\beta I + G)^{-m})$$

and the operator S_β has the form

$$S_\beta = I - R_\beta G = \beta^m (\beta I + G)^{-m}.$$

Remark 7.8. For $m = 1$ the iterated Lavrentiev's method equals Lavrentiev's regularization method, where

$$R_\beta := (G + \beta I)^{-1}.$$

This results in the auxiliary elements as in Section 3.3 when G is defined as $G = B^{-(2a+2)}$. The fact that assumption (58) holds only for with $\overline{p} = 1$ for Lavrentiev's regularization method is not restrictive then due to the choice of the exponent $2a + 2$ for the inverse of B. This adjustment replaces the exponent p of assumption (58) by $p/(2a + 2)$ such that the assumption is required to hold for $p/(2a+2) \leq \overline{p} = 1$. The condition $p/(2a+2) \leq 1$ holds for all applications of the assumption in our proofs if we consider $G = B^{-(2a+2)}$ instead of $G = B^{-1}$. In summary, we could have employed the auxiliary elements from Part I here as well; the necessity of assumption (58) clarifies the choice of $2a + 2$ as an exponent in Part I.

A useful consequence arising from the assumptions in (57)-(59) is the following lemma. The lemma and its proof are reproduced from [49] or [20].

Lemma 7.9. *There exists some positive and finite constant c such that for each $0 < p \leq 1$ it holds that*

$$\| R_\beta G^p \| \leq c\beta^{p-1} \qquad \textit{for} \;\; \beta > 0.$$

Proof. Since $R_\beta G^p = G^p R_\beta$, by (59), the assertion follows with the interpolation inequality (8), as well as the bounds in (57) and (58):

$$\begin{aligned} \| R_\beta G^p w \| &= \| G^p R_\beta w \| \leq \| G R_\beta w \|^p \| R_\beta w \|^{1-p} \\ &= \| (I - S_\beta) w \|^p \| R_\beta w \|^{1-p} \leq (1 + c_0)^p \| w \| c_*^{1-p} \beta^{p-1}, \qquad w \in X. \end{aligned}$$

□

Throughout this part, we assume that the operators $\{R_\beta\}_{\beta>0}$ satisfy the conditions (57)-(59), with saturation

$$\overline{p} \geq a + 1,$$

where a is given by item (e) of Assumption 7.5. By means of R_β, we define the second variant of auxiliary elements as follows:

$$\hat{u}_\beta := \overline{u} + R_\beta G(u^\dagger - \overline{u}) = u^\dagger - S_\beta(u^\dagger - \overline{u}) \qquad \text{for} \quad \beta > 0. \tag{62}$$

In our analysis, we specifically use their projections onto V_h, given by

$$\hat{u}_{h,\beta} := P_h \hat{u}_\beta \qquad \text{for} \quad \beta > 0 \quad \text{and} \quad h > 0. \tag{63}$$

Providing upper limits for norms involving these auxiliary elements (63), the following Lemma 7.10 gives an analogue to Lemma 3.10.

Lemma 7.10. *If source condition (46) applies for some $0 < p \leq 1$ then there exist positive constants c_i, $i = 1, 2, 3$, such that the following inequalities hold for $a \leq p_0$, with p_0 given in assumption (48):*

(i) $\|\hat{u}_{h,\beta} - u^\dagger\| \leq c_1(\beta^p + h^p)$,

(ii) $\|\hat{u}_{h,\beta} - u^\dagger\|_{-a} \leq c_2(\beta^{a+p} + h^{a+p} + h^a \beta^p)$,

(iii) $\|\hat{u}_{h,\beta} - \overline{u}\|_1 \leq c_3 \beta^{p-1}$,

for each $h > 0$ and each finite $\beta > 0$.

Proof. We begin with the proof of item (i). We use the representation of the auxiliary elements $\hat{u}_{h,\beta} = P_h \hat{u}_\beta$ with $\hat{u}_\beta$ in (62), source condition (46), assumption (58), and estimate (53) of Lemma 7.4:

$$\begin{aligned}
\|\hat{u}_{h,\beta} - u^\dagger\| &= \| P_h(u^\dagger - S_\beta(u^\dagger - \overline{u})) - u^\dagger \| = \| P_h S_\beta(\overline{u} - u^\dagger) - (I - P_h)u^\dagger \| \\
&\leq \| P_h S_\beta G^p w \| + \| (I - P_h)u^\dagger \| \leq c_p \beta^p \rho + k_p h^p \| u^\dagger \|_p.
\end{aligned}$$

The estimates in (58) and (53) can be applied because with $0 < p \leq 1$ it is clear that $p \leq \overline{p}$ and $p \leq p_0$. Further, since $u^\dagger - \overline{u} \in X_p$, it follows that $u^\dagger \in X_p$, thus $u^\dagger$ is bounded in X_p. Moreover, c_p and k_p are bounded as functions of p. Therefore, the estimate of item (i) follows for c_1 chosen appropriately. To show the estimate in item (ii), we first proceed similarly as in the proof of item (i):

$$\begin{aligned}
\| \hat{u}_{h,\beta} - u^\dagger \|_{-a} &= \| G^a (P_h S_\beta(\overline{u} - u^\dagger) - (I - P_h)u^\dagger) \| \\
&= \| G^a S_\beta(\overline{u} - u^\dagger) - G^a(I - P_h)u^\dagger - G^a(I - P_h)S_\beta(\overline{u} - u^\dagger)) \| \\
&\leq \| G^a S_\beta G^p w \| + \| G^a(I - P_h)u^\dagger \| + \| G^a(I - P_h)S_\beta G^p w \| \\
&\leq c_2(\beta^{a+p} + h^{a+p} + h^a \beta^p).
\end{aligned}$$

The last estimate requires some explanation. Therefore, we examine the three norms before the last inequality separately. Condition (59) implies that

$G^a S_\beta = S_\beta G^a$. Consequently, from (58) we obtain $\| G^a S_\beta G^p w \| \le c_{a+p} \rho \beta^{a+p}$ for $a + p \le \overline{p}$. The condition $a + p \le \overline{p}$ holds, because $p \le 1$ and $\overline{p}$ is assumed to satisfy $\overline{p} \ge a + 1$. To give an estimate for the second norm, we use Lemma 7.4 twice, taking advantage of the fact that the orthogonal projector $I - P_h$ is idempotent:

$$\begin{aligned} \| G^a (I - P_h) u^\dagger \| &= \| G^a (I - P_h)(I - P_h) u^\dagger \| \le \| G^a (I - P_h) \| \| (I - P_h) u^\dagger \| \\ &\le k_p k_a h^{a+p} \| u^\dagger \|_p. \end{aligned}$$

Another application of Lemma 7.4 and (58) provides an estimate for the third norm:

$$\| G^a (I - P_h) S_\beta G^p w \| \le k_a h^a c_p \beta^p \rho.$$

It remains to verify item (iii):

$$\begin{aligned} \| \hat{u}_{h,\beta} - \overline{u} \|_1 &= \| P_h R_\beta G(u^\dagger - \overline{u}) - (I - P_h)\overline{u} \|_1 \\ &= \| G^{-1}(P_h R_\beta G(u^\dagger - \overline{u}) - (I - P_h)\overline{u}) \| \\ &\le \| G^{-1} P_h R_\beta G(u^\dagger - \overline{u}) \| + \| G^{-1}(I - P_h)\overline{u} \|. \end{aligned}$$

We examine the summands in the latter expression separately. Using (59), (51), and Lemma 7.9, we get

$$\begin{aligned} \| G^{-1} P_h R_\beta G(u^\dagger - \overline{u}) \| &= \| G^{-1} P_h G R_\beta (u^\dagger - \overline{u}) \| = \| P_h R_\beta (u^\dagger - \overline{u}) \| \\ &= \| P_h R_\beta G^p w \| \le c \beta^{p-1}. \end{aligned}$$

Since $\overline{u} \in X_1$, we have $\overline{u} = G\overline{w}$ for some $\overline{w} \in X$ with $\| \overline{w} \| \le \overline{\rho}$. Using (51), which holds for $I - P_h$ instead of P_h as well, and Lemma 7.4, it follows

$$\| G^{-1}(I - P_h)\overline{u} \| = \| G^{-1}(I - P_h) G\overline{w} \| = \| (I - P_h)\overline{w} \| \le \overline{\rho}.$$

Thus

$$\| \hat{u}_{h,\beta} - \overline{u} \|_1 \le c\beta^{p-1} + \mathcal{O}(1) \qquad \text{for} \quad 0 < \beta < \infty.$$

□

Equipped with the auxiliary elements and the estimates of Lemma 7.10, we can make further steps towards the proof of this part's main theorem in Chapter 8. Beforehand, in Section 7.4, we provide an example that satisfies conditions (48) and (49).

7.4 Illustrative example of a projection meeting required conditions

This section presents an example that satisfies the conditions (48) and (49). Let $X = L^2(0,1)$. As in Chapter 6, we consider the integration operator

$$J : X \to X, \qquad [Ju](t) = \int_0^t u(x)\,\mathrm{d}x, \qquad 0 \le t \le 1$$

and define $B = (J^*J)^{-1/2}$. Then B is linear, densely defined, unbounded, selfadjoint, and positive definite. Moreover

$$\|\, u \,\|_1 = \|\, Bu \,\| = \|\, (J^*J)^{-1/2} u \,\| = \|\, u' \,\|, \qquad u \in \mathcal{D}(B),$$

where u' denotes the weak derivative of $u \in X$. To define a suitable finite-dimensional subspace V_h, we consider the class of linear splines. We outline a brief introduction to the space of linear spline functions. A comprehensive treatment of this topic is given in the monograph [47, Chapter 2], which serves as a starting point for this introductory part.

A *linear spline* $s : [a,b] \to \mathbb{R}$ is a continuous function, that is piecewise linear on given subintervals. These subintervals are determined through ordered so-called knots $x_i \in [a,b]$, for $i = 0, \ldots, N$ and some $N \in \mathbb{N}$ such that

$$\Delta := \{a = x_0 < x_1 < \cdots < x_N = b\} \quad \text{and}$$
$$[a,b] = [x_0, x_1) \cup [x_1, x_2) \cup \cdots \cup [x_{N-1}, x_N) \cup \{x_N\}.$$

For a function u and a given set of knots Δ, the corresponding interpolating linear spline s is determined uniquely through the local representations

$$s(x) = u(x_i) + \frac{u(x_{i+1}) - u(x_i)}{x_{i+1} - x_i}(x - x_i), \qquad x_i \le x \le x_{i+1},$$

for $i = 0, \ldots, N-1$. By

$$S_\Delta := \{s \in C[a,b] : s \text{ is linear on } [x_{i-1}, x_i],\ x_i \in \Delta \quad \text{for} \ \ i = 1, \cdots, N\}$$

we denote the space of linear splines for Δ.

Theorem 7.11. *Let $B = (J^*J)^{-1/2}$ be defined as above. Additionally, let $\Delta = \{0 = x_0 < x_1 < \cdots < x_N = 1\}$ be a partition of the interval $[0,1]$, with $h = \max_{i=1,\ldots,N}(x_i - x_{i-1})$ and $h_* = \min_{i=1,\ldots,N}(x_i - x_{i-1})$. Let P_h be the orthogonal projection onto the corresponding space $V_h = S_\Delta$ of linear splines. If the ratio h/h_* is bounded, then the conditions (48) and (49) are satisfied with $p_0 = 2$.*

Proof. For any $u \in L^2(0,1)$, let $P_h u \in S_\Delta$ denote the orthogonal projection of u onto S_Δ and let s represent the corresponding interpolating linear spline. The proof is structured in two steps. First, we demonstrate that (48) and (49) hold when $P_h u$ is replaced by s. More precisely, we prove that

$$\| s - u \| \leq \left(\frac{h}{\pi}\right)^2 \| u \|_2 \qquad \text{for all} \quad u \in X_2 \quad \text{and} \tag{64}$$

$$\| s \|_1 \leq \| u \|_1 \qquad \text{for all} \quad u \in X_1. \tag{65}$$

Subsequently, we verify that the original assertions of (48) and (49) are satisfied. In the first step of the proof, we follow the procedure of the proof of Theorem 6.2 in [9, Section 6.6]. We will utilize Friedrich's inequality, which can be found for example under Theorem 6.1 in [9, Section 6.6] or in [47, Lemma 9.24]. It states that for functions $u \in H^1(a,b)$, with $u(a) = u(b) = 0$, we have that

$$\| u \| \leq \frac{b-a}{\pi} \| u' \|. \tag{66}$$

To verify both of the estimates in (64) and (65), we use that for any $u \in X_1$, and consequently for $u \in X_2 \subset X_1$ the following equation holds:

$$\int_0^1 s'(x)(u'(x) - s'(x)) \, \mathrm{d}x = 0. \tag{67}$$

This follows because $s'(x)$ takes constant values s_i on each interval $[x_{i-1}, x_i]$ for $i = 1, \ldots, N$, and $u(x) = s(x)$ at $x = x_i$, $i = 0, \ldots, N$:

$$\begin{aligned} \int_0^1 s'(x)(u'(x) - s'(x)) \, \mathrm{d}x &= \sum_{i=1}^N \int_{x_{i-1}}^{x_i} s'(x)(u'(x) - s'(x)) \, \mathrm{d}x \\ &= \sum_{i=1}^N s_i (u(x) - s(x)) \Big|_{x=x_{i-1}}^{x_i} = 0. \end{aligned}$$

We first verify (65). Let $u \in X_1 = \mathcal{D}(B) = \{u \in H^1(0,1) : u(1) = 0\}$. It holds that

$$0 \leq \int_0^1 (s'(x) - u'(x))^2 \, \mathrm{d}x = \int_0^1 u'(x)^2 - s'(x)^2 - 2s'(x)(u'(x) - s'(x)) \, \mathrm{d}x.$$

Equation (67) implies

$$0 \leq \int_0^1 u'(x)^2 - s'(x)^2 \, \mathrm{d}x,$$

which is equivalent to

$$\| s \|_1^2 \leq \| u \|_1^2, \tag{68}$$

that is (65).

Now, we verify (64). Let $u \in X_2 = \{u \in H^2(0,1) : u'(0) = 0,\ u(1) = 0\}$. The equation in (67) can be expanded to

$$\int_0^1 (u'(x) - (u'(x) - s'(x)))(u'(x) - s'(x))\,\mathrm{d}x = 0,$$

which is equivalent to

$$\int_0^1 u'(x)(u'(x) - s'(x))\,\mathrm{d}x = \int_0^1 (u'(x) - s'(x))^2\,\mathrm{d}x. \tag{69}$$

By partial integration, the left-hand side of (69) can be formulated as

$$\begin{aligned}\int_0^1 u'(x)(u'(x) - s'(x))\,\mathrm{d}x &= \sum_{i=1}^N \int_{x_{i-1}}^{x_i} u'(x)(u'(x) - s'(x))\,\mathrm{d}x \\ = \sum_{i=1}^N \big(u'(x)(u(x) - s(x))\big|_{x=x_{i-1}}^{x_i} &- \int_{x_{i-1}}^{x_i} u''(x)(u(x) - s(x))\,\mathrm{d}x\big).\end{aligned} \tag{70}$$

The first term in the sum of (70) vanishes, therefore

$$\int_0^1 u'(x)(u'(x) - s'(x))\,\mathrm{d}x = -\int_0^1 u''(x)(u(x) - s(x))\,\mathrm{d}x. \tag{71}$$

By substitution of (71) into (69), we observe that

$$\int_0^1 (u'(x) - s'(x))^2\,\mathrm{d}x = -\int_0^1 u''(x)(u(x) - s(x))\,\mathrm{d}x.$$

The Cauchy-Schwarz inequality gives an upper bound for the right-hand side of the latter equation, such that

$$\| u' - s' \|^2 \leq \| u'' \| \| u - s \|. \tag{72}$$

Since $u(x) - s(x) = 0$ on the boundaries $\partial([x_{i-1}, x_i]) = \{x_{i-1}, x_i\}$ and $u(x) - s(x) \in H^1(x_{i-1}, x_i)$ for $i = 1, \ldots, N$, we can apply Friedrichs's inequality (66) piecewise to $\int_0^1 (u(x) - s(x))^2\,\mathrm{d}x$, which yields

$$\int_0^1 (u(x) - s(x))^2\,\mathrm{d}x = \sum_{i=1}^N \int_{x_{i-1}}^{x_i} (u(x) - s(x))^2\,\mathrm{d}x$$

$$
\begin{aligned}
&= \sum_{i=1}^{N} \| u - s \|^2_{L^2(x_{i-1},x_i)} \\
&\leq \sum_{i=1}^{N} \Big(\frac{x_i - x_{i-1}}{\pi}\Big)^2 \| u' - s' \|^2_{L^2(x_{i-1},x_i)} \\
&\leq \max_{i=1,\dots,N} \Big(\frac{x_i - x_{i-1}}{\pi}\Big)^2 \| u' - s' \|^2. \qquad (73)
\end{aligned}
$$

We can rewrite (73) and use (68) to conclude

$$\| u - s \| \leq \frac{h}{\pi} \| u' - s' \| \leq \frac{h}{\pi} (\| u' \| + \| s' \|) \leq \frac{2h}{\pi} \| u' \|, \qquad (74)$$

which is (48) with $P_h u$ replaced by s and for $p_0 = 1$. Inserting (72) into (73) yields

$$\| u - s \|^2 \leq \Big(\frac{h}{\pi}\Big)^2 \| u'' \| \| u - s \|,$$

which is equivalent to

$$\| u - s \| \leq \Big(\frac{h}{\pi}\Big)^2 \| u'' \|,$$

that is (64). We now proceed to prove the original assertions (48) and (49). The first estimate, (48), follows directly from (64) noting that $\| P_h u - u \| \leq \| \ell - u \|$ for any $\ell \in S_\Delta$. For the second assertion, (49), we remark that a more general result is presented in [6]. To verify (49), we use that

$$\| P_h u - s + s \|_1 \leq \| P_h (u - s) \|_1 + \| s \|_1.$$

Combining this with (65) and the estimate $\| P_h (u - s) \| \leq \| u - s \| \leq \frac{2h}{\pi} \| u \|_1$, given in (74), it remains to show that

$$\| P_h (u - s) \|_1 \leq \frac{c}{h_*} \| P_h (u - s) \|,$$

where c is a positive and finite constant. More generally, we show that for any linear spline $\ell \in S_\Delta$ it holds that

$$\| \ell \|_1 = \| \ell' \| \leq \frac{c}{h_*} \| \ell \|.$$

By $h_i = x_{i+1} - x_i$ for $i = 0, \dots, N-1$ we denote the spacing between consecutive grid points. A basis for V_h is formed by hat functions, defined as

$$\psi_0(x) = \begin{cases} 1 - \frac{x}{h_0}, & \text{if } x_0 \leq x \leq x_1, \\ 0, & \text{otherwise} \end{cases}$$

for the boundary basis function at $i = 0$ and

$$\psi_i(x) = \begin{cases} \frac{x - x_{i-1}}{h_{i-1}}, & \text{if } x_{i-1} \le x \le x_i, \\ \frac{x_{i+1} - x}{h_i}, & \text{if } x_i < x \le x_{i+1}, \\ 0, & \text{otherwise,} \end{cases}$$

for $i = 1, \dots, N-1$. Any $\ell \in V_h$ thus has the representation

$$\ell(x) = \sum_{i=0}^{N-1} b_i \psi_i(x),$$

where b_i are appropriate coefficients. The derivative of ℓ has the representation

$$\ell'(x) = \sum_{i=0}^{N-1} b_i \psi_i'(x),$$

where ψ_i' are the derivatives of ψ_i for each $i = 0, \dots, N-1$. That is

$$\psi_0'(x) = \begin{cases} -\frac{1}{h_0}, & \text{if } x_0 \le x \le x_1, \\ 0, & \text{otherwise} \end{cases}$$

and

$$\psi_i'(x) = \begin{cases} \frac{1}{h_{i-1}}, & \text{if } x_{i-1} \le x \le x_i, \\ -\frac{1}{h_i}, & \text{if } x_i < x \le x_{i+1} \\ 0, & \text{otherwise} \end{cases}$$

for $i = 1, \dots, N-1$.

The corresponding Gram matrices G and $G' \in \mathbb{R}^{N \times N}$ are symmetric tridiagonal matrices given by

$$\begin{aligned} G &= \begin{pmatrix} \langle \psi_0, \psi_0 \rangle & \cdots & \langle \psi_0, \psi_{N-1} \rangle \\ \vdots & & \vdots \\ \langle \psi_{N-1}, \psi_0 \rangle & \cdots & \langle \psi_{N-1}, \psi_{N-1} \rangle \end{pmatrix} \\ &= \frac{1}{6} \begin{pmatrix} 2h_0 & h_0 & 0 & \cdots & 0 \\ h_0 & 2(h_0 + h_1) & h_1 & \cdots & 0 \\ 0 & h_1 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & & h_{N-2} \\ 0 & \cdots & 0 & h_{N-2} & 2(h_{N-2} + h_{N-1}) \end{pmatrix} \end{aligned}$$

and

$$G' = \begin{pmatrix} \langle \psi_0', \psi_0' \rangle & \cdots & \langle \psi_0', \psi_{N-1}' \rangle \\ \vdots & & \vdots \\ \langle \psi_{N-1}', \psi_0' \rangle & \cdots & \langle \psi_{N-1}', \psi_{N-1}' \rangle \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{h_0} & -\frac{1}{h_0} & 0 & \cdots & 0 \\ -\frac{1}{h_0} & \frac{1}{h_0} + \frac{1}{h_1} & -\frac{1}{h_1} & \cdots & 0 \\ 0 & -\frac{1}{h_1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & -\frac{1}{h_{N-2}} \\ 0 & \cdots & 0 & -\frac{1}{h_{N-2}} & \frac{1}{h_{N-2}} + \frac{1}{h_{N-1}} \end{pmatrix}.$$

For the spectral norm $\| \cdot \|_2$ of G' it holds that

$$\| G' \|_2 \leq \| G' \|_\infty \leq \frac{4}{h_*},$$

where $\| \cdot \|_\infty$ is the maximum of the absolute row sums. The matrix G is strictly diagonal dominant, and for such a matrix, the following inequality holds:

$$\| y \|_\infty \leq \max_{i=0,\dots,N-1} \Big(\big(|\langle \psi_i, \psi_i \rangle| - \sum_{k=0, k\neq i}^{N-1} |\langle \psi_i, \psi_k \rangle| \big)^{-1} \Big) \| Gy \|_\infty,$$

for each $y \in \mathbb{R}^N$. In our case, this maximum leads to

$$\| y \|_\infty \leq \frac{h_*}{3} \| Gy \|_\infty.$$

Further details are given for example in [47, Lemma 2.13]. Consequently,

$$\| G^{-1} \|_2 \leq \| G^{-1} \|_\infty \leq \frac{3}{h_*}.$$

Moreover, G^{-1} is positive definite and G' is positive semidefinite, with both matrices being symmetric. Thus, $G^{-1/2}$ and $(G')^{1/2}$ are symmetric satisfying

$$\| G^{-\frac{1}{2}} \|^2 = \| G^{-1} \| \quad \text{and} \quad \| (G')^{\frac{1}{2}} \|^2 = \| G' \|.$$

Using these equalities as well as the estimates for the spectral norms of G' and G^{-1}, we obtain

$$\begin{aligned} \| \ell' \|^2 = b^T G' b = \| (G')^{\frac{1}{2}} b \|_2^2 \leq \| (G')^{\frac{1}{2}} \|_2^2 \| b \|_2^2 \leq \frac{4}{h_*} \| b \|_2^2 = \frac{4}{h_*} \| G^{-\frac{1}{2}} G^{\frac{1}{2}} b \|_2^2 \\ \leq \frac{4}{h_*} \| G^{-\frac{1}{2}} \|_2^2 \| G^{\frac{1}{2}} b \|_2^2 \leq \frac{12}{h_*^2} b^T G b = \frac{12}{h_*^2} \| \ell \|^2. \end{aligned}$$

Thus, the statement follows. □

8 Convergence analysis

In this chapter, we prove the main theorem of this part. As a preparation, we prove the next lemma that gives an analogue to Lemma 4.1. It will be used to prove the subsequent essential lemma. Throughout this chapter, we assume that Assumptions 7.1 and 7.5 hold.

Lemma 8.1. *It holds that*

$$\max\left\{\|F(u^{\delta}_{h,\alpha}) - f^{\delta}\|, \sqrt{\alpha}\|u^{\delta}_{h,\alpha} - \overline{u}\|_1\right\} \leq C_a c_2 3 h^{a+p} + \sqrt{\alpha} c_3 h^{p-1} + \delta$$

for each positive α and δ, and each positive and finite h. The constant C_a is given by Assumption 7.5 (e), and the constants c_2 and c_3 are given by Lemma 7.10. The a priori choices $h_* = \delta^{\frac{1}{a+p}}$ and $\alpha_* = \delta^{\frac{2(a+1)}{a+p}}$ lead to

$$\max\left\{\|F(u^{\delta}_{h_*,\alpha_*}) - f^{\delta}\|, \sqrt{\alpha_*}\|u^{\delta}_{h_*,\alpha_*} - \overline{u}\|_1\right\} \leq c_4\delta$$

for each $\delta > 0$, where c_4 is a positive and finite constant.

Proof. We consider the auxiliary elements (63), with $\beta = h$. By Assumption 7.5 (b), it holds that $\hat{u}_{h,\beta} \in V_h \subset D = \mathcal{D}(F) \cap X_1$, and we can follow the procedure as in the proof of Lemma 4.1. Specifically, we use the triangle inequality, Assumption 3.7 (e), and Lemma 7.10 in the given order to obtain the following estimate:

$$\begin{aligned}
T^{\delta}_{\alpha}(u^{\delta}_{h,\alpha})^{\frac{1}{2}} &\leq T^{\delta}_{\alpha}(\hat{u}_{h,\beta})^{\frac{1}{2}} \leq \|F(\hat{u}_{h,\beta}) - f^{\delta}\| + \sqrt{\alpha}\|\hat{u}_{h,\beta} - \overline{u}\|_1 \\
&\leq \|F(\hat{u}_{h,\beta}) - f^{\dagger}\| + \sqrt{\alpha}\|\hat{u}_{h,\beta} - \overline{u}\|_1 + \delta \\
&\leq C_a\|\hat{u}_{h,\beta} - u^{\dagger}\|_{-a} + \sqrt{\alpha}\|\hat{u}_{h,\beta} - \overline{u}\|_1 + \delta \\
&\leq C_a c_2 3 h^{a+p} + \sqrt{\alpha} c_3 h^{p-1} + \delta.
\end{aligned}$$

Note that (55) can be applied, because for finite $h > 0$ the norm $\|\hat{u}_{h,\beta} - u^{\dagger}\|_{-a}$ is bounded from above. Thus, the first statement of the lemma follows. The second statement is an immediate consequence. □

Notice the two terms h^{a+p} and h^{p-1} in the upper bound of the first estimate in Lemma 8.1: On the one hand, the term h^{a+p} decreases, as h approaches zero. On the other hand, the term h^{p-1} indicates that values of h close to zero are not advantageous. This promotes a good balance between fine and coarse discretization.

The next lemma establishes bounds for the difference $u^{\delta}_{h,\alpha} - \hat{u}_{h,\beta}$ measured under the strong norm $\|\cdot\|_1$ and the weaker norm $\|\cdot\|_{-a}$. The proof of this chapter's main theorem relies on these bounds.

Lemma 8.2. *There exist positive and finite constants c_5 and c_6, such that for each positive and finite α, δ, and $\beta = h$, it holds that*

$$\| u^\delta_{h,\alpha} - \hat{u}_{h,\beta} \|_{-a} \le c_5(h^{a+p} + \delta + \sqrt{\alpha} h^{p-1}) \quad \text{and} \tag{75}$$

$$\| u^\delta_{h,\alpha} - \hat{u}_{h,\beta} \|_1 \le \alpha^{-\frac{1}{2}} c_6(h^{a+p} + \delta + \sqrt{\alpha} h^{p-1}). \tag{76}$$

For the a priori choices $h_* = \delta^{\frac{1}{a+p}}$ and $\alpha_* = \delta^{\frac{2(a+1)}{a+p}}$, it follows

$$\| u^\delta_{h_*,\alpha_*} - \hat{u}_{h_*,h_*} \|_{-a} \le c_5 \delta \quad \text{and} \tag{77}$$

$$\| u^\delta_{h_*,\alpha_*} - \hat{u}_{h_*,h_*} \|_1 \le c_6 \delta^{\frac{p-1}{a+p}} \tag{78}$$

for each positive and finite δ.

Proof. We begin with the proof of (75). The triangle inequality, estimate (56) in item (e) of Assumption 7.5, item (ii) of Lemma 7.10, and Lemma 8.1 yield the following:

$$\begin{aligned}
\| u^\delta_{h,\alpha} - \hat{u}_{h,\beta} \|_{-a} &\le \| u^\delta_{h,\alpha} - u^\dagger \|_{-a} + \| u^\dagger - \hat{u}_{h,\beta} \|_{-a} \\
&\le \tfrac{1}{c_a} \| F(u^\delta_{h,\alpha}) - F(u^\dagger) \| + 3c_2 h^{a+p} \\
&\le \tfrac{1}{c_a} (\| F(u^\delta_{h,\alpha}) - f^\delta \| + \delta) + 3c_2 h^{a+p} \\
&\le \tfrac{1}{c_a} (C_a 3c_2 h^{a+p} + \sqrt{\alpha} c_3 h^{p-1} + 2\delta) + 3c_2 h^{a+p} \\
&\le c_5(h^{a+p} + \delta + \sqrt{\alpha} h^{p-1}),
\end{aligned}$$

for $c_5 > 0$ chosen appropriately. The latter bound is finite, thereby justifying our use of estimate (56).

We can give an estimate for $\| u^\delta_{h,\alpha} - \hat{u}_{h,\beta} \|_1$ by applying the triangle inequality, Lemma 8.1, and item (iii) of Lemma 7.10:

$$\begin{aligned}
\| u^\delta_{h,\alpha} - \hat{u}_{h,\beta} \|_1 &\le \| u^\delta_{h,\alpha} - \overline{u} \|_1 + \| \overline{u} - \hat{u}_{h,\beta} \|_1 \\
&\le \frac{1}{\sqrt{\alpha}} (C_a 3c_2 h^{a+p} + \sqrt{\alpha} c_3 h^{p-1} + \delta) + c_3 h^{p-1}.
\end{aligned}$$

The statement in (76) follows for $c_6 > 0$ chosen appropriately. The estimates in (77) and (78) are immediate consequences. □

We have now all the requisites to prove this chapter's main theorem that gives an error estimate for the discretized Tikhonov regularization under a Hölder-type source condition (46) and an a priori parameter choice.

Theorem 8.3. *There is a positive and finite constant C such that for $h_* = \delta^{\frac{1}{a+p}}$ and $\alpha_* = \delta^{\frac{2(a+1)}{a+p}}$, we have*

$$\| u^\delta_{h_*,\alpha_*} - u^\dagger \| \le C \delta^{\frac{p}{a+p}} \tag{79}$$

for each $\delta > 0$.

Proof. We consider the auxiliary elements $\hat{u}_{h,\beta}$, with parameters $h = \beta = h_*$. Applying the triangle inequality, the interpolation inequality (8), as well as item (i) of Lemma 7.10, and Lemma 8.2 to the left-hand side of (79) yields

$$\begin{aligned}
\| u^\delta_{h_*,\alpha_*} - u^\dagger \| &\leq \| u^\delta_{h_*,\alpha_*} - \hat{u}_{h_*,h_*} \| + \| \hat{u}_{h_*,h_*} - u^\dagger \| \\
&\leq \| u^\delta_{h_*,\alpha_*} - \hat{u}_{h_*,h_*} \|_{-a}^{\frac{1}{1+a}} \| u^\delta_{h_*,\alpha_*} - \hat{u}_{h_*,h_*} \|_1^{\frac{a}{1+a}} + c_1 \delta^{\frac{p}{a+p}} \\
&\leq (c_5 \delta)^{\frac{1}{1+a}} \big(c_6 \delta^{\frac{p-1}{a+p}} \big)^{\frac{a}{1+a}} + c_1 \delta^{\frac{p}{a+p}} = C \delta^{\frac{p}{a+p}}
\end{aligned}$$

for each $\delta > 0$ and $C = c_5^{\frac{1}{1+a}} c_6^{\frac{a}{1+a}} + c_1$. $\square$

9 Numerical results

To verify the results of Theorem 8.3 numerically, we recall the setting from Chapter 6. Specifically, our objective is to solve the operator equation $F(u) = f^\dagger$, for the operator $F : L^2(0,1) \to L^2(0,1)$ defined by

$$F(u) = \exp[Ju],$$

with

$$[Ju](x) = \int_0^x u(t)\,\mathrm{d}t \qquad \text{for } \ 0 \le x \le 1.$$

Moreover, $B = (J^*J)^{-1/2}$. As a finite-dimensional subspace V_h, we consider the space of linear splines, presented in Section 7.4. We examine two different solutions $u_1^\dagger$ and $u_2^\dagger$ that satisfy the Hölder-type source condition (46). First, we recall that $u^\dagger$ as in Chapter 6 satisfies $u^\dagger \in X_p = \mathcal{D}(B^p) = \mathcal{R}((J^*)^p)$ for either $p + \mu < 1/2$, or $p + \mu = 1/2$ and $\nu > 1/2$. Therefore, the solution

$$u_1^\dagger(t) = (1-\mu)t^{-\mu}(-\ln\theta t)^{-\nu} + \nu t^{-\mu}(-\ln\theta t)^{-\nu-1}, \qquad 0 < t \le 1,$$

with

$$\mu = \tfrac{1}{4}, \ \ \nu = 1, \ \text{and} \ \ \theta = 0.3$$

provides our first example with $u_1^\dagger \in X_p$ for $p < 1/4$ and $u_1^\dagger \notin X_p$ for $p \ge 1/4$. The image of $u_1^\dagger$ under F is given by

$$[F(u_1^\dagger)](t) = f_1^\dagger(t) = \exp(t^{1-\mu}(-\ln\theta t)^{-\nu}).$$

The solution

$$u_2^\dagger(t) = \mathrm{e}^{-t}, \qquad 0 \le t \le 1,$$

with

$$[F(u_2^\dagger)](t) = f_2^\dagger(t) = \exp(\mathrm{e}^{-t} - 1), \qquad 0 \le t \le 1,$$

provides our second example. We briefly verify that source condition (46) with some $0 < p < 1/2$ is satisfied for $u_2^\dagger$. The coefficients for $u_2^\dagger$ in the series representation (35), are given by

$$a_n = \int_0^1 e^{-t}\cos(n - \tfrac{1}{2})\pi t\,\mathrm{d}t = \frac{1 + \mathrm{e}^{-1}(n-\frac{1}{2})\pi\sin(n-\frac{1}{2})\pi}{1 + (n-\frac{1}{2})^2\pi^2}.$$

As in Chapter 6, the Picard criterion yields $u_2^\dagger \in X_p = \mathcal{R}((J^*)^p)$ if and only if

$$\sum_{n=1}^{\infty} \frac{a_n^2}{\sigma_n^2} < \infty,$$

where $\sigma_n = 1/((n-1/2)\pi)$ are the singular values of J^*. The series can be estimated as

$$\sum_{n=1}^{\infty} \frac{(n-\frac{1}{2})^{2p}\pi^{2p}(1+\mathrm{e}^{-1}(n-\frac{1}{2})\pi\sin(n-\frac{1}{2})\pi)^2}{(1+(n-\frac{1}{2})^2\pi^2)^2} < \sum_{n=1}^{\infty}(n-\tfrac{1}{2})^{2p-2}\pi^{2p-2}.$$

The dominating series converges if and only if $2p-2<-1$, that is $p<1/2$. Conversely, it can be shown that the series on the left-hand side diverges for $p \geq 1/2$, because it behaves asymptotically like the harmonic series for $p=1/2$ and dominates the harmonic series for $p>1/2$.

The numerical experiments were carried out in Rstudio [52] using the `fminunc` command provided in the `pracma` package [4]. We perturbed $f_1^\dagger$ and $f_2^\dagger$ in the following manner:

$$f_i^\delta(t) = f_i^\dagger(t) + \delta\sin(100\pi t) \qquad \text{for} \quad i=1,2.$$

δ	α_*	h_*	$\Vert u^\delta_{h_*,\alpha_*} - u_1^\dagger \Vert$	$\frac{\Vert u^\delta_{h_*,\alpha_*} - u_1^\dagger \Vert}{\Vert u_1^\dagger \Vert}$	$\delta^{\frac{p}{a+p}}$
0.0100	$2.15 \cdot 10^{-7}$	0.0215	0.1901	0.2200	0.4642
0.0050	$2.14 \cdot 10^{-8}$	0.0121	0.1151	0.1332	0.4135
0.0025	$2.12 \cdot 10^{-9}$	0.0068	0.0588	0.0680	0.3684
0.0012	$2.10 \cdot 10^{-10}$	0.0038	0.0403	0.0466	0.3282
0.0006	$2.09 \cdot 10^{-11}$	0.0021	0.0790	0.0914	0.2924
0.0003	$2.07 \cdot 10^{-12}$	0.0012	0.0099	0.0115	0.2605
0.0002	$2.05 \cdot 10^{-13}$	0.0007	0.0051	0.0059	0.2321

Table 3: Parameter selection and resulting errors for the regularized solutions for $u_1^\dagger$ for different noise levels and with $p=0.2$.

Tables 3 and 4 present the results for the oversmoothing Tikhonov regularization in the finite-dimensional setting. Table 3 presents results for the example $u_1^\dagger$, whit p chosen as $p=0.2$ and Table 4 presents results for $u_2^\dagger$ with p chosen as $p=0.45$. In each table, the second and the third column list values of the regularization parameters $\alpha=\alpha_*$ and $h=h_*$, respectively, for the different noise levels in column 1. The third and the fourth column show the error and the relative error of the regularized solution, respectively. The fifth column includes the values for the established converge rate $\delta^{p/(a+p)}$. The values for α are rounded to three significant digits in scientific notation. All other values are rounded to four decimal places. The tables confirm the established convergence rate.

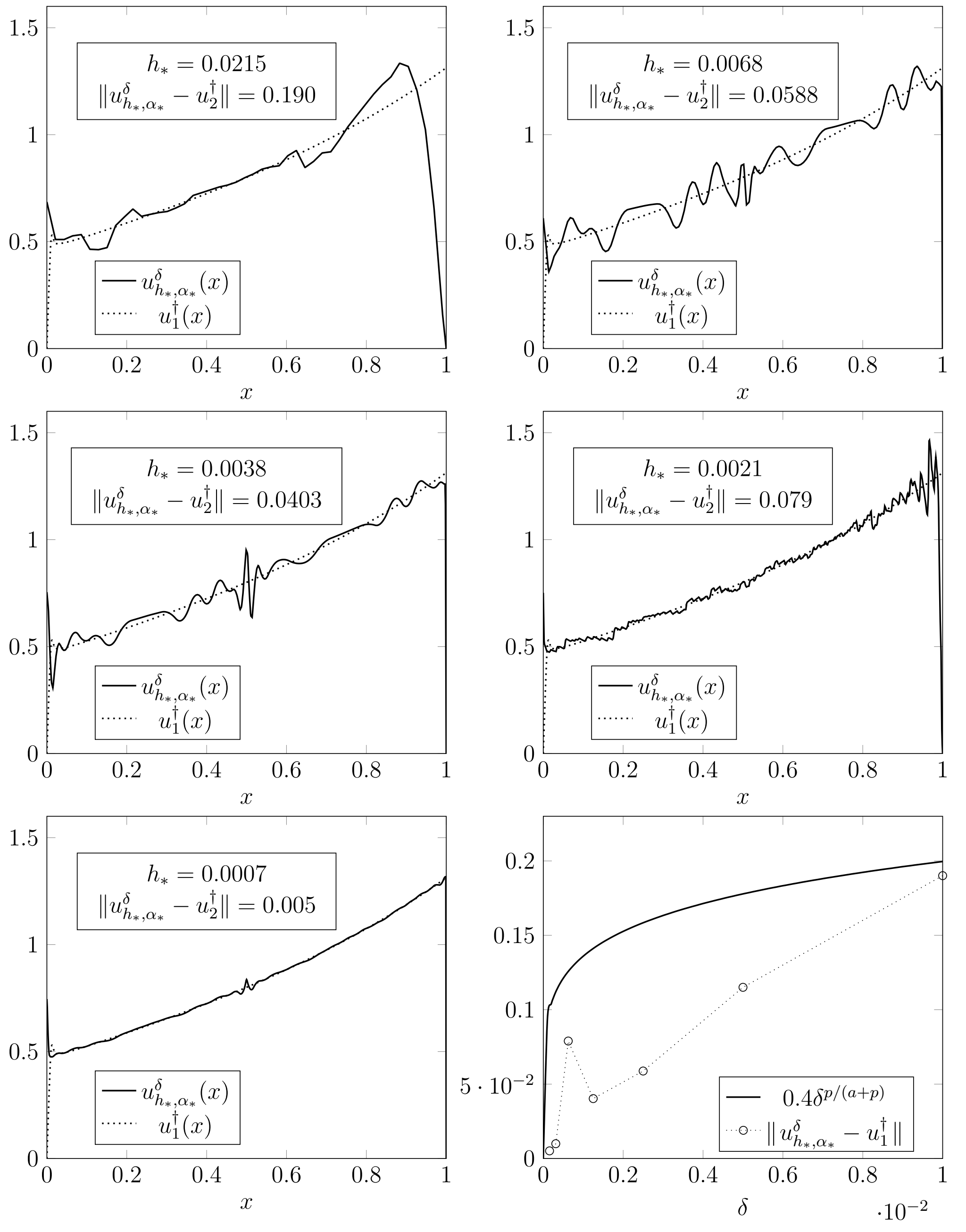

Figure 2: Regularized solutions for the oversmoothing Tikhonov regularization on the space of linear splines for $u_1^\dagger \in X_p$, $0 \leq p < 1/4$ and different values of δ. Bottom-right: Comparison of the errors to the established convergence rate for $p = 0.2$.

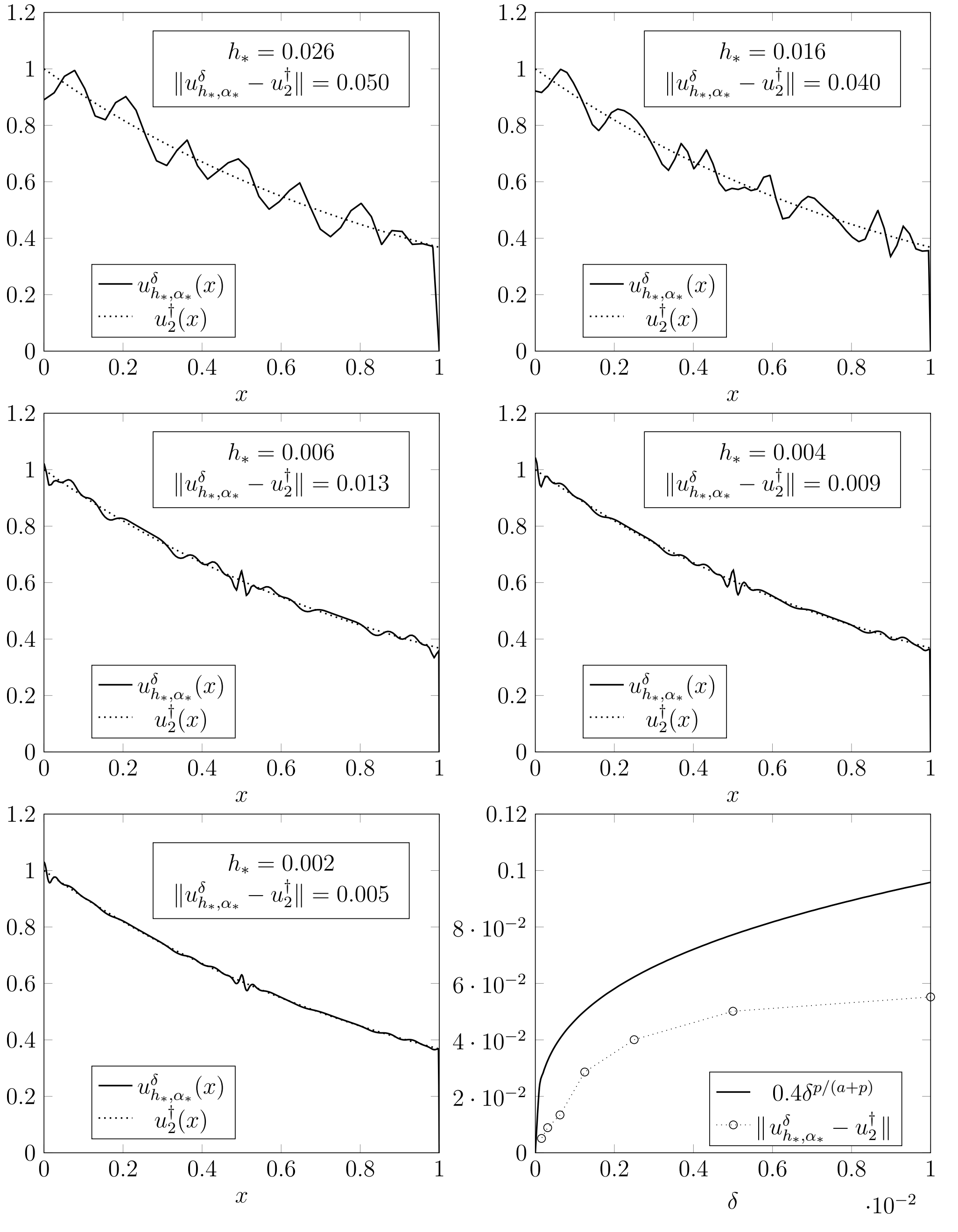

Figure 3: Regularized solutions for the oversmoothing Tikhonov regularization on the space of linear splines for $u_2^\dagger \in X_p$, $0 \leq p < 1/2$ and different values of δ. Bottom-right: Comparison of the errors to the established convergence rate for $p = 0.45$.

δ	α_*	h_*	$\| u^\delta_{h_*,\alpha_*} - u^\dagger_2 \|$	$\frac{\| u^\delta_{h_*,\alpha_*} - u^\dagger_2 \|}{\| u^\dagger_2 \|}$	$\delta^{\frac{p}{a+p}}$
0.0100	$3.04 \cdot 10^{-6}$	0.0418	0.0552	0.0840	0.2395
0.0050	$4.49 \cdot 10^{-7}$	0.0259	0.0502	0.0763	0.1931
0.0025	$6.64 \cdot 10^{-8}$	0.0161	0.0401	0.0609	0.1558
0.0012	$9.81 \cdot 10^{-9}$	0.0100	0.0286	0.0435	0.1256
0.0006	$1.45 \cdot 10^{-9}$	0.0062	0.0134	0.0204	0.1013
0.0003	$2.14 \cdot 10^{-10}$	0.0038	0.0089	0.0135	0.0817
0.0002	$3.16 \cdot 10^{-11}$	0.0024	0.0051	0.0077	0.0659

Table 4: Parameter selection and resulting errors for the regularized solutions for $u^\dagger_2$ for different noise levels and with $p = 0.45$.

Figures 2 and 3 show the regularized solutions for $u^\dagger_1$ and $u^\dagger_2$, respectively. Each figure consists of six graphs: the upper four and the bottom-left graphs show the regularized solutions as solid lines for different noise levels in comparison to the actual solution $u^\dagger$, which is displayed in a dotted line. In the bottom-right graph, the errors for different noise levels are compared to the established convergence rate. The figures illustrate how the accuracy of the regularized solutions increases as the noise level δ decreases, thereby visually confirming our theoretical findings. The boundary condition $u^\delta_\alpha(1) = 0$ does not affect the accuracy of the regularized solutions as much as in the continuous setting because of the rapid drop of the spline functions at the right endpoint.

10 Conclusion and outlook

In the infinite-dimensional setting, we have established convergence rates for the oversmoothing Tikhonov regularization under a mixed source condition, considering a priori as well as a posteriori parameter choice strategies. Identifying suitable functions that satisfy the source condition is challenging. However, after determining the asymptotic behavior of Fourier coefficients for a specific function, we were able to validate the source condition for this function. Numerical experiments involving this function confirmed our findings regarding the convergence rates.

Within the research area of discretization in regularization methods, oversmoothing has not yet been investigated. Therefore, our results provide an initial step towards incorporating the aspect of oversmoothing into the finite-dimensional setting. We have focused on establishing convergence rates under Hölder-type source conditions and a priori choices for the regularization parameter and the discretization level. We were able to replace the strong assumption of $u^\dagger$ being an interior point of the domain of F with the weaker assumption that $V_h \subset X_1$.

Future research encompasses promising tasks. First, extending our results to a more general source condition in terms of an index function as in [38] in the infinite-dimensional setting or in terms of a mixed source condition in the finite-dimensional setting is a promising task of future work. Second, further examples that satisfy these source conditions and the two-sided Lipschitz inequality (see Assumption 3.7 (f)) should be identified. This two-sided Lipschitz inequality has recently been relaxed to hold only locally. In earlier studies, such as [24], this inequality was formulated to hold globally, on the entire domain $\mathcal{D}$. Further relaxation of this inequality might be another future objective. Moreover, exploring a posteriori parameter choice strategies and extending results to the Banach space setting in the finite-dimensional setting are intriguing tasks for further investigation. Lastly, improving the numerical results of Chapter 6 by considering a solution that satisfies the boundary condition could be an additional objective.

References

[1] R. A. Adams and J. J. F. Fournier, *Sobolev Spaces*, 2nd ed., vol. 140 of *Pure and applied mathematics*, Amsterdam: Academic Press, 2003.

[2] H. Bateman and Bateman Manuscript Project, *Higher Transcendental Functions*. New York/Toronto/London: McGraw-Hill Book Company, 1953, vol. 1.

[3] J. Baumeister, *Stable Solution of Inverse Problems* (Advanced Lectures in Mathematics). Braunschweig/Wiesbaden: F. Vieweg & Sohn, 1987.

[4] H. W. Borchers, *Pracma: Practical Numerical Math Functions*, R package version 2.3.3, 2021.

[5] D.-H. Chen, B. Hofmann, and I. Yousept, "Oversmoothing Tikhonov regularization in Banach spaces," *Inverse Problems*, vol. 37, no. 8, p. 085 007, 2021.

[6] M. Crouzeix and V. Thomée, "The stability in L_p and W_p^1 of the L_2-projection onto finite element function spaces," *Mathematics of Computation*, vol. 48, no. 178, pp. 521–532, 1987.

[7] P. Deuflhard, H. W. Engl, and O. Scherzer, "A convergence analysis of iterative methods for the solution of nonlinear ill-posed problems under affinely invariant conditions," *Inverse Problems*, vol. 14, no. 5, pp. 1081–1106, 1998.

[8] H. Engl, M. Hanke, and A. Neubauer, *Regularization of Inverse Problems*, vol. 375 of *Mathematics and Its Applications*, Dordrecht/Boston/ London: Kluwer Academic Publishers, 1996.

[9] K. Finck von Finckenstein, *Einführung in die Numerische Mathematik*. München: Carl Hanser Verlag, 1977, vol. 1.

[10] J. Flemming, *Variational Source Conditions, Quadratic Inverse Problems, Sparsity Promoting Regularization, New Results in Modern Theory of Inverse Problems and an Application in Laser Optics* (Frontiers in Mathematics). Cham: Birkhauser, 2018.

[11] O. Forster, *Analysis 1, Differential- und Integralrechnung einer Veränderlichen* (Vieweg Studium Grundkurs Mathematik), 4th rev. ed. Braunschweig: F. Vieweg & Sohn, 1983.

[12] D. Gerth, B. Hofmann, and C. Hofmann, "Case Studies and a Pitfall for Nonlinear Variational Regularization Under Conditional Stability," in *Inverse Problems and Related Topics. ICIP2 2018. Springer Proceedings in Mathematics & Statistics,* J. Cheng, S. Lu, and M. Yamamoto, Eds., vol. 310, Singapore: Springer Nature, 2020, ch. 8, pp. 177–203.

[13] R. Gorenflo and M. Yamamoto, "Operator theoretic treatment of linear Abel integral equations of first kind," *Japan Journal of Industrial and Applied Mathematics*, vol. 16, no. 1, pp. 137–161, 1999.

[14] C. Groetsch, *Inverse Problems in the Mathematical Sciences.* Braunschweig/Wiesbaden: F. Vieweg & Sohn, 1993.

[15] C. Groetsch, *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, vol. 105 of *Research Notes in Mathematics*, Boston: Pitman, 1984.

[16] J. Hadamard, *Lectures on the Cauchy Problem in Linear Partial Differential Equations* (Silliman Memorial Lectures). New Haven: Yale University Press, 1923.

[17] G. Helmberg, *Introduction to Spectral Theory in Hilbert Space.* Amsterdam: North-Holland, 1975.

[18] B. Hofmann, *Mathematik inverser Probleme* (Mathematik für Ingenieure und Naturwissenschaftler). Stuttgart/Leipzig: Teubner, 1999.

[19] B. Hofmann and C. Hofmann, "The Impact of the Discrepancy Principle on the Tikhonov-Regularized Solutions with Oversmoothing Penalties," *Mathematics*, vol. 8, no. 3, p. 331, 2020.

[20] B. Hofmann, C. Klinkhammer, and R. Plato, "New results for variational regularization with oversmoothing penalty term in Banach spaces," *Communications on Analysis and Computation*, vol. 1, no. 2, pp. 94–115, 2023.

[21] B. Hofmann and P. Mathé, "A Priori Parameter Choice in Tikhonov Regularization with Oversmoothing Penalty for Non-linear Ill-Posed Problems," in *Inverse Problems and Related Topics. ICIP2 2018. Springer Proceedings in Mathematics & Statistics,* J. Cheng, S. Lu, and M. Yamamoto, Eds., vol. 310, Singapore: Springer Nature, 2020, ch. 8, pp. 169–176.

[22] B. Hofmann and P. Mathé, "Analysis of Profile Functions for General Linear Regularization Methods," *SIAM Journal on Numerical Analysis*, vol. 45, no. 3, pp. 1122–1141, 2007.

[23] B. Hofmann and P. Mathé, "Tikhonov regularization with oversmoothing penalty for non-linear ill-posed problems in Hilbert scales," *Inverse Problems*, vol. 34, no. 1, p. 015 007, 2017.

[24] B. Hofmann and R. Plato, "Convergence results and low-order rates for nonlinear Tikhonov regularization with oversmoothing penalty term," *ETNA - Electronic Transactions on Numerical Analysis*, vol. 53, pp. 313–328, 2020.

[25] B. Hofmann and R. Plato, "On ill-posedness concepts, stable solvability and saturation," *Journal of Inverse and Ill-posed Problems*, vol. 26, no. 2, pp. 287–297, 2017.

[26] B. Hofmann and O. Scherzer, "Local ill-posedness and source conditions of operator equations in Hilbert spaces," *Inverse Problems*, vol. 14, no. 5, pp. 1189–1206, 1998.

[27] T. Hohage, "Iterative Methods in Inverse Obstacle Scattering: Regularization Theory of Linear and Nonlinear Exponentially ill-posed Problems," Ph.D. dissertation, Johannes Kepler University Linz, 1999.

[28] T. Hohage, "Regularization of exponentially ill-posed problems," *Numerical Functional Analysis and Optimization*, vol. 21, no. 3-4, pp. 439–464, 2000.

[29] B. Kaltenbacher and A. Neubauer, "Convergence of projected iterative regularization methods for nonlinear problems with smooth solutions," *Inverse Problems*, vol. 22, no. 3, pp. 1105–1119, 2006.

[30] B. Kaltenbacher, "A Projection-Regularized Newton Method for Nonlinear Ill-Posed Problems and Its Application to Parameter Identification Problems with Finite Element Discretization," *SIAM Journal on Numerical Analysis*, vol. 37, no. 6, pp. 1885–1908, 2000.

[31] B. Kaltenbacher, "Regularization by projection with a posteriori discretization level choice for linear and nonlinear ill-posed problems," *Inverse Problems*, vol. 16, no. 5, pp. 1523–1539, 2000.

[32] J. T. King and A. Neubauer, "A variant of finite-dimensional Tikhonov regularization with a-posteriori parameter choice," *Computing*, vol. 40, no. 2, pp. 91–109, 1988.

[33] C. Klinkhammer and R. Plato, "Analysis of the discrepancy principle for Tikhonov regularization with oversmoothing penalty term under low order source conditions," *Inverse Problems*, vol. 39, p. 105 014, 2023.

[34] S. G. Krein and Y. I. Petunin, "Scales of Banach spaces," *Russian Mathematical Surveys*, vol. 21, no. 2, pp. 85–159, 1966.

[35] R. Kress, *Linear Integral Equations*, 1st ed., vol. 82 of *Applied Mathematical Sciences*, Berlin/Heidelberg: Springer, 1989.

[36] A. K. Louis, *Inverse und schlecht gestellte Probleme*. Stuttgart: Teubner, 1989.

[37] S. Lu and S. V. Pereverzev, *Regularization Theory for Ill-posed Problems*, vol. 58 of *Inverse and Ill-Posed Problems*, Berlin: De Gruyter, 2013.

[38] P. Mathé and S. V. Pereverzev, "Geometry of linear ill-posed problems in variable Hilbert scales," *Inverse Problems*, vol. 19, no. 3, pp. 789–803, 2003.

[39] P. Miller and T. Hohage, "Convergence rates for oversmoothing Banach space regularization," *ETNA - Electronic Transactions on Numerical Analysis*, vol. 57, pp. 101–126, 2022.

[40] V. A. Morozov, *Methods for Solving Incorrectly Posed Problems*. New York: Springer, 1984.

[41] F. Natterer, "Error bounds for Tikhonov regularization in Hilbert scales," *Applicable Analysis*, vol. 18, no. 1-2, pp. 29–37, 1984.

[42] F. Natterer, "Regularisierung schlecht gestellter Probleme durch Projektionsverfahren," *Numerische Mathematik*, vol. 28, no. 3, pp. 329–341, 1977.

[43] A. Neubauer, "An a Posteriori Parameter Choice for Tikhonov Regularization in Hilbert Scales Leading to Optimal Convergence Rates," *SIAM Journal on Numerical Analysis*, vol. 25, no. 6, pp. 1313–1326, 1988.

[44] A. Neubauer, "Tikhonov regularisation for non-linear ill-posed problems: Optimal convergence rates and finite-dimensional approximation," *Inverse Problems*, vol. 5, no. 4, pp. 541–557, 1989.

[45] A. Neubauer, "Tikhonov regularization of nonlinear Ill-posed problems in Hilbert scales," *Applicable Analysis*, vol. 46, no. 1–2, pp. 59–72, 1992.

[46] R. Plato and G. Vainikko, "On the regularization of projection methods for solving Ill-posed problems," *Numerische Mathematik*, vol. 57, no. 1, pp. 63–79, 1990.

[47] R. Plato, *Numerische Mathematik Kompakt, Grundlagenwissen Für Studium und Praxis*, 5th ed. Berlin: Springer Spektrum, 2021.

[48] R. Plato, "Optimal algorithms for linear ill-posed problems yield regularization methods," *Numerical Functional Analysis and Optimization*, vol. 11, pp. 111–118, 1990.

[49] R. Plato and B. Hofmann, “Variational regularization with oversmoothing penalty term in Banach spaces,” 2022, Paper submitted. Preprint availabe as arXiv:2206.02010v2, https://arxiv.org/abs/2206.02010.

[50] R. Ramlau, “A modified landweber method for inverse problems,” *Numerical Functional Analysis and Optimization*, vol. 20, no. 1–2, pp. 79–98, 1999.

[51] R. Remmert and G. Schumacher, *Funktionentheorie 2*, 3rd rev. ed. Berlin/Heidelberg: Springer, 2007.

[52] RStudio Team, *RStudio: Integrated Development Environment for R*, RStudio, PBC, Boston, MA, 2021.

[53] O. Scherzer, “An iterative multi level algorithm for solving nonlinear ill-posed problems,” *Numerische Mathematik*, vol. 80, no. 4, pp. 579–600, 1998.

[54] T. Schuster, B. Kaltenbacher, B. Hofmann, and K. S. Kazimierski, *Regularization Methods in Banach Spaces*, vol. 10 of *Radon Series on Computational and Applied Mathematics*, Berlin/Boston: De Gruyter, 2012.

[55] U. Tautenhahn, “Optimality for ill-posed problems under general source conditions,” *Numerical Functional Analysis and Optimization*, vol. 19, no. 3-4, pp. 377–398, 1998.

[56] A. E. Taylor, *Introduction to Functional Analysis.* New York: Wiley & Sons, 1958.

[57] G. Vainikko, “The discrepancy principle for a class of regularization methods,” *USSR Computational Mathematics and Mathematical Physics*, vol. 22, no. 3, pp. 1–19, 1982.

[58] G. Vainikko, *Regularisierung nichkorrekter Aufgaben*, Preprint No 200, Universität Kaiserslautern, 1991.

[59] F. P. Vasil’ev, *Methods for Solving Extremal Problems.* Moskau: Nauka, 1980, In Russian.

[60] A. Zygmund, *Trigonometric Series*, 2nd reprinted ed. Cambridge: Cambridge University Press, 1979, vol. 1 and 2.