Inverse problems with Poisson data: Tikhonov-type regularization and iteratively regularized Newton methods

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades

'Doctor rerum naturalium'

der Georg-August-Universität Göttingen

vorgelegt von

Frank Werner

aus Hofgeismar

Göttingen 2011

Referent:	Prof. Dr. T. Hohage
Korreferent:	Prof. Dr. A. Munk
Tag der mündlichen Prüfung:	23.01.2012

`Mathematik ist nicht alles,`
`aber ohne Mathematik ist alles nichts.`

Hans-Olaf Henkel
(ehem. Präsident der Leibniz-Gemeinschaft)

ACKNOWLEDGMENTS

First of all, I would like to thank Prof. Dr. Thorsten Hohage for the supervision of this thesis. He supported me with several helpful hints and suggestions when I needed them, and he always had time to discuss my work. He moreover introduced me into the very interesting field of statistical inverse problems and suggested the fruitful subject of the thesis at hand. It is a pleasure to have Prof. Dr. Axel Munk as co-referee for this thesis, and I want to express my gratitude to him. At this point I also want to thank my late upper school math teacher Jochen Desel, which I understand as my first mentor and without whom I would probably never have studied mathematics.

Several parts of my work would have been impossible without hints from and discussions with colleagues. To name the most important ones, I would like to thank Jens Flemming for his work on variational inequalities, and especially for providing his results on the optimality of those inequalities under logarithmic source conditions. Without his proof of Lemma 3.22, my theoretical results would have been much less satisfying. Moreover I am indebted to Patricia Reynaud-Bouret who helped me to prove Theorem 4.1 and hence provided the possibility to apply my results to the case of full random data.

I would like to thank Olha Ivanyshyn and Thorsten Hohage for source code concerning the forward implementation of the inverse obstacle scattering and phase retrieval problems and Robert Stück for providing his implementation of the 4Pi algorithm and several helpful hints and comments about it.

Furthermore, I especially wish to express my gratitude to the other members of the inverse problems working group in Göttingen for interesting remarks, a good working atmosphere and entertainment in several coffee breaks. Special thanks go to my officemate Christoph, who always had time for discussions not only on the topic of this thesis. I am moreover indebted to Robert and Carolin for careful proof-reading parts of this work.

During my whole PhD studies I was supported by the Research Training Group 1023 'Identification in Mathematical Models: Synergy of Stochastic and Numerical Methods' and I would like to acknowledge not only the financial support but moreover many helpful meetings and the exchange of information with other members of the RTG 1023. I also want to acknowledge support by the Federal Ministry of Education and Research (BMBF) through the project INVERS.

Last but not least I want to thank my wife Kirsten for her encouraging help during the whole three years and her never ending patience while listening to me explaining my work. As you sure know, explaining your own work to others is the most important step in understanding and updating it, and so the support I received from her was invaluable for the development of this thesis.

CONTENTS

CHAPTER
ONE

INTRODUCTION

In many practical problems one has to determine some model parameters from observed data. Whereas the forward problem to predict the data given the model parameters is in the most cases well-understood and stable solvable, the inverse problem to determine the parameters from usually noisy data is often instable and much less is known. Mathematically, this fact is described by the notion of *well- and ill-posedness* due to HADAMARD [Had52]: A problem is called *well-posed*, if

(a) a solution exists,

(b) the solution is unique and

(c) the solution depends continuously on the data.

If any of these conditions is violated, the problem is called *ill-posed*. Since in all applications the data will be measured only up to some noise, condition (c) is the most delicate one: If the solution does not depend continuously on the data and the data is erroneous, any ad hoc reconstruction must considered to be useless.

Let us assume that the problem under consideration can be formulated as an operator equation

$$F(u) = g \tag{I}$$

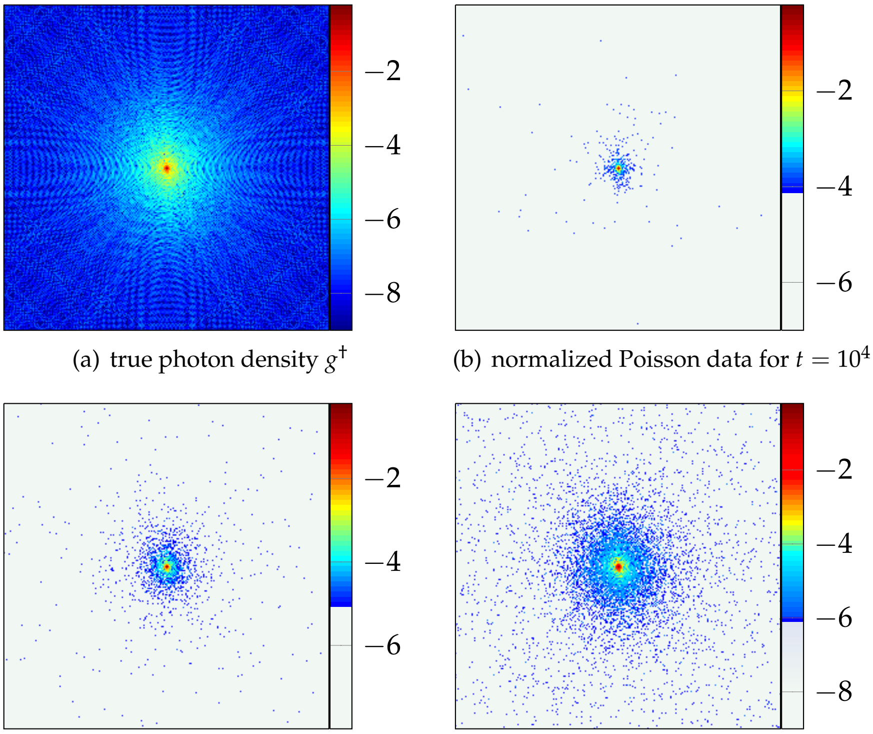
where g are the data, u the unknown parameters and the operator F describes the (possibly nonlinear) dependence of both. Considering the conditions (a)-(c) from above, the problem (I) is *well-posed* if and only if

(a) the operator is surjective,

(b) the operator is injective and

(c) the operator is continuously invertible.

This work is mainly concerned with applications from **photonic imaging**, where the observed data corresponds to counts of photons which have interacted with an unknown object $u^\dagger$ of interest and the aim is to reconstruct as much information as possible about $u^\dagger$. Due to fundamental physical and mathematical reasons, photon count data is Poisson distributed as we will see in Chapter 2. Hence, the observed data is not corrupted by deterministic noise but given by a stochastic process, namely a Poisson process G_t. Roughly spoken one observes a random set of points (corresponding to the photon detections) with a special structure (this will also be clarified in Chapter 2) and tries to reconstruct the underlying reason which causes the true photon density. The additional parameter t denotes the exposure time of the measurement procedure and is proportional to the total number of photon counts. As $t \to \infty$ we expect the normalized data $\frac{1}{t}G_t$ to tend to the true photon density $g^\dagger = F\left(u^\dagger\right)$ in a suitable sense. This is shown exemplarily in Figure 1.1.



(a) true photon density $g^\dagger$ (b) normalized Poisson data for $t = 10^4$

(c) normalized Poisson data for $t = 10^5$ (d) normalized Poisson data for $t = 10^6$

Figure 1.1: Logarithmic plots of the true photon density $g^\dagger$ and normalized Poisson data $\frac{1}{t}G_t$ for different values of t. The color values have been chosen such that regions without photon counts are white to illustrate the sparsity of the data.

Photonic imaging includes examples like Positron Emission Tomography, coherent X-ray imaging, astronomical imaging and fluorescence microscopy. In all these applications the parameter t is limited due to various reasons:

- In Positron emission tomography a large t corresponds to a stronger source of radiation inside the patient and causes hence more harm.
- In coherent X-ray and astronomical imaging the observation time is limited.
- In fluorescence microscopy t is limited due to the effect of photobleaching.

Moreover, a long observation time may always cause motion artifacts. The parameter t has different physical meanings in all aforementioned examples, but for simplicity we will call it *observation time* in the following. As per description one is interested in reconstructing u also for small t which corresponds to very few photon counts as shown in Figure 1.1, subplot (b). Especially in those cases the information about the data distribution must be incorporated into the reconstruction procedure.

It seems to be a promising approach to seek approximations u_{app} of $u^\dagger$ as

$$u_{\text{app}} \in \operatorname*{argmax}_u \mathscr{P}\left(G_t \,\middle|\, \text{the true intensity is } F(u)\right), \tag{II}$$

i.e. to choose u_{app} such that it maximizes the probability to observe G_t. The ansatz (II) is called *maximum-likelihood approach* and is widely used to gain estimators for distribution characteristics. Replacing the term $\mathscr{P}\left(G_t \,\middle|\, \text{the true intensity is } F(u)\right)$ in (II) by its negative logarithm and substituting argmax by argmin does not change u_{app}, and hence we will denote

$$\mathcal{S}(F(u); G_t) := -\ln\left(\mathscr{P}\left(G_t \,\middle|\, \text{the true intensity is } F(u)\right)\right)$$

in the following. This shortens the notation and allows us to analyze the approach (II) in a more general setup. Moreover, $\mathcal{S}$ will include the parameter t to ensure a proper scaling of the problem. For our example of Poisson data, the functional $\mathcal{S}$ is convex in its first argument, which simplifies the minimization of $\mathcal{S}$.
Unfortunately, the ill-posedness of (I) carries over to (II) in the sense that if F is not continuously invertible, then also u_{app} does in general not depend continuously on the data. To overcome this problem, one adds a weighted *penalty term* $\mathcal{R}$ to the functional and approximates $u^\dagger$ by

$$u_\alpha \in \operatorname*{argmin}_u \left[\mathcal{S}(F(u); G_t) + \alpha\mathcal{R}(u)\right]. \tag{III}$$

The additional term $\mathcal{R}$ usually includes *a priori* knowledge about $u^\dagger$ (it might for example be known from physics that $u^\dagger$ is smooth) and should enforce u_α not to oscillate too much. Regrettably, the approach (III) needs an additional *regularization parameter* $\alpha > 0$ which must be chosen in a proper way. It turns out that determining a good value for α is a difficult problem in practice, especially if the reconstructions should be done completely automatic.
The functional in (III) is called *Tikhonov-type functional* due to TIKHONOV [Tik63a,Tik63b]. He considered approximating solutions to (I) as minimizers of

$$u \mapsto \|F(u) - G_t\|_{\mathbb{Y}}^2 + \alpha \|u\|_{\mathbb{X}}^2$$

where the $\mathbb{X}$-norm includes derivatives of the function u.
The reconstruction procedure (III) leads to several questions which need to be answered:

- Does u_α in (III) exist?
- Does u_α in (III) depend continuously on the data?
- For a proper choice of α depending on t (and maybe G_t), does u_α converge in expectation or at least in probability to the true solution $u^\dagger$ as $t \to \infty$?
- Can the rate of convergence in the former item be specified?

It can be shown under reasonable assumptions on $\mathcal{S}$ and $\mathcal{R}$ that u_α as in (III) exists and depends continuously on the data. To establish convergence (and convergence rates) one usually needs a concentration inequality, which ensures that the difference between $\mathcal{S}(F(u); G_t)$ and $\mathcal{S}\left(F(u); g^\dagger\right)$ is small with overwhelming probability as $t \to \infty$. Such a concentration inequality will be established in Chapter 4. This leads to the first main result of this thesis, namely rates of convergence in expectation for u_α as the observation time t tends to ∞. One might argue that this result has no direct meaning for practical applications since t will be fixed or at least bounded in this context, but it indicates that the method (III) works and yields *reasonable* reconstructions. Furthermore, convergence rates provide a qualitative estimate how much the reconstructions will profit by an improved measurement procedure which yields higher values of t, i.e. how much an increase of the number of photons will increase the accuracy of the reconstruction. This is of interest in practice for example to decide if more harm for a patient should be accepted or if a more expensive device should be used.

We consider especially three applications from photonic imaging, namely an inverse obstacle scattering problem without phase, a phase retrieval problem from optics and a semiblind deconvolution problem. All three examples have in common that they are nonlinear (i.e. the describing operator F is nonlinear), and thus the functional from (III) is no longer convex in u. Unfortunately this implies that there might exist many local minimizers and thus u_α is difficult to calculate.
Usually the first choice for solving nonlinear equations is Newton's method. It consists in linearizing the problem (I) and solving the linearized equations iteratively.
For a nonlinear Fréchet differentiable operator $F : D(F) \subset \mathbb{X} \to \mathbb{Y}$ one ends up with the linearized equations

$$F(u_n) + F'[u_n](u - u_n) = g \tag{IV}$$

which need to be solved in every Newton step for $u = u_{n+1}$.
At this point, we are again able to incorporate our knowledge about the distribution of G_t, i.e. we approximate $u^\dagger$ by

$$u_{n+1} \in \operatorname*{argmin}_{u} \mathcal{S}\left(F(u_n) + F'[u_n](u - u_n); G_t\right). \tag{V}$$

Unfortunately, the ill-posedness of F carries over to $F'[u_n]$ under very mild conditions and hence also (V) is ill-posed. Thus, in every Newton step some sort of regularization is needed. This leads to our second approach to tackle (I), namely an iteratively regularized Newton method of the form

$$u_{n+1} := \operatorname*{argmin}_{u \in \mathfrak{B}} \left[\mathcal{S}\left(F(u_n) + F'[u_n](u - u_n); G_t\right) + \alpha_n \mathcal{R}(u)\right] \tag{VI}$$

where $u_0 \in \mathbb{X}$ is some (sufficiently good) initial guess and $\mathcal{R}$ is again a penalty term based on a priori knowledge. As Newton's method is known to converge very fast (at least locally) under reasonable conditions on F, we expect that only a few iterations of the method (VI) are necessary. As in (III), the regularization parameters $\alpha_n > 0$ need to be chosen in a proper way, but the more difficult issue in (VI) is the determination of a suitable stopping index N. The second main result of this thesis are rates of convergence for u_N in expectation at $t \to \infty$ for different choices of N, namely an *a priori* choice where N is determined **before** the reconstruction starts and an adaptive *a posteriori* choice where the index N is determined **during** the iteration procedure.

The existing theory on inverse problems with Poisson data is far away from being complete. Inverse problems with Poisson data have been studied by ANTONIADIS & BIGOT [AB06] for strongly restricted classes of linear operators F only. The method (III) has been studied by BARDSLEY [Bar10], BENNING & BURGER [BB11] and Flemming [Fle10, Fle11], but only under deterministic noise assumptions. The case of additive random noise was treated by BISSANTZ, HOHAGE & MUNK [BHM04], but this does not apply to the case of Poisson data. For the general Newton-type method (IV) no deterministic convergence analysis exists to the authors best knowledge. A convergence analysis for nonlinear inverse problems with additive random noise was presented by BAUER, HOHAGE & MUNK [BHM09], but the theory there does again not apply to the case of Poisson data. It is the main aim of this thesis to provide a proper setup for inverse problems with Poisson data and to present first convergence rates results for the aforementioned methods.

The detailed organization of this thesis is as follows.

In **Chapter 2** we will describe the data distribution and the model we use explicitly. Afterwards we will present the considered applications in detail.

Chapter 3 is dedicated to the analysis of (III) in a deterministic setup. We will repeat the known theory about the well-posedness of (III) before we present our assumptions which are used to obtain convergence rates. These assumptions will be linked to the smoothness of the unknown solution $u^\dagger$ and formulated in terms of variational inequalities, which are discussed in detail. Afterwards we present our convergence theorems for Tikhonov-type regularization (III) under an a priori as well as a Lepskiĭ-type parameter choice rule.

In **Chapter 4** we will apply the theory from Chapter 3 to the case of random data. Therefore we will first provide a concentration inequality and afterwards modify our proofs from the aforementioned chapter to gain convergence rates under an a priori as well as a Lepskiĭ-type parameter choice rule. This chapter closes part one of this thesis and completes our theoretical results on Tikhonov-type regularization.

Chapter 5 states the known results on the iteratively regularized Gauss-Newton method (IRGNM), which is given by (VI) with $\mathcal{S}(g;\hat{g}) = \|g-\hat{g}\|_{\mathbb{Y}}^p$ and $\mathcal{R}(u) = \|u-u_0\|_{\mathbb{X}}^q$ for Banach norms $\|\cdot\|_{\mathbb{Y}}$ and $\|\cdot\|_{\mathbb{X}}$. The presented results form the basis for the generalization (VI) of this method we aim for.

The generalization of the IRGNM - the proposed iteratively regularized Newton method (VI) - is analyzed in the deterministic setup in **Chapter 6**. To obtain rates of convergence for this method, we first describe generalized assumptions on the nonlinearity of F and combine them afterwards with the smoothness assumptions from Chapter 3 to gain convergence rates. Again, we consider an a priori as well as a Lepskiĭ-type stopping rule.

In **Chapter 7** we will apply the theory from Chapter 6 to the case of random data. This is done similarly to the case of Tikhonov-type regularization with the help of the concentration inequality from Chapter 4. This chapter completes the theory for (VI).

The last part of this thesis is contained in **Chapter 8**, where we present our simulations and calculations for the applications from Chapter 2.

CHAPTER
TWO

INVERSE PROBLEMS WITH POISSON DATA

In this chapter we will introduce the Poisson distribution and present some examples for inverse problems with Poisson data. In this field, ideal measurements (neglecting read-out errors and finite volume averaging of the detectors) are given by a **Poisson point process** with the true data $g^\dagger$ as mean, as we will motivate in Section 2.2. We will give a definition and fundamental properties of a Poisson point process.

2.1 Poisson data

2.1.1 The Poisson distribution

Before we are able to describe Poisson point processes, we need to recall the Poisson distribution. It is used to model the emission of radioactive particles over a fixed time interval or to account for the arrival times of telephone calls at an exchange. Roughly spoken, the Poisson distribution is used whenever one counts how often a phenomenon happens during a period of time or in a given area under the assumption that the probability of the phenomenon to happen is constant in time or space.
The Poisson distribution is defined as follows:

DEFINITION 2.1:
A random variable X is said to have the ***Poisson distribution*** $\mathscr{P}(\lambda)$ *with parameter* $\lambda \in (0, \infty)$, *if its range consists only of non-negative integers and it holds*

$$\mathbf{P}(X = k) = \exp(-\lambda) \frac{\lambda^k}{k!} \tag{2.1a}$$

for all $k \in \mathbb{N}_0$. *For such X we write* $X \sim \mathscr{P}(\lambda)$.

This definition can be extended easily to the limit cases $\lambda \in \{0, \infty\}$ by defining a Poisson distribution $\mathscr{P}(0)$ to be concentrated at 0, i.e.

$$\mathbf{P}(X = k) = \begin{cases} 1 & \text{if } k = 0, \\ 0 & \text{otherwise,} \end{cases} \tag{2.1b}$$

and a Poisson distribution $\mathscr{P}(\infty)$ to be concentrated at ∞, i.e.

$$\mathbf{P}(X = k) = \begin{cases} 1 & \text{if } k = \infty, \\ 0 & \text{otherwise.} \end{cases} \tag{2.1c}$$

One of the basic properties of the Poisson distribution is that both, mean and variance, are given by the parameter λ, i.e.

$$\mathbf{E}(X) = \lambda, \tag{2.2a}$$

$$\mathbf{V}(X) = \lambda, \tag{2.2b}$$

which can be obtained directly by summation.
Whenever $X \sim \mathscr{P}(\lambda)$ is observed, the question arises if the parameter λ can be estimated from the observation k. In the general theory of random variables there is a vast literature on suitable estimators of parameters or other distribution characteristics. We are mainly interested in the **maximum likelihood** property of an estimator which is defined as follows:

DEFINITION 2.2:
*Let $X \sim \mathscr{P}(\lambda)$ be a random variable and let $\bar{\lambda} = \bar{\lambda}(k)$ be an estimator for the parameter λ. $\bar{\lambda}$ is said to have the **maximum likelihood property**, if for any event $k \in \mathbb{N}_0$ it holds*

$$\mathbf{P}\left(X = k \mid X \sim \mathscr{P}(\bar{\lambda})\right) \geq \mathbf{P}\left(X = k \mid X \sim \mathscr{P}(\lambda^*)\right) \qquad \text{for all } \lambda^* \in [0, \infty). \tag{2.3}$$

Hence, a maximum likelihood estimator $\bar{\lambda}$ for λ maximizes the probability to observe k. Inserting the definition of the Poisson distribution into (2.3), it is easy to see that $\bar{\lambda}$ maximizes the so-called **maximum likelihood functional**

$$\lambda \mapsto \mathbf{P}\left(X = k \mid X \sim \mathscr{P}(\lambda)\right) = \exp(-\lambda)\frac{\lambda^k}{k!} \tag{2.4}$$

over $\lambda \in [0, \infty)$ for fixed $k \in \mathbb{N}_0$. Via differentiation of (2.4) we obtain easily that

$$\bar{\lambda} := k$$

is the only maximum likelihood estimator for the Poisson distribution. But moreover, the characterization of this estimator as the maximizer of (2.4) is quite helpful as already mentioned in the introduction. For simplicity, not the functional (2.4) itself is maximized, but the negative logarithm of (2.4) is minimized (which yields the same solution). The corresponding functional is known as the **negative log-likelihood functional** ρ_{poiss}. Up to a constant independent of λ it is given by

$$\rho_{\text{poiss}}(\lambda; k) := \lambda - k \ln(\lambda), \qquad k \in \mathbb{N}_0, \lambda \in [0, \infty)$$

where we use the convention $0 \ln(0) := 0$. In case of a vector $(X_1, ..., X_d)$ of independent Poisson distributed random variables $X_i \sim \mathscr{P}(\lambda_i)$ and an event vector $k = (k_1, ..., k_d) \in \mathbb{N}_0^d$, this generalizes to

$$\rho_{\text{poiss}}(\lambda; k) := \sum_{i=1}^{d} \left[\lambda_i - k_i \ln(\lambda_i)\right], \qquad k \in \mathbb{N}_0^d, \lambda \in [0, \infty)^d. \tag{2.5}$$

Using the convention $0 \ln(0) := 0$ as above and defining $\rho_{\text{poiss}}(\lambda; k) := \infty$ if $\lambda_i < 0$ for some $i \in \{1, ..., d\}$, the functional (2.5) is defined for any $\lambda \in \mathbb{R}^d$ and enforces its minimizers to fulfill $\lambda_i > 0$ if $k_i > 0$ and $\lambda_i \geq 0$ if $k_i = 0$. Note moreover that ρ_{poiss} is convex in its first argument.

The fact that the Poisson distribution $\mathscr{P}(\lambda)$ concentrates more and more around its expectation λ as $\lambda \to \infty$ is reflected by estimates in probability, which can be seen as prototypes of **concentration inequalities**. The most simple example of a concentration inequality is TSCHEBYSCHOW'S inequality

$$\mathbf{P}(|X - \mathbf{E}(X)| \geq a) \leq \frac{\mathbf{V}(X)}{a^2} \tag{2.6}$$

which holds for any $a > 0$ and any real valued random variable X with finite second moment. The concentration gets visible if we introduce a scaling parameter $t > 0$ which should be interpreted as an observation time and define for fixed $\lambda > 0$ the random variables $Y \sim \mathscr{P}(t\lambda)$ and the normalization $X := \frac{1}{t}Y$. Then (2.6) implies by $\mathbf{E}(X) = \lambda$ and $\mathbf{V}(X) = \frac{1}{t^2}\mathbf{V}(Y) = \frac{1}{t}\lambda$ that

$$\mathbf{P}(|X - \lambda| \geq a) \leq \frac{\lambda}{ta^2} \to 0 \qquad \text{as} \qquad t \to \infty.$$

This ensures that X is as more centered around its expectation $\mathbf{E}(X) = \lambda$ the larger t is.

2.1.2 Poisson point processes

In this section we will give a definition and present some fundamental properties of a Poisson point process. For more information about Poisson processes in general we refer to [Kin93] and the references therein. This section has been developed along [Kin93] and [RB03].

A Poisson point process can be described as a (countable) random set of points in the **state space** Ω. In our setup, Ω will be an observation manifold in $\mathbb{R}^d$, e.g. an open set or a sphere. The geometry of the space Ω is not of interest for the Poisson point process, the only thing required is a measurable space containing enough measurable sets. For such a random set of points to be a Poisson point process, two features are characteristic: statistical independence and the Poisson distribution. This is clarified in the following definition:

DEFINITION 2.3 (POISSON POINT PROCESS):
*Let (Ω, Σ) be a measurable space with $\{x\} \in \Sigma$ for all $x \in \Omega$. A random countable set of points G is said to be a **Poisson point process** (or short **Poisson process**) on (Ω, Σ) if*

(a) for all disjoint sets $A_1, ..., A_n \in \Sigma$ the random variables

$$N_j = N(A_j) := \#(G \cap A_j)$$

denoting the number of points from G lying in A_j are statistically independent and

*(b) there exists a **mean measure** ν such that for all $A \in \Sigma$ the random variable $N(A)$ obeys a Poisson distribution with parameter $\nu(A)$, i.e. $N(A) \sim \mathscr{P}(\nu(A))$.*

For more information about Poisson point processes and point processes in general we refer to [KMM78, Rei93]. A simulation of a Poisson process is shown in Figure 2.1.

First of all, the definition of the Poisson distribution (2.1) ensures $\mathbf{P}(N(A) = \infty) = 0$ if $\nu(A) < \infty$. Therefore, if the mean measure ν is finite, the Poisson process G consists of a finite set of points with probability 1. This allows for a representation of G as a sum of Dirac measures

$$G = \sum_{k=1}^{N} \delta_{x_i} \tag{2.7}$$

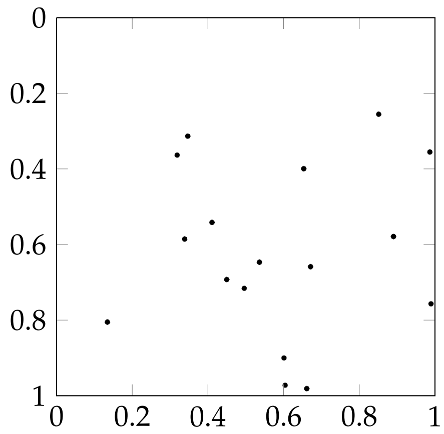


Figure 2.1: Simulated observation of a Poisson point process on $\Omega = [0,1]^2$ with mean measure $\mathrm{d}\nu = g\,\mathrm{d}x$ where $g(x_1, x_2) = \frac{1}{10000} x_1 \cdot x_2$.

where in the setup of photonic imaging the point x_k corresponds to the position of the k-th detected photon and N is the total number of detected photons. Note that $N \sim \mathscr{P}(\nu(\Omega))$ and especially $\mathbf{E}(N) = \nu(\Omega)$. The representation (2.7) allows for $x_i = x_j$ with $i \neq j$, i.e. it would be possible to count two (or more) photons at the same position. But our definition of the Poisson process as a random **set** of points prohibits that and hence in our continuous setup all points x_i need to be pairwise different. For real world measurements where some finite volume averaging needs to be applied, this is not necessarily the case since every detector measures photons occurring in an area of positive size, where hence arbitrarily many photons might be observed in that area.
It follows immediately from the property (b) of Definition 2.3 that the mean measure ν is always a measure without atoms, i.e.

$$\nu(\{x\}) = 0 \qquad \text{for all} \qquad x \in \Omega, \tag{2.8}$$

since otherwise (b) with $A = \{x\} \in \Sigma$ leads to

$$\begin{aligned}\mathbf{P}(N(A) \geq 2) &= \sum_{k=2}^{\infty} \exp(-\nu(A)) \frac{\nu(A)^k}{k!} \\ &= 1 - \exp(-\nu(\{x\})) - \nu(\{x\}) \exp(-\nu(\{x\})) \\ &> 0\end{aligned}$$

which contradicts $N(A) = \#G \cap \{x\} \leq 1$.
The representation (2.7) of G as a measure itself allows to define integrals w.r.t. G in an easy manner: For a measurable function $f : \Omega \to \mathbb{R}$ we set

$$\int_{\Omega} f \,\mathrm{d}G = \sum_{k=1}^{N} f(x_k).$$

With this, one obtains formulas for the expectation and variance of integrals w.r.t. G (if existent) of the following kind (cf. [Kin93, eq. (3.9) and (3.10)]):

$$\mathbf{E} \int_{\Omega} f \, \mathrm{d}G = \int_{\Omega} f \, \mathrm{d}\nu, \tag{2.9a}$$

$$\mathbf{V} \int_{\Omega} f \, \mathrm{d}G = \int_{\Omega} f^2 \, \mathrm{d}\nu. \tag{2.9b}$$

The integrals on the right-hand side simplify if the Poisson process has a special structure as given in the following definition:

DEFINITION 2.4:
Let G be a Poisson point process on (Ω, Σ) with mean measure ν where Ω is a subset of $\mathbb{R}^d$ and Σ the Borel σ-algebra on Ω.

- *If ν is absolutely continuous w.r.t. the Lebesgue measure, then the Radon-Nikodym derivative $g \in \mathbf{L}^1(\Omega)$ of ν w.r.t. the Lebesgue measure is called the **intensity** of G.*
- *If G has a constant intensity g, then G is called a **homogeneous** Poisson process.*

Note that for a Poisson process with intensity g we have $N(A) \sim \mathscr{P}\left(\int_A g \, \mathrm{d}x\right)$ for all Borel measurable sets A. Even though our definition of a Poisson process requires the Poisson distribution explicitly, the Poisson distribution is in some sense inevitable, as the following theorem states:

THEOREM 2.5:
Let $\Omega \subset \mathbb{R}^d$ be open and let Σ be the Borel σ-algebra on Ω. Assume that a random countable set of points G (i.e. a point process) is given such that

(a) for all disjoint sets $A_1, ..., A_n \in \Sigma$ the random variables N_j are statistically independent and

(b) there exists an intensity $g \in \mathbf{L}^1(\Omega)$ such that for all $A \in \Sigma$ we have $\mathbf{E}(N(A)) = \int_A g \, \mathrm{d}x$.

Then we have $N(A) \sim \mathscr{P}\left(\int_A g \, \mathrm{d}x\right)$ for all $A \in \Sigma$, i.e. G is a Poisson process with intensity g.

PROOF:
See e.g. [KMM78, Thm 1.11.8] or [Kal97, Thm. 10.11]. ∎

In the following we will due to Theorem 2.5 assume that the measure space (Ω, μ) from the definition of a Poisson process is given by the Lebesgue measure on Ω and the measurable sets are those contained in the Borel σ-algebra, i.e. $\mathrm{d}\mu = \mathrm{d}x$.
In a more general setup, we could also allow the measurement manifold Ω to include a time variable τ, i.e. $\Omega = \bar{\Omega} \times (0, t)$. This would allow for non-stationary intensities g of G, but due to our main application we are interested in the case that g does not change over the whole observation time t. This can be modeled by our setup from above as follows:

BASIC ASSUMPTIONS AND NOTATIONS 1:
We assume that our data consist of a possibly inhomogeneous Poisson process G_t with spatial intensity $tg^\dagger$ (which corresponds to a mean measure $\mathrm{d}\nu = tg^\dagger \, \mathrm{d}x$) where the parameter $t > 0$ can be interpreted as the **observation time** of the object.
For our theoretical analysis we will study the limit $t \to \infty$ using the **noise level**

$$\psi(t) := \frac{1}{\sqrt{t}}, \qquad t > 0. \tag{2.10}$$

If $\int_\Omega g^\dagger \,\mathrm{d}x = 1$, the expected number of total photons N is equal to t (i.e. $\mathbf{E}(N) = t$) and thus the data contains more information the larger t is.
If we consider the normalized Poisson process $\frac{1}{t}G_t$, then by (2.9) we have by the meaning of integrals

$$\text{“}\mathbf{E}\left(\tfrac{1}{t}G_t\right) = g^\dagger\text{”} \qquad \text{and} \qquad \text{“}\mathbf{V}\left(\tfrac{1}{t}G_t\right) = \tfrac{1}{t^2}tg^\dagger = \tfrac{1}{t}g^\dagger\text{”}.$$

Considering the standard deviation (i.e. the square-root of the variance) as noise level, this shows that $\psi(t)$ indeed describes the decay of the noise.

The fact that the normalized Poisson process $\frac{1}{t}G_t$ concentrates more and more around its “expectation” $g^\dagger$ for $t \to \infty$ is described more precisely by **concentration inequalities**. For a Poisson process (2.7), concentration inequalities are more difficult to handle than (2.6), since symbols like $\mathbf{E}(G_t)$ or $\mathbf{V}(G_t)$ do not exist. Moreover, a concentration inequality of the form (2.6) would not be sufficient since inequalities which control the supremum over all integrals of f w.r.t. the difference $G_t - \nu$ for f in a suitable family are needed. Concentration inequalities of this type usually base on the work of TALAGRAND, who contributed a lot to the theory of stochastic processes in Banach spaces. We will use the following concentration inequality, which has been proven by REYNAUD-BOURET [RB03] and reads as follows:

LEMMA 2.6 ([RB03, COROLLARY 2]):
Let G be a Poisson process with finite mean measure ν. Let $\{f_a\}_{a\in A}$ be a countable family of functions with values in $[-b, b]$. One considers

$$Z := \sup_{a\in A}\left|\int_\Omega f_a(x)\,(\,\mathrm{d}G - \,\mathrm{d}\nu)\right| \qquad \textit{and} \qquad v_0 := \sup_{a\in A}\int_\Omega f_a^2(x)\,\mathrm{d}\nu.$$

Then for all positive numbers ρ and ε it holds

$$\mathbf{P}\left(Z \geq (1+\varepsilon)\,\mathbf{E}(Z) + \sqrt{12v_0\rho} + \kappa(\varepsilon)\,b\rho\right) \leq \exp(-\rho) \tag{2.11}$$

where $\kappa(\varepsilon) = 5/4 + 32/\varepsilon$.

REYNAUD-BOURET [RB03] describes this result as an analogon “to Talagrand’s inequalities for empirical processes”.

As mentioned in the introduction, we want to use the negative log-likelihood functional to incorporate the data distribution into our reconstruction process. The negative log-likelihood functional has already been calculated in Section 2.1.1 for an elementwise Poisson distributed vector (2.5). It seems natural to generalize this functional to the case of data G_t as

$$\rho_{\text{poiss}}(g; G_t) = \begin{cases} \int_\Omega g\,\mathrm{d}x - \int_\Omega \ln(g)\,\mathrm{d}G_t & \text{if } g \geq 0 \text{ a.e.,} \\ \infty & \text{otherwise.} \end{cases}$$

By definition, we have $\int_\Omega \ln(g)\,\mathrm{d}G_t = \sum_{k=1}^N \ln(g(x_k))$ which is infinite if $g(x_k) = 0$ for any $k \in \{1, ..., N\}$. Nevertheless, since G_t consists of more photons the larger t is, the functional will tend to ∞ as $t \to \infty$.

Therefore, some normalization is needed and so we define

$$\mathcal{S}_t(g; G_t) := \begin{cases} \int_\Omega g\,\mathrm{d}x - \frac{1}{t}\int_\Omega \ln(g)\,\mathrm{d}G_t & \text{if } g \geq 0 \text{ a.e.,} \\ \infty & \text{otherwise} \end{cases} \tag{2.12a}$$

as data misfit functional for random data G_t. In case of exact data $g^\dagger$, the functional (2.5) is generalized to the continuous case by

$$\mathcal{S}\left(g; g^\dagger\right) := \begin{cases} \int\limits_\Omega g \,\mathrm{d}x - \int\limits_\Omega \ln(g)\, g^\dagger \,\mathrm{d}x & \text{if } g_{|_{\{g^\dagger>0\}}} > 0 \text{ a.e.}, \quad g_{|_{\{g^\dagger=0\}}} \geq 0 \text{ a.e.}, \\ \infty & \text{otherwise.} \end{cases} \tag{2.13}$$

The connection between (2.12a) and (2.13) is clear, since in case of noisy data the exact mean measure $\mathrm{d}\nu = tg^\dagger \,\mathrm{d}x$ has been replaced by the Poisson process $\mathrm{d}G_t$. The functional (2.13) is up to a constant the well-known **Kullback-Leibler divergence** $\mathbb{KL}\left(g^\dagger; g\right)$, where the constant is chosen such that $\mathbb{KL}\left(g^\dagger; g^\dagger\right) = 0$:

$$\mathbb{KL}\left(g^\dagger; g\right) = \begin{cases} \int\limits_\Omega \left(g - g^\dagger + g^\dagger \ln\left(\frac{g^\dagger}{g}\right)\right) \mathrm{d}x & \text{if } g_{|_{\{g^\dagger>0\}}} > 0 \text{ a.e.}, \quad g_{|_{\{g^\dagger=0\}}} \geq 0 \text{ a.e.}, \\ \infty & \text{otherwise.} \end{cases} \tag{2.12b}$$

In comparison to our previously introduced data fidelity $\mathcal{S}$, the arguments for the Kullback-Leibler divergence are interchanged due to general notation conventions. This is kept for the whole thesis, and if we say that $\mathbb{KL}$ is used as data fidelity term w.r.t. exact data, then this means $\mathcal{S}\left(g; g^\dagger\right) = \mathbb{KL}\left(g^\dagger; g\right)$ and so on.

The aforementioned concentration inequality (2.11) does not apply in case of $\mathcal{S}_t$ as discussed above, since the functions $\ln(g)$ will in general be unbounded. Therefore, we will introduce a shift $e > 0$ in the following way:

$$\mathbb{KL}_e\left(g^\dagger; g\right) = \begin{cases} \mathbb{KL}\left(g^\dagger + e; g + e\right) & \text{if } g \geq -\frac{e}{2} \text{ a.e.}, \\ \infty & \text{otherwise,} \end{cases} \tag{2.14a}$$

$$\mathcal{S}_{e,t}\left(g; G_t\right) = \begin{cases} \int\limits_\Omega g \,\mathrm{d}x - \frac{1}{t} \int\limits_\Omega \ln(g+e) \,\mathrm{d}G_t - e \int\limits_\Omega \ln(g+e) \,\mathrm{d}x & \text{if } g \geq -\frac{e}{2} \text{ a.e.}, \\ \infty & \text{otherwise.} \end{cases} \tag{2.14b}$$

Note that $\mathcal{S}_{e,t}$ is again convex in its first argument. By definition it holds $\mathbb{KL}_e(\hat{g}; g) = \mathbb{KL}(\hat{g} + e; g + e)$ if $g \geq -\frac{e}{2}$ a.e. and especially $\mathbb{KL}_0 = \mathbb{KL}$. The same is true for the data fidelity in case of noisy data, $\mathcal{S}_{0,t} = \mathcal{S}_t$ for any $t > 0$. In this sense, (2.14) is a generalization of (2.12).

We want to point out that the relaxation parameter $e > 0$ has been introduced due to the concentration inequality (2.11). If we had a concentration inequality for unbounded functions (say $\mathbf{L}^2$) at hand, then the convergence analysis would work also for the case $e = 0$. In principle, $e > 0$ can be interpreted as some background noise which always occurs and does not influence the data distribution. Nevertheless, for $e > 0$ the data fidelity $\mathcal{S}_{e,t}$ is not the true negative log-likelihood functional, but for small e it seems to be a reasonable approximation of it. Moreover, the numeric computations become easier for $e > 0$ due to the relaxation of the singularity at 0. We hope to find a concentration inequality for unbounded functions f_a in (2.11) in future, which will then lead to a convergence analysis for $e = 0$.

2.2 Examples of inverse problems with Poisson data

The most important examples of inverse problems with Poisson data are applications from photonic imaging. In photonic imaging the data consist of photon counts where the photons have interacted with some unknown object $u^\dagger$ of interest. If we assume that the interaction of the photons with $u^\dagger$ can be described by a possibly nonlinear operator equation

$$F\left(u^\dagger\right) = g^\dagger,$$

then the measurements are described by photon counts from the true photon density $g^\dagger$ on the measurement manifold $\Omega \subset \mathbb{R}^d$. If we assume to have a perfect photon detector, then due to fundamental physical reasons photon counts can be seen as a random set of points such that

- the counts on disjoint measurable sets are independent and
- the expected number of counts on a measurable set A is given by $t \int_A g^\dagger \,\mathrm{d}x$ with some parameter $t > 0$ (referred to as observation time).

Hence it follows directly from Theorem 2.5 that the measurements are mathematically described by a Poisson process with intensity $tg^\dagger$ if we neglect finite volume averaging and read-out noise.
In the following we will present three important examples from photonic imaging, which will also be treated in Chapter 8 by simulations and numerical tests.

2.2.1 An inverse obstacle scattering problem without phase

The scattering of polarized, transverse magnetic time harmonic electromagnetic waves by a perfect cylindrical conductor with smooth cross section $D \subset \mathbb{R}^2$ is described by the Helmholtz equation with homogeneous Neumann boundary conditions and the Sommerfeld radiation condition

$$\Delta u + k^2 u = 0, \qquad \text{in } \mathbb{R}^2 \setminus D, \tag{2.15a}$$

$$\frac{\partial u}{\partial n} = 0, \qquad \text{on } \partial D, \tag{2.15b}$$

$$\lim_{r\to\infty} \sqrt{r}\left(\frac{u_s}{r} - \mathrm{i}ku_s\right) = 0, \qquad \text{where } r := |x| \tag{2.15c}$$

with the **incident field** u_i, the **scattered field** u_s and the **total field** $u = u_i + u_s$. Here D is compact, $\mathbb{R}^2 \setminus D$ is connected, n is the outer normal vector on ∂D, and the incident field is considered to be a plane wave with direction $d \in S^2 = \left\{x \in \mathbb{R}^2 \;\middle|\; |x| = 1\right\}$ and wave number $k > 0$, i.e. $u_i(x) = \exp(\mathrm{i}kd \cdot x)$. This is a classical obstacle scattering problem, and we refer to the monograph [CK97] for further details and references. An example of u for a given shape D is shown in Figure 2.2.
The Sommerfeld radiation condition (2.15c) implies the asymptotic behavior

$$u_s(x) = \frac{\exp(\mathrm{i}k|x|)}{\sqrt{|x|}}\left(u_\infty\left(\frac{x}{|x|}\right) + \mathcal{O}\left(\frac{1}{|x|}\right)\right)$$

as $|x| \to \infty$, and u_∞ is called the **far field pattern** or **scattering amplitude** of u_s.
We consider the inverse problem to recover the shape of the obstacle D from photon counts of the scattered electromagnetic field far away from the obstacle. Since the photon density is proportional to the squared absolute value of the electric field, we have no immediate access to the phase of the electromagnetic field. Since at large distances the

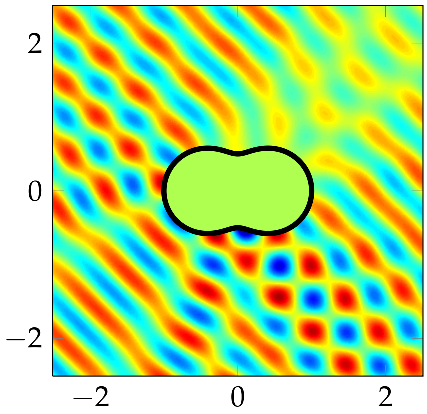


Figure 2.2: Real part $\Re(u)$ of the total field for the given peanut-shaped obstacle and an incident wave from 'South West'.

photon density is approximately proportional to $|u_\infty|^2$, our inverse problem is described by the operator equation

$$F(\partial D) = |u_\infty|^2. \tag{2.16}$$

A similar problem has been studied by Kress & Rundell [KR97] and with different methods and noise models also by Ivanyshyn & Kress [IK10]. Recall that $|u_\infty|$ is invariant under translations of ∂D. Therefore, it is only possible to recover the shape, but not the location of D. An example for photon counts from $|u_\infty|^2$ is shown in Figure 2.3.

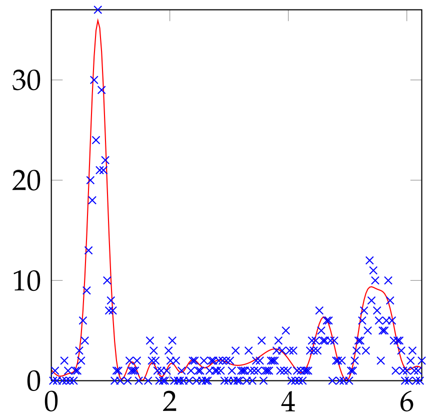


Figure 2.3: Exact and simulated Poisson data for the obstacle from Figure 2.2. The red line shows $|u_\infty|^2$, the blue crosses mark the simulated photon count data. The observation time was $t = 1000$, i.e. the expected number of total counts is 1000.

2.2.2 A phase retrieval problem

A problem which often occurs in optics is the following: Given the modulus $|\mathcal{F} f|$ of the Fourier transform of a function $f : \mathbb{R}^d \to \mathbb{C}$ we need to reconstruct the function f itself, or equivalently to *reconstruct the phase* $\mathcal{F} f / |\mathcal{F} f|$ of $\mathcal{F} f$. Problems of this type are hence called **phase retrieval problems**. Obviously, f is not uniquely determined by $|\mathcal{F} f|$ (for example a shift $\bar{f}(x) := f(x - y)$, $y \in \mathbb{R}^d$ does not change $|\mathcal{F} f|$) and hence additional a

priori information is needed. For phase retrieval problems in general see HURT [Hur89] or KLIBANOV, SACKS & TIKHONRAVOV [KST95] and the references therein.
Let us motivate this class of problems by a specific application. Consider the case that an incident plane wave in the $\mathbb{R}^3$ in x_3 direction is sent in and passes through a non-absorbing, weakly scattering object of interest in the lower half-space $\{x_3 < 0\}$ close to the plane $\{x_3 = 0\}$. Then the total field u solves the Helmholtz equation $\Delta u + k^2 u = 0$ and a radiation condition in the half-space $\{x_3 > 0\}$, and on the plane $\{x_3 = 0\}$ a representation $u(x', 0) = \exp(\mathrm{i}\varphi(x'))$ holds true. If the wave length is small compared to the length scale of the object, the **projection approximation**

$$\varphi(x') \approx \frac{k}{2} \int_{-\infty}^{0} \left(n^2(x', x_3) - 1\right) \mathrm{d}x_3, \qquad x' \in \mathbb{R}^2$$

is valid where n describes the refractive index of the object of interest (see e.g. [Pag06, Sec. 2.1]). A sketch of the experiment is shown in Figure 2.4. A priori information on φ concerning a jump at the boundary of its support can be obtained by placing a known transparent object before or behind the object of interest.

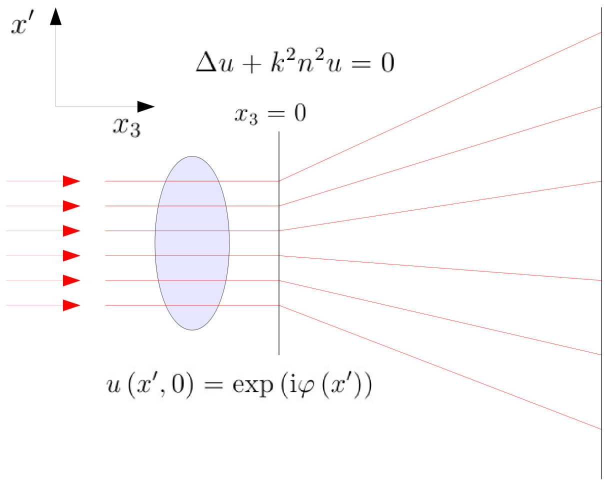


Figure 2.4: A sketch of our experiment leading to a phase retrieval problem. An incident plane wave is sent in in x_3-direction and scatters at an unknown object of interest D near the $\{x_3 = 0\}$ plane. Then the total field u solves the Helmholtz equation and a radiation condition in $\{x_3 > 0\}$. We aim to reconstruct information on the obstacle D by meanings of the function φ, which can also be seen as initial condition via $u(x', 0) = \exp(\mathrm{i}\varphi(x')), x' \in \mathbb{R}^2$.

Now assume that we are able to measure the intensity

$$g(x') = |u(x', \Delta)|^2, \qquad x' = (x_1, x_2) \in \mathbb{R}^2$$

of the electric field at a measurement plane $\{x_3 = \Delta\}$ far away, i.e. with large Δ. Applying the 2-dimensional Fourier transform $\mathcal{F}_2$ w.r.t. x_1 and x_2 to u, using the well-known properties of $\mathcal{F}$ and the Helmholtz equation, we find that

$$u(x', x_3) = \mathcal{F}_2^{-1}\left(\exp\left(\mathrm{i}x_3\sqrt{k^2 - |\cdot|^2}\right) \cdot \mathcal{F}_2(\exp(\mathrm{i}\varphi))\right)$$

since $u(\cdot,0) = \exp(\mathrm{i}\varphi(\cdot))$ can also be seen as an initial condition. Now replacing $\sqrt{k^2 - |\cdot|^2}$ by its second order Taylor expansion

$$\sqrt{k^2 - |\xi|^2} \approx k - \frac{|\xi|^2}{2k}, \qquad |\xi| \ll k$$

we obtain the **Fresnel approximation**

$$u(x', x_3) \approx \exp(\mathrm{i}kx_3)\, \mathcal{F}_2^{-1}\left(\exp\left(-\frac{\mathrm{i}x_3 |\cdot|^2}{2k}\right) \mathcal{F}_2(\exp(\mathrm{i}\varphi))\right). \tag{2.17}$$

Note that $|\xi| \ll k$ can be assumed in the Fourier domain if the wave length is small compared to the length scale of the object. For the next step we find that it holds

$$\exp\left(-\frac{\mathrm{i}x_3 |\cdot|^2}{2k}\right) = \mathcal{F}_2\left(-\frac{\mathrm{i}k}{x_3}\exp\left(\mathrm{i}\frac{k}{2x_3}|\cdot|^2\right)\right)$$

in a distributional sense. For any Schwartz function u_0 the convolution theorem yields now

$$\mathcal{F}_2^{-1}\left(\exp\left(-\frac{\mathrm{i}x_3 |\cdot|^2}{2k}\right) \cdot \mathcal{F}_2 u_0\right) = -\frac{\mathrm{i}k}{x_3}\exp\left(\mathrm{i}\frac{k}{2x_3}|\cdot|^2\right) * u_0$$

and this result can be extended to the case $u_0 = \exp(\mathrm{i}\varphi)$ by adding an imaginary part to φ which tends to 0.
Thus we obtain

$$\begin{aligned} u(x', x_3) &\approx \frac{1}{2\pi}\exp(\mathrm{i}kx_3)\left(\left(-\frac{\mathrm{i}k}{x_3}\exp\left(\mathrm{i}\frac{k}{2x_3}|\cdot|^2\right)\right) * \exp(\mathrm{i}\varphi)\right)(x') \\ &= -\frac{\mathrm{i}k}{2\pi x_3}\exp(\mathrm{i}kx_3)\int_{\mathbb{R}^2} \exp\left(\mathrm{i}\frac{k}{2x_3}|x'-y'|^2\right)\exp(\mathrm{i}\varphi(y'))\,\mathrm{d}y'. \end{aligned}$$

Some manipulations of the right-hand side show now that

$$u(x', x_3) \approx -\frac{\mathrm{i}k}{2\pi x_3}\exp\left(\mathrm{i}kx_3 + \frac{\mathrm{i}k}{2x_3}|x'|^2\right)\int_{\mathbb{R}^2}\exp\left(\frac{\mathrm{i}k|y'|^2}{2x_3} - \mathrm{i}\frac{k}{x_3}x'y'\right)\exp(\mathrm{i}\varphi(y'))\,\mathrm{d}y'.$$

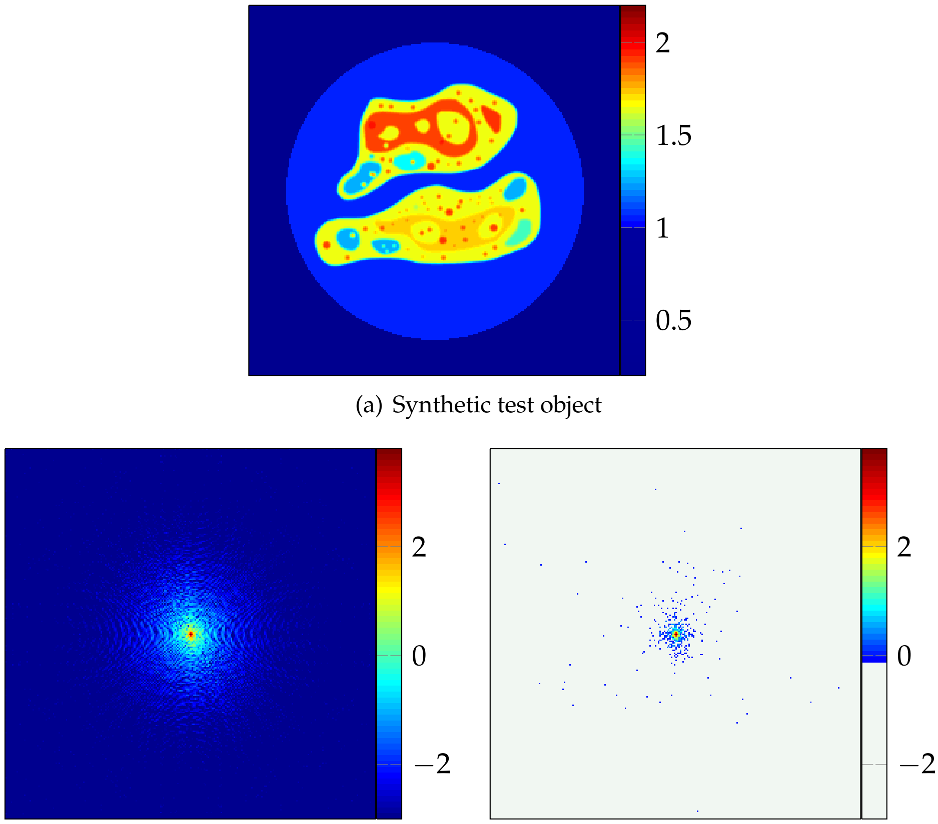
Finally, since we are interested in $g(x') = |u(x', \Delta)|^2$ for large Δ we use the **Fraunhofer approximation** (see e.g. BORN & WOLF [BW99, Sec. 8.3.3] or PAGANIN [Pag06, Sec. 1.5]) and drop the term $\frac{k|y'|^2}{2x_3}$ which leads to

$$\begin{aligned} |u(x', \Delta)|^2 &\approx \left|\frac{k}{2\pi\Delta}\int_{\mathbb{R}^2}\exp\left(-\mathrm{i}\left(\frac{k}{\Delta}x'\right)y'\right)\exp(\mathrm{i}\varphi(y'))\,\mathrm{d}y'\right|^2 \\ &= \frac{k^2}{\Delta^2}\left|\mathcal{F}(\exp(\mathrm{i}\varphi))\left(\frac{k}{\Delta}x'\right)\right|^2 \end{aligned} \tag{2.18}$$

for sufficiently large Δ. So the operator

$$F : \varphi \mapsto |\mathcal{F}(\exp(\mathrm{i}\varphi))|^2$$

is up to some rescaling a reasonable approximation to the true forward map. A simulated test object as well as corresponding true and simulated Poisson data for this problem are shown in Figure 2.5.



(a) Synthetic test object

(b) $\log_{10}$ of scaled exact data

(c) $\log_{10}$ of synthetic Poisson data

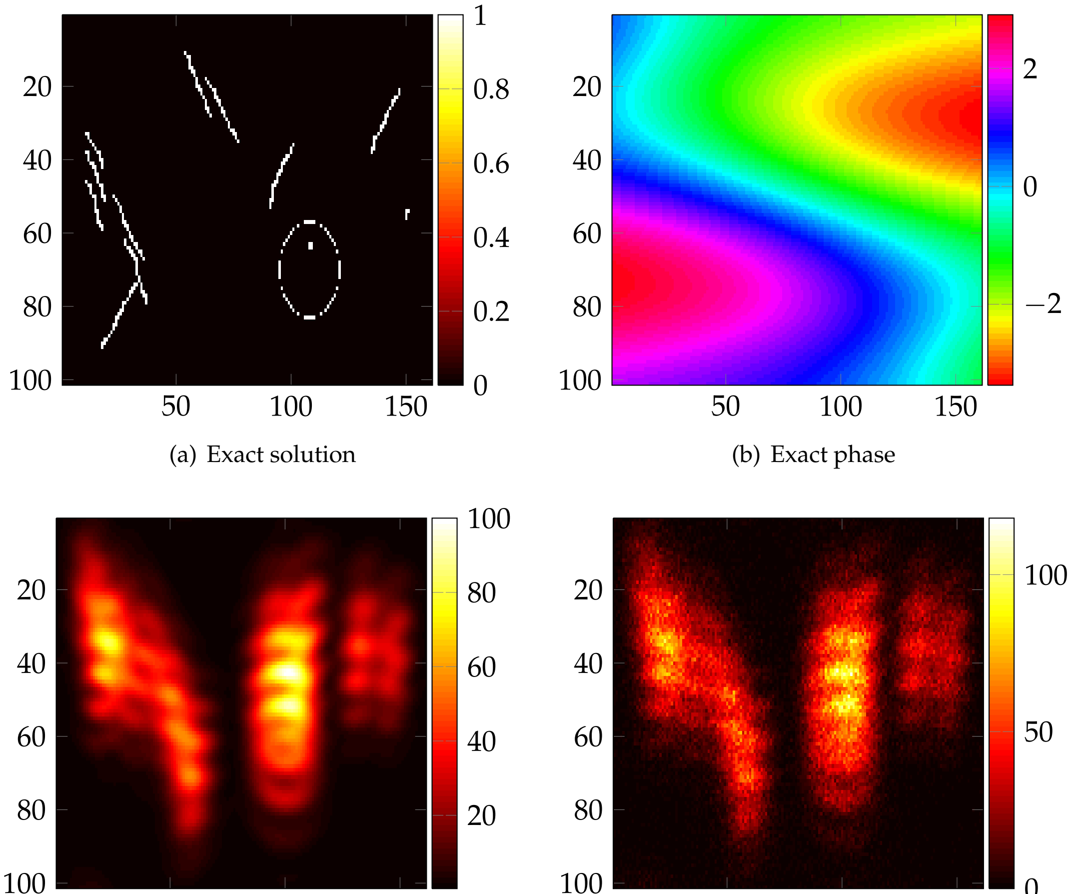
Figure 2.5: A synthetic test object, corresponding true and simulated Poisson data. The test object representing two cells has been taken from GIEWEKEMEYER ET AL. [GKK+11]. The observation time was $t = 10^4$, i.e. the expected number of total counts is 10.000.

2.2.3 A semiblind deconvolution problem in 4Pi microscopy

4Pi microscopy can be seen as an improvement of standard confocal microscopy. Both applications have in common that a fluorescent marker density is imaged by counting the emitted photons, which is for standard confocal microscopy modeled by a convolution equation

$$g(x) = F_{\mathrm{conf}}(f)(x) = \int_{[-R,R]^d} h(x-y)\, f(y)\, \mathrm{d}y, \qquad x \in [-R,R]^d. \tag{2.19}$$

Here f is a mathematical description of the fluorescent marker density on the domain $[-R,R]^d$, g is the measured intensity and h is called *point spread function*. Obviously, h depends on the special structure of the microscope, but it is nevertheless most often modeled as a Gaussian function.
The axial resolution of standard confocal microscopes is limited by the width of h along the optical axis. To overcome this limit, the 4Pi principle uses interference of coherent photons through two opposing objective lenses. This leads to an increase of the axial resolution by a factor of 3-7, but unfortunately the mathematical model (2.19) is no longer valid for this imaging process. The point spread function h depends now on the relative phase ϕ of the interfering photons, which is in general non-constant and hence has to be



(a) Exact solution (b) Exact phase (c) Exact data (d) Synthetic Poisson data

Figure 2.6: Synthetic test object and test phase as well as corresponding true and simulated Poisson data for the 4Pi problem. The expected number of total counts is 226829.49.

recovered from g together with f. The imaging process is now modeled by an equation of the form

$$g(x) = F_{4\text{Pi}}(f,\phi)(x) := \int_{[-R,R]^d} p(y-x,\phi(x))\, f(y)\, \mathrm{d}y, \qquad y \in [-R,R]^d. \tag{2.20}$$

Note that the operator $F_{4\text{Pi}}$ may be nonlinear in ϕ and that $f \mapsto F_{4\text{Pi}}(f,\phi)$ is not a convolution operator in general. The 4Pi point spread function p is approximately given by

$$p(x,\varphi) = h(x)\cos^n\left(cx_3 + \frac{\varphi}{2}\right)$$

where h is the point spread function of the corresponding standard confocal microscope and $n \in \{2,4\}$ (cf. [HS92]). The problem to reconstruct f and ϕ from observed data g^{obs} is in general underdetermined. Therefore, the information that f is a density and hence non-negative should be incorporated into the reconstruction procedure. This is done by choosing

$$\mathfrak{B} \subseteq \left\{(f,\phi) \in \mathbb{X} \;\middle|\; f \geq 0 \text{ a.e.}\right\}$$

and searching only for $(f, \phi) \in \mathfrak{B}$ with $F_{4\mathrm{Pi}}(f, \phi) \approx g^{\mathrm{obs}}$. For more details on 4Pi microscopy in general and a fast implementation of $F_{4\mathrm{Pi}}$ we refer to [Stü11].

2.3 General assumptions

As we have seen above, the measurements are mathematically described by a Poisson process with the true photon density as intensity. All the aforementioned examples from photonic imaging meet special properties for the forward operator F, which are collected in the following assumption and which we will use for inverse problems with Poisson data in general:

ASSUMPTION 2.7 (ASSUMPTION ON F FOR POISSON DATA):

(a) Let $\mathbb{Y} = \mathbf{L}^1(\Omega) \cap \mathbf{L}^\infty(\Omega)$ for some bounded observation domain $\Omega \subset \mathbb{R}^d$ with Lipschitz boundary $\partial\Omega$.

(b) $\mathfrak{B} \subset \mathbb{X}$ is a bounded convex set contained in the Banach space $\mathbb{X}$.

(c) The operator $F : \mathfrak{B} \to \mathbb{Y}$ is continuously Fréchet differentiable with derivative $F'[u] : \mathbb{X} \to \mathbb{Y}$ for all $u \in \mathfrak{B}$ w.r.t. all $\mathbf{L}^p(\Omega)$-norms.

(d) F is injective on $\mathfrak{B}$ and hence there exists a unique exact solution $u^\dagger \in \mathfrak{B}$ of the exact problem $F(u) = g^\dagger$.

(e) It holds

$$F(u) \geq 0 \quad a.e. \qquad \text{for all } u \in \mathfrak{B}.$$

The property (e) is required since data are drawn from a Poisson distribution with intensity $tg^\dagger = tF(u^\dagger)$. It is easy to see that the properties (a), (c) and (e) are fulfilled for the operators above if the set $\mathfrak{B}$ is properly chosen. If moreover $\mathfrak{B}$ is bounded, then also (b) is easy to fulfill. Assumption (d) is required to simplify the notation and the assertions of our theorems, but we can also formulate a similar theory if F is not injective on $\mathfrak{B}$ by replacing the exact solution $u^\dagger$ by a suitable generalization. This is not done here for simplicity.

Since $\mathfrak{B}$ is assumed to be bounded, we may define the finite quantity

$$\operatorname{diam}(\mathfrak{B}) := \sup_{u,v \in \mathfrak{B}} \|u - v\|_{\mathbb{X}}$$

which will be used frequently in our convergence analysis.

CHAPTER

THREE

TIKHONOV-TYPE REGULARIZATION

This chapter deals with Tikhonov-type regularization for the solution of an ill-posed problem

$$F(u) = g, \tag{3.1}$$

i.e. a minimization problem of the type

$$u_\alpha \in \operatorname*{argmin}_{u\in\mathbb{X}} \left[\mathcal{S}\left(F(u); g^{\mathrm{obs}}\right) + \alpha\,\mathcal{R}(u)\right] \tag{3.2}$$

where $\mathcal{S}$ is some suitable data misfit and $\mathcal{R}$ some penalty. The number $\alpha > 0$ determining the weight of $\mathcal{R}$ is called regularization parameter. We will call the method (3.2) *Tikhonov-type regularization* due to the Russian mathematician TIKHONOV [Tik63a, Tik63b], who proposed a stable way to approximate solutions of (3.1) via

$$u_\alpha \in \operatorname*{argmin}_{u\in\mathbb{X}} \left[\left\|F(u) - g^{\mathrm{obs}}\right\|_{\mathbb{Y}}^2 + \alpha\,\|u-u_0\|_{\mathbb{X}}^2\right] \tag{3.3}$$

in the case where $\mathbb{Y}$ is some Hilbert space and $\mathbb{X}$ is some Hilbert space of functions where the norm $\|\cdot\|_{\mathbb{X}}$ contains derivatives (e.g. $\mathbb{X}$ is some Sobolev space). By choosing some specific u_0, *a priori* information was incorporated into the regularization procedure. Our generalized setup (3.2) includes this classical setting if we choose $\mathcal{S}(g;\hat{g}) := \|g-\hat{g}\|_{\mathbb{Y}}^2$ and $\mathcal{R}(u) = \|u-u_0\|_{\mathbb{X}}^2$, which will be called **quadratic Hilbert space case** in the following. Due to the specific structure of the problem, the minimizer (3.3) can be computed by an explicit formula if F is linear (cf. Remark 3.7). This is no longer the case in the general setting (3.2), but for convex $\mathcal{S},\mathcal{R}$ and linear F, the problem (3.2) is convex and can hence be solved with tools from convex optimization. If F is nonlinear, the problem (3.2) might have many local minima and so u_α is in general difficult to find. To overcome this problem, we will consider an iteratively regularized Newton method, cf. Chapter 5 for the case of $\mathcal{S}$ and $\mathcal{R}$ given by norm powers and Chapter 6 for the most general case.

In this chapter we will start with an overview over the basic theory which shows that (3.2) is appropriate to provide stable approximations to (3.1) and link it to the classical theory of Tikhonov regularization (3.3) as presented by ENGL, HANKE & NEUBAUER [EHN96]. The theory for (3.2) has been described in detail in the PhD thesis [Pös08] by PÖSCHL, which has also been the main reference for Section 3.1.
In Section 3.2 we introduce the concept of source conditions, mention some motivation and interpretation and generalize this concept to the more recently developed one of variational source conditions. We briefly discuss the role of a suitable nonlinearity condition,

which will be done in more detail in Section 6.1. Afterwards we prove two general convergence rates results (see Theorem 3.28 and Theorem 3.30). These results are new in the presented formulation and cover known convergence rates results from the classical literature [EHN96] as well as more recent publications [BO04, Res05, SGG+08, FH10, Fle10]. The proofs use techniques introduced by KALTENBACHER & HOFMANN [KH10] and new ideas.

3.1 Regularization properties

For variational regularization methods like (3.2), one usually requires the following properties:

(a) Well-definedness, i.e. for any $\alpha > 0$ and any $g^{\mathrm{obs}} \in \mathbb{Y}$ there exists at least one minimizer.

(b) Stability, i.e. for fixed $\alpha > 0$ the minimizers u_α depend continuously on g^{obs}.

(c) Convergence, i.e. the regularized solutions u_α converge to a solution of (3.1) as the noise level and α tend to 0 in an appropriate manner.

Without the features (a) and (b), the approximation of solutions to (3.1) via (3.2) is not appropriate, since then the problem (3.2) is again ill-posed. Uniqueness of the minimizers is not needed, if (c) holds for any choice of a minimizer. Item (c) guarantees that the regularized solutions u_α indeed approximate solutions of the original problem (3.1). As already mentioned in the introduction, the choice of the regularization parameter α is a very delicate task. Moreover, we still have to define what is meant by noise level.

3.1.1 Notation

In the whole chapter we will use the following notations:

BASIC ASSUMPTIONS AND NOTATIONS 2:

- $F : D(F) \subset \mathbb{X} \to \mathbb{Y}$ is the forward operator, which is Fréchet differentiable with derivative $F'[u] : \mathbb{X} \to \mathbb{Y}$ for all $u \in \mathfrak{B}$ where $\mathfrak{B} \subset D(F)$ is some bounded convex set (in principle we can assume $D(F) = \mathfrak{B}$).
- The exact right-hand side is denoted by $g^\dagger \in \mathbb{Y}$.
- There exists a unique exact solution $u^\dagger \in \mathbb{X}$ of (3.1), i.e. $F(u^\dagger) = g^\dagger$.
- The measured data is denoted by $g^{\mathrm{obs}} \in \mathbb{Y}$.
- For erroneous data we assume to have an upper bound $\mathcal{S}\left(g^\dagger; g^{\mathrm{obs}}\right) \leq \delta$, where $\delta > 0$ is called the **noise level**.

To ensure the items (a)-(c) to hold, we need $\mathcal{S}$, $\mathcal{R}$ and F to fulfill certain conditions which are related to the structure of the spaces $\mathbb{X}$ and $\mathbb{Y}$ as well as to the topologies on these spaces w.r.t. which we want to have 'stability' and 'convergence'. We collect these assumptions as follows:

ASSUMPTION 3.1:

- *$\mathbb{X}$ and $\mathbb{Y}$ are vector spaces with associated topologies $\tau_\mathbb{X}$ and $\tau_\mathbb{Y}$.*
- *$F : D(F) \subset \mathbb{X} \to \mathbb{Y}$ is $\tau_\mathbb{X} - \tau_\mathbb{Y}$ continuous.*

- *$D(F)$ is $\tau_{\mathbb{X}}$ closed.*
- *$\mathcal{R} : \mathbb{X} \to \mathbb{Y}$ is proper, convex and $\tau_{\mathbb{X}}$ lower semicontinuous.*
- *$D := D(\mathcal{R}) \cap D(F) \neq \emptyset$.*
- *$\mathcal{S} : \mathbb{Y} \times \mathbb{Y} \to [0, \infty]$ fulfills the following conditions:*
  - ◇ *If $\mathcal{S}(g_k; g) \to 0$, then $g_k \to g$ w.r.t. $\tau_{\mathbb{Y}}$.*
  - ◇ *$\mathcal{S}$ is sequentially lower semi-continuous w.r.t. $\tau_{\mathbb{Y}}$, i.e. for $g_k \to v$ w.r.t. $\tau_{\mathbb{Y}}$ and $\hat{g}_k \to \hat{g}$ w.r.t. $\tau_{\mathbb{Y}}$ it holds*

$$\mathcal{S}(g; \hat{g}) \leq \liminf_{k\to\infty} S(g_k; \hat{g}_k).$$

  - ◇ *If $\lim_{k\to\infty} \mathcal{S}(g; g_k) = 0$, then for every $\hat{g} \in \mathbb{Y}$ with $\mathcal{S}(\hat{g}; g) < \infty$ we have*

$$S(\hat{g}; g_k) \to \mathcal{S}(\hat{g}; g).$$

  - ◇ *$\mathcal{S}(g; \hat{g}) = 0 \Leftrightarrow g = \hat{g}$.*
- *For every $\alpha > 0$, $g \in \mathbb{Y}$ and $M > 0$ the level sets*

$$\left\{u \in D \;\middle|\; \mathcal{S}(F(u); g) + \alpha \mathcal{R}(u) \leq M\right\}$$

  are $\tau_{\mathbb{X}}$ sequentially compact.

REMARK 3.2:
If one thinks of $\mathbb{X}$ and $\mathbb{Y}$ as Hilbert spaces and $\mathcal{S}(g; \hat{g}) = \|g - \hat{g}\|_{\mathbb{Y}}^2$, $\mathcal{R}(u) = \|u - u_0\|_{\mathbb{X}}^2$ for some $u_0 \in \mathbb{X}$, then choosing $\tau_{\mathbb{X}}$ and $\tau_{\mathbb{Y}}$ as the weak topologies on $\mathbb{X}$ and $\mathbb{Y}$ makes Assumption 3.1 for linear and bounded F fulfilled. In the case of a nonlinear F, one has to assume additionally that F is continuous and weakly closed (cf. [EHN96, Sec. 10.1]).

3.1.2 Well-definedness, stability and convergence

THEOREM 3.3 (WELL-DEFINEDNESS):
Let Assumption 3.1 hold true and $\alpha > 0$. Then there exists a solution of (3.2).

PROOF:
See [Pös08, Thm. 1.6]. ■

THEOREM 3.4 (STABILITY):
Let Assumption 3.1 hold true and assume $\alpha > 0$. Then the minimizers (3.2) *are stable with respect to the data g^{obs}, this is, for any sequence $(g_k)_{k\in\mathbb{N}} \subset \mathbb{Y}$ with $\mathcal{S}\left(g^{\text{obs}}; g_k\right) \to 0$ a sequence of corresponding minimizers*

$$u_k \in \operatorname*{argmin}_{u\in\mathbb{X}} \left[\mathcal{S}(F(u); g_k) + \alpha \mathcal{R}(u)\right]$$

has a subsequence, which converges w.r.t. $\tau_{\mathbb{X}}$ and each limit of a convergent subsequence is a minimizer (3.2). *Moreover, for each $\tau_{\mathbb{X}}$ convergent subsequence $(u_{k_m})_{m\in\mathbb{N}}$ it holds $\mathcal{R}(u_{k_m}) \to \mathcal{R}(u_\alpha)$ where u_α is the corresponding solution of* (3.2).

PROOF:
See [Pös08, Thm. 1.7]. ■

REMARK 3.5:

- In the Hilbert space case where $\mathcal{R}(u) = \|u - u_0\|_{\mathbb{X}}^2$ and $\tau_{\mathbb{X}}$ is the weak topology, we gain strong convergence of $(u_{k_m})_{m\in\mathbb{N}}$ to a minimizer u_α from $\mathcal{R}(u_{k_m}) \to \mathcal{R}(u_\alpha)$.
- If the minimizer u_α is unique (for example in the quadratic Hilbert space case for linear F), we find that the whole sequence $(u_k)_{k\in\mathbb{N}}$ converges to u_α w.r.t. $\tau_{\mathbb{X}}$.

THEOREM 3.6 (CONVERGENCE):
Let Assumption 3.1 hold true and assume $\alpha > 0$. Moreover let a sequence $(\delta_k)_{k\in\mathbb{N}}$ of noise levels be given which converges monotonically to 0 such that for corresponding data g_k we have $\mathcal{S}(g^\dagger; g_k) \leq \delta_k$.
If the regularization parameters $\alpha_k = \alpha(\delta_k)$ are chosen such that the sequence $(\alpha_k)_{k\in\mathbb{N}}$ is monotonically decreasing and fulfills

$$\alpha_k \to 0 \qquad \text{and} \qquad \frac{\delta_k}{\alpha_k} \to 0 \qquad \text{as} \qquad k \to \infty,$$

then any sequence of regularized solutions

$$u_k \in \operatorname*{argmin}_{u\in\mathbb{X}} \left[\mathcal{S}(F(u); g_k) + \alpha_k \mathcal{R}(u)\right]$$

is $\tau_{\mathbb{X}}$ convergent to $u^\dagger$.

PROOF:
See [Pös08, Thm. 1.9]. ∎

REMARK 3.7 (LINK TO CLASSICAL THEORY):
In the case where $\mathbb{X}$ and $\mathbb{Y}$ are Hilbert spaces, $\mathcal{S}(g;\hat{g}) = \|g - \hat{g}\|_{\mathbb{Y}}^2, \mathcal{R}(u) = \|u - u_0\|_{\mathbb{X}}^2$ for some $u_0 \in \mathbb{X}$ and F is a bounded linear operator mapping $\mathbb{X}$ to $\mathbb{Y}$, the properties (a) and (b) follow immediately from the representation (see [EHN96, Thm. 5.1])

$$u_\alpha = (F^*F + \alpha I)^{-1} F^* g^{\text{obs}}.$$

Property (c) has been formulated in a slightly different way by requiring that the **worst case error**

$$\sup\left\{\left\|u_{\alpha(\delta, g^{\text{obs}})} - u^\dagger\right\|_{\mathbb{X}} \mid g^{\text{obs}} \in Y, \left\|g^\dagger - g^{\text{obs}}\right\|_{\mathbb{Y}} \leq \delta\right\}$$

tends to 0 as the noise level tends to 0. It is easy to show that under the same conditions on the parameter choice Tikhonov regularization (3.3) leads to a convergent regularization method (cf. [EHN96, Thm. 5.2]) where also the worst case error tends to 0.

3.2 Source conditions and convergence rates

In this section we present the results on convergence rates of Tikhonov-type regularization (3.2). In the last decade, a vast number of publications on the analysis of generalizations of the classical Tikhonov regularization (3.3) have been published. As a first step only generalizations of the penalty term $\|u - u_0\|_{\mathbb{X}}^2$ have been considered (see for example [BO04, Res05, SGG+08]) and especially the case of a sparsity enforcing functional $\mathcal{R}$ is still a domain of active research (cf. [DDD04, GHS08, Lor08] and Remark 3.39). More recently, also the most general case (3.2) has been analyzed by several authors [Pös08, Bar10, FH10, Fle10, Fle11] but the theory cannot be considered to be complete.

To provide convergence rates, the noise level definition $\mathcal{S}\left(g^{\dagger};g^{\mathrm{obs}}\right)\leq\delta$ we chose at the beginning of this chapter is not sufficient, since $\mathcal{S}$ does not necessarily fulfill a triangle inequality. To overcome this problem, we will use the following noise level:

ASSUMPTION 3.8 (NOISE LEVEL):
The data misfit functionals $\mathcal{S}\left(\cdot;g^{\dagger}\right)$ w.r.t. exact data and $\mathcal{S}\left(\cdot;g^{\mathrm{obs}}\right)$ w.r.t. noisy data are connected as follows: There exists a constant $C_{\mathrm{err}}\geq 1$ and functionals $\mathbf{err}:\mathbb{Y}\to[0,\infty]$, $\mathfrak{s}:\mathbb{Y}\to(-\infty,\infty)$ such that

$$\mathcal{S}\left(g;g^{\mathrm{obs}}\right)+\mathfrak{s}\left(g^{\dagger}\right)\leq C_{\mathrm{err}}\mathcal{S}\left(g;g^{\dagger}\right)+C_{\mathrm{err}}\mathbf{err}\left(g\right)\tag{3.4a}$$

$$\frac{1}{C_{\mathrm{err}}}\mathcal{S}\left(g;g^{\dagger}\right)\leq\left(\mathcal{S}\left(g;g^{\mathrm{obs}}\right)+\mathfrak{s}\left(g^{\dagger}\right)\right)+\mathbf{err}\left(g\right)\tag{3.4b}$$

for all $g\in\mathbb{Y}$ and it holds $\mathcal{S}\left(g^{\dagger};g^{\dagger}\right)=0$.

This assumption needs some explanations:

- Obviously, (3.4) is always fulfilled with $\mathbf{err}\left(g\right)\equiv\infty$. But since $\mathbf{err}\left(g\right)$ will be used as noise level in the analysis, this case is useless. Therefore, we will require that $\mathbf{err}\left(g\right)$ is not only finite but sufficiently small for those g, to which we will apply (3.4). If for example $\mathcal{S}\left(g;\hat{g}\right)=\left\|g-\hat{g}\right\|_{\mathbb{Y}}^{r}$ with $r\in[1,\infty)$, then it follows from the simple inequalities $(a+b)^{r}\leq 2^{r-1}\left(a^{r}+b^{r}\right)$ and $|a-b|^{r}+b^{r}\geq 2^{1-r}a^{r}$ that (3.4) holds true with $\mathbf{err}\equiv\left\|g^{\mathrm{obs}}-g^{\dagger}\right\|_{\mathbb{Y}}^{r}$, $C_{\mathrm{err}}=2^{r-1}$ and $\mathfrak{s}\equiv 0$. In this sense, Assumption 3.8 can be seen as a generalization of the classical noise level $\left\|g^{\mathrm{obs}}-g^{\dagger}\right\|_{\mathbb{Y}}\leq\delta$. In the following Lemma, we will provide expressions for $\mathbf{err}\left(g\right)$ in case of the negative log-likelihood functionals for Poisson noise as introduced in the previous chapter, i.e. for $\mathcal{S}$ as in (2.14). Moreover in Chapter 4 we will describe bounds for $\mathbf{err}\left(g\right)$ in probability in case of the negative log-likelihood for Poisson data as fidelity term.
- The functional $\mathfrak{s}$ is used to overcome problems in estimating differences between $\mathcal{S}\left(\cdot;g^{\dagger}\right)$ and $\mathcal{S}\left(\cdot;g^{\mathrm{obs}}\right)$ (see the lemma below). In the following we will assume w.l.o.g. that $\mathfrak{s}\equiv 0$, since replacing $\mathcal{S}\left(\cdot;g^{\mathrm{obs}}\right)$ by $\mathcal{S}\left(\cdot;g^{\mathrm{obs}}\right)+\mathfrak{s}\left(g^{\dagger}\right)$ in any minimization problem does not change the minimizers. Nevertheless, note that $\mathfrak{s}\left(g^{\dagger}\right)$ will be unknown in general and hence $\mathcal{S}\left(\cdot;g^{\mathrm{obs}}\right)+\mathfrak{s}\left(g^{\dagger}\right)$ is not implementable.
- Note that the noise level $\mathcal{S}\left(g^{\dagger};g^{\mathrm{obs}}\right)\leq\delta$ which we used to prove the regularization properties (cf. Section 3.1) is somehow covered by Assumption 3.8 with $\delta=C_{\mathrm{err}}\mathbf{err}\left(g^{\dagger}\right)$, since

$$\mathcal{S}\left(g^{\dagger};g^{\mathrm{obs}}\right)\leq C_{\mathrm{err}}\mathcal{S}\left(g^{\dagger};g^{\dagger}\right)+C_{\mathrm{err}}\mathbf{err}\left(g^{\dagger}\right)=C_{\mathrm{err}}\mathbf{err}\left(g^{\dagger}\right)$$

  where we used $\mathcal{S}\left(g^{\dagger};g^{\dagger}\right)=0$.

LEMMA 3.9 (VERIFICATION OF ASSUMPTION 3.8 FOR POISSON DATA):
Let Assumption 2.7 hold true and define

$$\mathfrak{s}_e\left(g^{\dagger}\right):=\int_{\Omega}\left[\left(g^{\dagger}+e\right)\ln\left(\frac{e}{g^{\dagger}+e}\right)-g^{\dagger}\right]\mathrm{d}x,\qquad e>0.$$

Then for the data fidelity functionals $\mathcal{S}\left(\cdot,g^{\dagger}\right)$ and $\mathcal{S}\left(\cdot,g^{\mathrm{obs}}\right)$ given by $\mathbb{KL}_e$ and $\mathcal{S}_{e,t}$ as in (2.14) *with $e>0$ Assumption 3.8 holds true with $C_{\mathrm{err}}=1$ and*

$$\mathbf{err}\left(g\right)=\begin{cases}\frac{1}{t}\left|\int_{\Omega}\ln\left(g+e\right)\left(\mathrm{d}G_t-\mathrm{d}\nu\right)\right| & \text{if } g\geq-\frac{e}{2}\text{ a.e.,}\\ \infty & \text{otherwise.}\end{cases}\tag{3.5}$$

PROOF:
For the well-definedness of $\mathfrak{s}_e$ we refer to [HW11]. From (3.4) it can be seen that for $C_{\text{err}} = 1$ the choice

$$\mathbf{err}(g) = \left|\mathbb{KL}_e\left(g^\dagger; g\right) - \mathfrak{s}_e\left(g^\dagger\right) - \mathcal{S}_{e,t}\left(g; g^{\text{obs}}\right)\right|$$

is always sufficient. Plugging in the definitions of $\mathbb{KL}_e\left(g^\dagger;\cdot\right)$, $\mathfrak{s}_e$ and $\mathcal{S}_{e,t}\left(\cdot; G_t\right)$ respectively yields the claim. ■

In the general setting (3.2), a norm $\|\cdot\|_{\mathbb{X}}$ as distance measure is sometimes too restrictive to provide convergence rates and often it does not occur naturally in the variational context. Therefore, a widely used tool are Bregman distances, which coincide with the square of a norm in the Hilbert space case if $\mathcal{R}(u) = \|u - u_0\|_{\mathbb{X}}^2$ and can hence be seen as a generalization. The Bregman distance has been introduced by BURGER & OSHER [BO04] in the context of inverse problems, and has previously been used implicitly for maximum entropy regularization by EGGERMONT [Egg93]. The definition is as follows:

DEFINITION 3.10 (BREGMAN DISTANCE):
*Let $\mathcal{R} : \mathbb{X} \to (-\infty, \infty]$ be a convex and proper functional with subdifferential $\partial\mathcal{R}\left(u^\dagger\right) \subset \mathbb{X}^*$. For $u^\dagger \in \mathbb{X}$ and $u^* \in \partial\mathcal{R}\left(u^\dagger\right)$ the **Bregman distance** of $\mathcal{R}$ is defined as*

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u, u^\dagger\right) := \mathcal{R}(u) - \mathcal{R}\left(u^\dagger\right) - \left\langle u^*, u - u^\dagger\right\rangle, \qquad u \in \mathbb{X}$$

where $\langle\cdot,\cdot\rangle$ denotes the duality product of $\mathbb{X}^*$ and $\mathbb{X}$.

The Bregman distance can be visualized as the difference between the linearization of $\mathcal{R}$ around $u^\dagger$ evaluated at u (this is $\mathcal{R}\left(u^\dagger\right) + \left\langle u^*, u - u^\dagger\right\rangle$) and the true value $\mathcal{R}(u)$ (cf. Figure 3.1).
Depending on special properties of the Banach space $\mathbb{X}$ and the penalty functional $\mathcal{R}$ the Bregman distance is related to the norm. Some geometric properties of (special) Banach spaces are repeated in the following (see e.g. [BKM+08] and the references therein).

DEFINITION 3.11 (MODULUS OF CONVEXITY AND MODULUS OF SMOOTHNESS):
The function $\delta_{\mathbb{X}} : [0,2] \to [0,1]$ defined by

$$\delta_{\mathbb{X}}(\varepsilon) := \inf\left\{1 - \left\|\frac{1}{2}(u+v)\right\|_{\mathbb{X}} \;\middle|\; \|u\|_{\mathbb{X}} = \|v\|_{\mathbb{X}} = 1, \|u - v\|_{\mathbb{X}} \geq \varepsilon\right\}$$

*is called the **modulus of convexity of** $\mathbb{X}$ and the function $\rho_{\mathbb{X}} : [0,\infty) \to [0,\infty)$ defined by*

$$\rho_{\mathbb{X}}(\tau) := \frac{1}{2}\sup\left\{\|u+v\|_{\mathbb{X}} + \|u-v\|_{\mathbb{X}} - 2 \;\middle|\; \|u\|_{\mathbb{X}} = 1, \|v\|_{\mathbb{X}} \leq \tau\right\}$$

*is called the **modulus of smoothness of** $\mathbb{X}$.*

DEFINITION 3.12:
The Banach space $\mathbb{X}$ is said to be

- ***reflexive** if the canonic embedding $\mathbb{X} \ni u \mapsto Eu \in \mathbb{X}^{**}$ with $(Eu)(u^*) := \langle u^*, u\rangle$ for $u \in \mathbb{X}, u^* \in \mathbb{X}^*$ is an isomorphism,*

- ***strictly convex** if the functional $\mathbb{X} \ni u \mapsto \|u\|_{\mathbb{X}}^2 \in \mathbb{R}$ is strictly convex,*

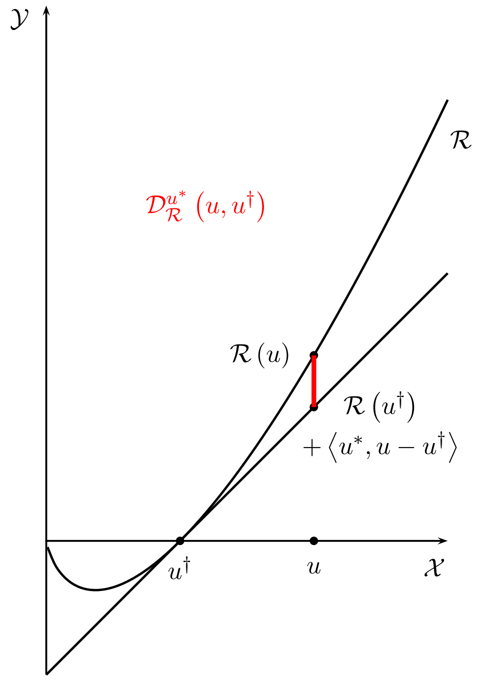


Figure 3.1: The Bregman distance of $\mathcal{R}(x) = x \cdot \ln(x)$ between u and $u^\dagger$.

- ***uniformly convex*** *if* $\delta_{\mathbb{X}}(\varepsilon) > 0$ *for all* $\varepsilon \in (0, 2]$,
- ***p-convex*** *for some* $p > 1$ *if there exists some* $C > 0$ *such that* $\delta_{\mathbb{X}}(\varepsilon) \geq C\varepsilon^p$ *for all* $\varepsilon \in [0, 2]$,
- ***smooth*** *if for any* $u \in \mathbb{X} \setminus \{0\}$ *there exists a unique* $u^* \in \mathbb{X}^*$ *such that* $\|u^*\|_{\mathbb{X}^*} = 1$ *and* $\langle u^*, u \rangle = \|u\|_{\mathbb{X}}$ *(*$\langle \cdot, \cdot \rangle$ *denotes the dual pairing between* $\mathbb{X}$ *and* $\mathbb{X}^*$*),*
- ***uniformly smooth*** *if* $\lim_{\tau \searrow 0} \frac{\rho_{\mathbb{X}}(\tau)}{\tau} = 0$,
- ***q-smooth*** *for some* $q > 1$ *if there exists some* $C > 0$ *such that* $\rho_{\mathbb{X}}(\tau) \leq C\tau^q$ *for all* $\tau \in [0, \infty)$.

From the polarization equality it can be seen that Hilbert spaces are 2-smooth and 2-convex and hence uniformly smooth and uniformly convex by definition. Moreover, for example the Banach space $\mathbf{L}^p(\Omega)$ with $1 < p \leq 2$ is 2-convex and p-smooth whereas $\mathbf{L}^q(\Omega)$ with $2 \leq q < \infty$ is q-convex and 2-smooth. In general, $\mathbb{X}$ is p-convex if and only if $\mathbb{X}^*$ is q-smooth where $\frac{1}{p} + \frac{1}{q} = 1$.
With the help of the above definitions we are now able to present a very helpful lemma on the Bregman distance, which in core goes back to work of XU & ROACH [XR91]:

LEMMA 3.13:
Let $\mathbb{X}$ *be a* p*-convex Banach space and* $\mathcal{R}(u) = \|u\|_{\mathbb{X}}^p$*. Then there exists some constant* $C_{\text{bd}} > 0$ *such that*

$$\left\| u - u^\dagger \right\|_{\mathbb{X}}^p \leq C_{\text{bd}} \mathcal{D}_{\mathcal{R}}^{u^*}\left(u, u^\dagger\right). \tag{3.6}$$

PROOF:
See [BKM+08, Lem. 2.7] ∎

Note that for a Hilbert space $\mathbb{X}$ and $p = 2$ we moreover have $C_{\text{bd}} = 1$ and equality in (3.6). For more information about the Bregman distance, we refer to [SGG+08, Sec. 3.2].

Now we are able to present the first step towards a convergence analysis:

LEMMA 3.14:
Let Assumption 3.8 hold and let $\alpha > 0$. Then the minimizers u_α fulfill the inequality

$$\alpha \mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) + \frac{1}{C_{\text{err}}}\mathcal{S}\left(F(u_\alpha); g^\dagger\right) \leq \alpha \left\langle u^*, u^\dagger - u_\alpha \right\rangle + \mathbf{err} \tag{3.7}$$

where

$$\mathbf{err} := \mathbf{err}\left(F(u_\alpha)\right) + C_{\text{err}}\mathbf{err}\left(g^\dagger\right). \tag{3.8}$$

PROOF:
The minimizing property leads to

$$\mathcal{S}\left(F(u_\alpha); g^{\text{obs}}\right) + \alpha\,\mathcal{R}(u_\alpha) \leq \mathcal{S}\left(g^\dagger; g^{\text{obs}}\right) + \alpha\,\mathcal{R}\left(u^\dagger\right)$$

and by rearranging terms we find

$$\alpha \mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) + \mathcal{S}\left(F(u_\alpha); g^{\text{obs}}\right) \leq \mathcal{S}\left(g^\dagger; g^{\text{obs}}\right) - \alpha\left\langle u^*, u_\alpha - u^\dagger\right\rangle$$

where we used the definition of the Bregman distance. Now adding $\mathfrak{s}_e\left(g^\dagger\right)$ on both sides and using (3.4) yields the claim. ■

3.2.1 Source conditions

From this result it is obvious that an appropriate estimate for $\langle u^*, u^\dagger - u_\alpha\rangle$ is sufficient to obtain convergence rates. On the other hand it is well-known from the classical theory (see e.g. [EHN96, Prop. 3.11]) that the convergence in Theorem 3.6 can be arbitrarily slow, i.e. without further assumptions on $u^\dagger$ no convergence rates can be obtained.

Spectral source conditions

In the classical setup where the underlying spaces $\mathbb{X}$ and $\mathbb{Y}$ are Hilbert spaces, range conditions of the form

$$u^\dagger - u_0 = \varphi\left(F'\left[u^\dagger\right]^* F'\left[u^\dagger\right]\right)\omega \tag{3.9}$$

are usually posed for some **index function** $\varphi : (0,\infty) \to (0,\infty)$, i.e. φ is continuous, strictly increasing and fulfills $\varphi(0) = 0$. The condition (3.9) uses the functional calculus and hence the underlying space structure explicitly. Such conditions were systematically studied in [Heg95, MP03]. The most common choices for φ are

$$\varphi_\nu(t) := t^\nu, \qquad \nu > 0, \tag{3.10a}$$

which is referred to as **Hölder-type source condition** and

$$\bar{\varphi}_p(t) := \begin{cases} (-\ln(t))^{-p} & \text{if } 0 < t \leq \exp(-p-1), \\ 0 & \text{if } t = 0, \end{cases} \qquad p > 0, \tag{3.10b}$$

which is known as **logarithmic source condition**. It is obvious that φ_ν is an index function, and via differentiation the same may be seen for $\bar{\varphi}_p$. If necessary, $\bar{\varphi}_p$ can be extended concavely to $(\exp(-p-1), \infty)$ via continuation by a suitable line.

For many interesting operators F range conditions of the form (3.9) can be interpreted as a smoothness condition in the sense that $u^\dagger$ belongs to some Sobolev space:

- In case of numerical differentiation (i.e. F is the bounded linear operator given by $(Fu)(y) = \int_0^y u(x)\,\mathrm{d}x$, $y \in [0,1]$ mapping $\mathbf{L}^2_\diamond([0,1])$ where the index $\diamond$ indicates that $\int_0^1 u\,\mathrm{d}y = 0$ into itself), we have (3.9) with $\varphi = \varphi_\nu$ if and only if $u^\dagger \in H^{2\nu}_{\mathrm{per}}([0,2\pi])$.
- For the backwards heat equation where one tries to determine the initial value $u(\cdot,0)$ from the heat distribution $u(\cdot,T)$ at some later date $T > 0$ it holds that

$$R\left(\bar{\varphi}_p\left(F^*F\right)\right) = D\left((I-\Delta)^p\right)$$

  where $F = c \cdot \exp(T\Delta)$ denotes the (linear) forward operator, $c > 0$ is some constant, Δ is the Laplace operator and I denotes the identity. The underlying spaces are chosen as $\mathbb{X} = \mathbb{Y} = \mathbf{L}^2(\Omega)$ (cf. [Hoh00, Sec. 8.1]).

  If $\Omega = \mathbb{R}^m$, then $D\left((I-\Delta)^p\right) = H^{2p}(\mathbb{R}^m)$, otherwise some boundary condition has to be included (see also [Hoh99, Sec. 3.71]). Hence, (3.9) with $\varphi = \bar{\varphi}_p$ is equivalent to $u^\dagger \in H$ where H is some subset of $H^{2p}(\Omega)$ depending on the boundary condition.
- If one tries to determine the potential of the earth from satellite measurements of the gravitational forces change rate in space this leads to a linear inverse problem where (3.9) with $\varphi = \bar{\varphi}_p$ is equivalent to $u^\dagger \in H^p\left(S^2\right)$ with the unit sphere S^2 (cf. [Hoh00, Sec. 8.2]).
- HOHAGE [Hoh97] has moreover shown that for an inverse potential as well as an inverse scattering problem (which are both nonlinear) the condition (3.9) with $\varphi = \varphi_\nu$ is too restrictive in the sense that even analyticity of $u^\dagger - u_0$ does not ensure (3.9) with $\varphi = \varphi_\nu$ to be fulfilled. Nevertheless, (3.9) with $\varphi = \bar{\varphi}_p$ has a suitable meaning for both examples.

For a compact linear operator $F = T$ we expect roughly spoken the following behavior: If the singular values σ_n of T decay to 0 at a polynomial rate, then (3.9) with $\varphi = \varphi_\nu$ seems reasonable. If σ_n decays exponentially to 0, then (3.9) with $\varphi = \varphi_\nu$ will in general be not fulfilled even for arbitrarily smooth $u^\dagger - u_0$, but (3.9) with $\varphi = \bar{\varphi}_p$ seems to be reasonable. Since the examples from Chapter 2 have all in common that the corresponding operators map arbitrary rough functions to analytic or at least arbitrary smooth functions, we expect for all these problems only weak source conditions (e.g. (3.9) with $\varphi = \bar{\varphi}_p$) to hold, but Hölder-type source conditions to be much too restrictive. Thus our theory focuses on the case of general φ including the weak cases.

Variational inequalities

A range condition like (3.9) is limited to the case where $\mathbb{X}$ and $\mathbb{Y}$ are Hilbert spaces, since otherwise no functional calculus is available. To overcome this lack, HOFMANN ET AL. [HKPS07] introduced **variational inequalities** in additive form, which means an assumption of the following type:

ASSUMPTION 3.15 (ADDITIVE VARIATIONAL INEQUALITY):
There exists $u^* \in \partial\mathcal{R}\left(u^\dagger\right) \subset \mathbb{X}'$, a parameter $\beta \in [0,1)$ and an index function φ where φ^2 is concave such that

$$\left\langle u^*, u^\dagger - u\right\rangle \leq \beta\mathcal{D}_{\mathcal{R}}^{u^*}\left(u, u^\dagger\right) + \varphi_{\mathrm{add}}\left(\mathcal{S}\left(F(u)\,; g^\dagger\right)\right) \qquad \text{for all } u \in \mathfrak{B}. \tag{3.11}$$

The original additive formulation in [HKPS07] deals only with the case $\varphi_{\text{add}} = \varphi_{\frac{1}{2}}$, which we will point out to be a case of special interest later on. The same assumption is used by SCHERZER ET AL. [SGG+08] to derive convergence rates for Tikhonov-type regularization. The case of a general index function φ_{add} was first treated by BOŢ & HOFMANN [BH10], who used a general form of Young's inequality to prove convergence rates in this general case. Moreover, HOFMANN & YAMAMOTO [HY10] prove equivalence of (3.11) with $\varphi_{\text{add}} = \varphi_{\frac{1}{2}}$ and (3.9) in the Hilbert space case for $\varphi = \varphi_{\frac{1}{2}}$ under a suitable nonlinearity condition and give a simple motivation of Assumption 3.15 for Hölder-type source conditions. Moreover, they show that (3.11) is somehow limited to the case that φ_{add}^2 is concave, since (3.11) with $\varphi_{\text{add}} = \varphi_\nu$ for some $\nu > \frac{1}{2}$ implies $u^* = 0$ in the quadratic Hilbert space setting. For non-quadratic norm powers the concavity of φ_{add}^2 is an additional restriction. The singular case that $\mathcal{S}(g; \hat{g}) = \|g - \hat{g}\|_{\mathbb{Y}}$ and $\varphi_{\text{add}}(s) = cs$ with some $c > 0$ is for example not covered by our work, but since the Kullback-Leibler divergence is bounded from below by the square of a norm as we will see in Chapter 4 this is not of interest for us. We refer to [Fle11, Prop. 4.14] for this situation.
PÖSCHL adapted the concept of variational inequalities to the case that $\mathcal{S}$ is not given by the power of some norm and proved convergence rates. Also GRASMAIR [Gra10a] used this concept. Finally, FLEMMING [Fle10] (see also [FH10, Fle11]) not only uses additive variational inequalities to prove convergence rates for Tikhonov-type regularization (3.2) with general $\mathcal{S}$ and $\mathcal{R}$, but moreover presents a general connection from spectral source conditions to variational inequalities which will be pointed out in the following.
Note that we do not use some constant $\tilde{\beta}$ in front of the second term in (3.11), since this can be overcome by redefining φ_{add}. Moreover, this would not change the rate of convergence, as we will see in our analysis. Thus, only the asymptotic behavior of φ_{add} as the argument tends to 0 is of interest. For the important cases of Hölder-type and logarithmic source conditions, we will often write that φ_{add} is given by φ_ν or $\bar{\varphi}_p$ as in (3.10a) or (3.10b) respectively, which means that there exists some constant $c > 0$ such that $\varphi_{\text{add}} = c \cdot \varphi_\nu$ and so on.
For the formulation of implications between different smoothness concepts as well as convergence rates, we need the Fenchel conjugate Φ^* of a real valued function Φ.

DEFINITION 3.16 (CONJUGATE FUNCTION):
*Let $\Phi : (\infty, \infty) \to (-\infty, \infty]$ be a function defined on the real line. Then the **Fenchel conjugate** Φ^* is defined by*

$$\Phi^*(s) = \sup_{\sigma \in \mathbb{R}} (\sigma s - \Phi(\sigma)), \qquad s \in (-\infty, \infty).$$

The function Φ^* is always convex as supremum over the affine linear (and hence convex) functions $s \mapsto \sigma s - \Phi(\sigma)$.

REMARK 3.17:
If a function $\chi : [0, \infty) \to (-\infty, \infty)$ is given, then we may extend χ by the value ∞ also to the negative real line and are hence able to calculate the Fenchel conjugate by

$$\chi^*(s) = \sup_{\sigma \in \mathbb{R}} (\sigma s - \chi(\sigma)) = \sup_{\sigma \geq 0} (\sigma s - \chi(\sigma)), \qquad s \in (-\infty, \infty).$$

If we consider especially the convex function $-\varphi_{\text{add}}$ (and its extension to $\mathbb{R}$), then the Fenchel conjugate is given by

$$(-\varphi_{\text{add}})^*(s) = \sup_{\sigma \geq 0} (\sigma s + \varphi_{\text{add}}(\sigma))$$

which attains finite values for $s < 0$.

In the following theorem we will give explicit formulas to calculate φ_{add} from φ. This is done for the sake of completeness and to convince the reader that both smoothness concepts are connected in the general case. Nevertheless, the following theorem and lemma are not necessary to understand this thesis and may hence be skipped in first reading.

THEOREM 3.18 (VALIDITY OF ADDITIVE VARIATIONAL INEQUALITIES):
Assume that $F = T : \mathbb{X} \to \mathbb{Y}$ is a bounded linear operator between Hilbert spaces $\mathbb{X}$ and $\mathbb{Y}$ and let $\mathcal{R}(u) = \|u - u_0\|_{\mathbb{X}}^2$ and $\mathcal{S}(g; \hat{g}) = \|g - \hat{g}\|_{\mathbb{Y}}^2$. Moreover let a spectral source condition (3.9) *hold true for some index function φ such that φ^2 is concave. Then Assumption 3.15 is fulfilled with*

$$\begin{aligned} d(r) &:= \min\left\{ \left\| u^\dagger - (T^*T)^{\frac{1}{2}} \omega \right\|_{\mathbb{X}} \mid \omega \in \mathbb{X}, \|\omega\|_{\mathbb{X}} \le r \right\}, \\ D_\beta(r) &= \frac{1}{2(1-\beta)} d^2(r), \\ \varphi_{\text{add}}(\sigma) &= -D_\beta^*(-\sigma). \end{aligned}$$

Note that the parameter $\beta \in [0, 1)$ can be chosen in this linear setup!

PROOF:
See [Fle11, Cor. 13.7 and 13.8] ■

The question is how to calculate the distance function d, which measures the degree of violation of the benchmark source condition $u^\dagger = (T^*T)^{\frac{1}{2}} \omega$ or equivalently $u^\dagger = T^*\bar{\omega}$. We will now give two possibilities to calculate d:

LEMMA 3.19 (SEE [FLE11, THM. 13.10]):
Let T be a compact and injective linear operator between Hilbert spaces $\mathbb{X}$ and $\mathbb{Y}$ and assume that (3.9) *is fulfilled, but $u^\dagger \notin R\left((T^*T)^{\frac{1}{2}}\right)$. Assume moreover that $\|\omega\| = 1$.*

- *If $f : t \mapsto \frac{\sqrt{t}}{\varphi(t)}$ (with the definition $f(0) := 0$) is an index function, then*

  $$d(r) \le r\sqrt{f^{-1}\left(\frac{1}{r}\right)}$$

  for all sufficiently large r.
- *If φ^2 is concave, then*

  $$d(r) \le \left(-\varphi \circ \cdot^2\right)^* (-r)$$

  for all $r \ge 0$.

PROOF:
This follows from [Fle11, Thm. 13.10] with $\psi(t) = \sqrt{t}$. ■

The function φ_{add} has been calculated explicitly in the most important cases of Hölder-type and logarithmic source conditions. We will present a calculation for Hölder-type source conditions and cite the result for logarithmic source conditions in the following. But next, we need an implication of spectral source conditions in Hilbert spaces:

LEMMA 3.20:
Let $\mathbb{X}$ and $\mathbb{Y}$ be Hilbert spaces, $\mathcal{S}\left(g; g^{\mathrm{obs}}\right) = \left\|g - g^{\mathrm{obs}}\right\|_{\mathbb{Y}}^2$ and $\mathcal{R}(u) = \|u - u_0\|_{\mathbb{X}}^2$ and assume that (3.9) *holds true with some index function φ such that $\phi := \left(\varphi^2\right)^{-1}$ is convex. Then the following variational inequality is valid:*

$$\left|\left\langle u^*, u^\dagger - u\right\rangle\right| \leq 2\|\omega\|_{\mathbb{X}} \left\|u - u^\dagger\right\|_{\mathbb{X}} \varphi\left(\frac{\left\|F'\left[u^\dagger\right]\left(u - u^\dagger\right)\right\|_{\mathbb{Y}}^2}{\left\|u - u^\dagger\right\|_{\mathbb{X}}^2}\right) \qquad \text{for all } u \in \mathbb{X}\,. \quad (3.12)$$

PROOF:
First note that $u^* = 2\left(u^\dagger - u_0\right)$ since $\mathcal{R}(u) = \|u - u_0\|_{\mathbb{X}}^2$. Let $\sigma = \sigma\left(F'\left[u^\dagger\right]^* F'\left[u^\dagger\right]\right)$ denote the spectrum and $(E_\lambda)_{\lambda\in\sigma}$ the spectral family of $F'\left[u^\dagger\right]^* F'\left[u^\dagger\right]$. Using the self-adjointness of $\varphi\left(F'\left[u^\dagger\right]^* F'\left[u^\dagger\right]\right)$ we find

$$\begin{aligned}\left|\left\langle u^*, u^\dagger - u\right\rangle\right| &= 2\left|\left\langle u^\dagger - u_0, u^\dagger - u\right\rangle\right| \\ &= 2\left|\left\langle \omega, \varphi\left(F'\left[u^\dagger\right]^* F'\left[u^\dagger\right]\right)\left(u^\dagger - u\right)\right\rangle\right| \\ &\leq 2\|\omega\|_{\mathbb{X}} \left\|\varphi\left(F'\left[u^\dagger\right]^* F'\left[u^\dagger\right]\right)\left(u^\dagger - u\right)\right\|_{\mathbb{X}}\end{aligned}$$

where we used the Cauchy-Schwarz inequality. Moreover, using the integral representations

$$\begin{aligned}\left\|\varphi\left(F'\left[u^\dagger\right]^* F'\left[u^\dagger\right]\right)\left(u^\dagger - u\right)\right\|_{\mathbb{Y}}^2 &= \int_\sigma \varphi^2(\lambda)\, \mathrm{d}\left\|E_\lambda\left(u^\dagger - u\right)\right\|^2, \\ \left\|u - u^\dagger\right\|_{\mathbb{X}}^2 &= \int_\sigma \mathrm{d}\left\|E_\lambda\left(u^\dagger - u\right)\right\|^2\end{aligned}$$

we find

$$\begin{aligned}\left|\left\langle u^*, u^\dagger - u\right\rangle\right| &\leq 2\|\omega\|_{\mathbb{Y}} \left(\int_\sigma \varphi^2(\lambda)\, \mathrm{d}\left\|E_\lambda\left(u^\dagger - u\right)\right\|^2\right)^{\frac{1}{2}} \\ &= 2\|\omega\|_{\mathbb{X}} \left\|u - u^\dagger\right\|_{\mathbb{X}} \left(\frac{\int_\sigma \varphi^2(\lambda)\, \mathrm{d}\left\|E_\lambda\left(u^\dagger - u\right)\right\|^2}{\int_\sigma \mathrm{d}\left\|E_\lambda\left(u^\dagger - u\right)\right\|^2}\right)^{\frac{1}{2}} \\ &= 2\|\omega\|_{\mathbb{X}} \left\|u - u^\dagger\right\|_{\mathbb{X}} \left(\phi^{-1}\left(\phi\left(\int_\sigma \varphi^2(\lambda)\, \frac{\mathrm{d}\left\|E_\lambda\left(u^\dagger - u\right)\right\|^2}{\int_\sigma \mathrm{d}\left\|E_\lambda\left(u^\dagger - u\right)\right\|^2}\right)\right)\right)^{\frac{1}{2}}.\end{aligned}$$

Since the measure $\frac{\mathrm{d}\left\|E_\lambda\left(u^\dagger - u\right)\right\|^2}{\int_\sigma \mathrm{d}\left\|E_\lambda\left(u^\dagger - u\right)\right\|^2}$ on the right-hand side is a probability measure, ϕ is convex by assumption and ϕ^{-1} is monotonically increasing, we can now apply Jensen's inequality to find

$$\left|\left\langle u^*, u^\dagger - u\right\rangle\right| \leq 2\,\|\omega\|_{\mathbb{Y}} \left\|u - u^\dagger\right\|_{\mathbb{X}} \varphi\left(\int_\sigma \phi\left(\varphi^2(\lambda)\right) \frac{\mathrm{d}\left\|E_\lambda\left(u^\dagger - u\right)\right\|^2}{\int_\sigma \mathrm{d}\left\|E_\lambda\left(u^\dagger - u\right)\right\|^2}\right)$$

$$= 2\,\|\omega\|_{\mathbb{X}} \left\|u - u^\dagger\right\|_{\mathbb{X}} \varphi\left(\frac{\int_\sigma \lambda\,\mathrm{d}\left\|E_\lambda\left(u^\dagger - u\right)\right\|^2}{\left\|u - u^\dagger\right\|_{\mathbb{X}}^2}\right).$$

Now the assertion follows from

$$\int_\sigma \lambda\,\mathrm{d}\left\|E_\lambda\left(u^\dagger - u\right)\right\|^2 = \left\|\left(F'\left[u^\dagger\right]^* F'\left[u^\dagger\right]\right)^{\frac{1}{2}}\left(u^\dagger - u\right)\right\|_{\mathbb{X}}^2$$

$$= \left\|F'\left[u^\dagger\right]\left(u - u^\dagger\right)\right\|_{\mathbb{Y}}^2. \qquad ■$$

With the help of this result, we can prove an optimal implication for the case of Hölder-type source conditions:

LEMMA 3.21 (SEE ALSO [HY10, PROP. 6.6] OR [FLE11, SEC. 13.5.1]):
Assume that $F = T$ is a linear operator between Hilbert spaces $\mathbb{X}$ and $\mathbb{Y}$ and let $\mathcal{S}\left(g; g^{\mathrm{obs}}\right) = \left\|g - g^{\mathrm{obs}}\right\|_{\mathbb{Y}}^2$ and $\mathcal{R}(u) = \|u - u_0\|_{\mathbb{X}}^2$. Then a range condition (3.9) *with $\varphi = \varphi_\nu$ from* (3.10a) *for $\nu \in \left(0, \frac{1}{2}\right]$ implies an additive variational inequality* (3.11) *with arbitrary $\beta \in (0,1)$ and $\varphi_{\mathrm{add}} = \bar{\beta}\varphi_\kappa$ where $\kappa = \frac{2\nu}{2\nu+1}$ and $\bar{\beta} > 0$ is some constant.*

PROOF:
If $\nu = \frac{1}{2}$, then the assertion is obvious by Lemma 3.20. For $\nu < \frac{1}{2}$ we insert the special structure $\varphi = \varphi_\nu$ into the result (3.12) of Lemma 3.20 to obtain

$$\left|\left\langle u^*, u^\dagger - u\right\rangle\right| \leq 2\,\|\omega\|_{\mathbb{X}} \left\|u - u^\dagger\right\|_{\mathbb{X}}^{1-2\nu} \left\|F(u) - g^\dagger\right\|_{\mathbb{Y}}^{2\nu}, \qquad \text{for all } u \in \mathbb{X}.$$

Using a modification of Young's inequality

$$ab = \left(\frac{\varepsilon}{2\,\|\omega\|_{\mathbb{X}}}\right)^{\frac{1}{p}} a \cdot \left(\frac{2\,\|\omega\|_{\mathbb{X}}}{\varepsilon}\right)^{\frac{1}{p}} b \leq \frac{\varepsilon}{2\,\|\omega\|_{\mathbb{X}}}\frac{a^p}{p} + \left(\frac{2\,\|\omega\|_{\mathbb{X}}}{\varepsilon}\right)^{\frac{q}{p}} \frac{b^q}{q}, \qquad \frac{1}{p} + \frac{1}{q} = 1$$

with $p = \frac{2}{1-2\nu}$, $q = \frac{2}{1+2\nu}$ this yields

$$\left|\left\langle u^*, u^\dagger - u\right\rangle\right| \leq \varepsilon\frac{1-2\nu}{2}\left\|u - u^\dagger\right\|_{\mathbb{X}}^2 + (1+2\nu)\left(\frac{2\,\|\omega\|_{\mathbb{X}}}{\varepsilon}\right)^{\frac{1+2\nu}{1-2\nu}} \|\omega\|_{\mathbb{X}} \left\|F(u) - g^\dagger\right\|^{\frac{4\nu}{1+2\nu}}$$

for all $u \in \mathbb{X}$. Now choose ε such that $\varepsilon\frac{1-2\nu}{2} = \beta$ and obtain the assertion. ■

In case of logarithmic source conditions, the result reads as follows:

LEMMA 3.22:
Assume that $F = T$ is a compact and injective linear operator between Hilbert spaces $\mathbb{X}$ and $\mathbb{Y}$ and let $\mathcal{S}\left(g; g^{\mathrm{obs}}\right) = \left\|g - g^{\mathrm{obs}}\right\|_{\mathbb{Y}}^2$ and $\mathcal{R}(u) = \|u - u_0\|_{\mathbb{X}}^2$. Then a range condition (3.9) *with $\varphi = \bar{\varphi}_p$ from* (3.10b) *for $p \in (0, \infty)$ implies an additive variational inequality* (3.11) *with arbitrary $\beta \in (0,1)$ and $\varphi_{\mathrm{add}} = \bar{\beta}\bar{\varphi}_{2p}$ where $\bar{\beta} > 0$ is some constant.*

PROOF:
This is done by using the result of Lemma 3.19 and Theorem 3.18. For the details see [Fle11, Sec. 13.5.2]. ■

The additive variational inequality (3.11) has the advantage that even for nonlinear F the given formulation is meaningful and no Fréchet derivative of F is required. If (3.9) holds true, then we find as above an additive variational inequality where $\mathcal{S}\left(F(u);g^\dagger\right) = \left\|F(u) - g^\dagger\right\|_{\mathbb{Y}}^2$ in the quadratic Hilbert space case is replaced by $\left\|F'\left[u^\dagger\right]\left(u-u^\dagger\right)\right\|_{\mathbb{Y}}^2$. This formulation makes explicit usage of the Fréchet derivative F' of F. It we want to derive (3.11) in this context, an additional nonlinearity condition on F is needed. A widely used assumption especially for the analysis of iterative methods is the **tangential cone condition**:

ASSUMPTION 3.23 (TANGENTIAL CONE CONDITION):
There exists a constant $\bar{\eta} > 0$ *such that*

$$\left\|F(v) - F(u) - F'[u](v-u)\right\|_{\mathbb{Y}} \leq \bar{\eta}\left\|F(v) - F(u)\right\|_{\mathbb{Y}} \tag{3.13}$$

for all $u, v \in \mathfrak{B}$.

If Assumption 3.23 holds true, then by the second triangle inequality

$$\left\|F'\left[u^\dagger\right]\left(u-u^\dagger\right)\right\|_{\mathbb{Y}} \leq (1+\bar{\eta})\left\|F(u) - F\left(u^\dagger\right)\right\|_{\mathbb{Y}} \tag{3.14}$$

and hence we may replace $\left\|F'\left[u^\dagger\right]\left(u-u^\dagger\right)\right\|_{\mathbb{Y}}$ by $\left\|F(u) - F\left(u^\dagger\right)\right\|_{\mathbb{Y}}$ only loosing some constant. This shows that in the nonlinear case Assumption 3.15 can be seen as a combination of a source and a nonlinearity condition which is weaker than both assumptions together.
We do not discuss the nonlinearity condition (3.13) here, this will be done in Section 6.1, where also a generalization will be presented.

As we have seen in Lemma 3.20, from a Hilbert space setting a multiplicative variational source condition is easier to derive. Therefore, also variational inequalities in multiplicative form have been proposed by KALTENBACHER & HOFMANN [KH10]. The following assumption is a variation of the one proposed in [KH10], which avoids Fréchet derivatives of F and hence additive assumptions as the tangential cone condition which links them to F:

ASSUMPTION 3.24 (MULTIPLICATIVE VARIATIONAL INEQUALITY):
There exists $u^* \in \partial\mathcal{R}\left(u^\dagger\right) \subset \mathbb{X}'$, $\beta \geq 0$ *and a concave index function* $\varphi_{\mathrm{mult}} : (0,\infty) \to (0,\infty)$ *such that*

$$\left\langle u^*, u^\dagger - u\right\rangle \leq \beta \mathcal{D}_{\mathcal{R}}^{u^*}\left(u,u^\dagger\right)^{\frac{1}{2}} \varphi_{\mathrm{mult}}\left(\frac{\mathcal{S}\left(F(u);g^\dagger\right)}{\mathcal{D}_{\mathcal{R}}^{u^*}\left(u,u^\dagger\right)}\right) \qquad \text{for all } u \in \mathfrak{B}. \tag{3.15}$$

Moreover assume that φ_{mult} *is such that*

$$\sigma \mapsto \frac{\varphi_{\mathrm{mult}}(\sigma)}{\sqrt{\sigma}} \qquad \text{is monotonically decreasing.} \tag{3.16}$$

To motivate Assumption 3.24, one simply inserts (3.13) into (3.12) and generalizes the data misfit and the penalty.

The additional condition (3.16) restricts the possible index functions φ_{mult} in a similar way as requiring that φ_{add}^2 is concave in Assumption 3.15 does. For example, if we think of Hölder-type variational inequalities, i.e. $\varphi_{\text{mult}}(t) = t^\nu$, ν must be less or equal $\frac{1}{2}$. For additive variational inequalities (3.11) it is known from HOFMANN & YAMAMOTO that in case of a norm power $\mathcal{S}(g;\hat{g}) = \|g - \hat{g}\|_{\mathbb{Y}}^p$ the exponent for Hölder-type functions $\varphi_{\text{add}} = \varphi_\nu$ in (3.11) is limited to $\nu \leq \frac{1}{p}$ if $u^* \neq 0$ (cf. [HY10, Prop. 4.3]). Note that this range is not covered completely by our requirements, since $\varphi_{\frac{1}{p}}^2$ is concave if and only if $p \geq 2$. Thus for higher norm powers we are restricted to weaker source conditions. Nevertheless our theory can be extended to those cases by some simple changes. But since we are mainly interested in the Kullback-Leibler divergence as data fidelity term we restricted ourselves to these cases which are sufficient as we will see in Chapters 4 and 7.
Our motivation for nonlinear operators uses only the tangential cone condition so far, but in case of $\varphi = \varphi_{\frac{1}{2}}$ we expect that a tangential cone condition is not necessary for proving convergence rates. In the Hilbert space case (cf. for example [EHN96, Sec. 10]), a Lipschitz estimate on $F'[\cdot]$ is sufficient. It has already been proven by SCHERZER ET AL [SGG+08, Prop 3.35] that Assumption (3.15) with $\varphi_{\text{add}} = \varphi_{\frac{1}{2}}$ is valid in the Hilbert space setup if (3.9) with $\varphi = \varphi_{\frac{1}{2}}$ and the **Lipschitz estimate**

$$\left\|F(v) - F(u) - F'(u; v-u)\right\|_{\mathbb{Y}} \leq \frac{L}{2}\|v-u\|_{\mathbb{X}}^2 \tag{3.17}$$

on a sufficiently large set with $\frac{L}{2}\|\omega\| < 1$ hold true. To be more specific, in this setup Assumption (3.15) holds true with $\varphi_{\text{add}} = \varphi_{\frac{1}{2}}$ and the set $\mathfrak{B}$ where (3.17) is valid. It has been moreover shown by FLEMMING AND HOFMANN [FH11] that for convex $\mathfrak{B}$ and (3.13) as nonlinearity assumption the condition (3.9) with $\varphi = \varphi_{\frac{1}{2}}$ can be relaxed to a projected source condition

$$u^\dagger = P_{\mathfrak{B}}\left(u_0 + F'\left[u^\dagger\right]^* \omega\right) \tag{3.18}$$

where $P_{\mathfrak{B}} : \mathbb{X} \to \mathbb{X}$ denotes the metric projector onto $\mathfrak{B}$. The converse implication also holds true if $\mathfrak{B}$ contains inner points. Moreover the nonlinearity condition can be relaxed to the Lipschitz assumption (3.17) if $\frac{L}{2}\|\omega\| < 1$.

Comparison and limitations

Let us discuss some limitations of source conditions. In the quadratic Hilbert space case is has been proven by MATHÉ & HOFMANN [MH08] that any $u^\dagger \in \mathbb{X}$ fulfills a source condition of the form (3.9) for a suitable index function φ (which depends on $u^\dagger$). Note that the decay of $\varphi(t)$ as $t \searrow 0$ might be arbitrary slow in accordance with [EHN96, Prop. 3.11]. This fact has two main implications: On the one hand, a convergence analysis for linear F under a spectral source condition (3.9) and hence under a variational inequality (3.15) or (3.11) with general φ provides rates of convergence **for any unknown solution** $u^\dagger$. If F is nonlinear, the crux of the matter lies in the fulfillment of a nonlinearity condition. On the other hand, since also the source element ω in (3.9) fulfills again a source condition, the equation (3.9) **is not able to cover the whole smoothness of $u^\dagger$**. In that sense, any proven convergence rate is elementwise[1] suboptimal.

So far it is unknown if a similar result holds true also in the general case where Assumption 3.15 is seen as source condition, i.e. it has not been proven that for arbitrary $u^\dagger$ and linear F an inequality like (3.11) holds true. Nevertheless, it has been shown by FLEMMING, HOFMANN & MATHÉ [FHM11] that Assumption 3.15 **covers the whole smoothness of $u^\dagger$** and leads to elementwise optimal convergence rates. This is the best possible result, but the proof so far uses concepts of approximate source conditions and approximate variational inequalities, which we do not introduce in this work.

An overview over the different smoothness concepts is shown in Table 3.1.

	linear operator	tangential cone condition (3.13)	Lipschitz condition (3.17)
(3.9) with general φ	Assumptions 3.15 and 3.24 hold true	Assumptions 3.15 and 3.24 hold true	Unknown
(3.9) with $\varphi = \varphi_{\frac{1}{2}}$	Assumptions 3.15 and 3.24 hold true	Assumptions 3.15 and 3.24 hold true	Assumption 3.15 holds true if $\frac{L}{2}\left\|\omega\right\| < 1$
Projected source condition (3.18)	Assumptions 3.15 and 3.24 hold true	Assumption 3.15 holds true	Assumption 3.15 holds true if $\frac{L}{2}\left\|\omega\right\| < 1$

Table 3.1: Sufficient conditions for the validity of variational inequalities in the quadratic Hilbert space case where $\mathcal{S}(g;\hat{g}) = \|g-\hat{g}\|_{\mathbb{Y}}^2$ and $\mathcal{R}(u) = \|u-u_0\|_{\mathbb{X}}^2$ for Hilbert norms $\|\cdot\|_{\mathbb{X}}$ and $\|\cdot\|_{\mathbb{Y}}$.

3.2.2 General convergence theorems

In this subsection we will present and prove two general results on convergence rates for Tikhonov-type regularization (3.2). The first result provides rates of convergence under Assumption 3.24, the second under Assumption 3.15. KALTENBACHER & HOFMANN [KH10] have proven rates of convergence for the iteratively regularized Gauss-Newton method under a condition similar to Assumption 3.24 (see Chapter 5), and our proof for Tikhonov-type regularization uses the same techniques. Before we are able to prove convergence rates under Assumption 3.24 we need to define the corresponding rate functions:

DEFINITION 3.25 (RATE FUNCTION):
Let $\varphi_{\mathrm{mult}} : (0,\infty) \to (0,\infty)$ be some index function. For multiplicative variational inequalities as in Assumption 3.24 the function

$$\Theta(t) := t \cdot \varphi_{\mathrm{mult}}^2(t), \qquad t > 0 \tag{3.19}$$

*is called **rate function**. Moreover we denote $\vartheta = \sqrt{\Theta}$.*

[1] Optimality is usually meant w.r.t. some class M of exact solutions $u^\dagger$, and the known convergence rates results obtain the best possible convergence rates under the assumption that $u^\dagger$ belongs to the class. Nevertheless, for any specific $u^\dagger$ there exists a better convergence rate, and this rate is obtained under (3.11). This is meant by 'elementwise' optimality and suboptimality.

Note that $\Theta, \vartheta, \Theta^{-1}$ and ϑ^{-1} are again index functions.

LEMMA 3.26 (ADDITIONAL PROPERTIES OF φ_{mult} AND Θ, CF. [KH10, REM. 2]):
Let φ_{mult} as in Assumption 3.24.

(a) We have

$$\varphi_{\mathrm{mult}}\left(\vartheta^{-1}(Ct)\right) \leq \max\left\{\sqrt{C}, 1\right\} \varphi_{\mathrm{mult}}\left(\vartheta^{-1}(t)\right) \tag{3.20}$$

$$\varphi^2_{\mathrm{mult}}\left(\Theta^{-1}(Ct)\right) \leq \max\left\{\sqrt{C}, 1\right\} \varphi^2_{\mathrm{mult}}\left(\Theta^{-1}(t)\right) \tag{3.21}$$

for all $t \geq 0$ and $C > 0$.

(b) It holds

$$\varphi_{\mathrm{mult}}(\lambda t) \leq \lambda \varphi_{\mathrm{mult}}(t) \qquad \text{for all } t > 0 \text{ and } \lambda \geq 1 \tag{3.22}$$

(c) For all $t > 0$ and $\lambda \geq 1$ the following inequality holds:

$$\Theta(\lambda t) \leq \lambda^3 \Theta(t) \tag{3.23}$$

PROOF:

(a) Let us denote $\sigma = \vartheta(t)$. Then

$$\frac{\varphi_{\mathrm{mult}}\left(\vartheta^{-1}(t)\right)}{\sqrt{t}} = \frac{\varphi_{\mathrm{mult}}(\sigma)}{\sqrt{\vartheta(\sigma)}} = \frac{\varphi_{\mathrm{mult}}(\sigma)}{\sqrt{\sqrt{\sigma}\varphi_{\mathrm{mult}}(\sigma)}} = \sqrt{\frac{\varphi_{\mathrm{mult}}(\sigma)}{\sqrt{\sigma}}}$$

and the right-hand side of this equation is by (3.16) monotonically decreasing. Therefore

$$\frac{\varphi_{\mathrm{mult}}\left(\vartheta^{-1}(Ct)\right)}{\sqrt{Ct}} \leq \frac{\varphi_{\mathrm{mult}}\left(\vartheta^{-1}(t)\right)}{\sqrt{t}}$$

whenever $C \geq 1$ and hence $\varphi_{\mathrm{mult}}\left(\vartheta^{-1}(Ct)\right) \leq \sqrt{C}\varphi_{\mathrm{mult}}\left(\vartheta^{-1}(t)\right)$. For $C \leq 1$, is follows from the monotonicity of φ_{mult} and ϑ^{-1} that (3.20) also holds true.

The second inequality is obtained in the same way.

(b) By Assumption 3.24 φ_{mult} is concave and hence the assertion follows directly from $\varphi_{\mathrm{mult}}(0) = 0$.

(c) Since $\Theta(t) = t \cdot \varphi^2_{\mathrm{mult}}(t)$ by definition (3.19), we obtain the assertion by using (3.22). ■

LEMMA 3.27:
Assume that an inequality of the type

$$\alpha a + \frac{1}{C_1} b \leq \alpha C_2 \sqrt{a} \varphi_{\mathrm{mult}}\left(\frac{b}{a}\right)$$

holds true for all $\alpha > 0$ and some concave index function φ_{mult} fulfilling (3.16)*, where $C_1, C_2 > 0$ are some constants and a and b are non-negative functions of $\alpha > 0$. Then a and b fulfill*

$$a \leq C_1^{\frac{3}{2}} C_2^2 \varphi^2_{\mathrm{mult}}(\alpha),$$

$$b \leq C_1^3 C_2^3 \alpha \varphi^2_{\mathrm{mult}}(\alpha)$$

for all $\alpha > 0$.

PROOF:
Multiplying the given inequality with $\sqrt{b}/a$ leads to

$$\alpha\sqrt{b} + \frac{1}{C_1}\frac{b}{a}\sqrt{b} \leq C_2 \alpha \vartheta\left(\frac{b}{a}\right).$$

By considering only the first and the second term on the left-hand side this yields

$$\vartheta^{-1}\left(\frac{\sqrt{b}}{C_2}\right) \leq \frac{b}{a}, \tag{3.24}$$

$$\Phi\left(\frac{b}{a}\right)\sqrt{b} \leq C_1 C_2 \alpha$$

respectively, where $\Phi(t) = t/\vartheta(t) = \sqrt{t}/\varphi_{\mathrm{mult}}(t)$. Since Φ is monotonically increasing by (3.16), we can combine these two inequalities to

$$\Phi\left(\vartheta^{-1}\left(\frac{\sqrt{b}}{C_2}\right)\right)\sqrt{b} \leq C_1 C_2 \alpha.$$

Now note that $\Phi\left(\vartheta^{-1}(t)\right) = \vartheta^{-1}(t)/t$ and find

$$\vartheta^{-1}\left(\frac{\sqrt{b}}{C_2}\right) \leq C_1 \alpha.$$

This shows by using (3.22) that

$$\sqrt{b} \leq C_2 \vartheta(C_1\alpha) \leq C_1^{\frac{3}{2}} C_2 \sqrt{\alpha \varphi_{\mathrm{mult}}(\alpha)}$$

and hence

$$b \leq C_1^3 C_2^2 \alpha \varphi_{\mathrm{mult}}^2(\alpha)$$

for all $\alpha > 0$. Moreover, using (3.24) and $s^2\left(\varphi_{\mathrm{mult}}\left(\vartheta^{-1}\left(\frac{\sqrt{t}}{s}\right)\right)\right)^2 = t/\vartheta^{-1}\left(\frac{\sqrt{t}}{s}\right)$ we get

$$a \leq \frac{b}{\vartheta^{-1}\left(\frac{\sqrt{b}}{C_2}\right)} = C_2^2\left(\varphi_{\mathrm{mult}}\left(\vartheta^{-1}\left(\frac{\sqrt{b}}{C_2}\right)\right)\right)^2.$$

Now we use the result on b and (3.20) to find

$$a \leq C_2^2\left(\varphi_{\mathrm{mult}}\left(\vartheta^{-1}\left(C_1^{\frac{3}{2}} C_2 \sqrt{\alpha\varphi_{\mathrm{mult}}(\alpha)}\right)\right)\right)^2 \leq C_1^{\frac{3}{2}} C_2^3 \varphi_{\mathrm{mult}}^2(\alpha).$$

This proves the assertion. ■

Now we are able to formulate and prove a convergence rates result under Assumption 3.24 for Tikhonov-type regularization (3.2):

THEOREM 3.28 (CONVERGENCE RATES UNDER ASSUMPTION 3.24):
Let Assumptions 3.8 and 3.24 hold. Then any choice of minimizers u_α from (3.2) *fulfills*

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) = \mathcal{O}\left(\varphi_{\mathrm{mult}}^2(\alpha)\right), \tag{3.25a}$$

$$\mathcal{S}\left(F(u_\alpha); g^\dagger\right) = \mathcal{O}\left(\Theta(\alpha)\right) \tag{3.25b}$$

as $\alpha \searrow 0$ for exact data and if we choose α such that $\mathbf{err} \sim \Theta(\alpha)$ with $\mathbf{err}$ as in (3.8) and the rate function Θ in case of noisy data the following convergence rates are valid:

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) = \mathcal{O}\left(\varphi_{\mathrm{mult}}^2\left(\Theta^{-1}\left(\mathbf{err}\right)\right)\right), \tag{3.26a}$$

$$\mathcal{S}\left(F\left(u_\alpha\right); g^\dagger\right) = \mathcal{O}\left(\mathbf{err}\right) \tag{3.26b}$$

as $\mathbf{err} \searrow 0$.

PROOF:
Inserting Assumption 3.24 into the result of Lemma 3.14 we find

$$\alpha \mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) + \frac{1}{C_{\mathrm{err}}}\mathcal{S}\left(F\left(u_\alpha\right); g^\dagger\right) \leq \alpha\beta \mathcal{D}_{\mathcal{R}}^{u^*}\left(u, u^\dagger\right)^{\frac{1}{2}} \varphi_{\mathrm{mult}}\left(\frac{\mathcal{S}\left(F\left(u_\alpha\right); g^\dagger\right)}{\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right)}\right) + \mathbf{err}.$$

The parameter choice $\mathbf{err} \sim \Theta(\alpha)$ implies that there exists some constant $C > 0$ such that $\mathbf{err} \leq C\Theta(\alpha)$. We insert this and distinguish between the following two cases:

- $\alpha\beta \mathcal{D}_{\mathcal{R}}^{u^*}\left(u, u^\dagger\right)^{\frac{1}{2}} \varphi_{\mathrm{mult}}\left(\frac{\mathcal{S}\left(F(u_\alpha); g^\dagger\right)}{\mathcal{D}_{\mathcal{R}}^{u^*}(u_\alpha, u^\dagger)}\right) \leq C\Theta(\alpha)$. This yields

  $$\alpha \mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) + \frac{1}{C_{\mathrm{err}}}\mathcal{S}\left(F\left(u_\alpha\right); g^\dagger\right) \leq 2C\alpha\varphi_{\mathrm{mult}}^2(\alpha)$$

  which immediately implies

  $$\begin{aligned} \mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) &= \mathcal{O}\left(\varphi_{\mathrm{mult}}^2(\alpha)\right) = \mathcal{O}\left(\varphi_{\mathrm{mult}}^2\left(\Theta^{-1}\left(\mathbf{err}\right)\right)\right) \\ \mathcal{S}\left(F\left(u_\alpha\right); g^\dagger\right) &= \mathcal{O}\left(\alpha\varphi_{\mathrm{mult}}^2(\alpha)\right) = \mathcal{O}\left(\mathbf{err}\right). \end{aligned}$$

- $\alpha\beta \mathcal{D}_{\mathcal{R}}^{u^*}\left(u, u^\dagger\right)^{\frac{1}{2}} \varphi_{\mathrm{mult}}\left(\frac{\mathcal{S}\left(F(u_\alpha); g^\dagger\right)}{\mathcal{D}_{\mathcal{R}}^{u^*}(u_\alpha, u^\dagger)}\right) \geq C\Theta(\alpha)$. In this case we find

  $$\alpha \mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) + \frac{1}{C_{\mathrm{err}}}\mathcal{S}\left(F\left(u_\alpha\right); g^\dagger\right) \leq 2\alpha\beta \mathcal{D}_{\mathcal{R}}^{u^*}\left(u, u^\dagger\right)^{\frac{1}{2}} \varphi_{\mathrm{mult}}\left(\frac{\mathcal{S}\left(F\left(u_\alpha\right); g^\dagger\right)}{\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right)}\right).$$

  Lemma 3.27 implies (3.25) and by again inserting $\mathbf{err} \sim \Theta(\alpha)$ and using (3.22) we find (3.26). ∎

As opposed to the case of multiplicative variational inequalities, the proof of convergence rates under Assumption 3.15 is not only easier, but also provides an error decomposition.

REMARK 3.29:
Consider the function $f_{\mathrm{app}}(s) := \left(-\varphi_{\mathrm{add}}\right)^*\left(-\frac{1}{s}\right), s > 0$. We will show that this function is a bound for the approximation error under Assumption 3.15, which has first been shown by GRASMAIR [Gra10a]. Before doing so, let us collect some properties of f_{app}:

(a) f_{app} is monotonically increasing.

(b) It holds

$$f_{\mathrm{app}}(Cs) \leq \max\{1, C\} f_{\mathrm{app}}(s) \tag{3.27}$$

for all $s > 0, C > 0$.

(c) We have $f_{\text{app}}(s) \searrow 0$ as $s \searrow 0$.

PROOF:

(a) The monotonicity follows from the fact that both $(-\varphi_{\text{add}})^*$ as well as $s \mapsto -\frac{1}{s}, s > 0$ are monotonically increasing.

(b) If $C < 1$, then the assertion is trivial due to (a). For $C \geq 1$ we obtain from the concavity of φ_{add}^2 and $\varphi_{\text{add}}(0) = 0$ that

$$\varphi_{\text{add}}(Cs) \leq \sqrt{C}\varphi_{\text{add}}(s) \qquad \text{for all} \qquad s > 0.$$

Now using this and the definition, we obtain

$$\begin{aligned}
(-\varphi_{\text{add}})^* \left(-\frac{1}{Cs}\right) &= \sup_{\sigma \geq 0} \left(-\frac{1}{Cs}\sigma - (-\varphi_{\text{add}})(\sigma)\right) \\
&= \sup_{\sigma \geq 0} \left(\varphi_{\text{add}}(\sigma) - \frac{1}{s}\frac{\sigma}{C}\right) \\
&= \sup_{\bar{\sigma} = \sigma/C^2 \geq 0} \left(\varphi_{\text{add}}\left(C^2\bar{\sigma}\right) - \frac{1}{s}C\bar{\sigma}\right) \\
&\leq \sup_{\bar{\sigma} \geq 0} \left(C\varphi_{\text{add}}(\bar{\sigma}) - C\frac{1}{s}\bar{\sigma}\right) \\
&= C \sup_{\bar{\sigma} \geq 0} \left(\varphi_{\text{add}}(\bar{\sigma}) - \frac{1}{s}\sigma\right) \\
&= C(-\varphi_{\text{add}})^* \left(-\frac{1}{s}\right).
\end{aligned}$$

(c) By Young's inequality $\langle s^*, s \rangle \leq \Phi(s) + \Phi^*(s^*)$ where we have equality if and only if $s^* \in \partial\Phi(s)$ we can write

$$f_{\text{app}}(s) = \varphi_{\text{add}}(\sigma(s)) - \frac{\sigma(s)}{s}$$

for any choice $\sigma(s) \in \partial(-\varphi_{\text{add}})^* \left(-\frac{1}{s}\right)$. Note that $s > 0$ and hence also $\sigma(s) > 0$. As above we have moreover $\varphi_{\text{add}}(\sigma) \leq \max\left\{\sqrt{\sigma}, 1\right\} \varphi_{\text{add}}(1)$ for all $\sigma > 0$ by the concavity of φ_{add}^2. This yields

$$0 \leq \varphi_{\text{add}}(0) \leq f_{\text{app}}(s) = \varphi_{\text{add}}(\sigma(s)) - \frac{\sigma(s)}{s} \leq \max\left\{\sqrt{\sigma(s)}, 1\right\} \varphi_{\text{add}}(1) - \frac{\sigma(s)}{s}$$

and hence $\frac{\sigma(s)}{\max\left\{\sqrt{\sigma(s)},1\right\}} \leq \varphi_{\text{add}}(1)\, s$. Thus we find $\sigma(s) \searrow 0$ as $s \searrow 0$. This implies finally

$$f_{\text{app}}(s) = \varphi_{\text{add}}(\sigma(s)) - \frac{\sigma(s)}{s} \leq \varphi_{\text{add}}(\sigma(s)) \searrow 0 \qquad \text{as} \qquad s \searrow 0$$

and proves the claim. ■

Now we are able to provide the main convergence theorem under Assumption 3.15 for Tikhonov-type regularization (3.2):

THEOREM 3.30 (CONVERGENCE RATES UNDER ASSUMPTION 3.15):
Let Assumptions 3.8 and 3.15 hold true. Then for any choice of minimizers u_α from (3.2) the error decomposition

$$(1-\beta)\,\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) + \frac{1}{2C_{\mathrm{err}}\alpha}\mathcal{S}\left(F\left(u_\alpha\right); g^\dagger\right) \le 2C_{\mathrm{err}}\left(-\varphi_{\mathrm{add}}\right)^*\left(-\frac{1}{\alpha}\right) + \frac{\mathbf{err}}{\alpha} \tag{3.28}$$

is valid for any $\alpha > 0$ where **err** *is defined by* (3.8) *and for exact data it holds*

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) = \mathcal{O}\left(\left(-\varphi_{\mathrm{add}}\right)^*\left(-\frac{1}{\alpha}\right)\right), \tag{3.29a}$$

$$\mathcal{S}\left(F\left(u_\alpha\right); g^\dagger\right) = \mathcal{O}\left(\alpha\left(-\varphi_{\mathrm{add}}\right)^*\left(-\frac{1}{\alpha}\right)\right) \tag{3.29b}$$

as $\alpha \searrow 0$. If we choose in case of noisy data α such that

$$\frac{\tau}{\alpha} \in -\partial\left(-\varphi_{\mathrm{add}}\right)(\mathbf{err}) \tag{3.30}$$

with a tuning parameter $\tau > 0$ then the following convergence rate is valid:

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) = \mathcal{O}\left(\varphi_{\mathrm{add}}(\mathbf{err})\right) \qquad as \qquad \mathbf{err} \searrow 0. \tag{3.31a}$$

Before proving this result, we want to comment on the parameter choice rule (3.30):

REMARK 3.31:

(a) Since φ_{add} is assumed to be finite and concave, $-\varphi_{\mathrm{add}}$ is finite and convex and thus $-\varphi_{\mathrm{add}}$ is continuous (see e.g. [ET76, Cor. 2.3]). Hence $\partial\left(-\varphi_{\mathrm{add}}\right)(s) \neq \emptyset$ for all $s > 0$ (see e.g. [ET76, Prop. 5.2]) and so a parameter $\alpha > 0$ fulfilling (3.30) exists.

(b) If φ_{add} is differentiable at $s > 0$, then $\partial\left(-\varphi_{\mathrm{add}}\right)(s) = \left\{-\varphi'_{\mathrm{add}}(s)\right\}$ and hence (3.30) is equivalent to

$$\alpha \sim \frac{1}{\varphi'_{\mathrm{add}}(\mathbf{err})}.$$

(c) For concave φ_{add} it is easy to see that the left- and right-hand sided derivatives $\delta_-\varphi_{\mathrm{add}}(s) = \lim_{h\searrow 0}\frac{\varphi_{\mathrm{add}}(s)-\varphi_{\mathrm{add}}(s-h)}{h}$ and $\delta_+\varphi_{\mathrm{add}}(s) = \lim_{h\searrow 0}\frac{\varphi_{\mathrm{add}}(s+h)-\varphi_{\mathrm{add}}(s)}{h}$ exist and that it holds

$$-\partial\left(-\varphi_{\mathrm{add}}\right)(s) = \left[\delta_+\varphi_{\mathrm{add}}(s), \delta_-\varphi_{\mathrm{add}}(s)\right] \tag{3.32a}$$

$$= \left[\sup_{\sigma\in(s,\infty)}\frac{\varphi_{\mathrm{add}}(\sigma)-\varphi_{\mathrm{add}}(s)}{\sigma - s}, \inf_{\sigma\in[0,s)}\frac{\varphi_{\mathrm{add}}(s)-\varphi_{\mathrm{add}}(\sigma)}{s-\sigma}\right] \tag{3.32b}$$

where the last equality follows from the monotonicity of $z \mapsto \frac{\varphi_{\mathrm{add}}(x)-\varphi_{\mathrm{add}}(z)}{x-z}$. Thus our parameter choice is similar to the one proposed in [Fle11, Thm 4.11], and the following properties are proven using FLEMMING's arguments.

(d) If we choose α by (3.30), then by (3.32b)

$$\frac{\mathbf{err}}{\alpha(\mathbf{err})} \le \frac{\mathbf{err}}{\tau}\inf_{\sigma\in[0,\mathbf{err})}\frac{\varphi_{\mathrm{add}}(\mathbf{err})-\varphi_{\mathrm{add}}(\sigma)}{\mathbf{err}-\sigma} \le \frac{\mathbf{err}}{\tau}\frac{\varphi_{\mathrm{add}}(\mathbf{err})}{\mathbf{err}}$$

and thus $\frac{\mathbf{err}}{\alpha(\mathbf{err})} \searrow 0$ as $\mathbf{err} \searrow 0$.

(e) Since φ_{add}^2 is concave we gain moreover $\alpha(\mathbf{err}) \searrow 0$ as $\mathbf{err} \searrow 0$. To see this it suffices to show that $\sup_{\sigma\in(s,\infty)} \frac{\varphi_{\text{add}}(\sigma)-\varphi_{\text{add}}(s)}{\sigma-s} \nearrow \infty$ as $s \searrow 0$. If we had

$$\sup_{\sigma\in(t,\infty)} \frac{\varphi_{\text{add}}(\sigma)-\varphi_{\text{add}}(t)}{\sigma-t} \leq c \qquad \text{for some } c \in (0,\infty) \text{ and } t \in (0,t_0]\,,\ t_0 > 0,$$

then for any $s > 0$ and $t \in (0, \min\{s,t_0\})$ we find

$$\frac{\varphi_{\text{add}}(s)-\varphi_{\text{add}}(t)}{s-t} \leq \sup_{\sigma\in(t,\infty)} \frac{\varphi_{\text{add}}(\sigma)-\varphi_{\text{add}}(t)}{\sigma-t} \leq c.$$

Thus $\varphi_{\text{add}}(s) \leq \varphi_{\text{add}}(t) + c(s-t)$ and letting $t \searrow 0$ we find $\varphi_{\text{add}}(s) \leq cs$ for all $s \in (0,\infty)$. But due to the concavity of φ_{add}^2 we have $\sqrt{s}\varphi_{\text{add}}(1) \leq \varphi_{\text{add}}(s)$ for all $s \in [0,1]$, which contradicts $c < \infty$. Thus the supremum tends to ∞ as $s \searrow 0$.

Let us conclude with the proof of Theorem 3.30:

PROOF (OF THEOREM 3.30):
Plugging (3.11) into the result of Lemma 3.14 we find

$$\alpha(1-\beta)\,\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) + \frac{1}{C_{\text{err}}}\mathcal{S}\left(F(u_\alpha);g^\dagger\right) \leq \alpha\varphi_{\text{add}}\left(\mathcal{S}\left(F(u_\alpha);g^\dagger\right)\right) + \mathbf{err}.$$

Rearranging and dividing by α yields

$$\begin{aligned}
&(1-\beta)\,\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) + \frac{1}{2C_{\text{err}}\alpha}\mathcal{S}\left(F(u_\alpha);g^\dagger\right) \\
\leq&\varphi_{\text{add}}\left(\mathcal{S}\left(F(u_\alpha);g^\dagger\right)\right) - \frac{1}{2C_{\text{err}}\alpha}\mathcal{S}\left(F(u_\alpha);g^\dagger\right) + \frac{\mathbf{err}}{\alpha} \\
=&\sup_{\sigma\geq 0}\left(\varphi_{\text{add}}(\sigma) - \frac{\sigma}{2C_{\text{err}}\alpha}\right) + \frac{\mathbf{err}}{\alpha} \\
=&(-\varphi_{\text{add}})^*\left(-\frac{1}{2C_{\text{err}}\alpha}\right) + \frac{\mathbf{err}}{\alpha}.
\end{aligned}$$

Using (3.27) this shows (3.28). In the case of exact data we have $\mathbf{err} \equiv 0$ and hence (3.29) follows immediately. In presence of noise note that the minimal value of the right-hand side of (3.28) without the factor $2C_{\text{err}}$ is given by

$$\begin{aligned}
\inf_{\alpha>0}\left(\frac{\mathbf{err}}{\alpha} + (-\varphi_{\text{add}})^*\left(-\frac{1}{\alpha}\right)\right) &= -\sup_{\bar{\alpha}<0}\left(\mathbf{err}\cdot\bar{\alpha} - (-\varphi_{\text{add}})^*(\bar{\alpha})\right) \\
&= -(-\varphi_{\text{add}})^{**}(\mathbf{err}) \\
&= \varphi_{\text{add}}(\mathbf{err})\,.
\end{aligned}$$

Since in Young's inequality $\langle s^*, s\rangle \leq \Phi(s) + \Phi^*(s^*)$ we have equality if and only if $s^* \in \partial\Phi(s)$, this infimum is attained for any α_{opt} fulfilling

$$\frac{1}{\alpha_{\text{opt}}} \in -\partial(-\varphi_{\text{add}})(\mathbf{err})\,.$$

So for α as in (3.30) we have $\alpha = \tau\alpha_{\text{opt}}$ and hence by (3.27)

$$2C_{\text{err}}\left(-\varphi_{\text{add}}\right)^*\left(-\frac{1}{\alpha}\right)+\frac{\mathbf{err}}{\alpha}\leq 2C_{\text{err}}\max\left\{\tau,\frac{1}{\tau}\right\}\left(\frac{\mathbf{err}}{\alpha_{\text{opt}}}+\left(-\varphi_{\text{add}}\right)^*\left(-\frac{1}{\alpha_{\text{opt}}}\right)\right)$$
$$=2C_{\text{err}}\max\left\{\tau,\frac{1}{\tau}\right\}\varphi_{\text{add}}\left(\mathbf{err}\right).$$

This proves the assertion. ∎

3.2.3 Special cases

Before we comment on the relations between our results and the classical results as well as the more recent results from the literature, we will discuss the optimality of our results and present the special cases of Hölder-type and logarithmic source conditions.

General optimality

We have already proven within Lemma 3.20 that a classical range condition (3.9) together with a tangential cone condition implies a multiplicative variational inequality (3.15). Consider the quadratic Hilbert space case where an upper bound $\delta \geq \left\|g^\dagger - g^{\text{obs}}\right\|_{\mathbb{Y}}$ for the noise is given. Then Mathé and Pereverzev [MP03] have shown that under (3.9) the best possible rate for the error w.r.t. the norm is given by

$$\left\|u_\alpha - u^\dagger\right\|_{\mathbb{X}} = \mathcal{O}\left(\varphi\left(\vartheta^{-1}\left(\delta\right)\right)\right) \tag{3.33}$$

with ϑ as in Definition 3.25 with φ_{mult} replaced by φ and that this rate can be achieved. Since Lemma 3.20 shows that a classical range condition (3.9) together with a tangential cone condition implies a multiplicative variational inequality (3.15), we obtain in the quadratic Hilbert space case from Theorem 3.28 the rate

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) = \mathcal{O}\left(\varphi_{\text{mult}}^2\left(\Theta^{-1}\left(\mathbf{err}\right)\right)\right)$$

where $\mathbf{err}$ is given as in (3.8). Since then $\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) = \left\|u_\alpha - u^\dagger\right\|_{\mathbb{X}}^2$ and $\mathbf{err} = 3\delta^2$, we find from

$$\Theta^{-1}\left(\delta^2\right) = \vartheta^{-1}\left(\delta\right)$$

that we also obtain the optimal rate (3.33) for $\mathcal{S}\left(g; g^{\text{obs}}\right) = \left\|g - g^{\text{obs}}\right\|_{\mathbb{Y}}^2$ and $\mathcal{R}\left(u\right) = \left\|u - u_0\right\|_{\mathbb{X}}^2$ in case of Hilbert norms $\left\|\cdot\right\|_{\mathbb{X}}$ and $\left\|\cdot\right\|_{\mathbb{Y}}$.
In the case of additive variational inequalities, the optimality is much more difficult. We will see in the following that we obtain the optimal rates (3.33) for the case of Hölder-type and logarithmic φ by calculating the obtained rates explicitly. Flemming [Fle11, Sec 12.6] states that the concept of additive variational inequalities (3.11) and classical range conditions (3.9) yield the same rates for linear operators $F = T$ in the sense that every classical range condition (3.9) implies an additive variational inequality (3.11) with the help of Theorem 3.18 and Lemma 3.19 such that the obtained convergence rates coincide. But to our best knowledge, there is no direct proof for this proposition in the most general case without using further smoothness concepts like approximate source conditions and approximate variational inequalities.

Hölder-type source conditions

Let us consider the special case of Hölder-type source conditions where the index function φ is given as in (3.10a). We will start by calculating the functions ϑ, Θ in this case. By definition, it holds $\Theta(t) = t\varphi_\nu^2(t) = t^{1+2\nu}$ for $t \geq 0$. Hence,

$$\begin{aligned} \vartheta(t) &= t^{\frac{1+2\nu}{2}}, \qquad && t \geq 0, \\ \Theta^{-1}(t) &= t^{\frac{1}{1+2\nu}}, \qquad && t \geq 0, \\ \vartheta^{-1}(t) &= t^{\frac{2}{1+2\nu}}, \qquad && t \geq 0. \end{aligned}$$

The optimal rate (3.33) mentioned above is therefore given by

$$\varphi_\nu\left(\vartheta^{-1}(\delta)\right) = \delta^{\frac{2\nu}{2\nu+1}}, \qquad \text{as } \delta \searrow 0. \tag{3.34}$$

Now we are able to formulate the obtained convergence rates:

THEOREM 3.32 (CONVERGENCE RATES FOR HÖLDER-TYPE SOURCE CONDITIONS):

- *Under the Assumptions of Theorem 3.28 with $\varphi_{\text{mult}} = \varphi_\nu$ we obtain the convergence rate*

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) = \mathcal{O}\left(\mathbf{err}^{\frac{2\nu}{1+2\nu}}\right), \qquad \mathbf{err} \searrow 0.$$

- *Under the Assumptions of Theorem 3.30 with $\varphi_{\text{add}} = \varphi_\kappa$ we obtain the convergence rate*

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) = \mathcal{O}\left(\mathbf{err}^{\kappa}\right), \qquad \mathbf{err} \searrow 0.$$

PROOF:
This is obtained by plugging in the functions calculated above. ■

To be more specific, the convergence rates in the classical Hilbert space case are mentioned as corollary:

COROLLARY 3.33 (QUADRATIC HILBERT SPACE CASE):
Assume that $\mathcal{S}(g;\hat{g}) = \|g-\hat{g}\|_{\mathbb{Y}}^2$ and $\mathcal{R}(u) = \|u-u_0\|_{\mathbb{X}}^2$ for Hilbert spaces $\mathbb{X}$ and $\mathbb{Y}$ and assume that $F = T : \mathbb{X} \to \mathbb{Y}$ is a bounded linear operator. Then a range condition (3.9) *with $\varphi = \varphi_\nu$ where $\nu \in \left(0, \frac{1}{2}\right]$ and a known upper bound $\left\|g^\dagger - g^{\text{obs}}\right\|_{\mathbb{Y}} \leq \delta$ imply for a parameter α chosen such that $\alpha = \delta^{\frac{2}{2\nu+1}}$ the following convergence rates:*

$$\left\|u^\dagger - u_\alpha\right\|_{\mathbb{X}} = \mathcal{O}\left(\delta^{\frac{2\nu}{1+2\nu}}\right), \qquad \delta \searrow 0.$$

PROOF:
As already mentioned, Assumption 3.8 is fulfilled with $\mathbf{err} \equiv \delta^2$, $\mathfrak{s} \equiv 0$ and $C_{\text{err}} = 2$. The term $\mathbf{err}$ from (3.8) is hence given by $\mathbf{err} = 3\delta^2$ and the assumed range condition (3.9) with $\varphi = \varphi_\nu$ yields by Lemma 3.20 a multiplicative variational inequality (3.15) with $\varphi_{\text{mult}} = \varphi_\nu$. To apply Theorem 3.28 we note that (3.16) is fulfilled for $\nu \in \left(0, \frac{1}{2}\right]$ and that the parameter choice $\alpha = \delta^{\frac{2}{2\nu+1}}$ coincides with the condition $\mathbf{err} \sim \Theta(\alpha)$ from the theorem. Therefore we obtain order optimal convergence rates (3.34).
By Lemma 3.21 also an additive variational inequality (3.11) with $\varphi_{\text{add}} = \varphi_\kappa$ where $\kappa = \frac{2\nu}{2\nu+1}$ is valid since $\nu \in \left(0, \frac{1}{2}\right]$ and we hence are able to apply Theorem 3.30. Note that φ_{add}^2 is concave. Since $\varphi_{\text{add}}'(t) = \kappa t^{\kappa-1} = \kappa t^{\frac{-1}{2\nu+1}}$ the parameter choice $\alpha \sim \frac{1}{\varphi_{\text{add}}'(\mathbf{err})} = \mathbf{err}^{\frac{1}{2\nu+1}}$ in that theorem also coincides with the parameter choice $\alpha = \delta^{\frac{2}{2\nu+1}}$. Therefore we obtain order optimal convergence rates (3.34). ■

Logarithmic source conditions

Let us consider the special case of logarithmic source conditions where the index function φ is given as in (3.10b). We will start by calculating the functions ϑ, Θ in this case. By definition, it holds $\Theta(t) = t\bar{\varphi}_p^2(t) = t\bar{\varphi}_{2p}(t)$ for $t \geq 0$. The function Θ^{-1} cannot be calculated analytically, but its asymptotic behavior can be calculated by a simple inversion argument:

$$\Theta^{-1}(t) \sim \frac{t}{\bar{\varphi}_{2p}(t)}(1+o(1)), \qquad t \searrow 0.$$

In conclusion we obtain

$$\vartheta(t) = \sqrt{t}\bar{\varphi}_p(t), \qquad t \geq 0,$$

$$\vartheta^{-1}(t) \sim \frac{t^2}{\bar{\varphi}_{2p}(t^2)}(1+o(1)), \qquad t \searrow 0.$$

The optimal rate (3.33) mentioned above is therefore given by

$$\bar{\varphi}_p\left(\vartheta^{-1}(\delta)\right) \sim \bar{\varphi}_p(\delta)(1+o(1)), \qquad \delta \searrow 0. \tag{3.35}$$

Now we are able to formulate the obtained convergence rates:

THEOREM 3.34 (CONVERGENCE RATES FOR LOGARITHMIC SOURCE CONDITIONS):

- *Under the Assumptions of Theorem 3.28 with $\varphi_{\mathrm{mult}} = \bar{\varphi}_p$ we obtain the convergence rate*

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) = \mathcal{O}\left(\bar{\varphi}_{2p}(\mathbf{err})\right), \qquad \mathbf{err} \searrow 0.$$

- *Under the Assumptions of Theorem 3.30 with $\varphi_{\mathrm{add}} = \bar{\varphi}_p$ we obtain the convergence rate*

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) = \mathcal{O}\left(\bar{\varphi}_{p}(\mathbf{err})\right), \qquad \mathbf{err} \searrow 0.$$

PROOF:
This is obtained by plugging in the functions calculated above. ■

To be more specific, the convergence rates in the classical Hilbert space case are mentioned as corollary:

COROLLARY 3.35 (QUADRATIC HILBERT SPACE CASE):
Assume that $\mathcal{S}(g;\hat{g}) = \|g-\hat{g}\|_{\mathbb{Y}}^2$ and $\mathcal{R}(u) = \|u-u_0\|_{\mathbb{X}}^2$ for Hilbert spaces $\mathbb{X}$ and $\mathbb{Y}$ and assume that $F = T : \mathbb{X} \to \mathbb{Y}$ is a bounded linear operator which is compact and injective. Then a range condition (3.9) *with $\varphi = \bar{\varphi}_p$ where $p \in (0,\infty)$ and a known upper bound $\left\|g^\dagger - g^{\mathrm{obs}}\right\|_{\mathbb{Y}} \leq \delta$ where the parameter α is chosen such that $\delta^2 = \alpha\bar{\varphi}_{2p}(\alpha)$ implies the following convergence rates:*

$$\left\|u^\dagger - u_\alpha\right\|_{\mathbb{X}} = \mathcal{O}\left(\bar{\varphi}_p(\delta)\right), \qquad \delta \searrow 0.$$

PROOF:
As already mentioned, Assumption 3.8 is fulfilled with $\mathbf{err} \equiv \delta^2$, $\mathfrak{s} \equiv 0$ and $C_{\mathrm{err}} = 2$. The term $\mathbf{err}$ from (3.8) is hence given by $\mathbf{err} = 3\delta^2$ and the assumed range condition (3.9) with $\varphi = \bar{\varphi}_p$ yields by Lemma 3.20 a multiplicative variational inequality (3.15) with $\varphi_{\mathrm{mult}} = \bar{\varphi}_p$. To apply Theorem 3.28 we note that (3.16) is fulfilled trivially for any p and that our parameter choice coincides with the one from the Theorem. Therefore we obtain order optimal convergence rates (3.35). ■

3.2.4 Related work

REMARK 3.36 (LINK TO CLASSICAL THEORY):
Convergence rates for Tikhonov regularization (3.3) with linear operator $F = T$ under Hölder-type source conditions are well-known and there exists a vast amount of literature on that. We refer to [EHN96, Sec. 5.1] for more information and mention that these results are covered by ours as pointed out in Corollary 3.33.

Convergence rates for nonlinear Tikhonov regularization (i.e. (3.3) with possibly nonlinear F) have first been considered by ENGL, KUNISCH & NEUBAUER [EKN89] and NEUBAUER [Neu89] in 1989. These results have been reformulated and collected in [EHN96], where we will now link our theory to. [EHN96, Thm. 10.4] covers the case of a Hölder-type source condition (3.9) with $\varphi = \varphi_{\frac{1}{2}}$ in combination with a Lipschitz estimate and guarantees the rate $\left\|u_\alpha - u^\dagger\right\| = \mathcal{O}\left(\sqrt{\delta}\right)$. Recalling that $\mathcal{D}_{\mathcal{R}}^{u^*}\left(u, u^\dagger\right) = \left\|u - u^\dagger\right\|_{\mathbb{X}}^2$ in that case and $\mathbf{err} = 3\delta^2$ if $\left\|g^{\text{obs}} - g^\dagger\right\|_{\mathbb{Y}} \leq \delta$ shows that our result from Theorem 3.30 coincides with that rate. Theorem 3.28 also yields $\left\|u_\alpha - u^\dagger\right\| = \mathcal{O}\left(\sqrt{\delta}\right)$ and the assumptions are valid due to Table 3.1 under the same conditions.

The case of logarithmic source conditions in the quadratic Hilbert space case has for example been treated by HOHAGE [Hoh99, Thm. 3.14] where the optimal convergence rate (3.35) is proven for linear operators $T = F$. General source conditions have been studied by HEGLAND [Heg95], where also the optimal rate is derived. Moreover, MATHÉ & PEREVERZEV [MP03] provide a general convergence rates result for linear operators and discuss the optimality.

REMARK 3.37 (TIKHONOV REGULARIZATION WITH GENERAL PENALTY TERM):
In the last decade one came up to replace the penalty $\|u - u_0\|_{\mathbb{X}}^2$ by some arbitrary convex functional $\mathcal{R} : \mathbb{X} \to (-\infty, \infty]$. In this context, one still considers $\mathcal{S}(g; \hat{g}) = \|g - \hat{g}\|_{\mathbb{Y}}^2$ for some Hilbert space $\mathbb{Y}$, but $\mathbb{X}$ is allowed to be a Banach space. Possibly the first choice was **maximum entropy regularization** where $\mathbb{X}$ is some space of non-negative functions on $[a, b]$ and

$$\mathcal{R}(u) = \int_a^b u(t) \log\left(\frac{u(t)}{u^*(t)}\right) \mathrm{d}t$$

where $u^* \in \mathbb{X}$ is some prior information (cf. [EHN96, Sec. 5.3]). In this case, the Bregman distance is given by the Kullback-Leibler divergence (2.12b), i.e.

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u, u^\dagger\right) = \mathbb{KL}\left(u; u^\dagger\right).$$

Since the Kullback-Leibler divergence can be bounded from below by the $\mathbf{L}^2$-norm as we will see in Lemma 4.5 it is not surprising that for this case convergence rates in norm rather than w.r.t. the Bregman distance have been proven.
Another interesting case is

$$\mathcal{R}(u) = \sup\left\{\int_\Omega u \operatorname{div} f \,\mathrm{d}x \mid f \in C_0^1(\Omega), \|f\|_\infty \leq 1\right\}$$

which is referred to as **total variation regularization** (see e.g. [AV94, SGG+08] and the references therein).

The first convergence rates for Tikhonov-type regularization with general convex penalty $\mathcal{R}$ were presented by BURGER & OSHER [BO04] who proved the rate $\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) = \mathcal{O}\left(\delta\right)$ under the source condition

$$\exists\, u^* \in \partial\,\mathcal{R}\left(u^\dagger\right) \qquad \text{such that} \qquad u^* = F^*\omega \qquad \text{for some } \omega \in \mathbb{Y}\,. \tag{3.36}$$

for linear F. Since F^* maps $\mathbb{Y}$ to $\mathbb{X}^*$ and the dual space $\mathbb{X}^*$ of $\mathbb{X}$ differs from $\mathbb{X}$, the condition (3.9) with $\varphi = \varphi_{\frac{1}{2}}$ had to be generalized. It is well-known that in a Hilbert space setting $R\left((F^*F)^{\frac{1}{2}}\right) = R\left(F^*\right)$ and hence (3.9) with $\varphi = \varphi_{\frac{1}{2}}$ has the equivalent formulation $u^\dagger - u_0 = F^*\bar{\omega}$ for some other source element $\bar{\omega} \in \mathbb{Y}$. Moreover, if $\mathcal{R}\left(u\right) = \left\|u - u_0\right\|_{\mathbb{X}}^2$ then $\partial\,\mathcal{R}\left(u^\dagger\right) = u^\dagger - u_0$ which shows that (3.36) is a suitable generalization and can be interpreted as a Hölder-type source condition with $\nu = \frac{1}{2}$. One can easily see, that (3.36) also implies a variational inequality with $\varphi_{\text{add}} = \varphi_{\frac{1}{2}}$ (cf. for example [SGG$^+$08, Prop. 3.35]). Hence, the results of Theorem 3.28 and Theorem 3.30 cover the ones by BURGER & OSHER.

In 2005, RESMERITA [Res05] proved higher order convergence rates for general convex $\mathcal{R}$ and linear F under the source condition

$$\exists\, u^* \in \partial\,\mathcal{R}\left(u^\dagger\right) \qquad \text{such that} \qquad u^* = F^*F\omega \qquad \text{for some } \omega \in \mathbb{X}$$

which corresponds to a Hölder-type source condition with $\nu = 1$. This case is not covered by our work.

As already mentioned before, the book by SCHERZER ET AL. [SGG$^+$08] also uses variational inequalities to provide the rate $\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) = \mathcal{O}\left(\delta\right)$ for nonlinear F under a Hölder-type source condition with $\nu = \frac{1}{2}$ and a generalized Lipschitz condition. Since in that case (3.11) and (3.15) coincide, this result is covered both by Theorem 3.28 and Theorem 3.30.

Note that in general multiplicative and additive variational inequalities to not coincide!

REMARK 3.38 (TIKHONOV REGULARIZATION WITH GENERAL DATA MISFIT TERM):
As already mentioned, the most general case (3.2) has been studied in detail by PÖSCHL [Pös08], where also the rate $\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) = \mathcal{O}\left(\delta\right)$ under the variational inequality

$$\left\langle u^*, u^\dagger - u\right\rangle \leq \beta\mathcal{D}_{\mathcal{R}}^{u^*}\left(u, u^\dagger\right) + \mathcal{S}\left(F\left(u\right); g^\dagger\right) \qquad \text{for all } u \in \mathfrak{B}$$

has been proven. In this publication, the more simple noise level $\mathcal{S}\left(g^\dagger; g^{\text{obs}}\right) \leq \delta$ has been considered, which is sufficient due to the fact that the author assumed $\mathcal{S}$ to fulfill a triangle inequality. Since our noise model generalizes this assumptions, Theorem 3.28 and Theorem 3.30 cover this result.

As a next step, FLEMMING & HOFMANN [FH10] provided convergence rates for more general additive variational source conditions (3.11) with $\varphi_{\text{add}} = \varphi_\nu$, but still $\mathcal{S}$ had to fulfill a triangle-type inequality. The case $\nu = \frac{1}{2}$ is again covered by Theorem 3.28 and Theorem 3.30, the more general cases $\nu \neq \frac{1}{2}$ are connected to our Theorem 3.30 up to differences in the used noise models.

Finally, FLEMMING [Fle10] generalizes the results from [FH10] to general data misfit terms $\mathcal{S}$ which do no longer have to fulfill a triangle-type inequality and general additive variational inequalities (3.11). These results are collected in [Fle11] and differ from ours in the definition of the noise level. We overcome the problem of having no triangle inequality at hand by generalizing the noise model to (3.4), whereas FLEMMING assumes that $\mathcal{S}$ is bounded from below by some data misfit functional fulfilling a triangle-type inequality and uses only the noise level

$$\mathcal{S}\left(g^{\dagger}; g^{\mathrm{obs}}\right) \leq \delta. \tag{3.37}$$

Obviously, our noise level (3.4) with **err** $\equiv \delta$ implies (3.37), but is even more restrictive. On the other hand, we use the true data misfit functional in the source conditions, whereas he has to assume that the variational inequality holds with $\mathcal{S}$ replaced by the lower bound. From that point of view, his source conditions are stronger.
In his PhD thesis [Fle11] he proves moreover that his convergence analysis leads to order optimal convergence rates especially for Hölder-type and logarithmic source conditions in the quadratic Hilbert space case (cf. Section 3.2.3).

REMARK 3.39 (REGULARIZATION WITH SPARSITY CONSTRAINTS):
The choice

$$\mathcal{R}^{\mathrm{sp}}(u) := \sum_{i \in I} w_i \left|\langle u, \phi_i \rangle\right|^q, \qquad u \in \mathbb{X} \tag{3.38}$$

where $\{\phi_i\}_{i \in I}$ is some orthonormal basis of the Hilbert space $\mathbb{X}$ is referred to as **regularization with sparsity constraints**. This type of regularization has been studied somehow independent of generalized Tikhonov regularization, but due to the obvious relations to the publications [BO04, Res05] and other ones, there have always been cross connections to the theory presented above. The functional (3.38) is known to be sparsity enforcing for $q \leq 1$ in the sense that the regularized solutions have only finitely many non-zero components w.r.t. to the orthonormal basis $\{\phi_i\}_{i \in I}$ (cf. [Gra09a]). An iterative approach to compute the minimizers of the Tikhonov functional with $\mathcal{S}(g; \hat{g}) = \|g - \hat{g}\|_{\mathbb{Y}}^2$ and $\mathcal{R} = \mathcal{R}^{\mathrm{sp}}$ where $q = 1$ has been considered in [DDD04], and another method for $q < 1$ has been proposed by ZARZER [Zar09]. Note that the functional (3.38) is convex if and only if $q \geq 1$, but the most interesting case is $q = 0$ or at least $q > 0$ as small as possible. An introductory overview on regularization with sparsity constraints is given in [SGG⁺08].
Since one is interested rather in convergence rates w.r.t. some norm than in rates w.r.t. the Bregman distance, the results [BO04, Res05] have not been considered satisfactory. In 2008 LORENZ [Lor08] and GRASMAIR, HALTMEIER & SCHERZER [GHS08] proved rates of convergence w.r.t. l^1-norm or the l^2-norm of the coefficients, respectively. Later on, these results where improved and generalized in the group of SCHERZER, especially by GRASMAIR:
In [Gra09b] the author studies regularization with $0 < q < 1$ in (3.38) and proves convergence rates w.r.t. the l^1-norm, in [Gra10b, Gra10a] the author generalizes the former results to non-convex $\mathcal{R}$ by introducing a notation of generalized convexity, which somehow applies to the 'sparsest' case $q = 0$ in (3.38) up to some changes. The results presented there are no longer limited to the sparse case, they also apply for Tikhonov-type regularization (3.2) with $\mathcal{S}$ fulfilling a triangle-type inequality and **non-convex** penalty $\mathcal{R}$; convergence rates are shown under a general additive source condition w.r.t. a generalized Bregman distance.
Moreover, in [GHS11] the authors study also necessary conditions for convergence rates in regularization with sparsity constraints, which has not been done in the general setting (3.2).

3.3 A Lepskiĭ-type balancing principle

The aim of this section is to provide a parameter choice rule α, such that α is determined in an adaptive way during the reconstruction procedure. This is of special interest in the case that φ_{add} (and hence the smoothness of $u^\dagger$) is unknown, but one still wants to achieve the optimal convergence rate.
For the whole section we will assume that there exists some uniform upper bound $\overline{\mathbf{err}} > 0$ for the errors defined in (3.8), i.e.

$$\overline{\mathbf{err}} \geq \sup_{\alpha>0} \left(\mathbf{err}\left(F\left(u_\alpha\right)\right) + C_{\text{err}} \mathbf{err}\left(g^\dagger\right)\right). \tag{3.39}$$

The Lepskiĭ principle was developed by LEPSKIĬ [Lep90] in a statistical context. It has first been used for inverse problems with deterministic noise by MATHÉ & PEREVERZEV [MP03]. The way we present it here follows the work of MATHÉ [Mat06], where it is exhibited from a very general point of view.
Let $x \in M$ be some unknown element in a metric space (M, d) which we try to approximate by several other elements $\{x_1, ..., x_m\}$. Assume moreover that the error decomposition

$$d\left(x, x_j\right) \leq \frac{1}{2}\left(\phi\left(j\right) + \psi\left(j\right)\right) \qquad \text{for all} \qquad j = 1, ..., m \tag{3.40}$$

with a **known** non-increasing function ψ and an **unknown** non-decreasing function ϕ fulfilling $\phi(1) \leq \psi(1)$ holds true. We are interested in choosing the parameter $j \in \{1, ..., m\}$ such that the right-hand side attains its minimum over $\{1, ..., m\}$, i.e. mathematically we are looking for

$$\underline{j} := \operatorname*{argmin}_{1 \leq j \leq m} \left[\phi\left(j\right) + \psi\left(j\right)\right].$$

Since ϕ is unknown, the index $\underline{j}$ is also unknown and cannot be determined. The aim is to choose an element $x_{\bar{j}}$ such that the index $\bar{j}$ is easy to compute and that the distance $d\left(x, x_{\bar{j}}\right)$ to the unknown element $x \in M$ is optimal up to some multiplicative constant.
Due to Lepskiĭ we set

$$\bar{j} = \max\left\{j \leq m \mid d\left(x_i, x_j\right) \leq 2\psi\left(i\right) \text{ for all } i < j\right\}. \tag{3.41}$$

Then the following **oracle inequality** can be shown:

LEMMA 3.40 (DETERMINISTIC ORACLE INEQUALITY):
If (3.40) *holds true,* $\phi(1) \leq \psi(1)$ *and there is some* $D < \infty$ *such that*

$$\psi\left(i\right) \leq D\psi\left(i+1\right) \qquad \textit{for all} \qquad 1 \leq i \leq m-1, \tag{3.42}$$

then for $\bar{j}$ *as in* (3.41) *we have*

$$d\left(x, x_{\bar{j}}\right) \leq 3D \min\left\{\phi\left(j\right) + \psi\left(j\right) \mid j \in \{1, ..., m\}\right\}. \tag{3.43}$$

PROOF:
See MATHÉ [Mat06, Cor. 1]. ■

The interpretation of (3.43) is as follows: for regularization methods one is often able to decompose the error into a part which is decreasing as the regularization parameter increases (called the **propagated data noise error**) and an increasing part (called the

approximation error). The function ψ in (3.40) should be seen as the former one, the function ϕ as the latter. The function ψ is in general known, whereas ϕ depends on the unknown 'smoothness' of x (resp. $u^\dagger$) and is hence unknown. Now the choice (3.41) is up to a multiplicative factor of $3D$ as good as the best possible choice.
In the classical theory for Tikhonov regularization (3.3) with linear F and noisy data g^{obs} fulfilling $\left\|g^\dagger - g^{\mathrm{obs}}\right\|_{\mathbb{Y}} \leq \delta$ it is known that an error estimate of the form

$$\left\|u^\dagger - u_\alpha\right\|_{\mathbb{X}} \leq \left\|u^\dagger - \left(F^*F + \alpha I\right)^{-1} F^*Fu^\dagger\right\|_{\mathbb{X}} + \left\|\left(F^*F + \alpha I\right)^{-1} F^*\left(g^\dagger - g^{\mathrm{obs}}\right)\right\|_{\mathbb{X}}$$

holds. The second term on the right-hand side can be bounded by $\frac{\delta}{2\sqrt{\alpha}}$, which is known as the propagated data noise error. The first term on the right-hand side can be bounded by some function only under smoothness conditions on $u^\dagger$. For example under a Hölder-type source condition with smoothness index $\nu \in (0,1)$ (i.e. (3.9) with $\varphi = \varphi_\nu$ as in (3.10a)), it can be shown that

$$\left\|u^\dagger - \left(F^*F + \alpha I\right)^{-1} F^*Fu^\dagger\right\|_{\mathbb{X}} \leq \|\omega\|_{\mathbb{X}}\, \alpha^\nu$$

holds (see e.g. [EHN96, Sec. 5.1]). To apply the Lepskiĭ principle we choose $r > 1$ and define due to MATHÉ [Mat06] $\alpha_1 = \delta^2, \alpha_j := \alpha_1 r^{2j-2}$ for $j = 2, ..., m$ where m is the smallest value such that $\alpha_m \geq 1$. Then the function

$$\psi\left(j\right) := \frac{\delta}{\sqrt{\alpha_j}}$$

is decreasing, fulfills (3.42) with $D = r$ and for the increasing function

$$\phi\left(j\right) = 2\,\|\omega\|_{\mathbb{X}}\, \alpha_j^\nu$$

the error decomposition (3.40) holds true with $x_j := u_{\alpha_j} \in \mathbb{X}$. The condition $\phi(1) \leq \psi(1)$ is fulfilled if $2\,\|\omega\|_{\mathbb{X}} \leq \delta^{-\nu}$. So the deterministic oracle inequality yields

$$\left\|u^\dagger - u_{\alpha_{\bar{j}}}\right\|_{\mathbb{X}} \leq 3r \min\left\{\phi\left(j\right) + \psi\left(j\right) \;|\; j \in \{1, ..., m\}\right\} = \mathcal{O}\left(\delta^{\frac{2\nu}{2\nu+1}}\right)$$

which is known to be of optimal order.
In our general setup we have been able to prove convergence rates only w.r.t. the Bregman distance $\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right)$. It is well-known, that $\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right)$ is in general not a metric, so the Lepskiĭ principle cannot be applied in the general case. Therefore we will assume that there exists some constant $C_{\mathrm{bd}} < \infty$ and some exponent $q > 1$ such that

$$\left\|u - u^\dagger\right\|_{\mathbb{X}}^q \leq C_{\mathrm{bd}} \mathcal{D}_{\mathcal{R}}^{u^*}\left(u, u^\dagger\right) \qquad \text{for all } u \in \mathfrak{B}. \tag{3.44}$$

If for example $\mathbb{X}$ is a Hilbert space and $\mathcal{R}(u) = \|u - u_0\|_{\mathbb{X}}^2$, then $\mathcal{D}_{\mathcal{R}}^{u^*}\left(u, u^\dagger\right) = \left\|u - u^\dagger\right\|_{\mathbb{X}}^2$ and (3.44) is fulfilled with $q = 2$ and $C_{\mathrm{bd}} = 1$. Moreover as we have seen in Lemma 3.13 (3.44) holds true whenever $\mathbb{X}$ is a q-convex Banach space and $\mathcal{R}(u) = \|u\|_{\mathbb{X}}^q$. Note that (3.44) is not limited to the case of q-convex Banach spaces, since also for maximum entropy regularization (3.44) holds true.
The estimate (3.44) ensures that we obtain convergence rates w.r.t. a metric, namely the norm in $\mathbb{X}$. As we have already seen in Theorem 3.30, under Assumption 3.15 also an error decomposition holds true. Let us reformulate this result to apply the Lepskiĭ principle:

COROLLARY 3.41 (ERROR DECOMPOSITION):
Let Assumptions 3.8 and 3.15 hold true with $\beta \in \left[0, \frac{1}{2}\right]$ *and suppose* $\mathcal{R}$ *fulfills* (3.44)*. If moreover* (3.39) *is valid, then for any choice of minimizers* u_α *from* (3.2) *the error decomposition*

$$\left\| u_\alpha - u^\dagger \right\|_{\mathbb{X}} \leq \left(4C_{\mathrm{err}}C_{\mathrm{bd}}\right)^{\frac{1}{q}} \left(-\varphi_{\mathrm{add}}\right)^* \left(-\frac{1}{\alpha}\right)^{\frac{1}{q}} + \left(2C_{\mathrm{bd}}\right)^{\frac{1}{q}} \frac{\overline{\mathbf{err}}^{\frac{1}{q}}}{\alpha^{\frac{1}{q}}} \tag{3.45}$$

is valid for any $\alpha > 0$*.*

PROOF:
Theorem 3.30 implies by (3.28) that

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right) \leq \frac{2C_{\mathrm{err}}}{1-\beta} \left(-\varphi_{\mathrm{add}}\right)^* \left(-\frac{1}{\alpha}\right) + \frac{1}{1-\beta}\frac{\overline{\mathbf{err}}}{\alpha}$$

holds true for all $\alpha > 0$. Now using (3.44) and taking the q-th root implies

$$\left\| u_\alpha - u^\dagger \right\|_{\mathbb{X}} \leq \left(\frac{2C_{\mathrm{bd}}C_{\mathrm{err}}}{1-\beta} \left(-\varphi_{\mathrm{add}}\right)^* \left(-\frac{1}{\alpha}\right) + \frac{C_{\mathrm{bd}}}{1-\beta}\frac{\overline{\mathbf{err}}}{\alpha}\right)^{\frac{1}{q}}.$$

Since the map $t \mapsto t^{\frac{1}{q}}$ is concave this yields by $\beta \leq \frac{1}{2}$ the assertion. ∎

To apply the Lepskiĭ principle to the general case, we proceed as before: For some $r > 1$ we set $\alpha_1 = \overline{\mathbf{err}}, \alpha_j := \alpha_1 r^{2j-2}$ for $j = 2, ..., m$ where m is the smallest value such that $\alpha_m \geq 1$ and denote $x_j := u_{\alpha_j} \in \mathbb{X}$. Then the estimate (3.45) implies (3.40) with the functions

$$\psi(j) := 2\left(2C_{\mathrm{bd}}\right)^{\frac{1}{q}} \left(\frac{\overline{\mathbf{err}}}{\alpha_j}\right)^{\frac{1}{q}},$$

$$\phi(j) := 2\left(4C_{\mathrm{err}}C_{\mathrm{bd}}\right)^{\frac{1}{q}} \left(-\varphi_{\mathrm{add}}\right)^* \left(-\frac{1}{\alpha_j}\right)^{\frac{1}{q}},$$

where the metric is given by the $\mathbb{X}$-norm. Moreover, ϕ is indeed increasing by Remark 3.29, ψ decreasing and fulfills (3.42) with $D = r^{\frac{2}{q}}$. The condition $\phi(1) \leq \psi(1)$ is fulfilled if and only if $1 \geq \left(2C_{\mathrm{err}}\right)^{\frac{1}{q}} \left(\left(-\varphi_{\mathrm{add}}\right)^* \left(-\frac{1}{\overline{\mathbf{err}}}\right)\right)^{\frac{1}{q}}$ which is due to Lemma 3.29 the case if $\overline{\mathbf{err}} > 0$ is small enough. Choosing

$$\bar{j} = \max\left\{ j \leq m \;\middle|\; \left\| u_{\alpha_i} - u_{\alpha_j} \right\|_{\mathbb{X}} \leq 4\left(2C_{\mathrm{bd}}\right)^{\frac{1}{q}} \left(\frac{\overline{\mathbf{err}}}{\alpha_i}\right)^{\frac{1}{q}} \text{ for all } i < j \right\} \tag{3.46}$$

in specification of (3.41) and denoting $\bar{\alpha} := \alpha_{\bar{j}}$, we find by Lemma 3.40 that

$$\left\| u_{\bar{\alpha}} - u^\dagger \right\|_{\mathbb{X}} \leq 3r^{\frac{2}{q}} \min\left\{\phi(j) + \psi(j) \;\middle|\; j \in \{1, ..., m\}\right\}.$$

To calculate the right-hand side of this expression, we need to balance the terms $\left(\frac{\overline{\mathbf{err}}}{\alpha_j}\right)^{\frac{1}{q}}$ and $\left(-\varphi_{\mathrm{add}}\right)^* \left(-\frac{1}{\alpha_j}\right)^{\frac{1}{q}}$. In the proof of Theorem 3.30 this has been done by the use of (3.30), but the typical case is that $\frac{1}{\alpha_j} \notin -\partial\left(-\varphi_{\mathrm{add}}\right)\left(\overline{\mathbf{err}}\right)$ for all $j = 1, ..., m$. The following lemma allows for some sloppiness in the choice of the optimal α_j to achieve optimal rates.

LEMMA 3.42:
Assume in the previously described setup that $\overline{\mathbf{err}}$ *is sufficiently small. Then for*

$$j_* = \min\left\{ j \in \{1, ..., m\} \mid \frac{1}{\alpha_j} \leq -\inf \partial\left(-\varphi_{\text{add}}\right)\left(\overline{\mathbf{err}}\right) \right\} > 1$$

the inequality

$$\varphi_{\text{add}}\left(\sigma\right) - \frac{r^2}{\alpha_{j_*}}\sigma \leq r^2 \varphi_{\text{add}}\left(\overline{\mathbf{err}}\right) - \frac{r^2}{\alpha_{j_*}}\overline{\mathbf{err}} \tag{3.47}$$

holds true for all $\sigma \geq 0$ *and hence*

$$\phi\left(j_* - 1\right) + \psi\left(j_* - 1\right) = \mathcal{O}\left(\varphi_{\text{add}}\left(\overline{\mathbf{err}}\right)^{\frac{1}{q}}\right).$$

PROOF:
As a first step we will show that it holds

$$1 \leq -\inf \partial\left(-\varphi_{\text{add}}\right)\left(\overline{\mathbf{err}}\right), \tag{3.48a}$$

$$1 > \overline{\mathbf{err}} \cdot \left(-\inf \partial\left(-\varphi_{\text{add}}\right)\left(\overline{\mathbf{err}}\right)\right), \tag{3.48b}$$

whenever $\overline{\mathbf{err}}$ is sufficiently small. Therefore note that

$$-\inf \partial\left(-\varphi_{\text{add}}\right)(s) = \delta_- \varphi_{\text{add}}(s) = \inf_{\sigma \in [0,s)} \frac{\varphi_{\text{add}}(s) - \varphi_{\text{add}}(\sigma)}{s - \sigma} \tag{3.49}$$

by Remark 3.31(c). Since

$$\sup_{\sigma \in (s,\infty)} \frac{\varphi_{\text{add}}(\sigma) - \varphi_{\text{add}}(s)}{\sigma - s} \leq \inf_{\sigma \in [0,s)} \frac{\varphi_{\text{add}}(s) - \varphi_{\text{add}}(\sigma)}{s - \sigma}, \tag{3.50}$$

(3.48a) follows directly from the fact that φ_{add}^2 is concave as in Remark 3.31(e). Due to (3.49) we have moreover

$$-\inf \partial\left(-\varphi_{\text{add}}\right)(s) \leq \frac{\varphi_{\text{add}}(s) - \varphi_{\text{add}}(\sigma)}{s - \sigma} \qquad \text{for all} \qquad 0 \leq \sigma < s.$$

With $\sigma = 0$ it follows $s \cdot \left(-\inf \partial\left(-\varphi_{\text{add}}\right)(s)\right) \leq \varphi_{\text{add}}(s)$ for all $s > 0$ which shows (3.48b). The properties (3.48) show that j_* exists and $j_* > 1$.

Now we will show (3.47). By (3.49) and the definition of j_* we have

$$\frac{1}{\alpha_{j_*}} \leq \frac{\varphi_{\text{add}}\left(\overline{\mathbf{err}}\right) - \varphi_{\text{add}}(\sigma)}{\overline{\mathbf{err}} - \sigma} \qquad \text{for all} \qquad 0 \leq \sigma < \overline{\mathbf{err}}$$

and due to (3.50) and the definition of α_j moreover

$$\frac{1}{\alpha_{j_*}} = \frac{1}{r^2}\frac{1}{\alpha_{j_*-1}} \geq \frac{1}{r^2}\frac{\varphi_{\text{add}}(\sigma) - \varphi_{\text{add}}\left(\overline{\mathbf{err}}\right)}{\sigma - \overline{\mathbf{err}}} \qquad \text{for all} \qquad \overline{\mathbf{err}} < \sigma < \infty.$$

This implies

$$\sigma \in \left[0, \overline{\mathbf{err}}\right): \qquad \frac{1}{\alpha_{j_*}}\left(\overline{\mathbf{err}} - \sigma\right) \leq \varphi_{\text{add}}\left(\overline{\mathbf{err}}\right) - \varphi_{\text{add}}(\sigma),$$

$$\sigma \in \left(\overline{\mathbf{err}}, \infty\right): \qquad \frac{1}{\alpha_{j_*}}\left(\overline{\mathbf{err}} - \sigma\right) \leq \frac{1}{r^2}\left(\varphi_{\text{add}}\left(\overline{\mathbf{err}}\right) - \varphi_{\text{add}}(\sigma)\right)$$

which proves by rearrangements (3.47). Using $\frac{r^2}{\alpha_{j^*}} = \frac{1}{\alpha_{j^*-1}}$ we find moreover

$$(-\varphi_{\mathrm{add}})^* \left(-\frac{1}{\alpha_{j_*-1}}\right) = \sup_{\sigma \geq 0} \left(\varphi_{\mathrm{add}}(\sigma) - \frac{1}{\alpha_{j_*-1}}\sigma\right) \leq r^2 \varphi_{\mathrm{add}}\left(\overline{\mathbf{err}}\right) - \frac{1}{\alpha_{j_*-1}}\overline{\mathbf{err}}$$

and hence

$$\begin{aligned}
&\phi(j_* - 1) + \psi(j_* - 1) \\
=&2(2C_{\mathrm{bd}})^{\frac{1}{q}} \left(\frac{\overline{\mathbf{err}}}{\alpha_{j_*-1}}\right)^{\frac{1}{q}} + 2(4C_{\mathrm{err}}C_{\mathrm{bd}})^{\frac{1}{q}} (-\varphi_{\mathrm{add}})^* \left(-\frac{1}{\alpha_{j_*-1}}\right)^{\frac{1}{q}} \\
\leq&2(2C_{\mathrm{bd}})^{\frac{1}{q}} \left(1 - (2C_{\mathrm{err}})^{\frac{1}{q}}\right) \left(\frac{\overline{\mathbf{err}}}{\alpha_{j_*-1}}\right)^{\frac{1}{q}} + 2(4C_{\mathrm{err}}C_{\mathrm{bd}})^{\frac{1}{q}} r^{\frac{2}{q}} \varphi_{\mathrm{add}}\left(\overline{\mathbf{err}}\right)^{\frac{1}{q}}
\end{aligned}$$

which proves by $(2C_{\mathrm{err}})^{\frac{1}{q}} > 1$ the assertion. ■

Let us now collect the result on our a posteriori Lepskiĭ-type stopping rule:

COROLLARY 3.43:
Let Assumptions 3.8 and 3.15 with $\beta \in \left[0, \frac{1}{2}\right]$ *hold true,* (3.39) *be valid and* $\mathcal{R}$ *satisfy* (3.44). *Then the Lepskiĭ-type balancing principle* (3.46) *yields the convergence rate*

$$\left\| u_{\bar{\alpha}} - u^\dagger \right\|_{\mathbb{X}}^q = \mathcal{O}\left(\varphi_{\mathrm{add}}\left(\overline{\mathbf{err}}\right)\right) \tag{3.51}$$

as $\overline{\mathbf{err}} \searrow 0$.

Note that the convergence rate (3.51) is in the quadratic Hilbert space case known to be optimal for Hölder-type or logarithmic source conditions as already discussed before.

CHAPTER
FOUR

TIKHONOV-TYPE REGULARIZATION WITH POISSON DATA

In this chapter we will apply the results on Tikhonov-type regularization (3.2) to the case of Poisson data as discussed in Chapter 2 where the data fidelity term $\mathcal{S}$ is chosen to be a variant of the negative log-likelihood. We will specially point out our results for the case of a linear operator F, because then all assumptions can (and will) easily be checked and verified.
For the whole chapter let the data fidelity terms $\mathcal{S}\left(\cdot; g^\dagger\right)$ and $\mathcal{S}\left(\cdot; g^{\mathrm{obs}}\right)$ w.r.t. exact and noisy data be given by (2.14) respectively with some fixed $e > 0$.

4.1 Bounding the error terms

In this subsection we will bound the error term (3.5). Since G_t is random, no uniform bound can be expected. But as already mentioned in Chapter 2, we will bound (3.5) in probability with the help of a concentration inequality. Unfortunately, the concentration inequality only applies for bounded functions, and hence $e > 0$ in (3.5) is required. The main concentration inequality states that the supremum over the integrals of f against $\frac{1}{t}\,\mathrm{d}G_t - g^\dagger\,\mathrm{d}x$ decays at least as fast as the noise level $\psi(t) = \frac{1}{\sqrt{t}}$ defined in (2.10):

THEOREM 4.1:
Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with Lipschitz boundary, $R \geq 1$ and suppose $s > \frac{d}{2}$. Define

$$B_s(R) := \left\{ \mathfrak{g} \in H^s(\Omega) \;\middle|\; \|\mathfrak{g}\|_{H^s(\Omega)} \leq R \right\}.$$

Then there exist constants $C_{\mathrm{conc}}, C_\rho \geq 1$ depending only on Ω and s such that

$$\mathbf{P}\left(\sup_{\mathfrak{g} \in B_s(R)} \left| \int_\Omega \mathfrak{g} \left(\frac{1}{t}\,\mathrm{d}G_t - g^\dagger\,\mathrm{d}x \right) \right| \leq \rho\psi(t) \right) \geq 1 - \exp\left(-\frac{\rho}{RC_{\mathrm{conc}}} \right) \tag{4.1}$$

for all $t \geq 1$ and $\rho \geq RC_\rho$.

PROOF:
First note that due to $s > \frac{d}{2}$ there exists a continuous embedding $E_\infty : H^s(\Omega) \hookrightarrow \mathbf{L}^\infty(\Omega)$, i.e. it holds

$$\|\mathfrak{g}\|_{\mathbf{L}^\infty(\Omega)} \leq \|E_\infty\| \cdot R \qquad \text{for all} \qquad \mathfrak{g} \in B_s(R).$$

Now choose a countable subset

$$\left\{\mathfrak{g}_j\right\}_{j\in\mathbb{N}} \subset B_s(R)$$

which is dense in $B_s(R)$ w.r.t. the H^s-norm and define

$$Z := \sup_{j\in\mathbb{N}} \left| \int_\Omega \mathfrak{g}_j(x) \left(\frac{1}{t} \mathrm{d}G_t - g^\dagger \,\mathrm{d}x \right) \right| \qquad \text{and} \qquad v_0 := \sup_{j\in\mathbb{N}} \int_\Omega \frac{1}{t} \left(\mathfrak{g}_j(x)\right)^2 g^\dagger \,\mathrm{d}x.$$

Since all $\mathfrak{g}_j$ are elements of $B_s(R)$ we obtain $\left\|\frac{1}{t}\mathfrak{g}_j\right\|_\infty \leq \psi^2(t)\,\|E_\infty\|\,R$ and so REYNAUD-BOURET's concentration inequality (see Lemma 2.6) yields

$$\mathbf{P}\left(Z \geq (1+\varepsilon)\,\mathbf{E}(Z) + \sqrt{12 v_0 \bar\rho} + \kappa(\varepsilon)\,\|E_\infty\|\,R\psi^2(t)\,\bar\rho \right) \leq \exp(-\bar\rho) \tag{4.2}$$

for all $\bar\rho > 0$ where $\kappa(\varepsilon) = 5/4 + 32/\varepsilon$. By $s > \frac{d}{2}$ the set $\left\{\mathfrak{g}_j\right\}_{j\in\mathbb{N}}$ is also dense in $B_s(R)$ w.r.t. the $\mathbf{L}^\infty$-norm and hence we may replace the supremum over $j \in \mathbb{N}$ by the supremum over all $\mathfrak{g} \in B_s(R)$. Furthermore, a simple calculation shows

$$v_0 = \sup_{j\in\mathbb{N}} \int_\Omega \frac{1}{t} \mathfrak{g}_j^2(x)\, g^\dagger \,\mathrm{d}x \leq \psi^2(t)\,\|E_\infty\|^2 R^2 \left\|g^\dagger\right\|_{\mathbf{L}^1(\Omega)}.$$

With $\varepsilon = 1$ this yields from (4.2) the estimate

$$\mathbf{P}\left(\sup_{\mathfrak{g}\in B_s(R)} \left| \int_\Omega \mathfrak{g} \left(\frac{1}{t} \mathrm{d}G_t - g^\dagger \,\mathrm{d}x \right) \right| \geq 2\mathbf{E}(Z) + C_2 R\psi(t)\sqrt{\bar\rho} + C_3 R\psi^2(t)\,\bar\rho \right) \leq \exp(-\bar\rho) \tag{4.3}$$

for all $\bar\rho, t > 0$ with $C_2 := \sqrt{12}\,\|E_\infty\| \sqrt{\|g^\dagger\|_{\mathbf{L}^1(\Omega)}}$ and $C_3 := \left(32 + \frac{5}{4}\right)\|E_\infty\|$. It remains to control $\mathbf{E}(Z)$. For this we need a technical lemma:

LEMMA 4.2:
Let $\Omega \subset \mathbb{R}^d$ be a bounded domain with Lipschitz boundary, $R > 0$ and suppose $s > \frac{d}{2}$. Then there exists a family of real-valued functions $(\phi_j)_{j\in\mathbb{N}}$, numbers $(\gamma_j)_{j\in\mathbb{N}}$ fulfilling

$$\sum_{j=0}^\infty \gamma_j^2 \int_\Omega \phi_j^2 g^\dagger \,\mathrm{d}x < \infty \tag{4.4}$$

and some constant $C_{\mathrm{ext}} > 0$ such that

$$\forall \mathfrak{g} \in B_s(R) \quad \exists\, (\beta_j)_{j\in\mathbb{N}} \text{ s.th. } \mathfrak{g} = \sum_{j=0}^\infty \beta_j \phi_j, \quad \sum_{j=0}^\infty \left(\frac{\beta_j}{\gamma_j}\right)^2 \leq C_{\mathrm{ext}} R. \tag{4.5}$$

PROOF:
Choose some $\kappa > 0$ such that $\Omega \subset\subset [-\kappa,\kappa]^d$, i.e. the bounded domain Ω is compactly included in $[-\kappa,\kappa]^d$. Then there exists a continuous extension operator $E : H^s(\Omega) \hookrightarrow H_0^s\left([-\kappa,\kappa]^d\right)$ (see e.g. [Wlo87, Cor. 5.1]). Moreover, $H_0^s\left([-\kappa,\kappa]^d\right)$ is a bounded subset of $H_{\mathrm{per}}^s\left([-\kappa,\kappa]^d\right)$ and in conclusion we have a bounded extension operator

$$E_{\mathrm{ext}} : H^s(\Omega) \hookrightarrow H_{\mathrm{per}}^s\left([-\kappa,\kappa]^d\right).$$

The norm on $H_{\mathrm{per}}^s\left([-\kappa,\kappa]^d\right)$ can be defined equivalently by meanings of the discrete Fourier transform. For details we refer to [Kre99, Sec. 8.1], the necessary properties are sketched in the following. Let

$$\Phi_k(x) := \begin{cases} \frac{1}{2\kappa} & k = 0, \\ \frac{1}{\sqrt{\kappa}} \sin\left(\frac{\pi k}{\kappa} x\right) & k < 0, \\ \frac{1}{\sqrt{\kappa}} \cos\left(\frac{\pi k}{\kappa} x\right) & k > 0, \end{cases} \qquad x \in [-\kappa, \kappa]$$

and furthermore

$$\phi_j(x) = \prod_{i=1}^{d} \Phi_{j_i}(x_i), \qquad j \in \mathbb{Z}^d, x \in [-\kappa, \kappa]^d.$$

Then we have

$$H^s_{\mathrm{per}}\left([-\kappa,\kappa]^d\right) = \left\{ \mathfrak{g} \in \mathbf{L}^2_{\mathrm{per}}\left([-\kappa,\kappa]^d\right) \;\middle|\; \sum_{j\in\mathbb{Z}} \left(1+|j|^2\right)^s \left|\left\langle \mathfrak{g}, \phi_j \right\rangle_{\mathbf{L}^2\left([-\kappa,\kappa]^d\right)}\right|^2 < \infty \right\}$$

and this space is a Hilbert space equipped with the norm

$$\|\mathfrak{g}\|^2_{H^s_{\mathrm{per}}\left([-\kappa,\kappa]^d\right)} = \sum_{j\in\mathbb{Z}} \left(1+|j|^2\right)^s \left|\left\langle \mathfrak{g}, \phi_j \right\rangle_{\mathbf{L}^2\left([-\kappa,\kappa]^d\right)}\right|^2.$$

We obtain especially

$$E_{\mathrm{ext}}\left(B_s(R)\right) \subset \left\{ \mathfrak{g} \in H^s_{\mathrm{per}}\left([-\kappa,\kappa]^d\right) \;\middle|\; \|\mathfrak{g}\|_{H^s_{\mathrm{per}}\left([-\kappa,\kappa]^d\right)} \leq \tilde{R} \right\}$$

with $\tilde{R} = C_{\mathrm{ext}} R$ for $C_{\mathrm{ext}} := \|E_{\mathrm{ext}}\|$.

Hence, we can choose $\gamma_j = \left(1+|j|^2\right)^{-\frac{s}{2}}$ and $\beta_j = \left\langle \mathfrak{g}, \phi_j \right\rangle$ to ensure (4.5). For (4.4) we calculate

$$\begin{aligned} \sum_{j\in\mathbb{Z}^d} \gamma_j^2 \int_\Omega \phi_j^2 g^\dagger \,\mathrm{d}x &\leq \sum_{j\in\mathbb{Z}^d} \left(1+|j|^2\right)^{-s} \int_\Omega g^\dagger \,\mathrm{d}x \\ &\leq \int_\Omega g^\dagger \,\mathrm{d}x \sum_{j\in\mathbb{Z}^d} \left(1+|j|^2\right)^{-s} \\ &< \infty. \end{aligned}$$

due to $s > \frac{d}{2}$ and hence (4.4) is valid. By renumbering $\mathbb{Z}^d$ to $\mathbb{N}$ we obtain the assertion. ■

With the help of this lemma we can now insert (4.5) and apply Hölder's inequality for sums to find

$$Z \leq \sup_{\sum_{j=0}^{\infty}\left(\frac{\beta_j}{\gamma_j}\right)^2 \leq C_{\text{ext}}R} \left| \sum_{j=0}^{\infty} \frac{\beta_j}{\gamma_j} \int_{\Omega} \gamma_j \phi_j \left(\frac{1}{t}\,\mathrm{d}G_t - g^{\dagger}\,\mathrm{d}x\right) \right|$$

$$\leq \sup_{\sum_{j=0}^{\infty}\left(\frac{\beta_j}{\gamma_j}\right)^2 \leq C_{\text{ext}}R} \sum_{j=0}^{\infty} \left|\frac{\beta_j}{\gamma_j}\right| \left| \int_{\Omega} \gamma_j \phi_j \left(\frac{1}{t}\,\mathrm{d}G_t - g^{\dagger}\,\mathrm{d}x\right) \right|$$

$$\leq \sqrt{C_{\text{ext}}R} \sqrt{\sum_{j=0}^{\infty} \gamma_j^2 \left| \int_{\Omega} \phi_j \left(\frac{1}{t}\,\mathrm{d}G_t - g^{\dagger}\,\mathrm{d}x\right) \right|^2}$$

$$= \sqrt{C_{\text{ext}}R} \sqrt{\sum_{j=0}^{\infty} \gamma_j^2 \left(\int_{\Omega} \phi_j \left(\frac{1}{t}\,\mathrm{d}G_t - g^{\dagger}\,\mathrm{d}x\right) \right)^2}$$

where we used that the functions ϕ_j are real-valued. Hence by Jensen's inequality,

$$\mathbf{E}(Z) \leq \sqrt{C_{\text{ext}}R}\,\mathbf{E}\left(\sqrt{\sum_{j=0}^{\infty} \gamma_j^2 \left(\int_{\Omega} \phi_j \left(\frac{1}{t}\,\mathrm{d}G_t - g^{\dagger}\,\mathrm{d}x\right) \right)^2} \right)$$

$$\leq \sqrt{C_{\text{ext}}R} \sqrt{\mathbf{E}\left(\sum_{j=0}^{\infty} \gamma_j^2 \left(\int_{\Omega} \phi_j \left(\frac{1}{t}\,\mathrm{d}G_t - g^{\dagger}\,\mathrm{d}x\right) \right)^2 \right)}$$

$$= \sqrt{C_{\text{ext}}R} \sqrt{\sum_{j=0}^{\infty} \gamma_j^2 \mathbf{E}\left(\left(\int_{\Omega} \phi_j \left(\frac{1}{t}\,\mathrm{d}G_t - g^{\dagger}\,\mathrm{d}x\right) \right)^2 \right)}. \tag{4.6}$$

Denote $X_j = \int_{\Omega} \phi_j \left(\frac{1}{t}\,\mathrm{d}G_t - g^{\dagger}\,\mathrm{d}x\right)$. Then $\mathbf{E}(X_j) = 0$ and hence $\mathbf{E}\left(X_j^2\right) = \mathbf{V}(X_j)$, which may be calculated by formula (2.9b), i.e.

$$\mathbf{E}\left(\left(\int_{\Omega} \phi_j \left(\frac{1}{t}\,\mathrm{d}G_t - g^{\dagger}\,\mathrm{d}x\right) \right)^2 \right) = \frac{1}{t^2}\mathbf{E}\left(\left(\int_{\Omega} \phi_j \left(\mathrm{d}G_t - tg^{\dagger}\,\mathrm{d}x\right) \right)^2 \right)$$

$$= \psi^2(t) \int_{\Omega} \phi_j^2 g^{\dagger}\,\mathrm{d}x.$$

Plugging this into (4.6), we find

$$\mathbf{E}(Z) \leq \psi(t)\sqrt{C_{\text{ext}}R} \sqrt{\sum_{j=0}^{\infty} \gamma_j^2 \int_{\Omega} \phi_j^2 g^{\dagger}\,\mathrm{d}x} = C_1\sqrt{R}\psi(t)$$

where $C_1 := \sqrt{C_{\text{ext}}}\sqrt{\sum_{j=0}^{\infty} \gamma_j^2 \int_{\Omega} \phi_j^2 g^\dagger \, dx}$ is finite by (4.4). Inserting this estimate into (4.3) and using $\mathbf{P}(A) = 1 - \mathbf{P}(A^c)$ yields

$$\mathbf{P}\left(\sup_{\mathfrak{g} \in B_s(R)} \left| \int_{\Omega} \mathfrak{g} \left(\frac{1}{t} \, dG_t - g^\dagger \, dx \right) \right| \leq C_1 \sqrt{R} \psi(t) + \sqrt{\bar{\rho}} \psi(t) C_2 R + \psi^2(t) C_3 R \bar{\rho} \right) \geq 1 - \exp(-\bar{\rho}) \tag{4.7}$$

with C_2 and C_3 defined after equation (4.3). Now note that for $t, \bar{\rho}, R \geq 1$ we have $\psi^2(t) \leq \psi(t)$ and hence

$$C_1 \sqrt{R} \psi(t) + \sqrt{\bar{\rho}} \psi(t) C_2 R + \psi^2(t) C_3 R \bar{\rho} \leq \bar{\rho} \psi(t) R (C_1 + C_2 + C_3).$$

Therefore we gain from (4.7) that

$$\mathbf{P}\left(\sup_{\mathfrak{g} \in B_s(R)} \left| \int_{\Omega} \mathfrak{g} \left(\frac{1}{t} \, dG_t - g^\dagger \, dx \right) \right| \leq \bar{\rho} \psi(t) R (C_1 + C_2 + C_3) \right) \geq 1 - \exp(-\bar{\rho})$$

whenever $\bar{\rho}, t \geq 1$. Define $C_{\text{conc}} = C_1 + C_2 + C_3$ and $\rho = \bar{\rho} R C_{\text{conc}}$. Then the assertion follows with $C_\rho = \max\{C_{\text{conc}}, 1\}$. ■

Now we are able to bound the error terms **err** as in (3.5) in probability. Remember that **err** was defined as

$$\mathbf{err}(g) = \left| \int_{\Omega} \ln(g + e) \left(\frac{1}{t} \, dG_t - g^\dagger \, dx \right) \right|$$

if $g \geq -\frac{e}{2}$ a.e., which is fulfilled for all $g \in F(u)$, $u \in \mathfrak{B}$ by Assumption 2.7(e).
To apply the concentration inequality to our case, we need to show that

$$\ln(F(u) + e) \in B_s(R) \qquad \text{for all} \qquad u \in \mathfrak{B}$$

with some constant $R \geq 1$.

COROLLARY 4.3:
Let Assumption 2.7 hold true and assume moreover that there exists $s > \frac{d}{2}$ such that $F : \mathfrak{B} \to H^s(\Omega)$ with

$$R := \sup_{u \in \mathfrak{B}} \|F(u)\|_{H^s(\Omega)} < \infty. \tag{4.8}$$

Then there exist $C_{\text{conc}}, C_\rho \geq 1$ depending only on Ω and s such that

$$\mathbf{P}\left(\sup_{u \in \mathfrak{B}} \mathbf{err}(F(u)) \leq \rho \psi(t) \right) \geq 1 - \exp\left(-\frac{\rho}{R \max\{e^{-s}, \ln(R)\} C_{\text{conc}}} \right) \tag{4.9}$$

for all $t \geq 1, \rho \geq R \max\{e^{-s}, \ln(R)\} C_\rho$.

PROOF:
W.l.o.g we may assume that $R \geq 1$. Again due to $s > \frac{d}{2}$ we have $\|F(u)\|_{\mathbf{L}^\infty(\Omega)} \leq \|E_\infty\| \cdot R$ for all $u \in \mathfrak{B}$.
It is well-known that for $\Omega \subset \mathbb{R}^d$ smooth, $g \in H^s(\Omega) \cap \mathbf{L}^\infty(\Omega)$ and $\Phi \in C^s(\mathbb{R})$ where $C^s(\mathbb{R})$ denotes the Hölder space of $\lfloor s \rfloor$-times differentiable functions on $\mathbb{R}$ one has $\Phi \circ$

$g \in H^s(\Omega)$ (see e.g. [BM01]). To apply this result, we first extend the function $x \mapsto \ln(x+e)$ from $(0, \|E_\infty\| \cdot R)$ (since we have $0 \le F(u) \le \|E_\infty\| \cdot R$ a.e.) to a function Φ on the whole real line, which can be done such that $\Phi \in C^s(\mathbb{R})$. Then for any fixed $u \in \mathfrak{B}$ we obtain $\Phi \circ F(u+e) \in H^s(\Omega)$ and since $\Phi_{|_{(e,\|E_\infty\|\cdot R)}}(\cdot) = \ln(\cdot + e)$ and $0 \le F(u) \le \|E_\infty\| \cdot R$ a.e. we have

$$\Phi \circ F(u+e) = \ln(F(u)+e) \quad \text{a.e.}$$

Since all derivatives up to order s of $x \mapsto \ln(x+e)$ and hence of Φ on $(0, \|E_\infty\| \cdot R)$ can be bounded by some constant of order $\max\{e^{-s}, \ln(\|E_\infty\| \cdot R)\}$, the extension and composition procedure described above is bounded, i.e. there exists $C > 0$ independent of u such that

$$\|\ln(F(u)+e)\|_{H^s(\Omega)} \le C \max\left\{e^{-s}, \ln(R)\right\} \|F(u)\|_{H^s(\Omega)} \le C \max\left\{e^{-s}, \ln(R)\right\} R$$

for all $u \in \mathfrak{B}$. Now the assertion follows from Theorem 4.1. ■

REMARK 4.4:
Note that the concentration inequality (4.9) holds true only if

$$\rho > R \max\left\{e^{-s}, \ln(R)\right\} C_\rho$$

which tends to ∞ as $e \searrow 0$. Thus, our convergence analysis will fail if we let $e \searrow 0$ as $t \to \infty$. To fill this gap we would need to accept worse convergence rates by letting $e \searrow 0$ coupled with $\alpha \searrow 0$.

4.2 Convergence rates results

4.2.1 Preliminaries

Before we present and prove the results, let us discuss variational inequalities for the case of $\mathcal{S}$ being the Kullback-Leibler divergence (2.12b) or its shifted version (2.14a). We will use a special property of the Kullback-Leibler divergence, which is based on the results of BORWEIN & LEWIS [BL91]:

LEMMA 4.5 (LOWER BOUND FOR $\mathbb{KL}_e$):
Let $e \ge 0$ and $g, \hat{g} \in \mathbf{L}^\infty(\Omega)$ with $\hat{g} \ge 0$ a.e. and $g \ge -\frac{e}{2}$ a.e. Then it holds

$$\|g - \hat{g}\|^2_{\mathbf{L}^2(\Omega)} \le \left(\frac{4}{3}\|g+e\|_{\mathbf{L}^\infty(\Omega)} + \frac{2}{3}\|\hat{g}+e\|_{\mathbf{L}^\infty(\Omega)}\right) \mathbb{KL}_e(\hat{g}, g). \tag{4.10}$$

PROOF:
Consider the function

$$f(\tau) := \left(\frac{4}{3} + \frac{2}{3}\tau\right)(\tau \ln(\tau) - \tau + 1) - (\tau - 1)^2, \qquad \tau \ge 0.$$

Simple calculations show

$$f'(\tau) = \frac{4}{3}(\tau \ln(\tau) + \ln(\tau)) - \frac{8}{3}\tau + \frac{8}{3},$$
$$f''(\tau) = \frac{4}{3}\left(\ln(\tau) + \frac{1}{\tau} - 1\right),$$
$$f'''(\tau) = \frac{4}{3}\left(\frac{1}{\tau} - \frac{1}{\tau^2}\right)$$

for $\tau > 0$. Since $f'''(\tau) < 0$ for $\tau \in (0,1)$ and $f'''(\tau) > 0$ for $\tau > 1$, it follows that

$$f''(\tau) \geq f''(1) = 4 > 0$$

and hence f is strictly convex on $[0,\infty)$. Therefore we conclude from $f'(1) = 0$ that $f(\tau) \geq f(1) = 0$ for all $\tau \geq 0$. Now let $\tau = \bar{u}/\bar{v}$ for $\bar{u} \in [0,\infty)$ and $\bar{v} \in (0,\infty)$ to obtain

$$0 \leq \left(\frac{4}{3} + \frac{2}{3}\frac{\bar{u}}{\bar{v}}\right)\left(\frac{\bar{u}}{\bar{v}}\ln\left(\frac{\bar{u}}{\bar{v}}\right) - \frac{\bar{u}}{\bar{v}} + 1\right) - \left(\frac{\bar{u}}{\bar{v}} - 1\right)^2$$

and hence by multiplication with $\bar{v}^2$ that

$$(\bar{u} - \bar{v})^2 \leq \left(\frac{4}{3}\bar{v} + \frac{2}{3}\bar{u}\right)\left(\bar{u}\ln\left(\frac{\bar{u}}{\bar{v}}\right) - \bar{u} + \bar{v}\right) \tag{4.11}$$

for all $\bar{u} \in [0,\infty)$ and $\bar{v} \in (0,\infty)$. For the inequality (4.10) observe that for $e = 0$ we are done if $g = 0$ and $\hat{g} > 0$ on a set of positive measure, since then $\mathbb{KL}(\hat{g}, g) = \infty$. For $e = 0$ we can hence restrict to $g > 0$ and $\hat{g} > 0$ a.e.
Now it follows from (4.11) with $\bar{u} = \hat{g} + e$ and $\bar{v} = g + e$ via integration that

$$\begin{aligned}
&\|g - \hat{g}\|^2_{\mathbf{L}^2(\Omega)} \\
\leq &\int_\Omega \left(\frac{4}{3}(g+e) + \frac{2}{3}(\hat{g}+e)\right)\left((\hat{g}+e)\ln\left(\frac{\hat{g}+e}{g+e}\right) - \hat{g} + g\right) \mathrm{d}x \\
\leq &\left(\frac{4}{3}\|g+e\|_{\mathbf{L}^\infty(\Omega)} + \frac{2}{3}\|\hat{g}+e\|_{\mathbf{L}^\infty(\Omega)}\right)\int_\Omega \left((\hat{g}+e)\ln\left(\frac{\hat{g}+e}{g+e}\right) - \hat{g} + g\right) \mathrm{d}x \\
= &\left(\frac{4}{3}\|g+e\|_{\mathbf{L}^\infty(\Omega)} + \frac{2}{3}\|\hat{g}+e\|_{\mathbf{L}^\infty(\Omega)}\right)\mathbb{KL}(\hat{g}+e; g+e) \\
= &\left(\frac{4}{3}\|g+e\|_{\mathbf{L}^\infty(\Omega)} + \frac{2}{3}\|\hat{g}+e\|_{\mathbf{L}^\infty(\Omega)}\right)\mathbb{KL}_e(\hat{g}; g)
\end{aligned}$$

which proves the assertion. ■

COROLLARY 4.6:
Let $F = T$ be a bounded linear operator between the Hilbert space $\mathbb{X}$ and $\mathbf{L}^2(\Omega)$ and let Assumption 2.7 be fulfilled. If the spectral source condition (3.9) holds true with some general φ such that $(\varphi^2)^{-1}$ is convex and (3.16) is fulfilled, then Assumption 3.24 with $\mathcal{R}(u) = \|u - u_0\|^2_{\mathbb{X}}$ and $\mathcal{S} = \mathbb{KL}_e$, $e \geq 0$ holds true with $\varphi_{\text{mult}} = \varphi$. Moreover, a spectral source condition (3.9) with φ implies Assumption 3.15 with φ_{add} as specified in Theorem 3.18 (up to a multiplicative constant as agreed), $\mathcal{R}(u) = \|u - u_0\|^2_{\mathbb{X}}$, $\mathcal{S} = \mathbb{KL}_e$ and arbitrary $\beta \in [0,1)$.

PROOF:
To prove that Assumption 3.24 holds true, we find from Lemma 3.20 that

$$\left|\left\langle u^*, u^\dagger - u\right\rangle\right| \leq 2\|\omega\|_{\mathbb{X}}\left\|u - u^\dagger\right\|_{\mathbb{X}}\varphi\left(\frac{\|T(u-u^\dagger)\|^2_{\mathbf{L}^2(\Omega)}}{\|u-u^\dagger\|^2_{\mathbb{X}}}\right) \qquad \text{for all } u \in \mathbb{X}\,.$$

By assumption, we have $\mathcal{D}^{u^*}_{\mathcal{R}}(u, u^\dagger) = \left\|u - u^\dagger\right\|^2_{\mathbb{X}}$ and now inserting (4.10) and using the concavity of φ we find

$$\left|\left\langle u^*, u^\dagger - u\right\rangle\right| \leq 2C(u)\|\omega\|_{\mathbb{X}}\,\mathcal{D}^{u^*}_{\mathcal{R}}\left(u, u^\dagger\right)^{\frac{1}{2}}\varphi\left(\frac{\mathbb{KL}_e(g^\dagger; g)}{\mathcal{D}^{u^*}_{\mathcal{R}}(u, u^\dagger)}\right)$$

for all $u \in \mathbb{X}$ and hence especially for all $u \in \mathfrak{B}$ where

$$C(u) = \left(\frac{4}{3}\|Tu + e\|_{\mathbf{L}^\infty(\Omega)} + \frac{2}{3}\left\|g^\dagger + e\right\|_{\mathbf{L}^\infty(\Omega)}\right).$$

Since $\mathfrak{B}$ is bounded we have $\sup_{u\in\mathfrak{B}}\|Tu\|_{\mathbf{L}^\infty(\Omega)} < \infty$ by Assumption 2.7 and hence

$$\beta := 2\left(\frac{4}{3}\sup_{u\in\mathfrak{B}}\|Tu + e\|_{\mathbf{L}^\infty(\Omega)} + \frac{2}{3}\left\|g^\dagger + e\right\|_{\mathbf{L}^\infty(\Omega)}\right)\|\omega\|_{\mathbb{X}} < \infty$$

which shows the assertion.
For Assumption 3.15 we use the results of Theorem 3.18 to find an additive variational inequality with $\mathcal{R}(u) = \|u - u_0\|_{\mathbb{X}}^2$, $\mathcal{S}(g;\hat{g}) = \|g - \hat{g}\|_{\mathbf{L}^2(\Omega)}^2$ and the specified φ_{add}. The same argument as above shows the assertion. ■

COROLLARY 4.7:
Let $\mathbb{X}$ be a Hilbert space, let Assumption 2.7 be fulfilled and let the tangential cone condition (3.13) *be valid. If the spectral source condition* (3.9) *holds true with some general φ such that $(\varphi^2)^{-1}$ is convex and* (3.16) *is fulfilled, then Assumption 3.24 with $\mathcal{R}(u) = \|u - u_0\|_{\mathbb{X}}^2$ and $\mathcal{S} = \mathbb{KL}_e$, $e \geq 0$ holds true with $\varphi_{\mathrm{mult}} = \varphi$. Moreover, a spectral source condition* (3.9) *with φ implies Assumption 3.15 with φ_{add} as specified in Theorem 3.18 and $\mathcal{R}(u) = \|u - u_0\|_{\mathbb{X}}^2$ and $\mathcal{S} = \mathbb{KL}_e$.*

PROOF:
This is proven similarly to Corollary 4.6, where the estimate (3.14) which holds true by the tangential cone condition is inserted before applying (4.10). ■

COROLLARY 4.8:
Let $\mathbb{X}$ be a Hilbert space, let Assumption 2.7 be fulfilled and let the Lipschitz condition (3.17) *be valid. Then Assumption 3.15 with $\mathcal{R}(u) = \|u - u_0\|_{\mathbb{X}}^2$ and $\mathcal{S} = \mathbb{KL}_e$ holds true if $\frac{L}{2}\|\omega\| < 1$ with $\varphi_{\mathrm{add}} = \tilde{\beta}\varphi_{\frac{1}{2}}$ for some constant $\tilde{\beta} > 0$ which is small if $\|\omega\|$ is small.*

PROOF:
The result by FLEMMING AND HOFMANN [FH11] already mentioned in Section 3.2.1 yields a variational inequality of the kind

$$\left\langle u^*, u^\dagger - u\right\rangle \leq \beta\mathcal{D}_{\mathcal{R}}^{u^*}\left(u, u^\dagger\right) + \tilde{\beta}\left\|F(u) - g^\dagger\right\|_{\mathbf{L}^2(\Omega)}$$

where the constant $\tilde{\beta}$ is of order $\|\omega\|$. Now inserting (4.10) yields the claim. ■

REMARK 4.9:
Beyond the aforementioned corollaries note that Assumption 3.15 and 3.24 may hold true for better index functions φ_{mult} and φ_{add} respectively than those from the corresponding spectral source conditions, especially if $g^\dagger$ is close to 0 in parts of the domain. On the other hand, if there exists a constant $c > 0$ such that $F(u) \geq c$ a.e., then the Kullback-Leibler divergence can also be bounded from above by the $\mathbf{L}^2$ distance of both functions. Moreover, both assumptions or at least one of them might be fulfilled for other reasons.

4.2.2 Convergence rates for a Poisson process

This section provides the first main result of this thesis, namely convergence rates in expectation for full random data. The result splits into two parts, namely rates under an a priori stopping rule and an a posteriori stopping rule. As source condition we will use only Assumption 3.15 for simplicity, but similar results can be obtained under Assumption 3.24. Afterwards, we will specify the result in case of a linear operator in a separate corollary. Due to the usage of the concentration inequality (4.9) we will always assume that $e > 0$ in this subsection. The following lemma turns out to be quite helpful:

LEMMA 4.10:
Let $(E_k)_{k\in\mathbb{N}}$ be a family of events and $(d_t)_{t>0}$ be a family of random variables such that

(a) $\bigcup_{k\in\mathbb{N}} E_k = \mathcal{E}$ *where* $\mathcal{E}$ *is the event space,*

(b) there exists $c > 0$ *such that* $\mathbf{P}\left(E_k^c\right) \leq \exp\left(-ck\right)$ *for all* $k \in \mathbb{N}$,

(c) $\max_{E_k} d_t \leq C\left(k\right)\Xi\left(t\right)$ *for all* $k \in \mathbb{N}$, $t > 0$ *and*

(d) $\sum_{k=2}^{\infty} \exp\left(-c\left(k-1\right)\right) C\left(k\right) < \infty$.

Then it holds

$$\mathbf{E}\left(d_t\right) = \mathcal{O}\left(\Xi\left(t\right)\right) \qquad \text{as } t \to \infty.$$

PROOF:
Due to (a) and (c) it holds with $E_0 = \emptyset$

$$\begin{aligned}
\mathbf{E}\left(d_t\right) &= \sum_{k=1}^{\infty} \mathbf{P}\left(E_k \setminus E_{k-1}\right) \mathbf{E}\left(d_{n(t)} \,\middle|\, E_k \setminus E_{k-1}\right) \\
&\leq \sum_{k=1}^{\infty} \mathbf{P}\left(E_k \setminus E_{k-1}\right) \max_{E_k} d_{n(t)} \\
&\leq \sum_{k=1}^{\infty} \mathbf{P}\left(E_k \setminus E_{k-1}\right) C\left(k\right)\Xi\left(t\right) \\
&\leq \mathbf{P}\left(E_1\right) C\left(1\right)\Xi\left(t\right) + \sum_{k=2}^{\infty} \mathbf{P}\left(E_{k-1}^c\right) C\left(k\right)\Xi\left(t\right).
\end{aligned}$$

Now inserting (b) and using $\mathbf{P}\left(E_1\right) \leq 1$ we find

$$\mathbf{E}\left(d_t\right) \leq \Xi\left(t\right)\left(C\left(1\right) + \sum_{k=2}^{\infty} \exp\left(-c\left(k-1\right)\right) C\left(k\right)\right).$$

Property (d) yields the assertion. ■

THEOREM 4.11:
Let the Assumptions 2.7 and 3.15 be satisfied and $F : \mathfrak{B} \to H^s\left(\Omega\right)$ *such that* (4.8) *is fulfilled with* $s > \frac{d}{2}$. *If we choose the parameter* α *such that*

$$\frac{1}{\alpha} \in -\partial\left(-\varphi_{\text{add}}\right)\left(\psi\left(t\right)\right) \tag{4.12}$$

we obtain the convergence rate

$$\mathbf{E}\left(\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha, u^\dagger\right)\right) = \mathcal{O}\left(\varphi_{\text{add}}\left(\psi\left(t\right)\right)\right), \qquad t \to \infty. \tag{4.13}$$

PROOF:
Theorem 3.30 implies

$$(1-\beta)\,\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha,u^\dagger\right)+\frac{1}{2\alpha}\,\mathbb{KL}_e\left(g^\dagger;F\left(u_\alpha\right)\right)\le 2\left(-\varphi_{\text{add}}\right)^*\left(-\frac{1}{\alpha}\right)+\frac{\mathbf{err}}{\alpha}$$

where **err** is defined by (3.8). To apply the concentration inequality (4.9) we estimate furthermore

$$(1-\beta)\,\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha,u^\dagger\right)\le 2\left(-\varphi_{\text{add}}\right)^*\left(-\frac{1}{\alpha}\right)+\frac{2}{\alpha}\sup_{u\in\mathfrak{B}}\mathbf{err}\left(F\left(u\right)\right)\tag{4.14}$$

due to the definition (3.8) of **err**. Now choose

$$\rho_k:=R\max\left\{e^{-s},\left|\ln\left(R\right)\right|\right\}C_\rho k,\qquad k\in\mathbb{N}$$

where C_ρ is the constant from Corollary 4.3 and define the events

$$E_k:=\left\{\sup_{u\in\mathfrak{B}}\mathbf{err}\left(F\left(u\right)\right)\le\rho_k\psi\left(t\right)\right\},\quad k\in\mathbb{N}.$$

From (4.9) it is known that

$$\mathbf{P}\left(E_k^c\right)\le\exp\left(-\frac{C_\rho k}{C_{\text{conc}}}\right).$$

Moreover, we find from (4.14) that

$$\begin{aligned}\max_{E_K}\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha,u^\dagger\right)&\le\frac{2}{1-\beta}\left(\left(-\varphi_{\text{add}}\right)^*\left(-\frac{1}{\alpha}\right)+\frac{\rho_k\psi\left(t\right)}{\alpha}\right)\\&\le\frac{2\rho_k}{1-\beta}\left(\left(-\varphi_{\text{add}}\right)^*\left(-\frac{1}{\alpha}\right)+\frac{\psi\left(t\right)}{\alpha}\right).\end{aligned}$$

Again due to Young's inequality we find as in the proof of Theorem 3.30 for $\alpha=\alpha\left(t\right)$ as in (4.12) that

$$\left(-\varphi_{\text{add}}\right)^*\left(-\frac{1}{\alpha}\right)+\frac{\psi\left(t\right)}{\alpha}=\varphi_{\text{add}}\left(\psi\left(t\right)\right)$$

and hence

$$\max_{E_K}\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_\alpha,u^\dagger\right)\le C\left(k\right)\varphi_{\text{add}}\left(\psi\left(t\right)\right)$$

for all $k\in\mathbb{N}$ with $C\left(k\right)=\frac{2\rho_k}{1-\beta}$. For $c=\frac{C_\rho}{C_{\text{conc}}}$ the sum $\sum_{k=2}^{\infty}\exp\left(-c\left(k-1\right)\right)C\left(k\right)$ is convergent (and hence finite) since $C\left(k\right)\sim k$ as $k\to\infty$. This allows us to apply Lemma 4.10 with $\Xi\left(t\right)=\varphi_{\text{add}}\left(\psi\left(t\right)\right)$ and $d_t=\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_{\alpha(t)},u^\dagger\right)$ which proves (4.13). ■

THEOREM 4.12:
Let the Assumptions 2.7 and 3.15 be satisfied, $F:\mathfrak{B}\to H^2\left(\Omega\right)$ *such that* (4.8) *is fulfilled with* $s>\frac{d}{2}$. *Assume moreover that* $\beta\in\left[0,\frac{1}{2}\right]$, (3.44) *is fulfilled and* φ_{add} *is such that*

$$\ln\left(t\right)\cdot\varphi_{\text{add}}\left(\psi\left(t\right)\right)\searrow 0\qquad\textit{as}\qquad t\to\infty.\tag{4.15}$$

Define

$$\rho\left(t\right):=-\tau\ln\left(\psi\left(t\right)\right)=\frac{\tau}{2}\ln\left(t\right),\tag{4.16a}$$

$$\Phi_{\text{noi}}^{\text{det}}\left(n\right):=2\left(4C_{\text{bd}}\right)^{\frac{1}{q}}\left(\frac{\rho\left(t\right)\psi\left(t\right)}{\alpha_n}\right)^{\frac{1}{q}},\tag{4.16b}$$

with a tuning parameter $\tau \geq \frac{1}{2} R \max\{e^{-s}, |\ln(R)|\} C_{\mathrm{conc}}$*, fix* $r > 1$ *and set* $\alpha_j := \rho(t)\psi(t) \cdot r^{2j-2}$ *for* $j = 1, ..., m$ *where* m *is the smallest value such that* $\alpha_m \geq 1$*. Then for the a posteriori Lepskiĭ-type stopping rule*

$$\bar{j} := \min\left\{ j \leq m \mid \left\| u_{\alpha_i} - u_{\alpha_j} \right\|_{\mathbb{X}} \leq 2\Phi_{\mathrm{noi}}^{\mathrm{det}}(i) \text{ for all } i < j \right\} \tag{4.17}$$

we denote $\bar{\alpha} = \alpha_{\bar{j}}$ *and find that for sufficiently large* t *the estimate*

$$\mathbf{E}\left(\left\| u_{\bar{\alpha}} - u^{\dagger} \right\|_{\mathbb{X}}^{q} \right) \leq \bar{C}\left(\rho(t) + \operatorname{diam}(\mathfrak{B})^{q} \right) \varphi(\psi(t))$$

holds true with a constant $\bar{C}$ *independent of* t *and hence*

$$\mathbf{E}\left(\left\| u_{\bar{\alpha}} - u^{\dagger} \right\|_{\mathbb{X}}^{q} \right) = \mathcal{O}\left(\ln(t) \cdot \varphi_{\mathrm{add}}(\psi(t)) \right), \qquad t \to \infty. \tag{4.18}$$

PROOF:
Corollary 3.41 implies under the posed assumptions the error decomposition

$$\left\| u_{\alpha} - u^{\dagger} \right\|_{\mathbb{X}} \leq (4C_{\mathrm{bd}})^{\frac{1}{q}} (-\varphi_{\mathrm{add}})^{*}\left(-\frac{1}{\alpha} \right)^{\frac{1}{q}} + (2C_{\mathrm{bd}})^{\frac{1}{q}} \frac{\overline{\mathbf{err}}^{\frac{1}{q}}}{\alpha^{\frac{1}{q}}}$$

where $\overline{\mathbf{err}} = 2 \sup_{u \in \mathfrak{B}} \mathbf{err}(F(u))$.
Let t be so large that the assumptions from Corollary 4.3 hold true. Now consider the event

$$A_{\rho} = \left\{ \sup_{u \in \mathfrak{B}} \mathbf{err}(F(u)) \leq \rho(t)\psi(t) \right\}$$

which fulfills $\mathbf{P}(A_{\rho}) \geq 1 - \exp(-c\rho(t))$ with $c = \frac{1}{R \max\{e^{-s}, |\ln(R)|\} C_{\mathrm{conc}}}$ by (4.9). Then on A_{ρ} the error decomposition (3.40) is fulfilled with

$$\phi(j) = 2(4C_{\mathrm{bd}})^{\frac{1}{q}} (-\varphi_{\mathrm{add}})^{*}\left(-\frac{1}{\alpha_j} \right)^{\frac{1}{q}},$$

$$\psi = \Phi_{\mathrm{noi}}^{\mathrm{det}}.$$

The functions ϕ and ψ meet the required properties from Section 3.3 and hence we find by Lemma 3.40 that

$$\left\| u_{\bar{\alpha}} - u^{\dagger} \right\|_{\mathbb{X}} \leq 3r^{\frac{2}{q}} \min\left\{ \phi(j) + \psi(j) \mid j \in \{1, ..., m\} \right\} \tag{4.19}$$

on A_{ρ}. By estimating

$$\phi(j) + \psi(j) \leq \rho(t)\, 2(4C_{\mathrm{bd}})^{\frac{1}{q}} \left((-\varphi_{\mathrm{add}})^{*}\left(-\frac{1}{\alpha_j} \right)^{\frac{1}{q}} + \left(\frac{\psi(t)}{\alpha_j} \right)^{\frac{1}{q}} \right),$$

it can be seen in the same way as in Lemma 3.42 where $\overline{\mathbf{err}}$ is replaced by $\psi(t)$, that there exists a constant $\bar{C}$ independent of t such that

$$\mathbf{E}\left(\left\| u_{\bar{\alpha}} - u^{\dagger} \right\|_{\mathbb{X}}^{q} \mid A_{\rho} \right) \leq \bar{C}\rho(t)\, \varphi_{\mathrm{add}}(\psi(t))$$

for all t sufficiently large. Hence

$$\begin{aligned}\mathbf{E}\left(\left\|u_{\bar{\alpha}}-u^{\dagger}\right\|_{\mathbb{X}}^{q}\right)&=\mathbf{P}\left(A_{\rho}\right)\mathbf{E}\left(\left\|u_{\bar{\alpha}}-u^{\dagger}\right\|_{\mathbb{X}}^{q}\mid A_{\rho}\right)+\mathbf{P}\left(A_{\rho}^{c}\right)\mathbf{E}\left(\left\|u_{\bar{\alpha}}-u^{\dagger}\right\|_{\mathbb{X}}^{q}\mid A_{\rho}^{c}\right)\\&\leq\bar{C}\rho\left(t\right)\varphi_{\mathrm{add}}\left(\psi\left(t\right)\right)+\exp\left(-c\rho\left(t\right)\right)\mathrm{diam}\left(\mathfrak{B}\right)^{q}.\end{aligned}\tag{4.20}$$

Due to the definition of $\rho\left(t\right)$ and the choice of τ we can for all sufficiently large t furthermore assume that

$$\exp\left(-c\rho\left(t\right)\right)=\psi\left(t\right)^{c\tau}\leq C\varphi_{\mathrm{add}}\left(\psi\left(t\right)\right).$$

for some constant $C>0$ independent of t since $\varphi_{\mathrm{add}}^{2}$ is concave and $c\tau\geq\frac{1}{2}$. Inserting this into (4.20) yields the claim. ■

REMARK 4.13 (CHOICE OF τ IN (4.16)):
One might argue that the requirement $\tau\geq\frac{1}{2}R\max\left\{e^{-s},\left|\ln\left(R\right)\right|\right\}C_{\mathrm{conc}}$ in the theorem above cannot be ensured in practice since C_{conc} is unknown. But as we have seen in the proof of Theorem 4.1, we may calculate C_{conc} by evaluating the operator norms of the extension and embedding operator from that proof. This can be done numerically in practice. Moreover, the occurring norm of $g^{\dagger}$ can be set to 1 by rescaling t, i.e. it seems natural to assume $\left\|g^{\dagger}\right\|_{\mathbf{L}^{1}(\Omega)}=1$. Thus, the Lepskiĭ-type balancing principle (4.17) can be implemented in practice.

The following corollary specifies the results for the case of a quadratic Penalty term and an operator fulfilling the tangential cone condition, which includes the case of a linear operator.

COROLLARY 4.14:
Let $\mathbb{X}$ be a Hilbert space, Assumption 2.7 be satisfied, $F:\mathfrak{B}\to H^{2}\left(\Omega\right)$ such that (4.8) *holds true and use $\mathcal{R}\left(u\right)=\left\|u-u_{0}\right\|_{\mathbb{X}}^{2}$. Assume moreover that the operator F fulfills the tangential cone condition* (3.13) *(which includes the case that F is linear).*
If a spectral source condition (3.9) *with $\varphi=\bar{\varphi}_{p}$, $p>0$ holds true, then the parameter choice $\alpha=\frac{1}{\bar{\varphi}_{p}'(\psi(t))}$ leads to the convergence rate*

$$\mathbf{E}\left(\left\|u_{\alpha}-u^{\dagger}\right\|_{\mathbb{X}}^{2}\right)=\mathcal{O}\left(\bar{\varphi}_{2p}\left(\frac{1}{\sqrt{t}}\right)\right),\qquad t\to\infty.$$

If the spectral source condition (3.9) *with $\varphi=\varphi_{\nu}$, $\nu\in\left(0,\frac{1}{2}\right]$ is fulfilled, then the parameter choice $\alpha=t^{-\frac{1}{4\nu+2}}$ leads to the convergence rate*

$$\mathbf{E}\left(\left\|u_{\alpha}-u^{\dagger}\right\|_{\mathbb{X}}^{2}\right)=\mathcal{O}\left(t^{\frac{-2\nu}{4\nu+2}}\right),\qquad t\to\infty$$

and the Lepskiĭ-type balancing principle from Theorem 4.12 leads to the convergence rates

$$\mathbf{E}\left(\left\|u_{\alpha_{\bar{j}}}-u^{\dagger}\right\|_{\mathbb{X}}^{2}\right)=\mathcal{O}\left(\ln\left(t\right)\cdot t^{\frac{-2\nu}{4\nu+2}}\right),\qquad t\to\infty.$$

PROOF:
Corollary 4.7 yields the validity of Assumption 3.15 with sufficiently small β and hence we are able to apply Theorems 4.11 and 4.12 to find the assertions. ■

4.3 General convergence

To finish this chapter, we want to comment on the general convergence of (3.2) under Poisson data. Assume therefore that $\mathbb{X}$ is a Hilbert space and $\mathcal{R}$ is given by the squared norm in $\mathbb{X}$. If F is a bounded linear operator, then the assumptions of the following remark even simplify since the tangential cone condition is fulfilled with $\bar{\eta} = 0$.

REMARK 4.15 (REGULARIZATION PROPERTIES):
Let $F : \mathbb{X} \to \mathbb{Y}$ be an operator fulfilling Assumption 2.7 and (3.13). As above we find the existence of some index function φ such that (3.9) is fulfilled. The function φ can be chosen such that φ^2 is concave by possibly changing ω. Then Corollary (4.7) implies the validity of Assumption 3.15. Thus for proper chosen $\alpha = \alpha(t)$ we have $\left\| u_\alpha - u^\dagger \right\|_{\mathbb{X}} \to 0$ as $t \to \infty$ for the regularized solutions u_α gained by (3.2) with Poisson data G_t as described in Chapter 2.

Moreover if the spectral source condition is strong enough (i.e. such that (4.15) is fulfilled), then the a posteriori choice of α given by the Lepskiĭ-type balancing principle (4.17) yields convergence $\left\| u_\alpha - u^\dagger \right\|_{\mathbb{X}} \to 0$ as $t \to \infty$.

CHAPTER

FIVE

THE ITERATIVELY REGULARIZED GAUSS-NEWTON METHOD

In this chapter we will recall the known theory for the iteratively regularized Gauss-Newton method which we will generalize and apply to the case of Poisson data in Chapter 6 and 7 respectively. Assume for the whole chapter that an ill-posed problem

$$F(u) = g, \tag{5.1}$$

with a nonlinear, Fréchet-differentiable operator F is given where only noisy data g^{obs} fulfilling

$$\left\| g^{\dagger} - g^{\text{obs}} \right\|_{\mathbb{Y}} \leq \delta \tag{5.2}$$

for a known upper bound $\delta > 0$ are available. The iteratively regularized Gauss-Newton method consists in calculating

$$u_{n+1} \in \operatorname*{argmin}_{u \in \mathfrak{B}} \left[\left\| F(u_n) + F'[u_n](u - u_n) - g^{\text{obs}} \right\|_{\mathbb{Y}}^{r} + \alpha_n \left\| u - u_0 \right\|_{\mathbb{X}}^{p} \right] \tag{5.3a}$$

where $u_0 \in \mathfrak{B}$ is some initial guess and $r, p \in (1, \infty)$. The regularization parameters $(\alpha_n)_{n \in \mathbb{N}}$ are usually chosen in a way such that

$$\alpha_0 \leq 1, \qquad \alpha_n \searrow 0, \qquad 1 \leq \frac{\alpha_n}{\alpha_{n+1}} \leq C_{\text{dec}} \qquad \text{for all} \qquad n \in \mathbb{N}. \tag{5.3b}$$

As already mentioned in Chapter 3, Tikhonov-type regularization for nonlinear operators F has the disadvantage that the minimizer of the Tikhonov functional (and hence the regularized solution) is difficult to find since the problem is non-convex. For the method (5.3) we have to solve a convex minimization problem in every iteration, which seems to be easier, but possibly time consuming if not only a few iterations are needed. Nevertheless we expect fast convergence, since the method (5.3) is of Newton type.

This chapter will start with a short motivation of (5.3) and a comparison to other iterative methods. Moreover, the regularization properties of (5.3) are discussed and linked to the classical case of Hilbert spaces $\mathbb{X}$ and $\mathbb{Y}$ and $r = p = 2$. Finally, we will give some results on convergence rates, which will be covered by the more general results in the next chapter. The results from this chapter are in core covered by the work of KALTENBACHER ET AL. [KNS08, KSS09, KH10].

5.1 Idea and connections to other iterative methods

Nonlinear operator equations of the kind (5.1) are usually solved by linearization. Replacing the exact operator evaluation $F(u)$ by the first order Taylor expansion $F(u_n) + F'[u_n](u - u_n)$ around some former approximation u_n of the exact solution $u^\dagger$, it remains to solve the linearized equation

$$F(u_n) + F'[u_n](u - u_n) = g^{\mathrm{obs}} \tag{5.4}$$

for the next approximation $u = u_{n+1}$. If the operator equation (5.1) is ill-posed, then also (5.4) is in general ill-posed. Therefore, regularization is needed.
The method (5.3) is obtained by applying Tikhonov regularization to the linearized equation (5.4) in every iteration. It has first been proposed in the quadratic Hilbert space case by BAKUSHINSKIĬ [Bak92]. Since the linearized equation (5.4) could also be formulated as

$$F(u_n) + F'[u_n](h) = g^{\mathrm{obs}} \tag{5.5}$$

where h is considered as an **update** and the next iterate is gained via $u_{n+1} = u_n + h$, also methods of the form

$$h \in \operatorname*{argmin}_{u \in \mathfrak{B}} \left[\left\| F(u_n) + F'[u_n]h - g^{\mathrm{obs}} \right\|_{\mathbb{Y}}^{r} + \alpha_n \|h\|_{\mathbb{X}}^{p} \right], \tag{5.6a}$$

$$u_{n+1} = u_n + h \tag{5.6b}$$

seem suitable. The method (5.6) is known as the **Levenberg-Marquard method** and is motivated by applying Tikhonov regularization to (5.5) with initial guess $h_0 = 0$. Opposed to (5.3) it enforces only smoothness of the update h, whereas the iteratively regularized Gauss-Newton method guarantees smoothness of the next iterate u_{n+1}. One would expect (5.3) to be more stable against rounding errors. Moreover, it turned out that (5.3) is easier to analyze. For a recent result on the analysis of the Levenberg-Marquard method in Hilbert spaces with $r = p = 2$ providing order-optimal convergence we refer to [Han10].
We should mention also other iterative methods for nonlinear operator equations of gradient type. Since the solution of (5.1) is (under exact data) also minimum of the functional

$$J(u) = \left\| F(u) - g^{\mathrm{obs}} \right\|_{\mathbb{Y}}^{r},$$

another natural approach is to find this minimum by a steepest decent method. This means that one iteratively obtains approximations by

$$u_{n+1} = u_n - \mu_n \nabla J(u_n) \tag{5.7}$$

where $\nabla J(u_n)$ denotes the gradient of J at u_n and μ_n is a step-length parameter. For a Hilbert space $\mathbb{Y}$ and $r = 2$ it holds $\nabla J(u_n) = F'[u_n]^* \left(F(u_n) - g^{\mathrm{obs}}\right)$. The method (5.7) is known as **Landweber iteration** and regularization is obtained by choosing an appropriate stopping parameter (for example by the discrepancy principle). Landweber iteration is easy to implement, but it must be said however that it often converges slowly. For a convergence analysis in Hilbert spaces we refer to [KNS08, Cpt. 2]. A generalized and accelerated version of (5.7) has been analyzed in Banach spaces in [HK10a, HK10b].

5.2 Regularization properties

For iterative methods of the form (5.3) one requires slightly different properties than for Tikhonov-type regularization. Namely we are interested in the following properties:

(a) Well-definedness of the nth iterate, i.e. for any $n \in \mathbb{N}$ and any $g^{\text{obs}} \in \mathbb{Y}$ there exists at least one minimizer of (5.3a).

(b) Convergence for exact data, i.e. for $g^{\text{obs}} = g^{\dagger}$ there exists either some $N \in \mathbb{N}$ such that $u_N = u^{\dagger}$ or the regularized solutions u_n converge to $u^{\dagger}$.

(c) Convergence for noisy data, i.e. the regularized solutions u_N converge to $u^{\dagger}$ as the noise level tends to 0 and the stopping index N is chosen in an appropriate manner.

Item (a) requires two things, namely that $u_n \in \mathfrak{B} \subset D(F)$, which is guaranteed by Assumption 2 and the side-condition $u \in \mathfrak{B}$, and the existence of a minimizer of (5.3a). This can be achieved in the same way as for Tikhonov-type regularization in Theorem 3.3 in the general case and we will assume in the following that (a) holds true. Features (b) and (c) guarantee that the regularized solutions u_n indeed approximate solutions of the original problem (5.1).

5.2.1 Comments on the quadratic Hilbert space case

Item (b) was first investigated by BAKUSHINSKIĬ [Bak92], who proved local convergence for exact data under a nonlinearity and a source condition. He even proved a rate of convergence under that source condition. For the quadratic Hilbert space case a convergence is of (5.3) is given in [KNS08, Sec. 4.2] under a suitable nonlinearity condition and a Hölder-type source condition including the case $\nu = 0$. This corresponds to the case of no source condition, which is of special interest for the items (b) and (c), since it provides the required convergence for all $u^{\dagger}$. The authors there considered the case of an **a priori** stopping criterion $N_* = N_*(\delta)$ chosen such that

$$\begin{cases} \eta \alpha_{N_*}^{\nu+\frac{1}{2}} \leq \delta < \eta \alpha_n^{\nu+\frac{1}{2}} & \text{for all} \quad 0 \leq n < N_* \quad \text{if} \quad 0 < \nu \leq 1, \\ \eta \geq \delta \alpha_{N_*}^{-\frac{1}{2}} \quad \text{and} \quad N_*(\delta) \to \infty \quad \text{as} \quad \delta \searrow 0 & \text{if} \quad \nu = 0 \end{cases}$$

for some $\eta > 0$. Moreover they proved convergence (and also convergence rates) for an **a posteriori** stopping rule by the discrepancy principle, i.e. $N_* = N_*(\delta)$ is chosen such that

$$\left\| g^{\text{obs}} - F(u_{N_*}) \right\|_{\mathbb{Y}} \leq \tau\delta < \left\| g^{\text{obs}} - F(u_n) \right\|_{\mathbb{Y}} \qquad \text{for all} \quad 0 \leq n < N_* \tag{5.8}$$

where $\tau > 1$ is some sufficiently large tuning parameter.
We will not repeat these results in detail here, since we are interested in the more general case (5.3a) where $\mathbb{X}$ and $\mathbb{Y}$ do not need to be Hilbert spaces and $p \neq 2$ and $r \neq 2$ are allowed.

5.2.2 The case of general norm powers

The case of general norm powers has been treated by KALTENBACHER ET AL. in [KSS09], where the following results are taken from. The authors provide convergence of (5.3a) where also the regularization parameters $(\alpha_n)_{n\in\mathbb{N}}$ are chosen **a posteriori**. Since the geometry of Banach spaces is much more difficult than those of Hilbert spaces, they needed

to restrict the class of Banach spaces $\mathbb{X}$. Thus assume in the following that $\mathbb{X}$ is reflexive and strictly convex. If $\mathbb{X}$ is moreover p-convex, then convergence rates w.r.t. the Bregman distance imply rates w.r.t. the norm by Lemma 3.13.
Note that for reflexive and uniformly smooth $\mathbb{X}$ the subdifferential $\partial \|\cdot\|_{\mathbb{X}}^{p}\left(u^{\dagger}\right)$ is single valued and the only subgradient can be expressed with the help of the duality mapping $J_p : \mathbb{X} \to \mathbb{X}^*$ (cf. [SGG+08, Sec. 10.3]) characterized by

$$u^* \in J_p(u) \qquad \Leftrightarrow \qquad \langle u^*, u\rangle = \|u\|_{\mathbb{X}}^{p} \qquad \text{and} \qquad \|u^*\|_{\mathbb{X}^*} = \|u\|_{\mathbb{X}}^{p-1}.$$

Since we use only norm powers $\|\cdot\|_{\mathbb{X}}^{p}$ as penalty term, we will now specify the notation of the Bregman distance for this chapter. Here we denote for a uniformly smooth Banach space $\mathbb{X}$ by

$$\mathcal{D}_p^{u^*}\left(u, u^{\dagger}\right) := \mathcal{D}_{\|\cdot\|_{\mathbb{X}}^{p}}^{u^*}\left(u, u^{\dagger}\right) = \frac{1}{p}\|u\|_{\mathbb{X}}^{p} - \frac{1}{p}\left\|u^{\dagger}\right\|_{\mathbb{X}}^{p} - \left\langle J_p\left(u^{\dagger}\right), u - u^{\dagger}\right\rangle$$

the Bregman distance w.r.t. the p-th power of the norm in $\mathbb{X}$.
Now let us continue with the a posteriori choice of the regularization parameters proposed by KALTENBACHER ET AL. [KSS09]. It has been shown (cf. [KSS09, Lem. 1]) that the following parameter choice rule is meaningful: Chose α_n such that

$$\underline{\theta}\left\|F(u_n) - g^{\text{obs}}\right\|_{|Y} \leq \left\|F'[u_n](u_{n+1}(\alpha_n) - u_n) + F(u_n) - g^{\text{obs}}\right\|_{Y} \leq \bar{\theta}\left\|F(u_n) - g^{\text{obs}}\right\|_{\mathbb{Y}} \tag{5.9}$$

where $u_{n+1}(\alpha_n)$ denotes the solution to (5.3a) with regularization parameter α_n and $0 < \underline{\theta} < \bar{\theta} < 1$ are some parameters. Together with the discrepancy principle as stopping rule KALTENBACHER ET AL. [KSS09] obtained the following result:

THEOREM 5.1:
Let F be a weakly sequential closed (see e.g. [KSS09, eq. (11)]) nonlinear operator satisfying a tangential cone condition (3.13) *with a parameter $\bar{\eta}$ such that*

$$\bar{\eta} < \underline{\theta} < \bar{\theta} < 1$$

and chose the tuning parameter τ in the discrepancy principle (5.8) *so large that*

$$\bar{\eta} + \frac{1+\bar{\eta}}{\tau} \leq \underline{\theta}, \qquad \bar{\eta} < \frac{1-\bar{\theta}}{2}.$$

Moreover assume that either $F'[u] : \mathbb{X} \to \mathbb{Y}$ is weakly closed for all $u \in \mathfrak{B}$ and $\mathbb{Y}$ reflexive or $\mathfrak{B}$ is weakly closed.
Then for all $n \leq N_*(\delta) - 1$ with $N_*(\delta)$ according to (5.8) *the iterates*

$$u_{n+1} := \begin{cases} u_{n+1}(\alpha_n) & \text{if } \left\|F'[u_n](u_0 - u_n) + F(u_n) - g^{\text{obs}}\right\|_{\mathbb{Y}} \geq \bar{\theta}\left\|F(u_n) - g^{\text{obs}}\right\|_{\mathbb{Y}}, \\ u_0 & \text{else} \end{cases} \tag{5.10}$$

are well-defined and $u_{N_*(\delta)}$ converges strongly to $u^{\dagger}$ as $\delta \searrow 0$.

PROOF:
See [KSS09, Thm. 3]. Note that by Assumption 2 the solution $u^{\dagger}$ to (5.1) is unique. ∎

5.3 Convergence rates

As already mentioned before, BAKUSHINSKIĬ [Bak92] proved already rates of convergence in the noise free case for (5.3) with $r = p = 2$ and Hilbert norms $\|\cdot\|_{\mathbb{X}}$ and $\|\cdot\|_{\mathbb{Y}}$. More general convergence rate results for this classical case where obtained by BLASCHKE (now KALTENBACHER), NEUBAUER AND SCHERZER [BNS97] for Hölder-type source conditions (3.9) with $\varphi = \varphi_\nu$ and $\nu \in \left(0, \frac{1}{2}\right)$ as well as HOHAGE [Hoh97] for logarithmic convergence rates (3.9) $\varphi = \varphi_p$ and $p > 0$. Some of these results have been collected in the monograph [KNS08]. We will now present the most general result for (5.3) known so far obtained by KALTENBACHER & HOFMANN [KH10] which includes the aforementioned results.
We have already seen in Theorem 5.1 that for convergence (and hence also convergence rates) the degree of nonlinearity of F has to be restricted. This is necessary due to the fact that one has to compare the first order Taylor expansions used in the method (5.3a) with the true operator evaluation $F(u)$. In the case of general source conditions this is usually done by the already known tangential cone condition (3.13), but this condition can be relaxed for Hölder-type source conditions. The following nonlinearity condition can be seen as an attempt to merge both assumptions:

ASSUMPTION 5.2 (NONLINEARITY CONDITION OF [KH10]):
Assume that there exists some constant $K > 0$ such that

$$\sup_{v,w\in\mathbb{X},\ u^\dagger+v\in\mathfrak{B},\ u^\dagger+w\in\mathfrak{B}} \frac{\left\|\left(F'\left[u^\dagger+v\right]-F'\left[u^\dagger\right]\right)w\right\|_{\mathbb{Y}}}{\left\|F'\left[u^\dagger\right]w\right\|_{\mathbb{Y}}^{c_1}\mathcal{D}_p^{u^*}\left(u^\dagger+w,u^\dagger\right)^{c_2}\left\|F'\left[u^\dagger\right]v\right\|_{\mathbb{Y}}^{c_3}\mathcal{D}_p^{u^*}\left(u^\dagger+v,u^\dagger\right)^{c_4}} \leq K \tag{5.11}$$

with exponents $c_1, c_2, c_3, c_4 \geq 0$.

If $c_1 = c_3 = 0$ and $c_2 = c_4 = \frac{1}{2}$, (5.11) can be seen as the usual Lipschitz condition on F' in terms of the Bregman distance in $\mathbb{X}$.
If $c_1 = c_3 = \frac{1}{2}$ and $c_2 = c_4 = 0$, then (5.11) leads to

$$\left\|\left(F'[u+v]-F'[u]\right)w\right\|_{\mathbb{Y}} \leq K\left\|F'\left[u^\dagger\right]v\right\|_{\mathbb{Y}}^{\frac{1}{2}}\left\|F'\left[u^\dagger\right]w\right\|_{\mathbb{Y}}^{\frac{1}{2}},$$

and hence we have

$$\begin{aligned}\left\|F\left(u^\dagger+w\right)-F\left(u^\dagger\right)-F'\left[u^\dagger\right]w\right\|_{\mathbb{Y}} &\leq \int_0^1\left\|\left(F'\left[u^\dagger+tw\right]-F'\left[u^\dagger\right]\right)w\right\|_{\mathbb{Y}}\,\mathrm{d}t\\ &\leq \frac{K}{2}\left\|F'\left[u^\dagger\right]w\right\|_{\mathbb{Y}}\end{aligned}$$

and on the other hand

$$\begin{aligned}\left\|F\left(u^\dagger+w\right)-F\left(u^\dagger\right)\right\|_{\mathbb{Y}} &= \left\|\int_0^1 F'\left[u^\dagger+tw\right]w\,\mathrm{d}t\right\|_{\mathbb{Y}}\\ &\geq \left\|F'\left[u^\dagger\right]w\right\|_{\mathbb{Y}} - \int_0^1\left\|\left(F'\left[u^\dagger+tw\right]-F'\left[u^\dagger\right]\right)w\right\|_{\mathbb{Y}}\,\mathrm{d}t\\ &\geq \left\|F'\left[u^\dagger\right]w\right\|_{\mathbb{Y}} - \frac{K}{2}\left\|F'\left[u^\dagger\right]w\right\|_{\mathbb{Y}} = \frac{2-K}{2}\left\|F'\left[u^\dagger\right]w\right\|_{\mathbb{Y}}.\end{aligned}$$

Putting both inequalities together shows that a local tangential cone condition (i.e. (3.13) for $u = u^\dagger$ fixed) with $\eta = \frac{K}{2-K}$ is valid. It can be seen moreover that (3.13) with $u = u^\dagger$

fixed does not imply (5.11) in general. Therefore, (5.11) is a stronger condition in this case.
The cases where $c_1, c_2, c_3, c_4 > 0$ represent nonlinearity conditions which are stronger than the Lipschitz condition, but still weaker than the tangential cone condition. They are limited to the case of Hölder-type source conditions which will be specified in Assumption 5.4. But before that we have to specify the used source condition:

ASSUMPTION 5.3 (MULTIPLICATIVE VARIATIONAL INEQUALITY OF [KH10]):
There exists $u^* \in \partial \mathcal{R}\left(u^\dagger\right) \subset \mathbb{X}'$, $\beta \geq 0$ and a concave index function $\varphi_{\text{mult}} : (0,\infty) \to (0,\infty)$ such that

$$\left|\left\langle u^*, u^\dagger - u\right\rangle\right| \leq \beta \mathcal{D}_p^{u^*}\left(u, u^\dagger\right)^{\frac{1}{2}} \varphi_{\text{mult}}\left(\frac{\left\|F'\left[u^\dagger\right]\left(u - u^\dagger\right)\right\|_{\mathbb{Y}}^2}{\mathcal{D}_p^{u^*}\left(u, u^\dagger\right)}\right) \qquad \text{for all } u \in \mathfrak{B}. \quad (5.12)$$

Moreover assume that φ_{mult} is such that

$$t \mapsto \frac{\varphi_{\text{mult}}(t)}{\sqrt{t}} \qquad \text{is monotonically decreasing.}$$

If is easy to see that (5.12) is still motivated by the Hilbert space setting. Using the result of Lemma 3.20 and generalizing the $\mathbb{X}$-norm by the Bregman distance $\mathcal{D}_p^{u^*}\left(u, u^\dagger\right)$, it is clear that the condition (5.12) belongs to a range condition of the form (3.9) with the same index function.
Comparing (5.12) with (3.15) shows that the only difference is the absence of derivatives in (3.15), which can be gained from (5.12) by applying a usual tangential cone condition (3.13). This has already been used to motivate (3.15) in Section 3.2.1.
Now we are able to pose the interplay condition which connects the source condition (5.12) to the nonlinearity condition (5.11):

ASSUMPTION 5.4 (INTERPLAY CONDITION OF [KH10]):

(a) In case of $\varphi_{\text{mult}} = \varphi_{\frac{\nu}{2}}$ in (5.12) *with $\nu \in (0,1)$ let the exponents $c_1, c_2, c_3, c_4 \geq 0$ from* (5.11) *satisfy*

$$c_1 + c_2 \frac{2\nu}{\nu+1} \geq \frac{1}{2}, \qquad c_3 + c_4 \frac{2\nu}{\nu+1} \geq \frac{1}{2}$$

as well as

$$c_1 + c_2 r \geq \frac{1}{2}, \qquad c_3 + c_4 r \geq \frac{1}{2}$$

where either the latter two inequalities are proper or K in (5.11) *is sufficiently small. Moreover assume that β is sufficiently small.*

(b) In case of a general source condition (5.12) *with arbitrary φ_{mult} assume that*

$$c_1 = c_3 = \frac{1}{2}, \qquad c_2 = c_4 = 0$$

and that K in (5.11) *is sufficiently small. Moreover, denote*

$$\varphi_r(t) := t^{r-2} \vartheta^{-1}(t), \qquad \vartheta(t) := \varphi_{\text{mult}}(t)\sqrt{t}$$

and assume that

$$\varphi_{\text{mult}}\left(\hat{C}_r t\right) \leq \hat{C}_{\varphi_{\text{mult}}} \varphi_{\text{mult}}(t) \qquad \text{for all } 0 \leq t \leq \hat{t}$$

with some in [KH10, eq. (30)] specified constants $\hat{C}_r$, $\hat{C}_{\varphi_{\text{mult}}}$ and $\hat{t}$.

Note that item (a) allows for a Lipschitz-type nonlinearity condition in the limit case of the most often used source condition $\varphi_{\mathrm{mult}} = \varphi_{\frac{1}{2}}$, and for mixed nonlinearity conditions if $\varphi_{\mathrm{mult}} = \varphi_{\frac{\nu}{2}}$ with $\nu < 1$. In the case of general source conditions, item (b) implies a tangential cone condition in accordance with the fact that no convergence rates result under weaker nonlinearity conditions with general φ_{mult} is known.
We are now able to present the main result on (5.3) with **a priori** parameter choices:

THEOREM 5.5:
Let the Assumptions 5.2, 5.3 and 5.4 hold true and let $\mathcal{D}_p^{u^*}\left(u_0, u^\dagger\right)$ be sufficiently small. Then the iterates $(u_n)_{n\in\mathbb{N}}$ obtained from (5.3) *with exact data $g^{\mathrm{obs}} = g^\dagger$ fulfill*

$$\mathcal{D}_p^{u^*}\left(u_{n+1}, u^\dagger\right) = \begin{cases} \mathcal{O}\left(\alpha_n^{\frac{2\nu}{r(\nu+1)-2\nu}}\right) & \text{if } \varphi_{\mathrm{mult}} = \varphi_{\frac{\nu}{2}} \text{ in (5.12)}, \\ \mathcal{O}\left(\varphi_{\mathrm{mult}}^2\left(\vartheta^{-1}\left(\varphi_r^{-1}(\alpha_2)\right)\right)\right) & \text{else}, \end{cases}$$

$$\left\|F'\left[u^\dagger\right]\left(u_{n+1} - u^\dagger\right)\right\|_{\mathbb{Y}} = \begin{cases} \mathcal{O}\left(\alpha_n^{\frac{\nu+1}{r(\nu+1)-2\nu}}\right) & \text{if } \varphi_{\mathrm{mult}} = \varphi_{\frac{\nu}{2}} \text{ in (5.12)}, \\ \mathcal{O}\left(\varphi_r^{-1}(\alpha_n)\right) & \text{else}. \end{cases}$$

In case of noisy data assume additionally that (5.2) *holds true. Then for a parameter choice*

$$N_*(\delta) := \begin{cases} \min\left\{n \in \mathbb{N} \mid \alpha_n^{\frac{\nu+1}{r(\nu+1)-2\nu}} \leq \tau\delta\right\} & \text{if } \varphi_{\mathrm{mult}} = \varphi_{\frac{\nu}{2}} \text{ in (5.12)}, \\ \min\left\{n \in \mathbb{N} \mid \alpha_n \leq \varphi_r(\tau\delta)\right\} & \text{else} \end{cases}$$

with sufficiently large tuning parameter $\tau \geq 1$ one obtains

$$\mathcal{D}_p^{u^*}\left(u_{N_*(\delta)}, u^\dagger\right) = \begin{cases} \mathcal{O}\left(\delta^{\frac{2\nu}{1+\nu}}\right) & \text{if } \varphi_{\mathrm{mult}} = \varphi_{\frac{\nu}{2}} \text{ in (5.12)}, \\ \mathcal{O}\left(\varphi_{\mathrm{mult}}^2\left(\vartheta^{-1}(\delta)\right)\right) & \text{else} \end{cases} \tag{5.13}$$

as $\delta \searrow 0$.

PROOF:
See [KH10, Thm. 1]. ■

Note that the proven rates (5.13) are optimal as pointed out in Section 3.2.3. Moreover, a similar convergence rates result for the **a posteriori** parameter choices (5.9) and (5.8) can be obtained:

THEOREM 5.6:
Let Assumption 5.3 and a tangential cone condition (3.13) *with η fulfilling*

$$\eta < \underline{\theta} < \bar{\theta} < 1$$

hold and choose the tuning parameter τ in the discrepancy principle (5.8) *so large that*

$$\eta + \frac{1+\eta}{\tau} \leq \underline{\theta}, \qquad \eta < \frac{1-\bar{\theta}}{2}.$$

Assume that the noise level $\delta > 0$ from (5.2) *fulfills*

$$\delta < \frac{\left\|F(u_0) - g^{\mathrm{obs}}\right\|_{\mathbb{Y}}}{\tau}.$$

Moreover let $\mathcal{D}_p^{u^*}\left(u_0, u^\dagger\right)$ *be sufficiently small and assume that either* $F'[u] : \mathbb{X} \to \mathbb{Y}$ *is weakly closed for all* $u \in \mathfrak{B}$ *and* $\mathbb{Y}$ *reflexive or* $\mathfrak{B}$ *is weakly closed.*
Then the iterates u_{n+1} *defined by* (5.10) *are well defined for* $n \leq N_* - 1$ *with* $N_* = N_*(\delta)$ *obtained by the discrepancy principle* (5.8) *and the* α_n*'s chosen according to* (5.9) *and they fulfill*

$$\mathcal{D}_p^{u^*}\left(u_{N_*(\delta)}, u^\dagger\right) = \begin{cases} \mathcal{O}\left(\delta^{\frac{2\nu}{1+\nu}}\right) & \text{if } \varphi_{\text{mult}} = \varphi_{\frac{\nu}{2}} \text{ in (5.12)}, \\ \mathcal{O}\left(\varphi_{\text{mult}}^2\left(\vartheta^{-1}(\delta)\right)\right) & \text{else} \end{cases} \tag{5.14}$$

as $\delta \searrow 0$.

PROOF:
See [KH10, Thm. 2]. ■

CHAPTER
SIX

GENERALIZATION OF THE IRGNM

As in Chapters 3 and 5 we want to tackle an ill-posed problem

$$F(u) = g, \tag{6.1}$$

where the operator F is assumed to be nonlinear and Fréchet-differentiable. In concurrence with Chapter 5 we want to consider and iterative scheme, but we want to generalize the data fidelity and penalty term. This leads to

$$u_{n+1} \in \operatorname*{argmin}_{u \in \mathfrak{B}} \left[\mathcal{S}\left(F(u_n) + F'[u_n](u - u_n) ; g^{\mathrm{obs}} \right) + \alpha_n \mathcal{R}(u) \right] \tag{6.2a}$$

where $u_0 \in \mathfrak{B}$ is some initial guess, g^{obs} denotes the observed data, $\mathcal{S}$ some suitable data misfit and $\mathcal{R}$ some penalty as in Chapter 3. As in Chapter 5 we will assume that the regularization parameters $(\alpha_n)_{n\in\mathbb{N}}$ are chosen in a way such that

$$\alpha_0 \leq 1, \qquad \alpha_n \searrow 0, \qquad 1 \leq \frac{\alpha_n}{\alpha_{n+1}} \leq C_{\mathrm{dec}} \qquad \text{for all} \qquad n \in \mathbb{N}. \tag{6.2b}$$

In this chapter we will present different convergence theorems for the method (6.2) as well as an error decomposition in case of an additive variational inequality. To the author's best knowledge, such convergence rates results do not exist so far. As a first step towards the convergence analysis we will generalize the classical nonlinearity conditions (namely the tangential cone condition and the Lipschitz condition) and motivate these generalizations. The motivation itself will together with the motivation of the variational inequalities in Section 3.2.1 show that all results are generalizations of the known convergence rates results for the iteratively regularized Gauss-Newton method (5.3a) discussed in Chapter 5 and therefore include somehow the previous analysis [Bak92,BNS97,Hoh97, KH10].

6.1 Nonlinearity conditions

In this section we present generalized nonlinearity conditions, which we will use to provide convergence rates for the iteratively regularized Newton method (6.2). As we have seen in the last chapter, the degree of nonlinearity has to be restricted to ensure convergence of the iteratively regularized Gauss-Newton method (cf. Theorem 5.1). We used there the tangential cone condition (3.13), i.e. we assumed that there exists some $\bar{\eta} \geq 0$ such that

$$\left\|F(v) - F(u) - F'[u](v-u)\right\|_{\mathbb{Y}} \leq \bar{\eta}\,\|F(v) - F(u)\|_{\mathbb{Y}} \tag{6.3}$$

for all $u, v \in \mathfrak{B}$. Moreover we had to assume that $\bar{\eta}$ is small. This condition has been introduced by HANKE, NEUBAUER AND SCHERZER [HNS95] and is frequently used for the analysis of regularization methods for nonlinear inverse problems. Nevertheless, for many problems of practical interest it is very difficult to show that this condition is satisfied or not. It is known that if (6.3) holds and $F'\left[u^\dagger\right]$ is singular, then F must be constant along a certain affine subspace, cf. [HNS95, Prop. 2.1].
Another frequently used assumption is Lipschitz continuity of F', namely one assumes that there exists some $L > 0$ such that

$$\left\|F'[u] - F'[v]\right\|_{\mathbb{Y}} \leq L\,\|u - v\|_{\mathbb{X}} \tag{6.4}$$

for all $u, v \in \mathfrak{B}$. It can be seen by integration that (6.4) implies immediately

$$\left\|F(v) - F(u) - F'(u; v-u)\right\|_{\mathbb{Y}} \leq \frac{L}{2}\,\|v - u\|_{\mathbb{X}}^2 \tag{6.5}$$

for all $u, v \in \mathfrak{B}$. Hence, (6.5) restricts the class of operators F to those, whose first order Taylor expansion does not only converge superlinearly to F (as $u \to v$) - which is ensured by the Fréchet-differentiability - but converges quadratically. If for example the operator F is twice Fréchet-differentiable with a continuous second Fréchet-derivative F'', then the condition (6.4) is fulfilled whenever $\mathfrak{B}$ is small enough. The Lipschitz assumption (6.4) is widely used in the context of inverse problems, see e.g. [EHN96, SGG+08, Stü11].
Since many operators in practical applications are smooth, the condition (6.5) seems to be more reasonable than (6.3). Moreover, taking the ill-posedness of F into account shows that (6.3) is much more restrictive than (6.5). Nevertheless, (6.5) provides too little information for weak source conditions to prove convergence rates.

In our general setup where $\mathcal{S}$ does not necessarily fulfill a triangle inequality, the conditions (6.3) and (6.5) must be considered useless since in the general case $\mathcal{S}$ is not connected to the $\mathbb{Y}$-norm. Therefore, we have to use generalized formulations, which read as follows:

ASSUMPTION 6.1 (GENERALIZED TANGENTIAL CONE CONDITION):

(A) *There exist constants η (later assumed to be sufficiently small) and $C_{\mathrm{tc}} \geq 1$ such that for all $g^{\mathrm{obs}} \in \mathbb{Y}$*

$$\begin{aligned}
&\frac{1}{C_{\mathrm{tc}}}\mathcal{S}\left(F(v); g^{\mathrm{obs}}\right) - \eta\mathcal{S}\left(F(u); g^{\mathrm{obs}}\right) \\
\leq&\mathcal{S}\left(F(u) + F'(u; v-u); g^{\mathrm{obs}}\right) \\
\leq&C_{\mathrm{tc}}\mathcal{S}\left(F(v); g^{\mathrm{obs}}\right) + \eta\mathcal{S}\left(F(u); g^{\mathrm{obs}}\right) \qquad \text{for all } u, v \in \mathfrak{B}.
\end{aligned} \tag{6.6a}$$

(B) *There exist constants η (later assumed to be sufficiently small) and $C_{\mathrm{tc}} \geq 1$ such that*

$$\begin{aligned}
&\frac{1}{C_{\mathrm{tc}}}\mathcal{S}\left(F(v); g^{\dagger}\right) - \eta\mathcal{S}\left(F(u); g^{\dagger}\right) \\
\leq&\mathcal{S}\left(F(u) + F'(u; v-u); g^{\dagger}\right) \\
\leq&C_{\mathrm{tc}}\mathcal{S}\left(F(v); g^{\dagger}\right) + \eta\mathcal{S}\left(F(u); g^{\dagger}\right) \qquad \text{for all } u, v \in \mathfrak{B}.
\end{aligned} \tag{6.6b}$$

ASSUMPTION 6.2 (GENERALIZED LIPSCHITZ CONDITION):

(A) There exist constants K (later assumed to be sufficiently small), $r > 1$ and $C_{\mathrm{lip}} \geq 1$ such that for all $g^{\mathrm{obs}} \in \mathbb{Y}$

$$\begin{aligned}
&\frac{1}{C_{\mathrm{lip}}}\mathcal{S}\left(F(v);g^{\mathrm{obs}}\right) - K\mathcal{D}_{\mathcal{R}}^{u^*}\left(v,u^\dagger\right)^r - K\mathcal{D}_{\mathcal{R}}^{u^*}\left(u,u^\dagger\right)^r \\
\leq&\mathcal{S}\left(F(u)+F'(u;v-u);g^{\mathrm{obs}}\right) \qquad (6.7a)\\
\leq&C_{\mathrm{lip}}\mathcal{S}\left(F(v);g^{\mathrm{obs}}\right) + K\mathcal{D}_{\mathcal{R}}^{u^*}\left(v,u^\dagger\right)^r + K\mathcal{D}_{\mathcal{R}}^{u^*}\left(u,u^\dagger\right)^r \qquad \textit{for all } u,v\in\mathfrak{B}.
\end{aligned}$$

(B) There exist constants K (later assumed to be sufficiently small), $r > 1$ and $C_{\mathrm{lip}} \geq 1$ such that

$$\begin{aligned}
&\frac{1}{C_{\mathrm{lip}}}\mathcal{S}\left(F(v);g^\dagger\right) - K\mathcal{D}_{\mathcal{R}}^{u^*}\left(v,u^\dagger\right)^r - K\mathcal{D}_{\mathcal{R}}^{u^*}\left(u,u^\dagger\right)^r \\
\leq&\mathcal{S}\left(F(u)+F'(u;v-u);g^\dagger\right) \qquad (6.7b)\\
\leq&C_{\mathrm{lip}}\mathcal{S}\left(F(v);g^\dagger\right) + K\mathcal{D}_{\mathcal{R}}^{u^*}\left(v,u^\dagger\right)^r + K\mathcal{D}_{\mathcal{R}}^{u^*}\left(u,u^\dagger\right)^r \qquad \textit{for all } u,v\in\mathfrak{B}.
\end{aligned}$$

Both conditions ensure that the nonlinearity of F fits together with the data misfit functionals $\mathcal{S}$. Obviously, both are fulfilled with $\eta = 0$ and $C_{\mathrm{tc}} = 1$ or $K = 0$ and $C_{\mathrm{lip}} = 1$ respectively if F is linear. The distinction between the cases (A) and (B) is necessary since one usually does not want to assume a nonlinearity condition depending on the data, which might be random. On the other hand, as we will see in the following, for the case of a norm power $\mathcal{S}(g;\hat{g}) = \|g-\hat{g}\|_{\mathbb{Y}}^p$, (A) as well as (B) are valid if the classical nonlinearity condition is valid.

Finally we want to mention that the exponent $r > 1$ in Assumption 6.2 depends only on the data misfit term $\mathcal{S}$, which is pointed out in Lemma 6.4.

In the following we want to investigate the relation of Assumption 6.1 to the standard tangential cone condition (6.3) and the relation of Assumption 6.2 to the standard Lipschitz assumption (6.5).

LEMMA 6.3 (TANGENTIAL CONE CONDITION):

Assume that $\mathcal{S}(g;\hat{g}) = \|g-\hat{g}\|_{\mathbb{Y}}^p$ for some $p \geq 1$. If F fulfills the tangential cone condition (6.3) *with $\bar{\eta} \geq 0$ sufficiently small, then Assumptions 6.1A and 6.1B are satisfied with $\eta = 2^{2p-2}\bar{\eta}^p$ and*

$$C_{\mathrm{tc}} = \max\left\{\frac{1}{2^{1-p}-2^{p-1}\bar{\eta}^p}, 2^{p-1}+\bar{\eta}^p 2^{2p-2}\right\}.$$

PROOF:

Using the inequality $(a+b)^p \leq 2^{p-1}(a^p+b^p)$, $a,b \geq 0$ we find that

$$\begin{aligned}
&\left\|F(u)+F'(u;v-u)-g\right\|_{\mathbb{Y}}^p \\
\leq&\left(\left\|F(u)+F'(u;v-u)-F(v)\right\|_{\mathbb{Y}} + \|F(v)-g\|_{\mathbb{Y}}\right)^p \\
\leq&2^{p-1}\bar{\eta}^p\|F(u)-F(v)\|_{\mathbb{Y}}^p + 2^{p-1}\|F(v)-g\|_{\mathbb{Y}}^p \\
\leq&2^{2p-2}\bar{\eta}^p\|F(u)-g\|_{\mathbb{Y}}^p + \left(2^{p-1}+\bar{\eta}^p 2^{2p-2}\right)\|F(v)-g\|_{\mathbb{Y}}^p.
\end{aligned}$$

Moreover, with $|a-b|^p \geq 2^{1-p}a^p - b^p$, $a,b \geq 0$ we get

$$\begin{aligned}
&\left\|F(u) + F'(u;v-u) - g\right\|_{\mathbb{Y}}^p \\
\geq &\left|\|F(v)-g\|_{\mathbb{Y}} - \left\|F(u) + F'(u;v-u) - F(v)\right\|_{\mathbb{Y}}\right|^p \\
\geq &2^{1-p}\|F(v)-g\|_{\mathbb{Y}}^p - \bar{\eta}^p \|F(u)-F(v)\|_{\mathbb{Y}}^p \\
\geq &2^{1-p}\|F(v)-g\|_{\mathbb{Y}}^p - 2^{p-1}\bar{\eta}^p \|F(u)-g\|_{\mathbb{Y}}^p - 2^{p-1}\bar{\eta}^p \|F(v)-g\|_{\mathbb{Y}}^p \\
= &\left(2^{1-p} - 2^{p-1}\bar{\eta}^p\right)\|F(v)-g\|_{\mathbb{Y}}^p - 2^{p-1}\bar{\eta}^p \|F(u)-g\|_{\mathbb{Y}}^p
\end{aligned}$$

for all $g \in \mathbb{Y}$. Hence, (6.6) holds true with $\eta = 2^{2p-2}\bar{\eta}^p$ and

$$C_{\mathrm{tc}} = \max\left\{\frac{1}{2^{1-p} - 2^{p-1}\bar{\eta}^p}, 2^{p-1} + \bar{\eta}^p 2^{2p-2}\right\}$$

if $\bar{\eta}$ is sufficiently small. ■

LEMMA 6.4 (LIPSCHITZ CONDITION):
Let $\mathcal{S}(g;\hat{g}) = \|g-\hat{g}\|_{\mathbb{Y}}^p$ for some $p > 1$ and $\mathcal{R}(u) = \|u-u_0\|_{\mathbb{X}}^2$ where $\mathbb{X}$ is a Hilbert space. If F fulfills the Lipschitz condition (6.5)*, then Assumptions 6.2A and 6.2B are satisfied with $K = 2^{2p-2}L^p$, $r = p$ and $C_{\mathrm{lip}} = 2^{p-1}$.*

PROOF:
First note that in the case of $\mathcal{R}(u) = \|u-u_0\|_{\mathbb{X}}^2$ for a Hilbert space norm we have $\mathcal{D}_{\mathcal{R}}^{u^*}(u,u^\dagger) = \|u-u^\dagger\|_{\mathbb{X}}^2$. Using the inequality $(a+b)^p \leq 2^{p-1}(a^p+b^p)$, $a,b \geq 0$ twice, we find that

$$\begin{aligned}
&\left\|F(u) + F'(u;v-u) - g\right\|_{\mathbb{Y}}^p \\
\leq &\left(\left\|F(u) + F'(u;v-u) - F(v)\right\|_{\mathbb{Y}} + \|F(v)-g\|_{\mathbb{Y}}\right)^p \\
\leq &2^{p-1}\frac{L^p}{2^p}\|v-u\|_{\mathbb{X}}^{2p} + 2^{p-1}\|F(v)-g\|_{\mathbb{Y}}^p \\
\leq &2^{p-1}\|F(v)-g\|_{\mathbb{Y}}^p + 2^{2p-2}L^p\left\|v-u^\dagger\right\|_{\mathbb{X}}^{2p} + 2^{2p-2}L^p\left\|u-u^\dagger\right\|_{\mathbb{X}}^{2p} \\
= &2^{p-1}\|F(v)-g\|_{\mathbb{Y}}^p + 2^{2p-2}L^p\mathcal{D}_{\mathcal{R}}^{u^*}\left(v,u^\dagger\right)^p + 2^{2p-2}L^p\mathcal{D}_{\mathcal{R}}^{u^*}\left(u,u^\dagger\right)^p.
\end{aligned}$$

Moreover, with $|a-b|^p \geq 2^{1-p}a^p - b^p$, $a,b \geq 0$ we get

$$\begin{aligned}
&\left\|F(u) + F'(u;v-u) - g\right\|_{\mathbb{Y}}^p \\
\geq &\left|\|F(v)-g\|_{\mathbb{Y}} - \left\|F(u) + F'(u;v-u) - F(v)\right\|_{\mathbb{Y}}\right|^p \\
\geq &2^{1-p}\|F(v)-g\|_{\mathbb{Y}}^p - \frac{L^p}{2^p}\|v-u\|_{\mathbb{X}}^{2p} \\
\geq &2^{1-p}\|F(v)-g\|_{\mathbb{Y}}^p - 2^{p-1}L^p\left\|u-u^\dagger\right\|_{\mathbb{X}}^{2p} - 2^{p-1}L^p\left\|v-u^\dagger\right\|_{\mathbb{X}}^{2p} \\
= &2^{1-p}\|F(v)-g\|_{\mathbb{Y}}^p - 2^{p-1}L^p\mathcal{D}_{\mathcal{R}}^{u^*}\left(u,u^\dagger\right)^p - 2^{p-1}L^p\mathcal{D}_{\mathcal{R}}^{u^*}\left(v,u^\dagger\right)^p
\end{aligned}$$

for all $g \in \mathbb{Y}$. Hence, (6.7) holds true with $K = 2^{2p-2}L^p$, $r = p$ and $C_{\mathrm{lip}} = 2^{p-1}$. ■

Note that the term $\frac{L}{2}\|u-v\|_{\mathbb{X}}^2$ in (6.5) has been replaced by $K\mathcal{D}_{\mathcal{R}}^{u^*}(v,u^\dagger) + K\mathcal{D}_{\mathcal{R}}^{u^*}(u,u^\dagger)$, which is necessary since the Bregman distance can not measure the distance between u and v for $u^* \in \partial\mathcal{R}(u^\dagger)$.

6.2 Convergence rates for a priori stopping rules

With the nonlinearity conditions provided in the last section we are now able to prove convergence rates for the general case of an iteratively regularized Newton method (6.2) under **a priori** stopping rules. We will split the developed theory into three main theorems (cf. Theorem 6.5, Theorem 6.6 and Theorem 6.8), depending on the used variational inequality and nonlinearity condition. Compared to our convergence theorems for Tikhonov-type regularization (3.2) in Chapter 3, one additional convergence result is necessary, namely Theorem 6.8 which uses the additive variational inequality (3.11) with $\varphi_{\mathrm{add}} = \varphi_{\frac{1}{p}}$ and the generalized Lipschitz assumption (6.7). Theorems 6.5 and 6.6 correspond to the results for Tikhonov-type regularization (3.2) (cf. Theorems 3.28 and 3.30) and use the generalized tangential cone condition (6.6) due to the general index functions φ_{mult} and φ_{add} respectively. Finally we present our results for the special cases of logarithmic and Hölder-type source conditions and comment on the relation to previous work of other authors.
For the whole section we will use the following abbreviations:

$$d_n := \mathcal{D}_{\mathcal{R}}^{u^*}\left(u_n, u^\dagger\right)^{\frac{1}{2}}, \tag{6.8}$$

$$s_n := \mathcal{S}\left(F\left(u_n\right); g^\dagger\right). \tag{6.9}$$

6.2.1 Rates under a tangential cone condition

THEOREM 6.5 (CONVERGENCE RATES UNDER ASSUMPTION 3.24):
Let the Assumptions 3.8, 6.1A or 6.1B and 3.24 hold true, and suppose that η, $\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_0, u^\dagger\right)$ *and* $\mathcal{S}\left(F\left(u_0\right); g^\dagger\right)$ *are sufficiently small. Then the iterates* u_n *defined by* (6.2) *with exact data* $g^{\mathrm{obs}} = g^\dagger$ *fulfill*

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_n, u^\dagger\right) = \mathcal{O}\left(\varphi_{\mathrm{mult}}^2\left(\alpha_n\right)\right), \tag{6.10a}$$

$$\mathcal{S}\left(F\left(u_n\right); g^\dagger\right) = \mathcal{O}\left(\Theta\left(\alpha_n\right)\right) \tag{6.10b}$$

as $n \to \infty$. *For noisy data define*

$$\mathbf{err}_n := \frac{1}{C_{\mathrm{err}}}\mathbf{err}\left(F\left(u_{n+1}\right)\right) + 2\eta C_{\mathrm{tc}}\mathbf{err}\left(F\left(u_n\right)\right) + C_{\mathrm{tc}}C_{\mathrm{err}}\mathbf{err}\left(g^\dagger\right) \tag{6.11a}$$

in case of Assumption 6.1A or

$$\begin{aligned} \mathbf{err}_n \quad := \quad & \mathbf{err}\left(F\left(u_n\right) + F'\left(u_n; u_{n+1} - u_n\right)\right) \\ & + C_{\mathrm{err}}\mathbf{err}\left(F\left(u_n\right) + F'\left(u_n; u^\dagger - u_n\right)\right) \end{aligned} \tag{6.11b}$$

under Assumption 6.1B, and choose the stopping index n_* *by*

$$n_* := \min\left\{n \in \mathbb{N} \;\middle|\; \Theta\left(\alpha_n\right) \leq \tau \mathbf{err}_n\right\} \tag{6.12}$$

with a tuning parameter $\tau \geq 1$. *Then* (6.10) *holds for* $n \leq n_*$ *and the following convergence rates are valid as* $\mathbf{err}_n \searrow 0$:

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_{n_*}, u^\dagger\right) = \mathcal{O}\left(\varphi_{\mathrm{mult}}^2\left(\Theta^{-1}\left(\mathbf{err}_{n_*}\right)\right)\right), \tag{6.13a}$$

$$\mathcal{S}\left(F\left(u_{n_*}\right); g^\dagger\right) = \mathcal{O}\left(\mathbf{err}_{n_*}\right). \tag{6.13b}$$

PROOF:
Due to (3.15) we have

$$\begin{aligned}\mathcal{R}\left(u_{n+1}\right)-\mathcal{R}\left(u^{\dagger}\right) &= \mathcal{D}_{\mathcal{R}}^{u^*}\left(u_{n+1},u^{\dagger}\right)-\left\langle u^*,u^{\dagger}-u_{n+1}\right\rangle \\ &\geq d_{n+1}^2-\beta d_{n+1}\varphi_{\mathrm{mult}}\left(\frac{s_{n+1}}{d_{n+1}^2}\right). \end{aligned} \tag{6.14}$$

From the minimality condition (6.2a) with $u=u^{\dagger}$ we obtain

$$\begin{aligned}&\alpha_n\left(\mathcal{R}\left(u_{n+1}\right)-\mathcal{R}\left(u^{\dagger}\right)\right)+\mathcal{S}\left(F\left(u_n\right)+F'\left(u_n;u_{n+1}-u_n\right);g^{\mathrm{obs}}\right)\\ \leq&\mathcal{S}\left(F\left(u_n\right)+F'\left(u_n;u^{\dagger}-u_n\right);g^{\mathrm{obs}}\right),\end{aligned} \tag{6.15}$$

and putting (6.14) and (6.15) together we find that

$$\begin{aligned}&\alpha_n d_{n+1}^2+\mathcal{S}\left(F\left(u_n\right)+F'\left(u_n;u_{n+1}-u_n\right);g^{\mathrm{obs}}\right)\\ \leq&\mathcal{S}\left(F\left(u_n\right)+F'\left(u_n;u^{\dagger}-u_n\right);g^{\mathrm{obs}}\right)+\alpha_n\beta d_{n+1}\varphi_{\mathrm{mult}}\left(\frac{s_{n+1}}{d_{n+1}^2}\right).\end{aligned} \tag{6.16}$$

- In the case of 6.1B we use (3.4), which yields

$$\begin{aligned}&\alpha_n d_{n+1}^2+\frac{1}{C_{\mathrm{err}}}\mathcal{S}\left(F\left(u_n\right)+F'\left(u_n;u_{n+1}-u_n\right);g^{\dagger}\right)\\ \leq&C_{\mathrm{err}}\mathcal{S}\left(F\left(u_n\right)+F'\left(u_n;u^{\dagger}-u_n\right);g^{\dagger}\right)+\alpha_n\beta d_{n+1}\varphi_{\mathrm{mult}}\left(\frac{s_{n+1}}{d_{n+1}^2}\right)+\mathbf{err}_n\end{aligned}$$

  and (6.6b) with $v=u^{\dagger}$, $u=u_n$ leads to

$$\begin{aligned}&\alpha_n d_{n+1}^2+\frac{1}{C_{\mathrm{err}}}\mathcal{S}\left(F\left(u_n\right)+F'\left(u_n;u_{n+1}-u_n\right);g^{\dagger}\right)\\ \leq&\eta C_{\mathrm{err}}s_n+\alpha_n\beta d_{n+1}\varphi_{\mathrm{mult}}\left(\frac{s_{n+1}}{d_{n+1}^2}\right)+\mathbf{err}_n.\end{aligned}$$

  By (6.6b) with $v=u_{n+1}$, $u=u_n$ we obtain

$$\alpha_n d_{n+1}^2+\frac{1}{C_{\mathrm{tc}}C_{\mathrm{err}}}s_{n+1}\leq\eta\left(C_{\mathrm{err}}+\frac{1}{C_{\mathrm{err}}}\right)s_n+\alpha_n\beta d_{n+1}\varphi_{\mathrm{mult}}\left(\frac{s_{n+1}}{d_{n+1}^2}\right)+\mathbf{err}_n$$

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

- In the case of 6.1A we are able to apply (6.6a) with $v=u^{\dagger}$, $u=u_n$ and (6.6a) with $v=u_{n+1}$ and $u=u_n$ to (6.16) to conclude

$$\begin{aligned}&\alpha_n d_{n+1}^2+\frac{1}{C_{\mathrm{tc}}}\mathcal{S}\left(F\left(u_{n+1}\right);g^{\mathrm{obs}}\right)\\ \leq&2\eta\mathcal{S}\left(F\left(u_n\right);g^{\mathrm{obs}}\right)+C_{\mathrm{tc}}\mathcal{S}\left(F\left(u^{\dagger}\right);g^{\mathrm{obs}}\right)+\alpha_n\beta d_{n+1}\varphi_{\mathrm{mult}}\left(\frac{s_{n+1}}{d_{n+1}^2}\right).\end{aligned}$$

Due to (3.4) and $\mathcal{S}\left(g^\dagger; g^\dagger\right) = 0$ this yields

$$\alpha_n d_{n+1}^2 + \frac{1}{C_{\mathrm{tc}} C_{\mathrm{err}}} s_{n+1} \leq 2\eta C_{\mathrm{err}} s_n + \alpha_n \beta d_{n+1} \varphi_{\mathrm{mult}}\left(\frac{s_{n+1}}{d_{n+1}^2}\right) + \mathbf{err}_n$$

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

Using $2C_{\mathrm{err}} \geq C_{\mathrm{err}} + \frac{1}{C_{\mathrm{err}}}$ we find that both in case of 6.1A and 6.1B a **recursive error estimate** of the form

$$\alpha_n d_{n+1}^2 + \frac{1}{C_{\mathrm{tc}} C_{\mathrm{err}}} s_{n+1} \leq 2\eta C_{\mathrm{err}} s_n + \alpha_n \beta d_{n+1} \varphi_{\mathrm{mult}}\left(\frac{s_{n+1}}{d_{n+1}^2}\right) + \mathbf{err}_n \tag{6.17}$$

is valid for all $n \in \mathbb{N}$.

Assume in the following that $\mathbf{err}_n$ is sufficiently small to ensure $n_* > 0$. We will now prove by induction that (6.17) implies

$$d_n \leq C_1 \varphi_{\mathrm{mult}}\left(\alpha_n\right), \tag{6.18}$$

$$s_n \leq C_2 \Theta\left(\alpha_n\right) \tag{6.19}$$

for all $n \leq n_*$ in case of noisy data and for all $n \in \mathbb{N}$ in case of exact data with suitable constants $C_1, C_2 \geq 0$. For $n = 0$ (6.18) and (6.19) are guaranteed by the assumption that d_0 and s_0 are small enough. For the induction step we observe that (6.17) together with (6.12) and the induction hypothesis for $n \leq n_* - 1$ implies

$$\alpha_n d_{n+1}^2 + \frac{1}{C_{\mathrm{tc}} C_{\mathrm{err}}} s_{n+1} \leq C_{\eta,\tau} \Theta\left(\alpha_n\right) + \alpha_n \beta d_{n+1} \varphi_{\mathrm{mult}}\left(\frac{s_{n+1}}{d_{n+1}^2}\right)$$

where $C_{\eta,\tau} = 2\eta C_2 C_{\mathrm{err}} + 1/\tau$. Now we distinguish between two cases:

Case 1: $\alpha_n \beta d_{n+1} \varphi_{\mathrm{mult}}\left(\frac{s_{n+1}}{d_{n+1}^2}\right) \leq C_{\eta,\tau} \Theta\left(\alpha_n\right)$.

In that case we find

$$\alpha_n d_{n+1}^2 + \frac{1}{C_{\mathrm{tc}} C_{\mathrm{err}}} s_{n+1} \leq 2 C_{\eta,\tau} \Theta\left(\alpha_n\right)$$

which by $\Theta\left(t\right)/t = \varphi_{\mathrm{mult}}^2\left(t\right)$, (3.22) and (3.23) implies

$$d_{n+1} \leq \sqrt{2C_{\eta,\tau}} \varphi_{\mathrm{mult}}\left(\alpha_n\right) = \sqrt{2C_{\eta,\tau}} \varphi_{\mathrm{mult}}\left(\frac{\alpha_n}{\alpha_{n+1}} \alpha_{n+1}\right) \leq \sqrt{2C_{\eta,\tau}} C_{\mathrm{dec}} \varphi_{\mathrm{mult}}\left(\alpha_{n+1}\right),$$

$$s_{n+1} \leq 2 C_{\mathrm{tc}} C_{\mathrm{err}} C_{\eta,\tau} \Theta\left(\alpha_n\right) \leq 2 C_{\mathrm{tc}} C_{\mathrm{err}} C_{\eta,\tau} C_{\mathrm{dec}}^3 \Theta\left(\alpha_{n+1}\right).$$

The assertions now follow by choosing C_1 and C_2 large enough such that $\sqrt{2C_{\eta,\tau}} C_{\mathrm{dec}} \leq C_1$ and $2 C_{\mathrm{tc}} C_{\mathrm{err}} C_{\eta,\tau} C_{\mathrm{dec}}^3 \leq C_2$.

Case 2: $\alpha_n \beta d_{n+1} \varphi_{\mathrm{mult}}\left(\frac{s_{n+1}}{d_{n+1}^2}\right) > C_{\eta,\tau} \Theta\left(\alpha_n\right)$.

In that case we find

$$\alpha_n d_{n+1}^2 + \frac{1}{C_{\mathrm{tc}} C_{\mathrm{err}}} s_{n+1} \leq 2 \alpha_n \beta d_{n+1} \varphi_{\mathrm{mult}}\left(\frac{s_{n+1}}{d_{n+1}^2}\right)$$

which allows us to use Lemma 3.27. This yields (6.18) and (6.19).

Finally, (6.10) is already proven and (6.13) follows from inserting (6.12) into (6.18) and (6.19) using (3.21). ■

Now we will prove convergence rates of the iteratively regularized Newton method (6.2) under an additive variational inequality (3.11). As opposed to the case of a multiplicative variational inequality (3.15) we will moreover gain an error decomposition as in the case of Tikhonov-type regularization (3.2), which will allow for an a posteriori stopping rule. As a nonlinearity condition we will again use the generalized tangential cone condition (6.6), which due to Table 3.1 also helps to motivate the general form of (3.11). The special case $\varphi_{\text{add}}(t) = t^{\frac{1}{2}}$ will be considered in Section 6.2.2 in combination with the generalized Lipschitz assumption (6.7).

THEOREM 6.6 (CONVERGENCE RATES UNDER ASSUMPTION 3.15):
Let Assumptions 3.8, 6.1A or 6.1B and 3.15 hold true. Then the error of the iterates u_n defined by (6.2) *can be bounded by*

$$(1-\beta)\, d_{n+1}^2 + \frac{1}{2C_{\text{err}}C_{\text{tc}}\alpha_n} s_{n+1} \leq 2\eta C_{\text{err}} \frac{s_n}{\alpha_n} + 2C_{\text{tc}}C_{\text{err}} \left(-\varphi_{\text{add}}\right)^* \left(-\frac{1}{\alpha_n}\right) + \frac{\mathbf{err}_n}{\alpha_n} \tag{6.20}$$

for all $n \in \mathbb{N}$ where $\mathbf{err}_n$ is given as in (6.11)*. If there exists moreover an upper bound $\overline{\mathbf{err}} \geq \mathbf{err}_n$ for all $n \in \mathbb{N}$ and η and $\mathcal{S}\left(F\left(u_0\right); g^\dagger\right)$ are sufficiently small, the estimate*

$$s_n \leq \gamma_{\text{nl}} \left(2C_{\text{tc}}C_{\text{err}}\alpha_n \left(-\varphi_{\text{add}}\right)^* \left(-\frac{1}{\alpha_n}\right) + \overline{\mathbf{err}}\right) \tag{6.21}$$

holds true for all $n \in \mathbb{N}_0$ with

$$\gamma_{\text{nl}} = \frac{2C_{\text{tc}}C_{\text{err}}C_{\text{dec}}^2}{1 - 4\eta C_{\text{tc}}C_{\text{err}}^2 C_{\text{dec}}^2}.$$

This yields for exact data the rates

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_n, u^\dagger\right) = \mathcal{O}\left(\left(-\varphi_{\text{add}}\right)^* \left(-\frac{1}{\alpha_n}\right)\right), \tag{6.22a}$$

$$\mathcal{S}\left(F\left(u_n\right); g^\dagger\right) = \mathcal{O}\left(\alpha_n \left(-\varphi_{\text{add}}\right)^* \left(-\frac{1}{\alpha_n}\right)\right) \tag{6.22b}$$

as $n \to \infty$. In case of noisy data we choose

$$n_* = \min \left\{ n \in \mathbb{N} \;\middle|\; \frac{1}{\alpha_n} \geq -\inf \partial\left(-\varphi_{\text{add}}\right)\left(\overline{\mathbf{err}}\right) \right\} \tag{6.23}$$

and find that the estimates (6.22) *are valid for $n \leq n_*$ and it holds*

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_{n_*}, u^\dagger\right) = \mathcal{O}\left(\varphi_{\text{add}}\left(\overline{\mathbf{err}}\right)\right) \tag{6.24}$$

as $\overline{\mathbf{err}} \searrow 0$.

REMARK 6.7:
The stopping rule (6.23) is due to the same problem we had in Section 3.3: The optimal choice requires $\frac{1}{\alpha_{n_*}} \in -\partial\left(-\varphi_{\text{add}}\right)\left(\overline{\mathbf{err}}\right)$, but the typical case is $\frac{1}{\alpha_n} \notin -\partial\left(-\varphi_{\text{add}}\right)\left(\overline{\mathbf{err}}\right)$ for all $n \in \mathbb{N}$. It can be shown similarly to Lemma 3.42 that under (6.23) it holds

$$\varphi_{\text{add}}(\sigma) - \frac{C_{\text{dec}}}{\alpha_{n_*-1}}\sigma \leq C_{\text{dec}}\varphi_{\text{add}}\left(\overline{\mathbf{err}}\right) - \frac{C_{\text{dec}}}{\alpha_{n_*-1}}\overline{\mathbf{err}} \tag{6.25}$$

for all $\sigma \geq 0$.

For (6.25) to be meaningful, we assume here and in the following proof that $\overline{\mathbf{err}}$ is sufficiently small to ensure $n_* > 0$.

PROOF (OF THEOREM 6.6):
Similar to the proof of Theorem 6.5 the assumptions imply the iterative estimate

$$\alpha_n (1-\beta) d_{n+1}^2 + \frac{1}{C_{\mathrm{tc}} C_{\mathrm{err}}} s_{n+1} \leq 2\eta C_{\mathrm{err}} s_n + \alpha_n \varphi_{\mathrm{add}} (s_{n+1}) + \mathbf{err}_n$$

for all $n \in \mathbb{N}$. Rearranging terms and dividing by α_n yields

$$\begin{aligned}(1-\beta) d_{n+1}^2 + \frac{1}{2C_{\mathrm{tc}} C_{\mathrm{err}} \alpha_n} s_{n+1} &\leq 2\eta C_{\mathrm{err}} \frac{s_n}{\alpha_n} + \varphi_{\mathrm{add}} (s_{n+1}) - \frac{1}{2C_{\mathrm{tc}} C_{\mathrm{err}} \alpha_n} s_{n+1} + \frac{\mathbf{err}_n}{\alpha_n} \\ &\leq 2\eta C_{\mathrm{err}} \frac{s_n}{\alpha_n} + \sup_{\sigma \geq 0} \left(\varphi_{\mathrm{add}} (\sigma) - \frac{1}{2C_{\mathrm{tc}} C_{\mathrm{err}} \alpha_n} \sigma \right) + \frac{\mathbf{err}_n}{\alpha_n} \\ &= 2\eta C_{\mathrm{err}} \frac{s_n}{\alpha_n} + (-\varphi_{\mathrm{add}})^* \left(-\frac{1}{2C_{\mathrm{tc}} C_{\mathrm{err}} \alpha_n} \right) + \frac{\mathbf{err}_n}{\alpha_n} \\ &\leq 2\eta C_{\mathrm{err}} \frac{s_n}{\alpha_n} + 2C_{\mathrm{tc}} C_{\mathrm{err}} (-\varphi_{\mathrm{add}})^* \left(-\frac{1}{\alpha_n} \right) + \frac{\mathbf{err}_n}{\alpha_n}\end{aligned}$$

where we used the definition of the Fenchel conjugate and (3.27). Therefore, (6.20) is proven. From (6.20) we conclude using the bound $\mathbf{err}_n \leq \overline{\mathbf{err}}$ that

$$s_{n+1} \leq 4C_{\mathrm{tc}} C_{\mathrm{err}}^2 \eta s_n + 4C_{\mathrm{tc}}^2 C_{\mathrm{err}}^2 \alpha_n (-\varphi_{\mathrm{add}})^* \left(-\frac{1}{\alpha_n} \right) + 2C_{\mathrm{tc}} C_{\mathrm{err}} \overline{\mathbf{err}}$$

for all $n \in \mathbb{N}$. Now we prove (6.21) by induction: For $n = 0$ the assertion is true since s_0 was assumed to be sufficiently small. Now let (6.21) hold for some n. Then by the inequality above we find that

$$\begin{aligned}s_{n+1} &\leq 4C_{\mathrm{tc}} C_{\mathrm{err}}^2 \eta s_n + 4C_{\mathrm{tc}}^2 C_{\mathrm{err}}^2 \alpha_n (-\varphi_{\mathrm{add}})^* \left(-\frac{1}{\alpha_n} \right) + 2C_{\mathrm{tc}} C_{\mathrm{err}} \overline{\mathbf{err}} \\ &\leq 2C_{\mathrm{tc}} C_{\mathrm{err}} (2C_{\mathrm{err}} \eta \gamma_{\mathrm{nl}} + 1) \left(2C_{\mathrm{tc}} C_{\mathrm{err}} \alpha_n (-\varphi_{\mathrm{add}})^* \left(-\frac{1}{\alpha_n} \right) + \overline{\mathbf{err}} \right) \\ &\leq 2C_{\mathrm{tc}} C_{\mathrm{err}} C_{\mathrm{dec}}^2 (2C_{\mathrm{err}} \eta \gamma_{\mathrm{nl}} + 1) \left(2C_{\mathrm{tc}} C_{\mathrm{err}} \alpha_{n+1} (-\varphi_{\mathrm{add}})^* \left(-\frac{1}{\alpha_{n+1}} \right) + \overline{\mathbf{err}} \right)\end{aligned}$$

where we used (6.2b), the monotonicity of $\sigma \mapsto (-\varphi_{\mathrm{add}})^* \left(-\frac{1}{\sigma}\right)$ (cf. Remark 3.29) and (3.27). The definition of γ_{nl} implies $2C_{\mathrm{tc}} C_{\mathrm{err}} C_{\mathrm{dec}}^2 (2C_{\mathrm{err}} \eta \gamma_{\mathrm{nl}} + 1) \leq \gamma_{\mathrm{nl}}$ and hence the assertion is shown.

Plugging (6.20) and (6.21) together we find

$$(1-\beta) d_{n+1}^2 + \frac{1}{C_{\mathrm{tc}} C_{\mathrm{err}}} \frac{s_{n+1}}{\alpha_n} \leq (1 + 2\eta C_{\mathrm{err}} \gamma_{\mathrm{nl}}) \left(2C_{\mathrm{tc}} C_{\mathrm{err}} (-\varphi_{\mathrm{add}})^* \left(-\frac{1}{\alpha_n} \right) + \frac{\overline{\mathbf{err}}}{\alpha_n} \right) \quad (6.26)$$

for all $n \in \mathbb{N}$. In the noise-free case it holds $\overline{\mathbf{err}} = 0$ and hence we obtain (6.22) by using (3.27). Otherwise we find from (3.27) and (6.25) that

$$\begin{aligned}(-\varphi_{\mathrm{add}})^* \left(-\frac{1}{\alpha_{n_*-1}} \right) &\leq C_{\mathrm{dec}} (-\varphi_{\mathrm{add}})^* \left(-\frac{C_{\mathrm{dec}}}{\alpha_{n_*-1}} \right) \\ &= C_{\mathrm{dec}} \sup_{\sigma \geq 0} \left(\varphi_{\mathrm{add}} (\sigma) - \frac{C_{\mathrm{dec}}}{\alpha_{n_*-1}} \sigma \right) \\ &\leq C_{\mathrm{dec}}^2 \varphi_{\mathrm{add}} (\overline{\mathbf{err}}) - \frac{C_{\mathrm{dec}}^2}{\alpha_{n_*-1}} \overline{\mathbf{err}}.\end{aligned}$$

Inserting this into (6.26) with $n = n_* - 1$ and dropping the second term on the left-hand side we have

$$d_{n_*}^2 \leq (1 + 2\eta C_{\mathrm{err}}\gamma_{\mathrm{nl}}) \left(2C_{\mathrm{tc}}C_{\mathrm{err}}C_{\mathrm{dec}}^2\varphi_{\mathrm{add}}(\overline{\mathbf{err}}) + \frac{\overline{\mathbf{err}}}{\alpha_{n_*-1}}\left(1 - 2C_{\mathrm{tc}}C_{\mathrm{err}}C_{\mathrm{dec}}^2\right)\right)$$
$$\leq (1 + 2\eta C_{\mathrm{err}}\gamma_{\mathrm{nl}})\, 2C_{\mathrm{tc}}C_{\mathrm{err}}C_{\mathrm{dec}}^2\varphi_{\mathrm{add}}(\overline{\mathbf{err}})$$

using $2C_{\mathrm{tc}}C_{\mathrm{err}}C_{\mathrm{dec}}^2 \geq 1$ which shows (6.24). ■

6.2.2 Rates under a Lipschitz assumption

In case of Assumption (3.15) with $\varphi_{\mathrm{add}}(t) = t^{\frac{1}{2}}$ and $\mathcal{S}(g;\hat{g}) = \|g - \hat{g}\|_{\mathbb{Y}}^2$ we expect to obtain convergence rates also under a Lipschitz assumption instead of the tangential cone condition. We will show in the next theorem, that this is also true for more general $\mathcal{S}$.

THEOREM 6.8 (CONVERGENCE RATES UNDER A LIPSCHITZ ASSUMPTION):
Let Assumptions 3.8, 6.2A or 6.2B and 3.15 with $\varphi_{\mathrm{add}} = \tilde{\beta}\varphi_{\frac{1}{p}}$ where $p \in (1, r]$ hold true and assume that $\mathcal{D}_{\mathcal{R}}^{u^*}(u_0, u^\dagger)$, $\tilde{\beta}$ and K are sufficiently small. In this setup, φ_{add}^2 does not need to be concave. Then for exact data we obtain for the minimizers of (6.2) *the convergence rate*

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_{n_*}, u^\dagger\right) = \mathcal{O}\left(\alpha_n^{\frac{1}{p-1}}\right)$$

and in case of noisy data with the a-priori stopping rule

$$n_* = \min\left\{n \in \mathbb{N} \mid \alpha_n^{\frac{p}{p-1}} \leq \tau \mathbf{err}_n\right\} \tag{6.27}$$

for sufficiently large $\tau > 0$ where

$$\mathbf{err}_n = \mathbf{err}(F(u_{n+1})) + C_{\mathrm{lip}}C_{\mathrm{err}}\mathbf{err}\left(g^\dagger\right)$$

in case of 6.2A and

$$\mathbf{err}_n = \mathbf{err}(F(u_n + F'(u_n, u_{n+1} - u_n))) + C_{\mathrm{err}}\mathbf{err}\left(F\left(u^\dagger\right) + F'\left(u_n, u^\dagger - u_n\right)\right)$$

in case of 6.2B we obtain

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_{n_*}, u^\dagger\right) = \mathcal{O}\left(\mathbf{err}_{n_*}^{\frac{1}{p}}\right).$$

PROOF:
Similarly to the proof of Theorem 6.5 we obtain a recursive error estimate of the form

$$\alpha_n(1-\beta)d_{n+1}^2 + \frac{1}{C_{\mathrm{err}}C_{\mathrm{lip}}}s_{n+1} \leq 2C_{\mathrm{err}}Kd_n^{2r} + Kd_{n+1}^{2r} + \alpha_n\tilde{\beta}s_{n+1}^{\frac{1}{p}} + \tau^{-1}\alpha_n^{\frac{p}{p-1}}$$

for all $n \leq n_* - 1$. Applying Young's inequality and rearranging terms leads to

$$\alpha_n(1-\beta)d_{n+1}^2 + \left(\frac{1}{C_{\mathrm{err}}C_{\mathrm{lip}}} - \varepsilon\frac{\tilde{\beta}}{p}\right)s_{n+1} \leq 2C_{\mathrm{err}}Kd_n^{2r} + Kd_{n+1}^{2r} + \left(\frac{\tilde{\beta}}{q}\left(\frac{1}{\varepsilon}\right)^{\frac{q}{p}} + \tau^{-1}\right)\alpha_n^{\frac{p}{p-1}}.$$

for all $n \leq n_* - 1$ where $q = \frac{1}{1-\frac{1}{p}} = \frac{p}{p-1}$. Now choose $\varepsilon > 0$ such that the second term on the left-hand side vanishes, i.e. $\frac{1}{C_{\mathrm{err}} C_{\mathrm{lip}}} = \varepsilon \frac{\tilde{\beta}}{p}$. Moreover we divide by $\alpha_{n+1}^{\frac{p}{p-1}}$ to find

$$\begin{aligned}
\left(\frac{\alpha_n}{\alpha_{n+1}}\right) \frac{d_{n+1}^2}{\alpha_{n+1}^{\frac{1}{p-1}}} \leq& \frac{2C_{\mathrm{err}} K}{1-\beta} \frac{d_n^{2r}}{\alpha_{n+1}^{\frac{p}{p-1}}} + \frac{K}{1-\beta} \frac{d_{n+1}^{2r}}{\alpha_{n+1}^{\frac{p}{p-1}}} + \frac{1}{1-\beta}\left(\frac{\tilde{\beta}}{q}\left(\frac{1}{\varepsilon}\right)^{\frac{q}{p}} + \frac{1}{\tau}\right)\left(\frac{\alpha_n}{\alpha_{n+1}}\right)^{\frac{p}{p-1}} \\
=& \frac{2C_{\mathrm{err}} K}{1-\beta}\left(\frac{d_n^2}{\alpha_n^{\frac{1}{p-1}}}\right)^r \left(\frac{\alpha_n}{\alpha_{n+1}}\right)^{\frac{r}{p-1}} \alpha_{n+1}^{\frac{r-p}{p-1}} + \frac{K}{1-\beta}\left(\frac{d_{n+1}^2}{\alpha_{n+1}^{\frac{1}{p-1}}}\right)^r \alpha_{n+1}^{\frac{r-p}{p-1}} \\
&+ \frac{1}{1-\beta}\left(\frac{p-1}{p}\left(\frac{C_{\mathrm{err}} C_{\mathrm{lip}}}{p}\right)^{\frac{1}{p-1}} \tilde{\beta}^{\frac{p}{p-1}} + \frac{1}{\tau}\right)\left(\frac{\alpha_n}{\alpha_{n+1}}\right)^{\frac{p}{p-1}}
\end{aligned}$$

for all $n \leq n_* - 1$. Due to (6.2b) as well as $\frac{r-p}{p-1} \geq 0$ which holds true by $r \geq p$ and $p > 1$ we hence find for $\gamma_n = \alpha_n^{-\frac{1}{p-1}} d_n^2$ the estimate

$$\left(1 - \frac{K}{1-\beta}\gamma_{n+1}^{r-1}\right)\gamma_{n+1} \leq \frac{2C_{\mathrm{err}} K}{1-\beta} C_{\mathrm{dec}}^{\frac{r}{p-1}} \gamma_n^r + \frac{C_{\mathrm{dec}}^{\frac{p}{p-1}}}{1-\beta}\left(\frac{p-1}{p}\left(\frac{C_{\mathrm{err}} C_{\mathrm{lip}}}{p}\right)^{\frac{1}{p-1}} \tilde{\beta}^{\frac{p}{p-1}} + \frac{1}{\tau}\right) \tag{6.28}$$

for all $n \leq n_* - 1$. This estimate is of the same form as [KH10, eq. (45)] and hence we can deduce convergence rates in the same manner. For sufficiently small $A, \bar{\gamma}, \bar{\xi} > 0$ it can by seen via differentiation that the function $h : (0, \bar{\gamma}) \to (0, \bar{\xi}), h(\gamma) := (1 - A\gamma^{r-1})\gamma$ is strictly monotonically increasing and invertible with

$$h^{-1}(\xi) \leq 2\xi, \qquad \xi \in (0, \bar{\xi}). \tag{6.29}$$

Now let $\bar{\gamma}$ be so small that the right-hand side of (6.28) with γ_n replaced by $\bar{\gamma}$ is smaller than $\bar{\xi}$. Then $\gamma_0 \leq \bar{\gamma}$ holds true if $\mathcal{D}_{\mathcal{R}}^{u^*}(u_0, u^\dagger)$ is sufficiently small as given in the assumption. Now consider the function h from above with $A = (1-\beta)^{-1} K$ which is sufficiently small to ensure (6.29) since K was assumed to be sufficiently small. If $\gamma_n \leq \bar{\gamma}$ holds true for some $n \leq n_* - 1$ then by applying (6.29) to (6.28) we find

$$\begin{aligned}
\gamma_{n+1} &\leq \frac{4C_{\mathrm{err}} K}{1-\beta} C_{\mathrm{dec}}^{\frac{r}{p-1}} \gamma_n^r + \frac{2C_{\mathrm{dec}}^{\frac{p}{p-1}}}{1-\beta}\left(\frac{p-1}{p}\left(\frac{C_{\mathrm{err}} C_{\mathrm{lip}}}{p}\right)^{\frac{1}{p-1}} \tilde{\beta}^{\frac{p}{p-1}} + \frac{1}{\tau}\right) \\
&\leq \frac{4C_{\mathrm{err}} K}{1-\beta} C_{\mathrm{dec}}^{\frac{r}{p-1}} \bar{\gamma}^r + \frac{1}{2}\bar{\gamma}
\end{aligned}$$

if $\tilde{\beta}$ is sufficiently small and τ is sufficiently large. Moreover the first term on the right-hand side is less or equal $\frac{1}{2}\bar{\gamma}$ if K is sufficiently small. Thus it holds $\gamma_{n+1} \leq \bar{\gamma}$ and hence $\gamma_n \leq \bar{\gamma}$ for all $n \leq n_*$. This yields

$$d_{n_*}^2 = \gamma_{n_*} \alpha_{n_*}^{\frac{1}{p-1}} \leq \bar{\gamma} \alpha_{n_*}^{\frac{1}{p-1}} = \mathcal{O}\left(\alpha_{n_*}^{\frac{1}{p-1}}\right)$$

Due to (6.27) we get

$$d_{n_*}^2 = \mathcal{O}\left(\alpha_{n_*}^{\frac{1}{p-1}}\right) = \mathcal{O}\left(\mathbf{err}_{n_*}^{\frac{1}{p}}\right)$$

which proves the assertion. ∎

REMARK 6.9:
Note that the above result is the only one where we needed to assume that the constant in the second term of the additive variational inequality is sufficiently small.

REMARK 6.10 (NONLINEARITY CONDITIONS):
In the case of a norm power $\mathcal{S}(g;\hat{g}) = \|g-\hat{g}\|_{\mathbb{Y}}^r$, $r>1$, Theorem 6.8 shows that we can weaken the nonlinearity condition if the source condition is sufficiently strong. This fact is known from the classical theory. In case of a general data fidelity term $\mathcal{S}$ it remains unclear if the generalized Lipschitz assumption (6.7) is indeed weaker than the generalized tangential cone condition (6.6).
Especially for the case of Poisson data and $\mathcal{S}$ as defined in (2.14) we do not know if those assumptions are appropriate and either of them is fulfilled for the nonlinear operators occurring in our applications.

6.2.3 Special cases and connections to previous results

Before we comment on the relations between our results and the results for norm powers mentioned in Chapter 5 we will present the implications of our results for the special cases of Hölder-type and logarithmic source conditions.

Hölder-type source conditions

If the index function φ in (3.9) is given by $\varphi = \varphi_\nu$ as in (3.10a) then the corresponding functions φ_{mult} and φ_{add} have already been calculated in Section 3.2.3. This leads to the following convergence rates:

THEOREM 6.11 (CONVERGENCE RATES FOR HÖLDER-TYPE SOURCE CONDITIONS):

(a) Under the assumptions of Theorem 6.5 with $\varphi_{\text{mult}} = \varphi_\nu$ we obtain the convergence rate

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_{n_*}, u^\dagger\right) = \mathcal{O}\left(\mathbf{err}_{n_*}^{\frac{2\nu}{2\nu+1}}\right), \qquad \mathbf{err}_{n_*} \searrow 0.$$

(b) Under the assumptions of Theorem 6.6 with $\varphi_{\text{add}} = \varphi_\kappa$ we obtain the convergence rate

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_{n_*}, u^\dagger\right) = \mathcal{O}\left(\overline{\mathbf{err}}^{\kappa}\right), \qquad \overline{\mathbf{err}} \searrow 0.$$

(c) Under the assumptions of Theorem 6.8 we obtain the convergence rate

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_{n_*}, u^\dagger\right) = \mathcal{O}\left(\mathbf{err}_{n_*}^{\frac{1}{p}}\right), \qquad \mathbf{err}_{n_*} \searrow 0.$$

PROOF:
This follows directly from the mentioned Theorems and by plugging in the calculations from Section 3.2.3. ■

Now we will specify this result for the quadratic Hilbert space case:

COROLLARY 6.12 (QUADRATIC HILBERT SPACE CASE):
Assume that $\mathcal{S}(g;\hat{g}) = \|g-\hat{g}\|_{\mathbb{Y}}^2$ and $\mathcal{R}(u) = \|u-u_0\|_{\mathbb{X}}^2$ for Hilbert spaces $\mathbb{X}$ and $\mathbb{Y}$ and assume that $F : \mathbb{X} \to \mathbb{Y}$ is Fréchet differentiable and fulfills the tangential cone condition (6.3) *with sufficiently small $\bar{\eta}$. Let furthermore u_0 be sufficiently close to $u^\dagger$. Then a range condition* (3.9) *with $\varphi = \varphi_\nu$ where $\nu \in \left(0, \frac{1}{2}\right]$ and a known upper bound $\left\|g^\dagger - g^{\text{obs}}\right\|_{\mathbb{Y}} \leq \delta$ imply for*

a stopping parameter n_* chosen such that n_* is the first natural number with $\alpha_n \leq \delta^{\frac{2}{2\nu+1}}$ the following convergence rate:

$$\left\| u^\dagger - u_{n_*} \right\|_{\mathbb{X}} = \mathcal{O}\left(\delta^{\frac{2\nu}{1+2\nu}}\right), \qquad \delta \searrow 0.$$

PROOF:
First of all, the tangential cone condition (6.3) implies by Lemma 6.3 the generalized form (6.6b) with some constant C_{tc} and a parameter η which is small if $\bar{\eta}$ from (6.3) is small. As already mentioned, Assumption 3.8 is fulfilled with $\mathbf{err} \equiv \delta^2$, $\mathfrak{s} \equiv 0$ and $C_{\mathrm{err}} = 2$. The term $\mathbf{err}_n$ from (6.11b) is hence given by $\mathbf{err} = 3\delta^2$ and the assumed range condition (3.9) with $\varphi = \varphi_\nu$ yields by Lemma 3.20 a multiplicative variational inequality (3.15) with $\varphi_{\mathrm{mult}} = \varphi_\nu$. To apply Theorem 3.28 we note that (3.16) is fulfilled for $\nu \in \left(0, \frac{1}{2}\right]$ and that our stopping rule coincides with (6.12) used in the Theorem. Therefore we obtain convergence rates as in the assertion, which are known to be optimal in the linear case. ■

Convergence rates for the iteratively regularized Gauss-Newton method for Hölder-type source conditions with $\nu < \frac{1}{2}$ have been obtained initially in [BNS97] under a slightly different nonlinearity condition.

Moreover, Theorem 6.8 also states rates of convergence under a possibly weaker nonlinearity condition. This applies to the case of norm powers as follows:

REMARK 6.13 (APPLICATION TO THE CASE OF NORM POWERS):
Let us briefly apply Theorem 6.8 to the case where $\mathcal{S}$ is given by the p-th power of a Hilbert norm, i.e. let $\mathcal{S}(g_1; g_2) = \|g_1 - g_2\|_{\mathbb{Y}}^p$ for some $p > 1$ and $\mathcal{R}$ be the square of a Hilbert norm. Moreover assume that an upper bound $\left\|g^\dagger - g^{\mathrm{obs}}\right\|_{\mathbb{Y}} \leq \delta$ is given. Then Lemma 6.4 shows that the usual Lipschitz condition (6.5) implies Assumption 6.2A (as well as 6.2B) with $r = p$, $C_{\mathrm{lip}} = 2^{p-1}$ and $K = 2^{2p-2}L^p$. Moreover, Assumption 3.15 with $\varphi_{\mathrm{add}} = \tilde{\beta}\varphi_{\frac{1}{p}}$ holds true if the spectral source condition (3.9) with $\varphi = \varphi_{\frac{1}{2}}$ is valid. This can be seen by the help of Lemma 3.20 with $\varphi = \varphi_{\frac{1}{2}}$ and a similar argument as in the proof of Lemma 3.21.
By again choosing $\mathbf{err} \equiv \left\|g^\dagger - g^{\mathrm{obs}}\right\|_{\mathbb{Y}}^p$ and $C_{\mathrm{err}} = 2^{p-1}$, we obtain for

$$n_* = \min\left\{ n \in \mathbb{N} \mid \alpha_n \leq \bar{\tau}\delta^{p-1} \right\}$$

with sufficiently large $\bar{\tau} \geq 1$ the following convergence rate provided that $\|\omega\|$ is sufficiently small:

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_{n_*}, u^\dagger\right) = \mathcal{O}(\delta)$$

This result covers in parts the initial work of BAKUSHINSKIĬ [Bak92] on convergence rates for the iteratively regularized Gauss-Newton method.

Logarithmic source conditions

If the index function φ in (3.9) is given by $\varphi = \bar{\varphi}_p$ as in (3.10a) then the corresponding functions φ_{mult} and φ_{add} have already been calculated in Section 3.2.3. This leads to the following convergence rates:

THEOREM 6.14 (CONVERGENCE RATES FOR LOGARITHMIC SOURCE CONDITIONS):
- *Under the Assumptions of Theorem 3.28 with $\varphi_{\mathrm{mult}} = \bar{\varphi}_p$ we obtain the convergence rate*

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_{n_*}, u^\dagger\right) = \mathcal{O}\left(\bar{\varphi}_{2p}\left(\mathbf{err}_{n_*}\right)\right), \qquad \mathbf{err}_{n_*} \searrow 0.$$

- *Under the Assumptions of Theorem 3.30 with* $\varphi_{\text{add}} = \bar{\varphi}_p$ *we obtain the convergence rate*

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_{n_*}, u^\dagger\right) = \mathcal{O}\left(\bar{\varphi}_p\left(\overline{\mathbf{err}}\right)\right), \qquad \overline{\mathbf{err}} \searrow 0.$$

PROOF:
This is obtained by plugging in the functions calculated in Section 3.2.3. ■

Now we will specify this result for the quadratic Hilbert space case:

COROLLARY 6.15 (QUADRATIC HILBERT SPACE CASE):
Assume that $\mathcal{S}(g;\hat{g}) = \|g-\hat{g}\|_{\mathbb{Y}}^2$ *and* $\mathcal{R}(u) = \|u-u_0\|_{\mathbb{X}}^2$ *for Hilbert spaces* $\mathbb{X}$ *and* $\mathbb{Y}$ *and assume that* $F : \mathbb{X} \to \mathbb{Y}$ *is Fréchet differentiable and fulfills the tangential cone condition* (6.3) *with sufficiently small* $\bar{\eta}$*. Let furthermore* u_0 *be sufficiently close to* $u^\dagger$*. Then a range condition* (3.9) *with* $\varphi = \bar{\varphi}_p$ *where* $p > 0$ *and a known upper bound* $\left\|g^\dagger - g^{\text{obs}}\right\|_{\mathbb{Y}} \leq \delta$ *imply for a stopping parameter* n_* *chosen such that* n_* *is the first natural number with* $\alpha_n \bar{\varphi}_{2p}(\alpha_n) \leq \delta$ *the following convergence rate:*

$$\left\|u_{n_*} - u^\dagger\right\|_{\mathbb{X}} = \mathcal{O}\left(\bar{\varphi}_p(\delta)\right), \qquad \delta \searrow 0.$$

PROOF:
As before the tangential cone condition (6.3) implies by Lemma 6.3 the generalized form (6.6b) with some constant C_{tc} and a parameter η which is small if $\bar{\eta}$ from (6.3) is small. As already mentioned, Assumption 3.8 is fulfilled with $\mathbf{err} \equiv \delta^2$, $\mathfrak{s} \equiv 0$ and $C_{\text{err}} = 2$. The term $\mathbf{err}_n$ from (6.11b) is hence given by $\mathbf{err} = 3\delta^2$ and the assumed range condition (3.9) with $\varphi = \bar{\varphi}_p$ yields by Lemma 3.20 a multiplicative variational inequality (3.15) with $\varphi_{\text{mult}} = \bar{\varphi}_p$. To apply Theorem 3.28 we note that (3.16) is fulfilled trivially and that our stopping rule coincides with (6.12) used in the Theorem. Therefore we obtain convergence rates as in the assertion, which are known to be optimal in the linear case. ■

Convergence rates of this type have first been proven by HOHAGE [Hoh97] under a slightly different nonlinearity condition.

The general case for norm powers

Finally we will now relate our results to the case of general norm powers and general source conditions as we presented them in Chapter 5.

COROLLARY 6.16 (CASE OF GENERAL NORM POWERS):
Assume that $\mathcal{S}(g;\hat{g}) = \|g-\hat{g}\|_{\mathbb{Y}}^r$ *and* $\mathcal{R}(u) = \|u-u_0\|_{\mathbb{X}}^r$ *for Banach spaces* $\mathbb{X}$ *and* $\mathbb{Y}$ *and* $r, p > 1$*. Assume moreover that* $F : \mathbb{X} \to \mathbb{Y}$ *is Fréchet differentiable and fulfills the tangential cone condition* (6.3) *with sufficiently small* $\bar{\eta}$*. Let furthermore* u_0 *be sufficiently close to* $u^\dagger$ *and Assumption 3.24 hold true. Then a known upper bound* $\left\|g^\dagger - g^{\text{obs}}\right\|_{\mathbb{Y}} \leq \delta$ *implies for a proper chosen stopping parameter* n_* *the convergence rate*

$$\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_{n_*}, u^\dagger\right) = \mathcal{O}\left(\varphi_{\text{mult}}^2\left(\Theta^{-1}\left(\delta^r\right)\right)\right), \qquad \delta \searrow 0$$

with Θ *as in* (3.19).

PROOF:
In view of the fact that Assumption 3.8 is fulfilled with $\mathbf{err} \equiv \delta^r$, this follows from Lemma 6.3 and Theorem 6.5. ■

In comparison with Theorem 5.5, our nonlinearity condition is weaker as discussed in Section 6.1. Moreover our variational inequality (3.15) does not contain derivatives of F as (5.12) does. Nevertheless, our variational inequality contains different powers of the $\mathbb{Y}$-norm depending on the exponent r in the data fidelity term. This weakens or tightens the condition depending if $r < 2$ or $r > 2$. The variational inequalities coincide for $r = 2$, and in this case also the obtained rates do.

6.3 Convergence rates with an a posteriori stopping rule

As in Section 3.3 we now want to apply the Lepskiĭ balancing principle in case of an additive variational inequality (3.11). The result of this section (cf. Theorem 6.19) is a slight generalization of [HW11, Thm. 4.2] and is based on the more general representation of the approximation error introduced by GRASMAIR [Gra10a] and FLEMMING [Fle11].
The way we introduced the Lepskiĭ principle in Section 3.3 covers only an error decomposition with two terms, where in case of an iteratively regularized Newton method (6.2) usually three terms occur, namely the well-known approximation and propagated data noise error and additionally a **nonlinearity** error. As in Section 3.3 we need to restrict $\mathcal{R}$ such that the error measure in $\mathbb{X}$ is a metric. We will again assume that there exists some constant $C_{\text{bd}} < \infty$ and some exponent $q > 1$ such that

$$\left\|u - u^{\dagger}\right\|_{\mathbb{X}}^{q} \leq C_{\text{bd}} \mathcal{D}_{\mathcal{R}}^{u^*}\left(u, u^{\dagger}\right) \qquad \text{for all } u \in \mathfrak{B}. \tag{6.30}$$

The next Lemma generalizes the Lepskiĭ-type balancing principle as done by BAUER ET AL. [BHM09] to the aforementioned case of three error terms:

LEMMA 6.17 (LEPSKIĬ-TYPE BALANCING PRINCIPLE WITH THREE TERMS):
Assume that a sequence $(u_n)_{n\in\mathbb{N}}$ is given such that

$$\left\|u_n - u^{\dagger}\right\|_{\mathbb{X}} \leq \Phi_{\text{nl}}(n) + \Phi_{\text{app}}(n) + \Phi_{\text{noi}}(n) \tag{6.31}$$

holds true for all $n \in \mathbb{N}$ where

(a) Φ_{noi} is a non-decreasing known function fulfilling

$$\Phi_{\text{noi}}(n+1) \leq D\Phi_{\text{noi}}(n) \qquad \textit{for all} \qquad n \in \mathbb{N}, \tag{6.32}$$

(b) Φ_{app} is a non-increasing unknown function and

(c) Φ_{nl} is an unknown function fulfilling

$$\Phi_{\text{nl}}(n) \leq \gamma_{\text{nl}}\left(\Phi_{\text{noi}} + \Phi_{\text{app}}\right) \qquad \textit{for all} \qquad n \in \mathbb{N} \tag{6.33}$$

with some constants $D \geq 1$, $\gamma_{\text{nl}} > 0$.
Then the Lepskiĭ-type balancing principle

$$N_{\max} := \min\left\{n \in \mathbb{N} \mid \Phi_{\text{noi}}(n) \geq 1\right\}, \tag{6.34a}$$

$$n_{\text{bal}} := \min\left\{n \leq N_{\max} \mid \|u_n - u_m\|_{\mathbb{X}} \leq 4\left(1+\gamma_{\text{nl}}\right)\Phi_{\text{noi}}(m) \quad \textit{for all} \quad n < m \leq N_{\max}\right\} \tag{6.34b}$$

leads under the additional condition $\Phi_{\text{app}}(N_{\max}) \leq 1$ to the estimate

$$\left\|u_{n_{\text{bal}}} - u^{\dagger}\right\|_{\mathbb{X}} \leq 6D\left(1+\gamma_{\text{nl}}\right) \min_{n=1,\ldots,N_{\max}} \left(\Phi_{\text{app}}(n) + \Phi_{\text{noi}}(n)\right). \tag{6.35}$$

PROOF:
Inserting (6.33) into (6.31) yields

$$\left\| u_n - u^\dagger \right\|_{\mathbb{X}} \leq (1+\gamma_{\mathrm{nl}}) \left(\Phi_{\mathrm{app}}(n) + \Phi_{\mathrm{noi}}(n)\right) \qquad \text{for all} n \in \mathbb{N}. \tag{6.36}$$

Now define $m := N_{\max}$ and for $j = 1, ..., m$ set

$$\psi(j) := 2(1+\gamma_{\mathrm{nl}}) \Phi_{\mathrm{noi}}(N_{\max} + 1 - j),$$
$$\phi(j) := 2(1+\gamma_{\mathrm{nl}}) \Phi_{\mathrm{app}}(N_{\max} + 1 - j).$$

Then

- ϕ is unknown and non-decreasing and
- ψ is known and non-increasing and fulfills by (6.32)

$$\begin{aligned}\psi(i) &= 2(1+\gamma_{\mathrm{nl}}) \Phi_{\mathrm{noi}}(N_{\max}+1-i) \\ &= 2(1+\gamma_{\mathrm{nl}}) \Phi_{\mathrm{noi}}(N_{\max}+1-(i-1)+1) \\ &\leq D2(1+\gamma_{\mathrm{nl}}) \Phi_{\mathrm{noi}}(N_{\max}+1-(i-1)) \\ &= D\psi(i+1)\end{aligned}$$

  for all $1 \leq i \leq m-1$.

Moreover, by $\Phi_{\mathrm{app}}(N_{\max}) \leq 1$ we find $\phi(1) \leq \psi(1)$ and due to (6.31) it holds with $x_j := u_{N_{\max}+1-j}$, $j = 1, ..., N_{\max}$ the estimate

$$\left\| x_j - u^\dagger \right\|_{\mathbb{X}} \leq \frac{1}{2}(\psi(j) + \phi(j)), \qquad 1 \leq j \leq m.$$

Therefore, all assumptions of Lemma 3.40 are fulfilled and we find for the index

$$\bar{j} = \max\left\{ j \leq m \mid \left\| x_i - x_j \right\|_{\mathbb{X}} \leq 2\psi(i) \text{ for all } i < j \right\}. \tag{6.37}$$

the estimate

$$\left\| x_{\bar{j}} - u^\dagger \right\|_{\mathbb{X}} \leq 3D \min\left\{ \phi(j) + \psi(j) \mid j \in \{1, ..., m\} \right\}. \tag{6.38}$$

Now some easy index manipulations show that the indices from (6.37) and (6.34) are related via $\bar{j} = N_{\max} + 1 - n_{\mathrm{bal}}$ and hence we have

$$x_{\bar{j}} = u_{n_{\mathrm{bal}}}.$$

Therefore, (6.38) implies (6.35) and the assertion is proven. ■

To apply this principle to our case, let us rewrite the error decomposition already obtained in Theorem 6.6:

LEMMA 6.18 (ERROR DECOMPOSITION UNDER ASSUMPTION 3.15):
Let Assumptions 3.8, 6.1A or 6.1B and 3.15 with $\beta \in \left[0, \frac{1}{2}\right]$ *hold true. Moreover let* (6.30) *be fulfilled and assume that there exists a uniform upper bound* $\mathbf{err}_n \leq \overline{\mathbf{err}}$ *for all* $n \in \mathbb{N}$ *where* $\mathbf{err}_n$ *is given as in* (6.11)*. Then the error of the iterates* u_n *defined by* (6.2) *can be bounded by the sum of an* approximation error *bound* $\Phi_{\mathrm{app}}(n)$*, a* propagated data noise error *bound* $\Phi_{\mathrm{noi}}(n)$ *and a* nonlinearity error *bound* $\Phi_{\mathrm{nl}}(n)$*,*

$$\left\| u_n - u^\dagger \right\|_{\mathbb{X}} \leq \Phi_{\mathrm{nl}}(n) + \Phi_{\mathrm{app}}(n) + \Phi_{\mathrm{noi}}(n) \tag{6.39}$$

where

$$\Phi_{\mathrm{nl}}(n) := (4\eta C_{\mathrm{err}} C_{\mathrm{bd}})^{\frac{1}{q}} \left(\frac{s_n}{\alpha_n}\right)^{\frac{1}{q}},$$

$$\Phi_{\mathrm{app}}(n) := (4C_{\mathrm{tc}} C_{\mathrm{err}} C_{\mathrm{bd}})^{\frac{1}{q}} (-\varphi_{\mathrm{add}})^* \left(-\frac{1}{\alpha_n}\right)^{\frac{1}{q}},$$

$$\Phi_{\mathrm{noi}}(n) := (2C_{\mathrm{bd}})^{\frac{1}{q}} \left(\frac{\overline{\mathbf{err}}}{\alpha_n}\right)^{\frac{1}{q}}.$$

Moreover, if η and $\mathcal{S}\left(F(u_0); g^\dagger\right)$ are sufficiently small, the estimate

$$\Phi_{\mathrm{nl}}(n) \leq \bar{\gamma}_{\mathrm{nl}} \left(\Phi_{\mathrm{noi}}(n) + \Phi_{\mathrm{app}}(n)\right) \tag{6.40}$$

holds true for all $n \in \mathbb{N}_0$ with

$$\bar{\gamma}_{\mathrm{nl}} = (2\eta C_{\mathrm{err}} \gamma_{\mathrm{nl}})^{\frac{1}{q}} = \left(\frac{4\eta C_{\mathrm{tc}} C_{\mathrm{err}}^2 C_{\mathrm{dec}}^2}{1 - 4\eta C_{\mathrm{tc}} C_{\mathrm{err}}^2 C_{\mathrm{dec}}^2}\right)^{\frac{1}{q}}. \tag{6.41}$$

PROOF:
Under the made assumptions Theorem 6.6 implies by (6.20) that

$$(1-\beta) d_{n+1}^2 + \frac{1}{2C_{\mathrm{err}} C_{\mathrm{tc}} \alpha_n} s_{n+1} \leq 2\eta C_{\mathrm{err}} \frac{s_n}{\alpha_n} + 2C_{\mathrm{tc}} C_{\mathrm{err}} (-\varphi_{\mathrm{add}})^* \left(-\frac{1}{\alpha_n}\right) + \frac{\mathbf{err}_n}{\alpha_n}$$

for all $n \in \mathbb{N}$. Now by rearranging, using (6.30), $\beta \in \left[0, \frac{1}{2}\right]$ and $\mathbf{err}_n \leq \overline{\mathbf{err}}$ we find especially

$$\left\|u_n - u^\dagger\right\|_{\mathbb{X}}^q \leq 4\eta C_{\mathrm{err}} C_{\mathrm{bd}} \frac{s_n}{\alpha_n} + 4C_{\mathrm{tc}} C_{\mathrm{err}} C_{\mathrm{bd}} (-\varphi_{\mathrm{add}})^* \left(-\frac{1}{\alpha_n}\right) + 2C_{\mathrm{bd}} \frac{\overline{\mathbf{err}}}{\alpha_n}$$

for all $n \in \mathbb{N}$. Taking the q-th root and using the concavity of $t \mapsto t^{\frac{1}{q}}$ this shows (6.39). Moreover by Theorem 6.6 it holds (6.21), i.e.

$$s_n \leq \gamma_{\mathrm{nl}} \left(2C_{\mathrm{tc}} C_{\mathrm{err}} \alpha_n (-\varphi_{\mathrm{add}})^* \left(-\frac{1}{\alpha_n}\right) + \overline{\mathbf{err}}\right)$$

for all $n \in \mathbb{N}_0$. Multiplying this estimate with $\frac{4\eta C_{\mathrm{err}} C_{\mathrm{bd}}}{\alpha_n}$ and taking the q-th root yields

$$\Phi_{\mathrm{nl}}(n) \leq (2\eta C_{\mathrm{err}} \gamma_{\mathrm{nl}})^{\frac{1}{q}} \left(4C_{\mathrm{tc}} C_{\mathrm{err}} C_{\mathrm{bd}} (-\varphi_{\mathrm{add}})^* \left(-\frac{1}{\alpha_n}\right) + 2C_{\mathrm{bd}} \frac{\overline{\mathbf{err}}}{\alpha_n}\right)^{\frac{1}{q}}$$

for all $n \in \mathbb{N}_0$. Again using the concavity of $t \mapsto t^{\frac{1}{q}}$ this shows (6.40). ■

THEOREM 6.19 (CONVERGENCE RATES FOR A LEPSKIĬ-TYPE STOPPING RULE):
Let the assumptions of Lemma 6.18 hold true. Then the Lepskiĭ balancing principle (6.34) *leads to the convergence rate*

$$\left\|u_{n_{\mathrm{bal}}} - u^\dagger\right\|_{\mathbb{X}}^q = \mathcal{O}\left(\varphi_{\mathrm{add}}\left(\overline{\mathbf{err}}\right)\right), \qquad \mathbf{err} \searrow 0.$$

PROOF:
Assume in the following that $\overline{\mathbf{err}}$ is sufficiently small. Then $N_{\max}$ is sufficiently large to ensure $\Phi_{\text{app}}(N_{\max}) \leq 1$. Moreover, from Lemma 6.18 and Remark 3.29 we find that all assumptions of Lemma 6.17 are fulfilled and hence it holds

$$\left\| u_{n_{\text{bal}}} - u^{\dagger} \right\|_{\mathbb{X}} \leq 6 \left(1 + \gamma_{\text{nl}}\right) C_{\text{dec}}^{\frac{1}{q}} \min_{1 \leq n \leq N_{\max}} \left(\Phi_{\text{app}}(n) + \Phi_{\text{noi}}(n)\right).$$

Taking the q-th exponent and using $(a+b)^q \leq 2^q \left(a^q + b^q\right)$ it follows that

$$\begin{aligned}\left\| u_{n_{\text{bal}}} - u^{\dagger} \right\|_{\mathbb{X}}^{q} &\leq 12^q \left(1 + \gamma_{\text{nl}}\right)^q C_{\text{dec}} \min_{1 \leq n \leq N_{\max}} \left(\Phi_{\text{app}}(n)^q + \Phi_{\text{noi}}(n)^q\right) \\ &\leq 4 \cdot 12^q \left(1 + \gamma_{\text{nl}}\right)^q C_{\text{tc}} C_{\text{err}} C_{\text{bd}} C_{\text{dec}} \min_{1 \leq n \leq N_{\max}} \left(\left(-\varphi_{\text{add}}\right)^* \left(-\frac{1}{\alpha_n}\right) + \frac{\overline{\mathbf{err}}}{\alpha_n}\right).\end{aligned}$$

If we can show that $n_* - 1$ with n_* as in (6.23) is an element of $\{1, ..., N_{\max}\}$ then the assertion follows as in the proof of Theorem 6.6 from (6.25). Note that by definition it holds

$$\frac{1}{\alpha_{n_*-1}} < -\inf \partial \left(-\varphi\right) \left(\overline{\mathbf{err}}\right)$$

and hence

$$\begin{aligned}\Phi_{\text{noi}}\left(n_* - 1\right) &= \left(2C_{\text{bd}}\right)^{\frac{1}{q}} \left(\frac{\overline{\mathbf{err}}}{\alpha_{n_*-1}}\right)^{\frac{1}{q}} \\ &< \left(2C_{\text{bd}}\right)^{\frac{1}{q}} \left(\overline{\mathbf{err}} \cdot \left(-\inf \partial \left(-\varphi\right) \left(\overline{\mathbf{err}}\right)\right)\right)^{\frac{1}{q}}.\end{aligned}$$

As in the proof of Lemma 3.42 it can be seen that the right-hand side tends to 0 as $\overline{\mathbf{err}} \searrow 0$ and so we may assume that $\Phi_{\text{noi}}\left(n_* - 1\right) < 1$, which proves $n_* - 1 \leq N_{\max}$ and hence the assertion. ■

CHAPTER
SEVEN

ITERATIVELY REGULARIZED NEWTON METHODS WITH POISSON DATA

In this chapter we will apply the results on iteratively regularized Newton methods (6.2a) to the case of Poisson data as discussed in Chapter 2 where the data fidelity term $\mathcal{S}$ is chosen to be a variant of the negative log-likelihood. These results for the case of a Poisson process as data are new in the contents as well as in the presented way to our best knowledge.
For the whole chapter let the data fidelity terms $\mathcal{S}\left(\cdot; g^\dagger\right)$ and $\mathcal{S}\left(\cdot; g^{\mathrm{obs}}\right)$ w.r.t. exact and noisy data again be given by (2.14) respectively with some fixed $e > 0$. As in Chapter 4 we will use only Assumption 3.15 as smoothness condition on $u^\dagger$, but similar results can be obtained under Assumption 3.24.

7.1 Preliminaries

As in Chapter 4 we want to motivate our smoothness conditions on $u^\dagger$ in terms of variational inequalities before presenting the results. This is done in the same manner as Corollaries 4.6, 4.7 and 4.8 do, but a motivation for a linear operator is not necessary in this section. Moreover, we will not use the classical tangential cone condition (6.3), but the generalized version (6.6b), which is used anyway in our convergence analysis.

COROLLARY 7.1:
Let $\mathbb{X}$ be a Hilbert space and let Assumptions 2.7 and 6.1B be fulfilled. If the spectral source condition (3.9) holds true with some general φ such that φ^2 is concave, then Assumption 3.15 is valid with φ_{add} as specified in Theorem 3.18 (up to a multiplicative constant) and $\mathcal{R}(u) = \|u - u_0\|_{\mathbb{X}}^2$ and $\mathcal{S} = \mathbb{KL}_e$.

PROOF:
Theorem 3.18 yields an additive variational inequality of the form

$$\left\langle u^*, u^\dagger - u\right\rangle \leq \beta \left\|u - u^\dagger\right\|_{\mathbb{X}}^2 + \varphi_{\mathrm{add}}\left(\left\|F'\left[u^\dagger\right]\left(u - u^\dagger\right)\right\|_{\mathbf{L}^2(\Omega,\mu)}^2\right) \qquad \text{for all } u \in \mathbb{X}\,.$$

Moreover due to (4.10) we can bound the term $\left\|F'\left[u^\dagger\right]\left(u - u^\dagger\right)\right\|_{\mathbf{L}^2(\Omega,\mu)}^2$ from above up to some constant by $\mathbb{KL}_e\left(g^\dagger; F\left(u^\dagger\right) + F'\left[u^\dagger\right]\left(u - u^\dagger\right)\right)$, which is due to Assumption 6.1B bounded by $C_{\mathrm{tc}}\,\mathbb{KL}_e\left(g^\dagger; F(u)\right)$. Hence we find as in the proof of Corollary (4.7) that

$$\left\langle u^*, u^\dagger - u\right\rangle \leq \beta \left\|u - u^\dagger\right\|_{\mathbb{X}}^2 + c\varphi_{\mathrm{add}}\left(\mathcal{S}\left(F(u); g^\dagger\right)\right) \qquad \text{for all } u \in \mathfrak{B}$$

for some constant $c > 0$, i.e. Assumption 3.15 is valid. ■

The Corollary above shows that our smoothness assumption (3.11) is implied by (3.9) provided the nonlinearity condition (6.6b) holds true.

7.2 Convergence rates

This section provides the second main result of this thesis, namely convergence rates in expectation for iteratively regularized Newton methods with full random data. The result splits into two parts, namely rates under an a priori stopping rule and an a posteriori stopping rule. Due to the usage of the concentration inequality (4.1) the assumption $e > 0$ made at the beginning of this chapter is unavoidable.
The first step is to bound the error terms $\mathbf{err}_n$ as in (6.11b) in probability. Remember that $\mathbf{err}_n$ was defined as the sum of two error terms $\mathbf{err}(g)$ with g being the first order Taylor expansion of F around u_n evaluated at $u^\dagger$ and u_n respectively where

$$\mathbf{err}(g) = \left| \int_\Omega \ln(g+e)\left(\frac{1}{t}\,\mathrm{d}G_t - g^\dagger\,\mathrm{d}x\right)\right|$$

if $g \geq -\frac{e}{2}$ a.e. To apply the concentration inequality we need hence to show that

$$\ln\left(F(u_n) + F'(u_n; v-u_n) + e\right) \in B_s(R) \qquad \text{for all} \qquad n \in \mathbb{N} \text{ and } v \in \left\{u_{n+1}, u^\dagger\right\}$$

with some constant R and **moreover** that $F(u_n) + F'(u_n; v-u_n) \geq -\frac{e}{2}$ a.e. for all $n \in \mathbb{N}$ and $v \in \{u_{n+1}, u^\dagger\}$. Both things are done by an analog to Corollary 4.3:

COROLLARY 7.2:
Let Assumptions 2.7 and 6.1B hold true and assume moreover that F maps $\mathfrak{B}$ into the Sobolev space $H^s(\Omega)$ with $s > \frac{d}{2}$ such that

$$R := \sup_{u,v\in\mathfrak{B}} \left\|F(u) + F'[u](v-u)\right\|_{H^s(\Omega)} < \infty. \tag{7.1}$$

Then there exist $C_{\mathrm{conc}}, C_\rho \geq 1$ depending only on Ω and s such that

$$\begin{aligned} &\mathbf{P}\left(\sup_{u,v\in\mathfrak{B}} \mathbf{err}\left(F(u) + F'(u; v-u)\right) \leq \rho\psi(t)\right) \\ &\geq 1 - \exp\left(-\frac{\rho}{R\max\{e^{-s}, \ln(R)\}\,C_{\mathrm{conc}}}\right) \end{aligned} \tag{7.2}$$

for all $t \geq 1, \rho \geq R\max\{e^{-s}, \ln(R)\}\,C_\rho$.

PROOF:
From Assumption 2.7 we conclude that $\mathcal{S}_{e,t}\left(F(u); g^\dagger\right) < \infty$ for all $u \in \mathfrak{B}$. Hence, Assumption 6.1B implies $\mathcal{S}_{e,t}\left(F(u) + F'(u; v-u); g^\dagger\right) < \infty$ for all $u, v \in \mathfrak{B}$ and hence by definition $F(u) + F'(u; v-u) \geq -\frac{e}{2}$ a.e. for all $u, v \in \mathfrak{B}$. Since $s > \frac{d}{2}$ we have moreover $F(u) + F'(u; v-u) \leq \|E_\infty\|\, R$ a.e. for all $u, v \in \mathfrak{B}$. Now the assertion is proven similarly to Corollary 4.3. ■

THEOREM 7.3:
Let the Assumptions 2.7, 6.1B and 3.15 be satisfied and $F : \mathfrak{B} \to H^2(\Omega)$ such that (7.1) is fulfilled with $s > \frac{d}{2}$. Moreover assume that η and $\mathbb{KL}_e(g^\dagger; F(u_0))$ are sufficiently small. If we choose the stopping index n_* by

$$n_* = \min\left\{ n \in \mathbb{N} \;\middle|\; \frac{1}{\alpha_n} \geq -\inf \partial(-\varphi)(\psi(t)) \right\} \tag{7.3}$$

we obtain the convergence rate

$$\mathbf{E}\left(\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_{n_*}, u^\dagger\right)\right) = \mathcal{O}(\varphi(\psi(t))), \qquad t \to \infty. \tag{7.4}$$

PROOF:
Under the posed conditions we find from Theorem 6.6 by plugging (6.20) and (6.21) together the estimate

$$(1-\beta)\mathcal{D}_{\mathcal{R}}^{u^*}\left(u_{n_*}, u^\dagger\right) \leq (1+2\eta\gamma_{\mathrm{nl}})\left(2C_{\mathrm{tc}}(-\varphi_{\mathrm{add}})^*\left(-\frac{1}{\alpha_{n_*-1}}\right) + \frac{\overline{\mathbf{err}}}{\alpha_{n_*-1}}\right) \tag{7.5}$$

where $\overline{\mathbf{err}}$ is due to the definition of $\mathbf{err}_n$ in (6.11b) given by

$$\overline{\mathbf{err}} = 2 \sup_{u,v\in\mathfrak{B}} \mathbf{err}\left(F(u) + F'(u; v-u)\right).$$

As in the proof of Theorem 4.11 we choose

$$\rho_k := R \max\left\{e^{-s}, |\ln(R)|\right\} C_\rho k, \qquad k \in \mathbb{N}$$

where C_ρ is the constant from Corollary 7.2 and define the events

$$E_k := \left\{ \sup_{u,v\in\mathfrak{B}} \mathbf{err}\left(F(u) + F'(u; v-u)\right) \leq \rho_k \psi(t) \right\}, \quad k \in \mathbb{N}.$$

From (7.2) it is known that

$$\mathbf{P}\left(E_k^c\right) \leq \exp\left(-\frac{C_\rho}{C_{\mathrm{conc}}} k\right).$$

Moreover, we find from (7.5) that

$$\begin{aligned}\max_{E_k} \mathcal{D}_{\mathcal{R}}^{u^*}\left(u_{n_*}, u^\dagger\right) &\leq \frac{2C_{\mathrm{tc}}(1+2\eta\gamma_{\mathrm{nl}})}{1-\beta}\left((-\varphi_{\mathrm{add}})^*\left(-\frac{1}{\alpha_{n_*-1}}\right) + \frac{\rho_k\psi(t)}{\alpha_{n_*-1}}\right) \\ &\leq \frac{2C_{\mathrm{tc}}(1+2\eta\gamma_{\mathrm{nl}})\rho_k}{1-\beta}\left((-\varphi_{\mathrm{add}})^*\left(-\frac{1}{\alpha_{n_*-1}}\right) + \frac{\psi(t)}{\alpha_{n_*-1}}\right).\end{aligned}$$

Now (6.25) implies for $\overline{\mathbf{err}}$ replaced by $\psi(t)$ that under the stopping rule (7.3) we have

$$\max_{E_k} \mathcal{D}_{\mathcal{R}}^{u^*}\left(u_{n_*}, u^\dagger\right) \leq C(k)\,\varphi_{\mathrm{add}}(\psi(t))$$

for all $k \in \mathbb{N}$ with $C(k) = \frac{2C_{\mathrm{tc}}(1+2\eta\gamma_{\mathrm{nl}})\rho_k}{1-\beta}$. The sum $\sum_{k=2}^{\infty} \exp\left(-\frac{C_\rho}{C_{\mathrm{conc}}}(k-1)\right) C(k)$ converges since $C(k) \sim k$ as $k \to \infty$. This allows us to apply Lemma 4.10 with $\Xi(t) = \varphi_{\mathrm{add}}(\psi(t))$, $c = \frac{C_\rho}{C_{\mathrm{conc}}}$ and $d_{n(t)} = \mathcal{D}_{\mathcal{R}}^{u^*}\left(u_{n_*}, u^\dagger\right)$ which proves (7.4). ■

THEOREM 7.4:
Let the Assumptions 2.7, 6.1B and 3.15 be satisfied and $F : \mathfrak{B} \to H^2(\Omega)$ *such that* (7.1) *is fulfilled with* $s > \frac{d}{2}$*. Moreover assume that* η *and* $\mathbb{KL}_e\left(g^\dagger; F(u_0)\right)$ *are sufficiently small. If* $\beta \in \left[0, \frac{1}{2}\right]$, (3.44) *is fulfilled and* φ_{add} *is such that*

$$\ln(t) \cdot \varphi_{\text{add}}(\psi(t)) \searrow 0 \qquad as \qquad t \to \infty$$

we define

$$\rho(t) := -\tau \ln(\psi(t)) = \frac{\tau}{2}\ln(t),$$

$$\Phi_{\text{noi}}^{\text{det}}(n) := (4C_{\text{bd}})^{\frac{1}{q}} \left(\frac{\rho(t)\,\psi(t)}{\alpha_n}\right)^{\frac{1}{q}}, \tag{7.6a}$$

$$N_{\max} := \min\left\{n \in \mathbb{N} \mid \Phi_{\text{noi}}^{\text{det}}(n) \geq 1\right\}, \tag{7.6b}$$

$$n_{\text{bal}} := \min\left\{n \in \{1, \dots, N_{\max}\} \mid \forall m \geq n \;\; \|u_n - u_m\| \leq 4\left(1 + \bar{\gamma}_{\text{nl}}\right)\Phi_{\text{noi}}^{\text{det}}(m)\right\} \tag{7.6c}$$

with a tuning parameter $\tau \geq \frac{1}{2}R\max\left\{e^{-s}, |\ln(R)|\right\}C_{\text{conc}}$ *and* $\bar{\gamma}_{\text{nl}}$ *as specified in* (6.41) *for* $C_{\text{err}} = 1$*. Then this a posteriori Lepskiĭ-type stopping rule implies for all sufficiently large t the estimate*

$$\mathbf{E}\left(\left\|u_{n_{\text{bal}}} - u^\dagger\right\|_{\mathbb{X}}^q\right) \leq \left(\bar{C}\rho(t) + \operatorname{diam}(\mathfrak{B})^q\right)\varphi(\psi(t)).$$

with a constant $\bar{C}$ *independent of t and hence*

$$\mathbf{E}\left(\left\|u_{n_{\text{bal}}} - u^\dagger\right\|_{\mathbb{X}}^q\right) = \mathcal{O}\left(\ln(t) \cdot \varphi_{\text{add}}(\psi(t))\right), \qquad t \to \infty. \tag{7.7}$$

PROOF:
Lemma 6.18 implies under the given assumptions the error decomposition

$$\left\|u_n - u^\dagger\right\|_{\mathbb{X}} \leq \Phi_{\text{nl}}(n) + \Phi_{\text{app}}(n) + \Phi_{\text{noi}}(n) \tag{7.8}$$

where

$$\Phi_{\text{nl}}(n) := (4\eta C_{\text{err}} C_{\text{bd}})^{\frac{1}{q}} \left(\frac{s_n}{\alpha_n}\right)^{\frac{1}{q}},$$

$$\Phi_{\text{app}}(n) := (4C_{\text{tc}} C_{\text{err}} C_{\text{bd}})^{\frac{1}{q}} (-\varphi_{\text{add}})^* \left(-\frac{1}{\alpha_n}\right)^{\frac{1}{q}},$$

$$\Phi_{\text{noi}}(n) := (4C_{\text{bd}})^{\frac{1}{q}} \left(\frac{\sup\limits_{u,v \in \mathfrak{B}} \mathbf{err}\left(F(u) + F'(u; v-u)\right)}{\alpha_n}\right)^{\frac{1}{q}}$$

and Φ_{nl} fulfills

$$\Phi_{\text{nl}}(n) \leq \bar{\gamma}_{\text{nl}}\left(\Phi_{\text{noi}}(n) + \Phi_{\text{app}}(n)\right)$$

for all $n \in \mathbb{N}_0$. Let t be so large, that the assumptions from Corollary 7.2 hold true. Now consider the event

$$A_\rho = \left\{\sup_{u,v \in \mathfrak{B}} \mathbf{err}\left(F(u) + F'(u; v-u)\right) \leq \rho(t)\,\psi(t)\right\}$$

which fulfills $\mathbf{P}\left(A_\rho\right) \geq 1 - \exp\left(-c\rho\left(t\right)\right)$ with $c = \frac{1}{R\max\{e^{-s},|\ln(R)|\}C_{\mathrm{conc}}}$ by (7.2). Then on A_ρ we may replace Φ_{noi} in the error decomposition (7.8) by $\Phi_{\mathrm{noi}}^{\mathrm{det}}$ from (7.6a). Now the Lepskiĭ-type balancing principle (7.6) yields as in the proof Theorem 6.19 the estimate

$$\left\|u_{n_{\mathrm{bal}}} - u^\dagger\right\|_{\mathbb{X}} \leq 6\left(1+\bar{\gamma}_{\mathrm{nl}}\right) C_{\mathrm{dec}}^{\frac{1}{q}} \min_{n\leq N_{\max}} \left(\Phi_{\mathrm{app}}\left(n\right) + \Phi_{\mathrm{noi}}^{\mathrm{det}}\left(n\right)\right) \tag{7.9}$$

on A_ρ with $\bar{\gamma}_{\mathrm{nl}}$ as in (6.41). It can be seen in the same way as in the proof of Theorem 4.12 that

$$\mathbf{E}\left(\left\|u_{n_{\mathrm{bal}}} - u^\dagger\right\|_{\mathbb{X}}^q \;\middle|\; A_\rho\right) \leq \bar{C}\rho\left(t\right)\varphi_{\mathrm{add}}\left(\psi\left(t\right)\right)$$

for all t sufficiently large with some constant $\bar{C}$ independent of t. Hence

$$\begin{aligned}\mathbf{E}\left(\left\|u_{n_{\mathrm{bal}}} - u^\dagger\right\|_{\mathbb{X}}^q\right) &= \mathbf{P}\left(A_\rho\right)\mathbf{E}\left(\left\|u_{n_{\mathrm{bal}}} - u^\dagger\right\|_{\mathbb{X}}^q \mid A_\rho\right) + \mathbf{P}\left(A_\rho^c\right)\mathbf{E}\left(\left\|u_{n_{\mathrm{bal}}} - u^\dagger\right\|_{\mathbb{X}}^q \mid A_\rho^c\right) \\ &\leq \bar{C}\rho\left(t\right)\varphi\left(\psi\left(t\right)\right) + \exp\left(-\rho\left(t\right)\right)\operatorname{diam}\left(\mathfrak{B}\right)^q. \end{aligned} \tag{7.10}$$

Due to the definition of $\rho\left(t\right)$ and the choice of τ we can for all sufficiently large t furthermore assume that

$$\exp\left(-c\rho\left(t\right)\right) = \psi\left(t\right)^{c\tau} \leq C\varphi_{\mathrm{add}}\left(\psi\left(t\right)\right).$$

for some constant $C > 0$ independent of t since φ_{add}^2 is concave and $c\tau \geq \frac{1}{2}$. Inserting this into (7.10) yields the claim. ■

7.3 General convergence

To finish this chapter, we want to comment on the general convergence of (6.2) under Poisson data and compare them with the results for Tikhonov-type regularization in Section 4.3. As opposed to the result obtained there, we need to ensure that the nonlinearity of F fits together with the data fidelity term $\mathcal{S}$, which can no longer be guaranteed by the variational inequality. So we need to suppose additionally that Assumption 6.1B holds true. Assume again that $\mathbb{X}$ is a Hilbert space and $\mathcal{R}$ is given by the squared norm in $\mathbb{X}$.

REMARK 7.5 (REGULARIZATION PROPERTIES):
Let $F : \mathbb{X} \to \mathbb{Y}$ be an operator fulfilling Assumptions 2.7 and 6.1B. As before (cf. Section 4.3) we find the existence of some index function φ such that (3.9) is fulfilled. The function φ can be chosen such that φ^2 is concave by possibly changing ω. As in the proof of Corollary (7.1) we find that Assumption 3.15 is valid.
Thus for sufficiently small η and an initial guess which is close to $u^\dagger$ we find for a proper chosen stopping index n_* convergence $\left\|u_{n_*} - u^\dagger\right\|_{\mathbb{X}} \to 0$ as $t \to \infty$ for the regularized solutions u_n gained by (6.2) with Poisson data G_t as described in Chapter 2.

Moreover if the valid spectral source condition is strong enough (i.e. such that (4.15) is fulfilled), then the a posteriori choice of n_{bal} given by the Lepskiĭ-type balancing principle (7.6c) yields convergence $\left\|u_{n_{\mathrm{bal}}} - u^\dagger\right\|_{\mathbb{X}} \to 0$ as $t \to \infty$.

CHAPTER
EIGHT

NUMERICAL EXAMPLES

In this chapter we will apply the proposed iteratively regularized Newton method for Poisson data to three problems from photonic imaging. Note that all examples from Chapter 2 lead to a nonlinear forward operator F and hence the Tikhonov-type functional in (3.2) is not convex. Thus, we do not use Tikhonov-type regularization. The implementation of the iteratively regularized Newton method in a continuous setup as well as the first two examples of an inverse scattering problem without phase and a phase retrieval problem in optics (cf. Sections 2.2.1 and 2.2.2) have already been described in [HW11]. For these two examples, we present detailed numeric simulations and tests to illustrate the performance of our method. Finally we present a result for a third problem, namely a semiblind deconvolution problem from 4Pi microscopy (cf. Section 2.2.3).

8.1 Implementation of the algorithm

For a numerical realization of our proposed method (6.2) with $\mathcal{S} = \mathcal{S}_{e,t}$ as defined in (2.14) we need to solve a convex minimization problem in every Newton step, i.e. we need to compute

$$u_{n+1} = \underset{u\in\mathfrak{B}}{\operatorname{argmin}} \left[\int_\Omega \left(F(u_n) + F'(u_n; u-u_n)\right) \,\mathrm{d}x - \int_\Omega \ln\left(F(u_n) + F'(u_n; u-u_n) + e\right) \left(\frac{1}{t}\,\mathrm{d}G_t + e\,\mathrm{d}x\right) + \alpha_n \mathcal{R}(u) \right] \tag{8.1}$$

subject to $F(u_n) + F'(u_n; u-u_n) \geq -\frac{e}{2}$. There are several algorithms for solving convex minimization methods, e.g. primal-dual methods, sequential programming or the BFGS method. Our implementation does not solve the exact problem (8.1), but a quadratic approximation of it, which reduces the computational effort. We expect even better results if the problem (8.1) is solved exactly in each Newton step, which will be addressed in the future.

To obtain a discrete approximation of (8.1), we subdivide $\Omega = \bigcup_{j=1}^{d} \Omega_j$ into d disjoint subdomains Ω_j and generate a random data vector $\underline{g}^{\text{obs}} \in [0,\infty)^d$ with mutually independent components such that

$$\underline{g}_j^{\text{obs}} \sim \mathscr{P}\left(t \int_{\Omega_j} g^\dagger \, \mathrm{d}x\right) \qquad \text{for all} \qquad j = 1, ..., d.$$

Moreover define the mapping

$$\underline{F} : \mathbb{X} \to \mathbb{R}^d, \qquad u \mapsto \left(\int_{\Omega_j} F(u) \, \mathrm{d}x\right)_{j=1}^d$$

as a finite-dimensional approximation of F. Now (8.1) is approximated by the finite dimensional problem

$$\begin{aligned} \underline{u}^{n+1} = \operatorname*{argmin}_{\underline{u} \in V_k} \Bigg[& \sum_{j=1}^d \left(\underline{g}_j^n + \left(\underline{T}_n\left(\underline{u} - \underline{u}^n\right)\right)_j\right) \\ & - \sum_{j=1}^d \ln\left(\underline{g}_j^n + \left(\underline{T}_n\left(\underline{u} - \underline{u}^n\right)\right)_j + e\right)\left(\frac{1}{t}\underline{g}_j^{\text{obs}} + e\right) + \alpha_n \mathcal{R}\left(\underline{u}\right)\Bigg] \end{aligned} \tag{8.2}$$

subject to $\underline{g}_j^n + \left(\underline{T}_n\left(\underline{u} - \underline{u}^n\right)\right)_j \geq -\frac{e}{2}$ for all $j = 1, ..., d$ and $\underline{u} \in P_k(\mathfrak{B})$ where $V_k \subset \mathbb{X}$ is a k-dimensional subspace, $\underline{g}^n = \underline{F}(\underline{u}^n)$, $\underline{T}_n = \underline{F}'[\underline{u}^n]$ and $P_k : \mathbb{X} \to V_k$ denotes the projector. In the following we will identify V_k with $\mathbb{R}^k$.
Let us for a moment neglect both side conditions and assume that $\mathcal{R}$ is quadratic, e.g. $\mathcal{R}(u) = \|u - u_0\|_{\mathbb{X}}^2$ for some Hilbert space $\mathbb{X}$. We approximate the data fidelity term

$$\mathcal{S}_{e,t}\left(\underline{g} + \underline{h}; \underline{g}^{\text{obs}}\right) = \sum_{j=1}^d \left(\underline{g} + \underline{h}\right)_j - \sum_{j=1}^d \ln\left(\underline{g} + \underline{h} + e\right)_j \left(\frac{1}{t}\underline{g}_j^{\text{obs}} + e\right)$$

by the second order Taylor expansion

$$\mathcal{S}_{e,t}^{(2)}\left[\underline{g}; \underline{g}^{\text{obs}}\right](\underline{h}) := \mathcal{S}_{e,t}\left(\underline{g}; \underline{g}^{\text{obs}}\right) + \sum_{j=1}^d \left[\left(1 - \frac{\frac{1}{t}\underline{g}_j^{\text{obs}} + e}{\underline{g}_j + e}\right)\underline{h}_j + \frac{1}{2}\frac{\frac{1}{t}\underline{g}_j^{\text{obs}} + e}{(\underline{g}_j + e)^2}\underline{h}_j^2\right]$$

and define an inner iteration

$$\underline{h}^{n,l} := \operatorname*{argmin}_{\underline{h} \in \mathbb{R}^k} \left[\mathcal{S}_e^{(2)}\left[\underline{g}^n + \underline{T}_n\left(\underline{u}^{n,l} - \underline{u}^n\right); \underline{g}^{\text{obs}}\right](\underline{T}_n \underline{h}) + \alpha_n \mathcal{R}\left(\underline{u}^{n,l} + \underline{h}\right)\right] \tag{8.3}$$

for $l = 0, 1, ...$ with $\underline{u}^{n,0} := \underline{u}^n$ and $\underline{u}^{n,l+1} := \underline{u}^{n,l} + s_{n,l}\underline{h}^{n,l}$. The step-length parameter $s_{n,l}$ is introduced to ensure that $\underline{g}^n + \underline{T}_n\left(\underline{u}^{n,l+1} - \underline{u}^n\right) \geq -\eta e$ in each step. This can be guaranteed by

$$s_{n,l} = \max\left\{s \in [0,1] \;\middle|\; s\underline{T}_n\underline{h}^{n,l} \geq -\eta e - \underline{g}^n - \underline{T}_n\left(\underline{u}^{n,l} - \underline{u}^n\right)\right\} \tag{8.4}$$

with some tuning parameter $\eta \in [0, 1)$. In the computations we use $\eta = 0.9$, and the choice $\eta = 1/2$ would ensure that $\underline{u}^{n,l+1}$ satisfies the side condition in (2.14a). With these settings, (8.3) can be seen as a reasonable approximation to (8.1).

It follows from the first order optimality conditions, which are necessary and sufficient due to strict convexity here, that $\underline{u}^{n,l} = \underline{u}^{n,l+1}$ is the exact solution $\underline{u}^{n+1}$ of (8.2) if $\underline{h}^{n,l} = 0$. Therefore, we stop the inner iteration if $\frac{|\underline{h}^{n,l}|}{|\underline{h}^{n,0}|}$ is sufficiently small. We also stop the inner iteration if $s_{n,l}$ is 0 or too small.
Now note that $\underline{g}_j^{\text{obs}} + e > 0$ for all $j = 1, ..., d$ and hence by omitting terms independent of $\underline{h}$ and some simplifications we can write (8.3) as a least squares problem

$$\underline{h}^{n,l} = \underset{\underline{h} \in \mathbb{R}^k}{\operatorname{argmin}} \left[\sum_{j=1}^{d} \frac{1}{2} \left(\frac{\sqrt{\frac{1}{t}\underline{g}_j^{\text{obs}} + e}}{\underline{g}_j^{n,l} + e} \left(\underline{T}_n \underline{h}\right)_j + \frac{\underline{g}_j^{n,l} - \frac{1}{t}\underline{g}_j^{\text{obs}}}{\sqrt{\frac{1}{t}\underline{g}_j^{\text{obs}} + e}} \right)^2 + \alpha_n \mathcal{R}\left(\underline{u}^{n,l} + \underline{h}\right) \right] \tag{8.5}$$

with $\underline{g}^{n,l} := \underline{g}^n + \underline{T}_n \left(\underline{u}^{n,l} - \underline{u}^n\right)$. This quadratic problem can now be solved by the CG method.

```
INPUT: u_start, (α_n)_{n∈N}, e > 0, η ∈ (0,1), F_{|R^k}
n = 0;
u^n = u^start;
REPEAT
    l = 0;
    u^{n,0} = u^n;
    REPEAT
        Calculate h^{n,l} according to (8.5)
        Calculate s_{n,l} according to (8.4)
        u^{n,l+1} := u^{n,l} + s_{n,l} h^{n,l}
        l = l + 1;
    UNTIL |h^{n,l}| / |h^{n,0}| sufficiently small OR s_{n,l} too small
    u^{n+1} = u^{n,l};
    n = n + 1;
UNTIL STOP
OUTPUT: u^n
```

Figure 8.1: Our strategy for the implementation of (8.1).

In the examples below we observed fast convergence of the inner iteration (8.3). However, if the offset parameter e becomes too small or if $e = 0$ convergence deteriorates in general. This is not surprising since the iteration (8.3) cannot be expected to converge to the exact solution $\underline{u}^{n+1}$ of (8.2) if the side condition $\underline{g}^n + \underline{T}_n \left(\underline{u}^{n+1} - \underline{u}^n\right) \geq -e/2$ is active at $\underline{u}^{n+1}$.
To ensure that the data misfit term $\mathcal{S}_{e,t}$ indeed approximates the negative log-likelihood functional, we let moreover $e \searrow 0$ in our computations, i.e. the offset parameter e_n is reduced by a fixed factor in every iteration step. Since a small parameter e_n causes problems in the numerical realization of our algorithm, e_n should not tend to 0 to fast. Note that our convergence analysis only includes the case that $e_n \equiv e$ is fixed over the whole reconstruction procedure (cf. Remark 4.4). Nevertheless, a convergence analysis with $e_n \searrow 0$ is also possible by coupling the decay of e_n, α_n and the noise level and accepting worse convergence rates.

In the following we will compare our algorithm with the usual IRGNM (i.e. (5.3a) with $r = p = 2$) and a weighted version for the examples described in Chapter 2. The weighted version has already been proposed in [SBH11, Stü11] where the data fidelity is chosen to be

$$\Phi^2\left(g; g^{\text{obs}}\right) := \int_\Omega \frac{\left|g - g^{\text{obs}}\right|^2}{\max\left\{g^{\text{obs}}, c\right\}} \,\mathrm{d}x, \tag{8.6}$$

i.e. a weighted $\mathbf{L}^2$ norm. The weight approximates $1/g^\dagger$, which turns out to realize the second order Taylor approximation of $g \mapsto \mathbb{KL}\left(g^\dagger; g\right)$ around $g^\dagger$. The parameter $c > 0$ is included to avoid divisions by 0. The fidelity term in (8.6) with $c = 0$ is also known as *Pearson's* distance. If Φ^2 is used as data fidelity term, then in every Newton step a weighted least-squares problem needs to be solved, which is done with the help of the CG algorithm.

To compare our algorithm with the quadratic ones described above we perform 100 experiments with simulated Poisson data and use an oracle stopping rule to eliminate the influence of this choice. This is done by using

$$N := \underset{1 \leq n \leq N_{\max}}{\operatorname{argmin}} \; \mathbf{E}\|\underline{u}_n - \underline{u}^\dagger\|_{\mathbf{L}^2}^2 \tag{8.7}$$

as stopping index with the empirical version of the expectation. Moreover we compare the choice (8.7) for our algorithm with the choice provided by the Lepskiĭ-type balancing principle, i.e.

$$\begin{aligned}
\Phi_{\text{noi}}^{\text{det}}(n) &:= c_1\sqrt{2}\sqrt{\frac{\ln(t)\,\psi(t)}{\alpha_n}}, \\
N_{\max} &:= \min\left\{n \in \mathbb{N} \;\middle|\; \Phi_{\text{noi}}^{\text{det}}(n) \geq 1\right\}, \\
n_{\text{bal}} &:= \min\left\{n \in \{1, \ldots, N_{\max}\} \;\middle|\; \forall m \geq n : \; \|u_n - u_m\|_{\mathbf{L}^2} \leq c_2 \Phi_{\text{noi}}^{\text{det}}(m)\right\}
\end{aligned}$$

with suitable constants $c_1, c_2 > 0$.

8.2 An inverse obstacle scattering problem without phase

In this section we want to apply our method to the case of an inverse obstacle scattering problem without phase as presented in Chapter 2. As pointed out in Section 2.2.1 it is impossible to reconstruct the center of gravity of D since $|u_\infty|$ is invariant under translations. For plots we always shift the center of gravity of ∂D to the origin. We assume that D is star-shaped and represent ∂D by a periodic function q such that $\partial D = \{q(t)(\cos t, \sin t)^\top : t \in [0, 2\pi]\}$. For details on the implementation of F, its derivative and adjoint we refer to [Hoh98] where the mapping $q \mapsto u_\infty$ is considered as forward operator.

It seems reasonable to consider $F : q \to |u_\infty|^2$ as an operator

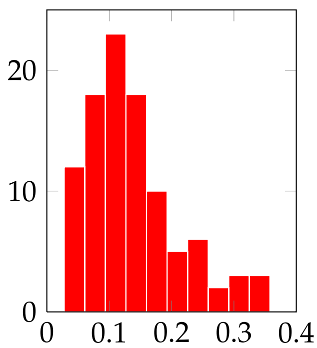
$$F : H_{\text{per}}^s([0, 2\pi]) \to \mathbf{L}^\infty([0, 2\pi]),$$

but it can be seen easily that $|u_\infty|^2$ is also an element of $\mathbf{L}^p([0, 2\pi])$ for any p in concurrence with Assumption 2.7(a). One could also choose any Sobolev space $H_{\text{per}}^s([0, 2\pi])$ with arbitrary s as image space, since the function $|u_\infty|^2$ is analytic. The set $\mathfrak{B} \subset \mathbb{X} =$

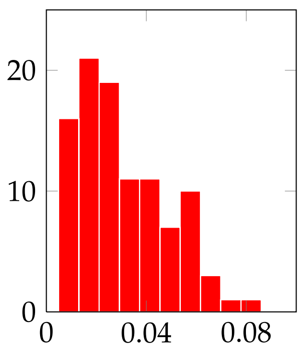
$H^s_{\text{per}}([0, 2\pi])$ can be chosen arbitrary in this setup, for our numerical simulations we use $\mathfrak{B} = \mathbb{X}$ even if this contradicts Assumption 2.7(b). For the Fréchet differentiability we refer to [Hoh98]. Further note that Assumption 2.7(e) is fulfilled for free by the definition of F.

As a test example we choose the peanut-shaped obstacle shown in Figure 2.2 described by $q^\dagger(t) = \frac{1}{2}\sqrt{3\cos^2 t + 1}$ with two incident waves from 'South West' and from 'East' with wave number $k = 10$. The incident directions are marked by two arrows in Figure 8.3. The initial guess for the Newton iteration is the unit circle described by $q_0 \equiv 1$, and we choose the Sobolev norm $\mathcal{R}(q) = \|q - q_0\|^2_{H^s}$ with $s = 1.6$ as penalty functional. The regularization parameters are chosen as $\alpha_n = 0.5 \cdot (2/3)^n$. Moreover, we choose an initial offset parameter $e = 0.002$, which is reduced by $\frac{4}{5}$ in each iteration step. The inner iteration (8.3) is stopped when $\|\underline{h}^{n,l}\| / \|\underline{h}^{n,0}\| \leq 0.1$, which was usually the case after about three iterations (or about five iterations for $\|\underline{h}^{n,l}\| / \|\underline{h}^{n,0}\| \leq 0.01$).

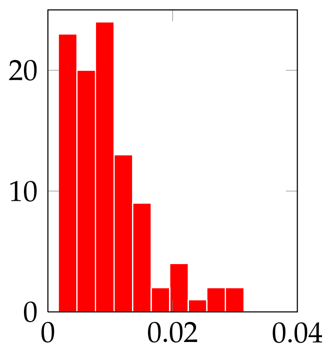
To illustrate the decay of noise as $t \to \infty$ we also calculated the error terms $\mathbf{err}_n$ as in (6.11b) for every iteration n. The result is shown in Figure 8.2 and one can see that the decay is indeed of order $\mathcal{O}\left(\frac{1}{\sqrt{t}}\right)$ as theoretically ensured by Corollary 7.2.



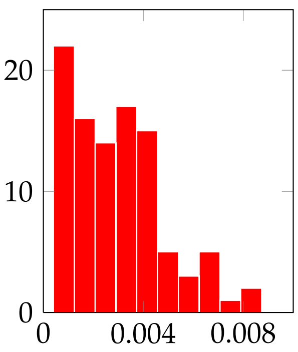
(a) histogram for $t = 10^2$



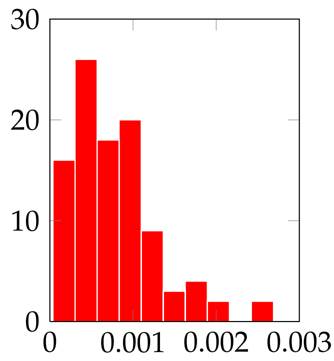
(b) histogram for $t = 10^3$



(c) histogram for $t = 10^4$



(d) histogram for $t = 10^5$



(e) histogram for $t = 10^6$

t	$\mathbf{E}\left(\max_{n\leq 20} \mathbf{err}_n\right)$	factor
10^2	0.1383	
		4.53
10^3	0.0305	
		3.18
10^4	0.0096	
		3.28
10^5	0.0029	
		3.70
10^6	0.0008	

(f) means for different t

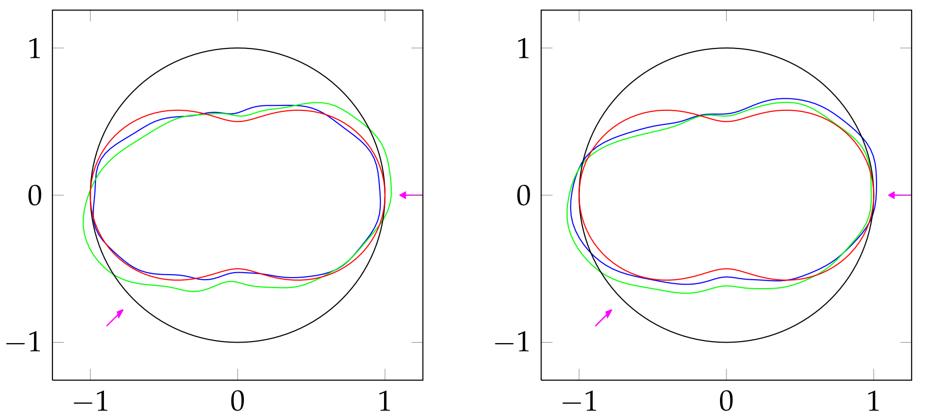
Figure 8.2: Overview for the error terms (6.11b) for the scattering problem. For different values of the expected total number of counts the value $\max_{n\leq 20} \mathbf{err}_n$ has been calculated in 100 experiments. Above the corresponding histograms and means are shown. The decay of order $\frac{1}{\sqrt{t}}$, i.e. reduction by a factor of $\sqrt{10} \approx 3.16$ in the table is clearly visible. Moreover, since all values are finite we note that $F(u_n) + F'(u_n, u^\dagger - u_n) \geq -e/2$ holds always true (this has theoretically been gained by the generalized tangential cone condition).

We gain convergence of Tikhonov-type regularization from Remark 4.15 if the tangential cone condition (6.3) is fulfilled with small $\bar{\eta}$. Moreover, a spectral source condition (3.9) implies under (6.3) due to Corollary (4.7) an additive variational inequality (3.11) with some index function φ_{add}. Thus for a proper chosen regularization parameter α Theorem 4.11 yields convergence rates.
Supposed the nonlinearity condition (6.6b) holds true we gain convergence of (8.1) from Remark 7.5 if u_0 is sufficiently good. Moreover, a spectral source condition (3.9) implies under (6.6b) due to Corollary (7.1) an additive variational inequality (3.11) with some index function φ_{add}. Thus for a sufficiently good initial guess and a proper chosen stopping index n_* Theorem 7.3 yields convergence rates.
Since F maps arbitrarily rough functions to analytic functions, we expect only a weak source condition (i.e. (3.9) with logarithmic φ) to hold. Therefore, we do not discuss the implications under weaker nonlinearity conditions here.

t	$\mathcal{S}(g;\underline{g}^{\mathrm{obs}})$	Stopping rule	N	$\sqrt{\mathbf{E}\lVert q_N - q^\dagger\rVert^2_{\mathbf{L}^2}}$	$\sqrt{\mathbf{V}\lVert q_N - q^\dagger\rVert_{\mathbf{L}^2}}$
100	$\left\lVert g-\underline{g}^{\mathrm{obs}}\right\rVert^2_{\mathbf{L}^2}$	Oracle	7	0.124	0.033
	$\Phi^2\left(g;\underline{g}^{\mathrm{obs}}\right)$	Oracle	2	0.122	0.018
	$\mathcal{S}_{e,t}\left(g;\underline{g}^{\mathrm{obs}}\right)$	Oracle	3	0.091	0.025
	$\mathcal{S}_{e,t}\left(g;\underline{g}^{\mathrm{obs}}\right)$	Lepskiĭ	4.4 ± 1.32	0.105	0.030
1000	$\left\lVert g-\underline{g}^{\mathrm{obs}}\right\rVert^2_{\mathbf{L}^2}$	Oracle	9	0.106	0.014
	$\Phi^2\left(g;\underline{g}^{\mathrm{obs}}\right)$	Oracle	7	0.091	0.012
	$\mathcal{S}_{e,t}\left(g;\underline{g}^{\mathrm{obs}}\right)$	Oracle	5	0.070	0.017
	$\mathcal{S}_{e,t}\left(g;\underline{g}^{\mathrm{obs}}\right)$	Lepskiĭ	4.8 ± 0.95	0.078	0.019
10000	$\left\lVert g-\underline{g}^{\mathrm{obs}}\right\rVert^2_{\mathbf{L}^2}$	Oracle	9	0.105	0.004
	$\Phi^2\left(g;\underline{g}^{\mathrm{obs}}\right)$	Oracle	23	0.076	0.048
	$\mathcal{S}_{e,t}\left(g;\underline{g}^{\mathrm{obs}}\right)$	Oracle	5	0.050	0.005
	$\mathcal{S}_{e,t}\left(g;\underline{g}^{\mathrm{obs}}\right)$	Lepskiĭ	5.6 ± 1.08	0.060	0.014

Table 8.1: $\mathbf{L}^2$-error statistics for the inverse obstacle scattering problem (2.16). We compare the data fidelity choice $\mathcal{S}_{e,t}$ with the standard $\mathbf{L}^2$ distance $\lVert g - g^{\mathrm{obs}}\rVert^2_{\mathbf{L}^2}$ and Pearson's distance Φ^2 given in (8.6) for different values of the expected total number of counts t each with 100 experiments. The error of the initial guess is $\lVert q_0 - q^\dagger\rVert_{\mathbf{L}^2} = 0.288$. All parameters as described in Section 8.2.

Error statistics of shape reconstructions from 100 experiments are shown in Table 8.1. It turned out that for $t = 100$ and $t = 1000$ the parameter c in (8.6) should be chosen small whereas for $t = 10000$ c should be much larger. We decided for $c = 0.2$ in all cases to produce comparable examples. Note that the mean square error is significantly smaller for the Kullback-Leibler divergence than for the $\mathbf{L}^2$-distances. For comparison of the oracle stopping rule (8.7) with the Lepskiĭ-type balancing principle we chose $c_1 = 0.1$ and $c_2 = 0.2$, which is necessary due to norm scaling in our algorithm. In Table 8.1 it can be seen that the Lepskiĭ principle yields reasonable results which are comparable to those obtained by the oracle stopping rule.



(a) results for $\mathcal{S} = \mathcal{S}_{e,t}$. blue: best, green: median, black: initial guess

(b) results for $\mathcal{S}(g_1; g_2) = \|g_1 - g_2\|^2_{\mathbf{L}^2}$. blue: best, green: median, black: initial guess

Figure 8.3: Numerical results for the inverse obstacle scattering problem (2.16). Panels (a) and (b) show best and median reconstruction from 100 experiments with $t = 1000$ expected counts where the stopping parameter N is chosen by the oracle rule (8.7). See also Table 8.1.

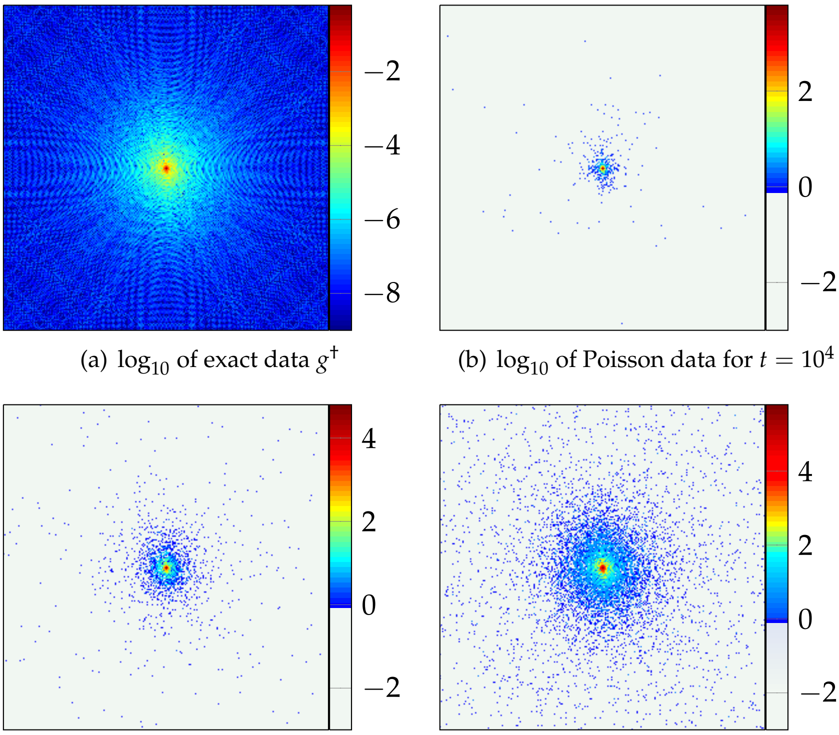
8.3 A phase retrieval problem

In this section we will consider a **phase retrieval problem** as described in Section 2.2.2. In the following we assume more specifically that $f : \mathbb{R}^2 \to \mathbb{C}$ is of the form $f(x) = \exp(i\varphi(x))$ with an unknown real-valued function φ where the support $\operatorname{supp}(\varphi)$ is known to be compact in $B_\rho = \left\{x \in \mathbb{R}^2 : |x| \leq \rho\right\}$. For a uniqueness result we refer to KLIBANOV [Kli06], although not all assumptions of this theorem are satisfied in the example below. It turns out to be particularly helpful if φ has a jump of known magnitude at the boundary of its support. We will assume for simplicity that $\operatorname{supp}(\varphi) = B_\rho$ and that $\varphi \approx \chi_{B_\rho}$ close to the boundary ∂B_ρ (here χ_{B_ρ} denotes the characteristic function of B_ρ). This leads as motivated in Section 2.2.2 to an inverse problem where the forward operator is given by

$$F : H^s(B_\rho) \to \mathbf{L}^\infty\left(\left[-\kappa,\kappa\right]^2\right), \quad (F(\varphi))(\xi) := \left| \int_{B_\rho} \exp(-i\xi \cdot x) \exp(i\varphi(x))\, dx \right|^2. \quad (8.8)$$

The a priori information on φ can be incorporated in the form of an initial guess $\varphi_0 \equiv 1$. As for the scattering problem it follows from the fact that the range of F consists of analytic functions that F maps also bounded into $\mathbf{L}^p\left(\left[-\kappa,\kappa\right]^2\right)$ for any p in concurrence with Assumption 2.7(a). Moreover, the choice of any Sobolev space $H^s\left(\left[-\kappa,\kappa\right]^2\right)$ with arbitrary s as image space is possible. The set $\mathfrak{B} \subset \mathbb{X} = H^s(B_\rho)$ can be chosen arbitrary

since the support constraint is already included by considering only functions on B_ρ in the space $\mathbb{X}$. For our numeric simulations we again dropped the constraint $u_n \in \mathfrak{B}$ which corresponds to the case $\mathfrak{B} = \mathbb{X}$, even if this contradicts Assumption 2.7(b). This is possible since Assumption 2.7(e) is again fulfilled for free by the definition of F. The Fréchet differentiability of F is obvious and the derivative as well as its adjoint can be calculated easily.
For our numeric calculations we choose the Sobolev index $s = \frac{1}{2}$ and the regularization parameters $\alpha_n = \frac{5}{10^6} \cdot (2/3)^n$. The offset parameter e is initially set to $2 \cdot 10^{-6}$ and reduced by a factor $\frac{4}{5}$ in each iteration step. The corresponding data for different t can be seen in Figure 8.4, which illustrates similarly to Figure 8.2 the decay of noise.



(a) $\log_{10}$ of exact data $g^\dagger$

(b) $\log_{10}$ of Poisson data for $t = 10^4$

(c) $\log_{10}$ of Poisson data for $t = 10^5$

(d) $\log_{10}$ of Poisson data for $t = 10^6$

Figure 8.4: Exact and corresponding Poisson data for different observation times t for the phase retrieval problem. Note that in comparison to Figure 1.1 the Poisson data has not been normalized here.

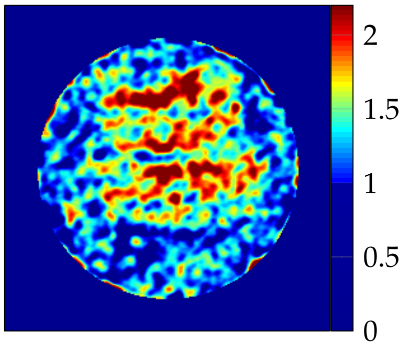
The convergence results obtained by our theoretical results are similar as for the scattering problem, i.e. supposed the tangential cone condition (6.3) is fulfilled with small $\bar{\eta}$ we gain convergence of Tikhonov-type regularization from Remark 4.15. If the spectral source condition (3.9) can be specified, we gain the convergence rate (4.13) as $t \to \infty$ for a proper chosen regularization parameter α. Supposed the nonlinearity condition (6.6b) holds true we gain convergence of (8.1) from Remark 7.5 if u_0 is sufficiently good. Moreover, a spectral source condition (3.9) yields for a sufficiently good initial guess and a proper chosen stopping index n_* the convergence rate (7.4). Note that we would expect only a weak source condition (i.e. (3.9) with logarithmic φ) to hold since the operator maps arbitrarily rough functions (with compact support) to analytic functions. Therefore, we do not discuss the implications under weaker nonlinearity conditions here.
In Table 8.2 the convergence of (8.5) is illustrated. As for the scattering problem we ob-

t	$\mathcal{S}\left(g;\underline{g}^{\mathrm{obs}}\right)$	Stopping rule	N	$\sqrt{\mathbf{E}\|\varphi_N-\varphi^\dagger\|^2_{\mathbf{L}^2}}$	$\sqrt{\mathbf{V}\|\varphi_N-\varphi^\dagger\|_{\mathbf{L}^2}}$
10^4	$\left\|g-\underline{g}^{\mathrm{obs}}\right\|^2_{\mathbf{L}^2}$	Oracle	4	49.29	4.83
	$\Phi^2\left(g;\underline{g}^{\mathrm{obs}}\right)$	Oracle	1	66.56	4.33
	$\mathcal{S}_{e,t}\left(g;\underline{g}^{\mathrm{obs}}\right)$	Oracle	8	42.78	2.13
	$\mathcal{S}_{e,t}\left(g;\underline{g}^{\mathrm{obs}}\right)$	Lepskiĭ	3.9 ± 0.72	46.78	4.21
10^5	$\left\|g-\underline{g}^{\mathrm{obs}}\right\|^2_{\mathbf{L}^2}$	Oracle	24	31.22	4.91
	$\Phi^2\left(g;\underline{g}^{\mathrm{obs}}\right)$	Oracle	3	37.09	1.23
	$\mathcal{S}_{e,t}\left(g;\underline{g}^{\mathrm{obs}}\right)$	Oracle	9	32.47	1.35
	$\mathcal{S}_{e,t}\left(g;\underline{g}^{\mathrm{obs}}\right)$	Lepskiĭ	9.1 ± 0.58	32.44	1.41
10^6	$\left\|g-\underline{g}^{\mathrm{obs}}\right\|^2_{\mathbf{L}^2}$	Oracle	30	16.83	2.31
	$\Phi^2\left(g;\underline{g}^{\mathrm{obs}}\right)$	Oracle	7	16.44	0.32
	$\mathcal{S}_{e,t}\left(g;\underline{g}^{\mathrm{obs}}\right)$	Oracle	14	18.60	0.88
	$\mathcal{S}_{e,t}\left(g;\underline{g}^{\mathrm{obs}}\right)$	Lepskiĭ	13.8 ± 0.63	18.60	0.80

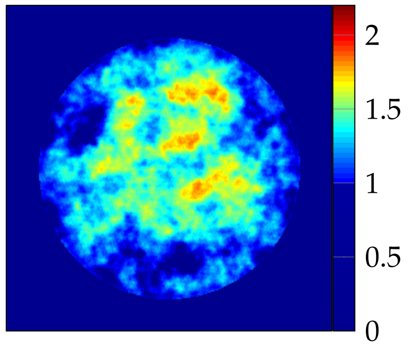
Table 8.2: $\mathbf{L}^2$-error statistics for the phase retrieval problem (8.8). We compare the data fidelity choices $\mathcal{S}_{e,t}$ with the standard $\mathbf{L}^2$ distance $\left\|g-g^{\mathrm{obs}}\right\|^2_{\mathbf{L}^2}$ and Pearson's distance Φ^2 given in (8.6) for different values of the expected total number of counts t each with 100 experiments. The initial error is $\|\varphi_0-\varphi^\dagger\|_{\mathbf{L}^2} = 75.05$. All parameters as described in Section 8.3. Note that it is unclear how close the reconstructions of our method for $t = 10^6$ are to those of (8.1), since for large t the system (8.5) is highly ill-conditioned and could hence not be solved reliably by the CG method!

served a quite inconsistent behavior for the choice of c in (8.6): For small t we require a small cutoff c, but for $t = 10^6$ the results are better for $c > 1$, which is not a reasonable choice. We decided again for $c = 0.2$ in all cases. It can be seen from Table 8.2 that especially for low count data the expected square error $\mathbf{E}\|\varphi_N-\varphi^\dagger\|^2_{\mathbf{L}^2}$ is significantly smaller. For high count data the expected square error for both methods is comparable with a slight advantage for the $\mathbf{L}^2$-distances. It must be said, however, that for high n (i.e. small regularization parameters α_n), which are necessary for high-count data, the system becomes so ill-conditioned that it can no longer be solved reliably by the CG method. For such n the CG iteration is typically terminated at the maximum number of iterations (e.g. 50 or 200) before the stopping criterion guaranteeing an accurate solution of the regularized Newton equations was met. **Hence, it remains unclear at the moment how close our results are to those of an accurate application of the respective methods and how those would compare.**

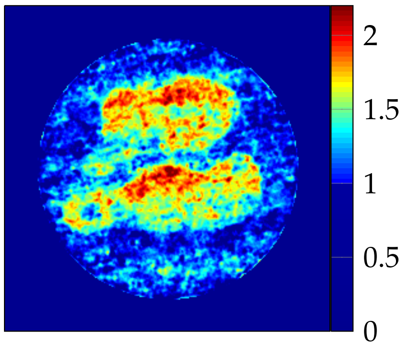
For comparison of the oracle stopping rule (8.7) with the Lepskiĭ-type balancing principle we chose $c_1 = 0.1$ and $c_2 = 20$, which seems to be a reasonable choice. In Table 8.1 it can



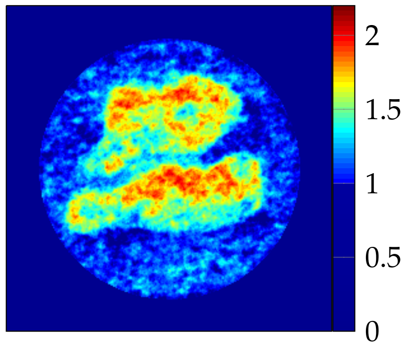
(a) median reconstruction for $t = 10^4$ of the weighted IRGNM



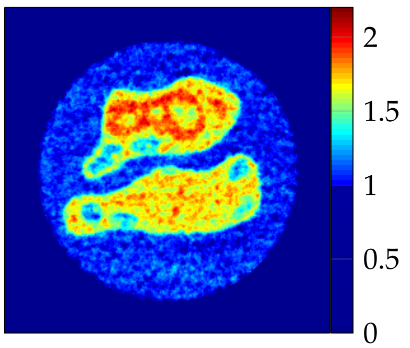
(b) median reconstruction for $t = 10^4$ of our method (8.5)



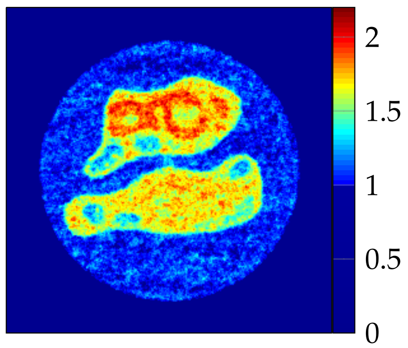
(c) median reconstruction for $t = 10^5$ of the weighted IRGNM



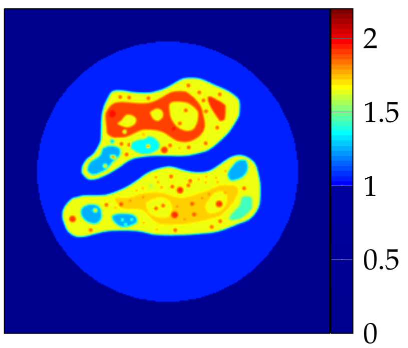
(d) median reconstruction for $t = 10^5$ of our method (8.5)



(e) median reconstruction for $t = 10^6$ of the weighted IRGNM



(f) median reconstruction for $t = 10^6$ of our method (8.5)



(g) exact solution

Figure 8.5: Median reconstructions for the phase retrieval problem with different observation times t in comparison to the exact solution. See also Table 8.2.

be seen that the Lepskiĭ principle yields similar results as the oracle stopping rule and the convergence rates seem to coincide.
Comparing the subplots in Figure 8.5 the previously described behavior becomes visible. For low count data our algorithm performs better and is able to reconstruct the shape of the cells much more clearly. In the case of high count data ($t = 10^6$), both reconstructions look similar.

8.4 A semiblind deconvolution problem

In this section we will consider the problem which arises from the application in 4Pi microscopy as described in Section 2.2.3. As mentioned there, the forward operator has the form

$$F_{4\text{Pi}}(f, \phi)(x) = \int_{\Omega} p(y - x, \phi(x)) f(y)\, \mathrm{d}y, \qquad y \in \Omega.$$

Since the point spread function p depends on the (unknown) phase ϕ and is hence partially unknown, this problem is called a **semiblind** deconvolution problem. For physical reasons, the unknown phase ϕ can be assumed to be smooth. Hence it seems natural to consider $F_{4\text{Pi}}$ as an operator

$$F_{4\text{Pi}} : \mathbb{X} := \mathbf{L}^2(\Omega) \times H^2(\Omega) \to \mathbf{L}^2(\Omega).$$

Using the special structure (see [Stü11])

$$p(x, \varphi) = h(x) \cos^n \left(c x_3 + \frac{\varphi}{2} \right) \tag{8.9}$$

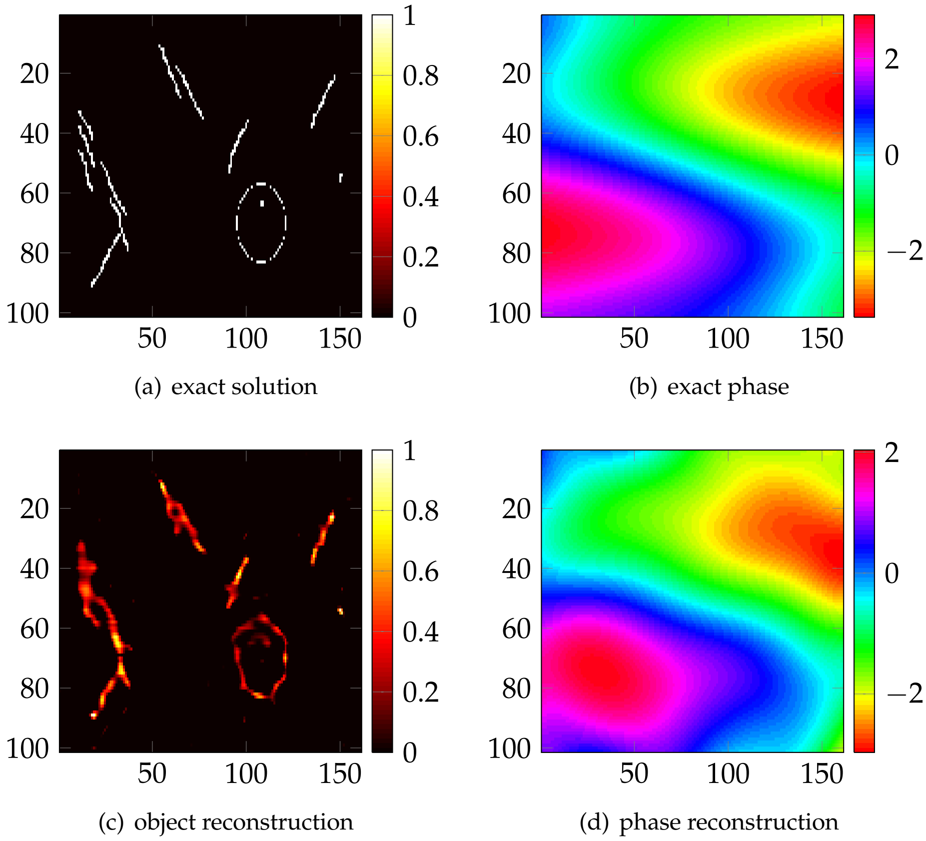
allows us to write $F_{4\text{Pi}}$ as a sum of convolutions with smooth kernels, and hence the functions in the image of F are also smooth. Thus F maps also bounded into $\mathbf{L}^p(\Omega)$ for any p in concurrence with Assumption 2.7(a) and into any Sobolev space $H^s(\Omega)$ with arbitrary s. To ensure that Assumption 2.7(e) is fulfilled, we need to choose $\mathfrak{B} \subseteq \{(f, \phi) \in \mathbb{X} \mid f \geq 0 \text{ a.e.}\}$. The boundedness of $\mathfrak{B}$ is dropped in our numeric simulations, but the side constraint $f \geq 0$ is implemented with the help of the Semi Smooth Newton Method and we refer again to [Stü11, Sec. 5.2] for details. The Fréchet differentiability of F and explicit representations of F' and its adjoint are investigated by Stück in [Stü11].

For a fast implementation of $F_{4\text{Pi}}$ using the special structure (8.9) we also refer to [Stü11]. The idea is to write $F_{4\text{Pi}}$ with the help of (8.9) as a sum of convolutions, which can be implemented using the fast Fourier transform (FFT). Therefore, the unknown function f as well as the different convolution kernels need to be approximated by periodic functions which is done by zero padding to a larger domain and periodic extension. The corresponding periodic function f is afterwards approximated by a trigonometric polynomial of finite order. The smooth phase is represented by a polynomial of small degree. Both together yield the k-dimensional approximation space V_k as described in Section 8.1. The incorporation of the side condition $f \geq 0$ a.e. with the help of the Semi Smooth Newton Method corresponds again to an iterative method for calculating $\underline{h}^{\mathrm{n,l}}$ in (8.5) and causes a high numerical effort. For our numeric tests we hence set the maximal number of inner iterations to 1.
It has been shown in [SBH11] that the operator F fulfills a Lipschitz condition (6.5). Moreover, convergence rates for the iteratively regularized Gauss Newton method (5.3a) in the

quadratic Hilbert space case with the side constraint $\mathfrak{B} = \{(f,\phi) \in \mathbb{X} \mid f \geq 0 \text{ a.e.}\}$ under the projected source condition (3.18) have been obtained. We want to mention here, that also the case of operator noise has been included in this study.
Those results are reproduced by our theory. Since the projected source condition (3.18) together with (6.5) and $\frac{L}{2}\|\omega\| < 1$ implies the additive variational inequality (3.11) with $\varphi_{\text{add}} = \varphi_{\frac{1}{2}}$, our Theorem 6.8 applies and yields as in Remark 6.13 the convergence rate

$$\left\| u_{n_*} - u^\dagger \right\|_{\mathbb{X}} = \mathcal{O}\left(\sqrt{\delta}\right) \tag{8.10}$$

where $u^\dagger = (f^\dagger, \phi^\dagger)$ and $u_{n_*} = (f_{n_*}, \phi_{n_*})$. Nevertheless, this rate holds true only if the operator F is given exactly. Operator noise can be included since Assumption 3.8 is formulated quite generally, but this is beyond the scope of this thesis.



(a) exact solution (b) exact phase

(c) object reconstruction (d) phase reconstruction

Figure 8.6: Exemplary reconstruction for the 4Pi problem. Panels (a) and (b) show exact object and phase, (c) and (d) show the reconstructions.

The convergence results for using the negative log-likelihood functional obtained by our theoretical results are similar as for other two problems, i.e. supposed the tangential cone condition (6.3) is fulfilled with small $\bar{\eta}$ we gain convergence of Tikhonov-type regularization from Remark 4.15. If the spectral source condition (3.9) can be specified, we gain the convergence rate (4.13) as $t \to \infty$ for a proper chosen regularization parameter α. Supposed the nonlinearity condition (6.6b) holds true we gain convergence of (8.1) from Remark 7.5 if u_0 is sufficiently good. Moreover, a spectral source condition (3.9) yields for a sufficiently good initial guess and a proper chosen stopping index n_* the convergence rate (7.4).

As for the other two aforementioned problems we would expect only a weak source condition (i.e. (3.9) with logarithmic φ) to hold since the operator maps arbitrarily rough functions to smooth functions. Nevertheless, for the comparability with [Stü11] we will discuss the usage of the stronger source condition (3.18) in combination with a weaker nonlinearity condition in the following.
If Assumption 6.2 is fulfilled with $r = 2$, then the result from Theorem 6.8 with $p = 2$ also applies to the case of $\mathcal{S}$ being the Kullback-Leibler divergence. This can be seen by using Corollary 4.8, which provides the variational inequality (3.11) with $\varphi_{\text{add}} = \beta' \varphi_{\frac{1}{2}}$ under (3.18). Thus, we are able to enhance the results from [Stü11] to the case of Poisson data under a similar source condition.

Due to the fact that the computation time for a single example is still very high (especially for 3D data sets as they occur in practice) we performed no statistical tests as for the first two examples. Numeric tests with a data-weighted $\mathbf{L}^2$-distance as fidelity term has been presented in [Stü11, SBH11], and different algorithms for the reconstruction of f have been considered and tested in [BCC06, VSE$^+$10]. In Figure 8.6 an exemplary reconstruction of our algorithm is shown. First comparisons with the algorithm from [Stü11] have shown that both algorithms yield similar results, which might be caused by setting the number of inner iterations to one.

BIBLIOGRAPHY

[AB06] A. ANTONIADIS AND J. BIGOT. *Poisson inverse problems*. *Ann. Statist.*, 34(5):2132–2158, **2006**.

[AV94] R. ACAR AND C. R. VOGEL. *Analysis of bounded variation penalty methods for ill-posed problems*. *Inverse Probl.*, 10(6):1217–1229, **1994**.

[Bak92] A. B. BAKUSHINSKIĬ. *On a convergence problem of the iterative-regularized Gauss-Newton method*. *Comput. Math. Math. Phys.*, 32(9):1353–1359, **1992**.

[Bar10] J. M. BARDSLEY. *A Theoretical Framework for the Regularization of Poisson Likelihood Estimation Problems*. *Inverse Probl. Imag.*, 4:11–17, **2010**.

[BB11] M. BENNING AND M. BURGER. *Error estimates for general fidelities*. *Electron. T. Numer. Ana.*, 38:44–68, **2011**.

[BCC06] D. BADDELEY, C. CARL AND C. CREMER. *4pi microscopy deconvolution with a variable point-spread function*. *Appl. Opt.*, 45(27):7056–7064, **2006**.

[BH10] R. I. BOŢ AND B. HOFMANN. *An extension of the variational inequality approach for nonlinear ill-posed problems*. *J. Integral Equations Appl.*, 22(3):369–392, **2010**.

[BHM04] N. BISSANTZ, T. HOHAGE AND A. MUNK. *Consistency and rates of convergence of nonlinear Tikhonov regularization with random noise*. *Inverse Probl.*, 20(6):1773–1789, **2004**.

[BHM09] F. BAUER, T. HOHAGE AND A. MUNK. *Iteratively regularized Gauss-Newton method for nonlinear inverse problems with random noise*. *SIAM J. Numer. Anal.*, 47(3):1827–1846, **2009**.

[BKM+08] T. BONESKY, K. S. KAZIMIERSKI, P. MAASS, F. SCHÖPFER AND T. SCHUSTER. *Minimization of Tikhonov functionals in Banach spaces*. *Abstr. Appl. Anal.*, pages Art. ID 192679, 19, **2008**.

[BL91] J. M. BORWEIN AND A. S. LEWIS. *Convergence of best entropy estimates*. *SIAM J. Optim.*, 1:191–205, **1991**.

[BM01] H. BREZIS AND P. MIRONESCU. *Composition in fractional Sobolev spaces*. *Discrete Contin. Dynam. Systems*, 7(2):241–246, **2001**.

[BNS97] B. BLASCHKE, A. NEUBAUER AND O. SCHERZER. *On convergence rates for the iteratively regularized Gauss-Newton method*. *IMA J. Numer. Anal.*, 17(3):421–436, **1997**.

[BO04] M. BURGER AND S. OSHER. *Convergence rates of convex variational regularization*. *Inverse Probl.*, 20(5):1411–1421, **2004**.

[BW99] M. BORN AND E. WOLF. *Principles of Optics*. Cambridge University Press, Cambridge, seventh edition, **1999**.

[CK97] D. COLTON AND R. KRESS. *Inverse Acoustic and Electromagnetic Scattering Theory*. Springer, Berlin, Heidelberg, New York, second edition, **1997**.

[DDD04] I. DAUBECHIES, M. DEFRISE AND C. DEMOL. *An iterative thresholding algorithm for linear inverse problems with a sparsity constraint*. *Comm. Pure Appl. Math.*, 57(11):1413–1457, **2004**.

[Egg93] P. P. B. EGGERMONT. *Maximum entropy regularization for Fredholm integral equations of the first kind*. *SIAM J. Math. Anal.*, 24(6):1557–1576, **1993**.

[EHN96] H. ENGL, M. HANKE AND A. NEUBAUER. *Regularization of Inverse Problems*. Springer, **1996**.

[EKN89] H. ENGL, K. KUNISCH AND A. NEUBAUER. *Convergence rates for Tikhonov regularisation of non-linear ill-posed problems*. *Inverse Probl.*, 5(4):523, **1989**.

[ET76] I. EKELAND AND R. TÉMAM. *Convex analysis and variational problems*. Studies in mathematics and its applications. North Holland, **1976**.

[FH10] J. FLEMMING AND B. HOFMANN. *A new approach to source conditions in regularization with general residual term*. *Numer. Funct. Anal. Optim.*, 31:254–284, **2010**.

[FH11] J. FLEMMING AND B. HOFMANN. *Convergence rates in constrained tikhonov regularization: equivalence of projected source conditions and variational inequalities*. *Inverse Probl.*, 27(8):085001, **2011**.

[FHM11] J. FLEMMING, B. HOFMANN AND P. MATHÉ. *Sharp converse results for the regularization error using distance functions*. *Inverse Probl.*, 27(2):025006, **2011**.

[Fle10] J. FLEMMING. *Theory and examples of variational regularisation with non-metric fitting functionals*. *J. Inverse Ill-Posed P.*, 18(6):677–699, **2010**.

[Fle11] J. FLEMMING. *Generalized Tikhonov regularization - Basic theory and comprehensive results on convergence rates*. Ph.D. thesis, Chemnitz University of Technology, **2011**.

[GHS08] M. GRASMAIR, M. HALTMEIER AND O. SCHERZER. *Sparse regularization with ℓ^q penalty term*. *Inverse Probl.*, 24(5):055020, **2008**.

[GHS11] M. GRASMAIR, M. HALTMEIER AND O. SCHERZER. *Necessary and sufficient conditions for linear convergence of ℓ^1-regularization*. *Comm. Pure Appl. Math.*, 64(2):161–182, **2011**.

[GKK$^+$11] K. GIEWEKEMEYER, S. P. KRÜGER, S. KALBFLEISCH, M. BARTELS, C. BETA AND T. SALDITT. *X-ray propagation microscopy of biological cells using waveguides as a quasipoint source*. *Phys. Rev. A*, 83:023804, **2011**.

[Gra09a] M. GRASMAIR. *Non-convex Sparse Regularization*. Technical report, NRN No. 86, **2009**.

[Gra09b] M. GRASMAIR. *Well-posedness and convergence rates for sparse regularization with sublinear l^q penalty term*. *Inverse Probl. Imag.*, 3(3):383–387, **2009**.

[Gra10a] M. GRASMAIR. *Generalized Bregman distances and convergence rates for non-convex regularization methods*. *Inverse Probl.*, 26(11):115014, **2010**.

[Gra10b] M. GRASMAIR. *Non-convex sparse regularisation*. *J. Math. Anal. Appl.*, 365(1):19–28, **2010**.

[Had52] J. HADAMARD. *Lectures on Cauchy's Problem in Linear Partial Differential Equations*. Dover, **1952**.

[Han10] M. HANKE. *The regularizing Levenberg-Marquardt scheme is of optimal order*. *J. Integral Equations Appl.*, 22(2):259–283, **2010**.

[Heg95] M. HEGLAND. *Variable Hilbert scales and their interpolation inequalities with applications to Tikhonov regularization*. *Appl. Anal.*, 59(1-4):207–223, **1995**.

[HK10a] T. HEIN AND K. S. KAZIMIERSKI. *Accelerated Landweber iteration in Banach spaces*. *Inverse Probl.*, 26(5):055002, 17, **2010**.

[HK10b] T. HEIN AND K. S. KAZIMIERSKI. *Modified Landweber iteration in Banach spaces - convergence and convergence rates*. *Numer. Funct. Anal. Optim.*, 31(10):1158–1184, **2010**.

[HKPS07] B. HOFMANN, B. KALTENBACHER, C. PÖSCHL AND O. SCHERZER. *A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators*. *Inverse Probl.*, 23(3):987–1010, **2007**.

[HNS95] M. HANKE, A. NEUBAUER AND O. SCHERZER. *A convergence analysis of the Landweber iteration for nonlinear ill-posed problems*. *Numer. Math.*, 72:21–37, **1995**.

[Hoh97] T. HOHAGE. *Logarithmic convergence rates of the Iteratively Regularized Gauss-Newton Method for an inverse potential and an inverse scattering problem*. *Inverse Probl.*, 13(5):1279–1299, **1997**.

[Hoh98] T. HOHAGE. *Convergence rates of a regularized Newton method in sound-hard inverse scattering*. *SIAM J. Numer. Anal.*, 36:125–142, **1998**.

[Hoh99] T. HOHAGE. *Iterative Methods in Inverse Obstacle Scattering: Regularization Theory of Linear and Nonlinear Exponentially Ill-Posed Problems*. Ph.D. thesis, Johannes Kepler Universität Linz, **1999**.

[Hoh00] T. HOHAGE. *Regularization of exponentially ill-posed problems*. *Numer. Funct. Anal. Optim.*, 21:439–464, **2000**.

[HS92] S. HELL AND E. H. K. STELZER. *Properties of a 4pi confocal fluorescence microscope*. *J. Opt. Soc. Am. A*, 9(12):2159–2166, **1992**.

[Hur89] N. E. HURT. *Phase retrieval and zero crossings*, volume 52 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, **1989**.

[HW11] T. HOHAGE AND F. WERNER. *Iteratively regularized Newton methods with general data misfit functionals and applications to Poisson data*. *http://arxiv.org/abs/1105.2690v1*, **2011**.

[HY10] B. HOFMANN AND M. YAMAMOTO. *On the interplay of source conditions and variational inequalities for nonlinear ill-posed problems*. *Appl. Anal.*, 89(11):1705 – 1727, **2010**.

[IK10] O. IVANYSHYN AND R. KRESS. *Identification of sound-soft 3D obstacles from phaseless data*. *Inverse Probl. Imag.*, 4(1):131–149, **2010**.

[Kal97] O. KALLENBERG. *Foundations of modern probability*. Probability and its applications. Springer, **1997**.

[KH10] B. KALTENBACHER AND B. HOFMANN. *Convergence Rates for the Iteratively Regularized Gauss-Newton Method in Banach Spaces*. *Inverse Probl.*, 26(3):035007, **2010**.

[Kin93] J. F. C. KINGMAN. *Poisson processes*, volume 3 of *Oxford Studies in Probability*. The Clarendon Press Oxford University Press, New York, **1993**.

[Kli06] M. V. KLIBANOV. *On the recovery of a 2-D function from the modulus of its Fourier transform*. *J. Math. Anal. Appl.*, 323(2):818–843, **2006**.

[KMM78] J. KERSTAN, K. MATTHES AND J. MECKE. *Infinitely divisible point processes*. Wiley series in probability and mathematical statistics. Wiley, **1978**.

[KNS08] B. KALTENBACHER, A. NEUBAUER AND O. SCHERZER. *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*, volume 6 of *Radon Series on Computational and Applied Mathematics*. de Gruyter, **2008**.

[KR97] R. KRESS AND W. RUNDELL. *Inverse obstacle scattering with modulus of the far field pattern as data*. In *Inverse problems in medical imaging and nondestructive testing (Oberwolfach, 1996)*, pages 75–92. Springer, Vienna, **1997**.

[Kre99] R. KRESS. *Linear integral equations*. Applied mathematical sciences. Springer, **1999**.

[KSS09] B. KALTENBACHER, F. SCHÖPFER AND T. SCHUSTER. *Iterative methods for nonlinear ill-posed problems in Banach spaces: convergence and applications to parameter identification problems*. *Inverse Probl.*, 25(6):065003, **2009**.

[KST95] M. V. KLIBANOV, P. E. SACKS AND A. V. TIKHONRAVOV. *The phase retrieval problem*. *Inverse Probl.*, 11(1):1–28, **1995**.

[Lep90] O. V. LEPSKIĬ. *A problem of adaptive estimation in Gaussian white noise*. *Teor. Veroyatnost. i Primenen.*, 35(3):459–470, **1990**.

[Lor08] D. A. LORENZ. *Convergence rates and source conditions for Tikhonov regularization with sparsity constraints*. *J. Inverse Ill-Posed P.*, 16(5):463–478, **2008**.

[Mat06] P. MATHÉ. *The Lepskiĭ principle revisited*. *Inverse Probl.*, 22(3):L11–L15, **2006**.

[MH08] P. MATHÉ AND B. HOFMANN. *How general are general source conditions?* *Inverse Probl.*, 24(1):015009, **2008**.

[MP03] P. MATHÉ AND S. PEREVERZEV. *Geometry of linear ill-posed problems in variable Hilbert scales*. *Inverse Probl.*, 19(3):789–803, **2003**.

[Neu89] A. NEUBAUER. *Tikhonov regularisation for non-linear ill-posed problems: optimal convergence rates and finite-dimensional approximation*. *Inverse Probl.*, 5(4):541–557, **1989**.

[Pag06] D. PAGANIN. *Coherent X-Ray Optics*. Oxford University Press, **2006**.

[Pös08] C. PÖSCHL. *Tikhonov Regularization with General Residual Term*. Ph.D. thesis, Universität Innsbruck, **2008**.

[RB03] P. REYNAUD-BOURET. *Adaptive estimation of the intensity of inhomogeneous Poisson processes via concentration inequalities*. *Probab. Theory Rel.*, 126(1):103–153, **2003**.

[Rei93] R.D. REISS. *A course on point processes*. Springer series in statistics. Springer, **1993**.

[Res05] E. RESMERITA. *Regularization of ill-posed problems in Banach spaces: convergence rates*. *Inverse Probl.*, 21(4):1303–1314, **2005**.

[SBH11] R. STÜCK, M. BURGER AND T. HOHAGE. *The Iteratively Regularized Gauß-Newton Method with Convex Constraints and Applications in 4Pi-Microscopy*. *Inverse Probl.*, to appear, **2011**.

[SGG+08] O. SCHERZER, M. GRASMAIR, H. GROSSAUER, M. HALTMEIER AND F. LENZEN. *Variational Methods in Imaging*. Applied Mathematical Sciences. Springer, **2008**.

[Stü11] R. STÜCK. *Semi-blind Deconvolution in 4Pi–Microscopy*. Ph.D. thesis, University of Göttingen, **2011**.

[Tik63a] A. N. TIKHONOV. *On the solution of incorrectly formulated problems and the regularization method*. *Soviet Math. Doklady*, 4:1035–1038, **1963**.

[Tik63b] A. N. TIKHONOV. *Regularization of incorrectly posed problems*. *Soviet Math. Doklady*, 4:1624–1627, **1963**.

[VSE+10] G. VICIDOMINI, R. SCHMIDT, A. EGNER, S. W. HELL AND A. SCHÖNLE. *Automatic deconvolution in 4pi-microscopy with variable phase*. *Opt. Exp.*, 18(8):10154 – 10167, **2010**.

[Wlo87] J. WLOKA. *Partial differential equations*. Cambridge University Press, **1987**.

[XR91] Z. B. XU AND G. F. ROACH. *Characteristic inequalities of uniformly convex and uniformly smooth Banach spaces*. *J. Math. Anal. Appl.*, 157(1):189–210, **1991**.

[Zar09] C. ZARZER. *On Tikhonov regularization with non-convex sparsity constraints*. *Inverse Probl.*, 25(2):025006, **2009**.

Name	Frank Werner
Address	Institute for Numerical and Applied Mathematics University of Göttingen Lotzestraße 16-18 37083 Göttingen / Germany
Email:	f.werner@math.uni-goettingen.de

Personal Details

Gender	Male
Date of birth	11th of March, 1985
Place of birth	Hofgeismar, Germany
Citizenship	German

Education

02/2009-01/2012	Ph.D. student of mathematics at the University of Göttingen (Germany). Supervisor: T. Hohage.
10/2004-01/2009	Student of mathematics (for diploma) at the University of Göttingen (Germany). Apr. 17, 2006 Intermediate examination for diploma in mathematics. Jan. 23, 2009 Examination (Diplom) in mathematics (passed with distinction). Thesis title: *Ein neuer numerischer Ansatz zur* $\mathbf{L}^p$*-Regularisierung*, Advisor: T. Hohage.
07/2001-06/2004	Secondary school 'Albert Schweitzer Schule' in Hofgeismar (Germany). June 25, 2004 Final exam.
07/1995-06/2001	Comprehensive school 'HGS Grebenstein' in Grebenstein (Germany).
07/1991-06/1995	Primary school 'Burgbergschule' in Grebenstein (Germany).

Research Experience

01/2011-01/2012	Scholarship holder of the Research Training Group 1023: Identification in Mathematical Models at the Institute for Numerical and Applied Mathematics, University of Göttingen (Germany).
09/2009-12/2010	Research assistant at the Institute for Mathematical Stochastics, University of Göttingen (Germany) in the BMBF-project INVERS: Deconvolution with sparsity constraints.
02/2009-08/2009	Scholarship holder of the Research Training Group 1023: Identification in Mathematical Models at the Institute for Numerical and Applied Mathematics, University of Göttingen (Germany).