

## DEAUTOCONVOLUTION IN THE TWO-DIMENSIONAL CASE\*

YU DENG<sup>†</sup>, BERND HOFMANN<sup>†</sup>, AND FRANK WERNER<sup>‡</sup>

**Abstract.** There is extensive mathematical literature on the inverse problem of deautoconvolution for a function with support in the unit interval  $[0, 1] \subset \mathbb{R}$ , but little is known about the multidimensional situation. This article tries to fill this gap with analytical and numerical studies on the reconstruction of a real function of two real variables over the unit square from observations of its autoconvolution on  $[0, 2]^2 \subset \mathbb{R}^2$  (full data case) or on  $[0, 1]^2$  (limited data case). In an  $L^2$ -setting, twofoldness and uniqueness assertions are proven for the deautoconvolution problem in 2D. Moreover, its ill-posedness is characterized and illustrated. Extensive numerical case studies give an overview of the behaviour of stable approximate solutions to the two-dimensional deautoconvolution problem obtained by Tikhonov-type regularization with different penalties and the iteratively regularized Gauss–Newton method.

**Key words.** deautoconvolution, inverse problem, ill-posedness, case studies in 2D, Tikhonov-type regularization, iteratively regularized Gauss–Newton method

**AMS subject classifications.** 47J06, 65R32, 45Q05, 47A52, 65J20

**1. Introduction.** The object of research in this work is the problem of *deautoconvolution*, where our focus is on the *two-dimensional case*. A square integrable real function of two variables  $x(t_1, t_2)$  ( $0 \leq t_1, t_2 \leq 1$ ) is to be identified from the function  $y = x * x$ , i.e., its autoconvolution. If we consider  $x$  as an element of the Hilbert space  $L^2(\mathbb{R}^2)$  with support  $\text{supp}(x) \subseteq [0, 1]^2$ , then, it is well-known that  $x * x$  belongs to  $L^2(\mathbb{R}^2)$  as well, with support  $\text{supp}(x * x) \subseteq [0, 2]^2$ . In this context, the elements  $x$  and  $x * x$  can be both considered as tempered distributions with compact support, where  $\text{supp}(\cdot)$  is defined as the essential support with respect to the Lebesgue measure  $\lambda$  in  $\mathbb{R}^2$ . Instead of  $y$  itself, only a noise corrupted version of it, denoted by  $y^\delta \in L^2(\mathbb{R}^2)$ , is available, where  $\delta \geq 0$  denotes the noise level. Since the inverse problem of deautoconvolution tends to be ill-posed, the aim of the recovery process is to find *stable approximate solutions* of  $x$  based on the data  $y^\delta$ . We are going to distinguish the *full data case*, where noisy data are available for  $y(s_1, s_2)$  ( $0 \leq s_1, s_2 \leq 2$ ), and the *limited data case*, where data are given for  $y(s_1, s_2)$  ( $0 \leq s_1, s_2 \leq 1$ ). Since the scope of the data in the limited data case is only 25%, when compared to the full data case, the effect of ill-posedness is stronger in this case. As a consequence, the chances for an accurate recovery of  $x$  are more restricted in the limited data case, than in the full case.

The simplest application of our deautoconvolution problem in two dimensions is the recovery of the density function  $x$ , with support in the unit square  $[0, 1]^2$ , of a two-dimensional random variable  $\mathfrak{X}$  from observations of the density function  $y$  of the two-dimensional random variable  $\mathfrak{Y} := \hat{\mathfrak{X}} + \bar{\mathfrak{X}}$ , where  $\mathfrak{X}$ ,  $\hat{\mathfrak{X}}$ , and  $\bar{\mathfrak{X}}$  are assumed to be independent and identically distributed (i.i.d.).

The deautoconvolution problem *in one dimension* has been considered extensively in the literature motivated by physical applications in spectroscopy; see, e.g., [5, 29]. Its mathematical analysis has been comprehensively implemented in the last decades with focus on properties of the specific forward operator, ill-posedness, and regularization based on the seminal paper [20]. In this context, we refer to [8, 10, 11, 13, 15, 16, 17, 23] for investigations concerning the stable identification of real functions  $x$  on the unit interval  $[0, 1]$  from noisy data of its autoconvolution

---

\*Received October 25, 2022, Accepted March 06, 2023, Published online April 20, 2023. Recommended by A. Buccini. Yu Deng and Bernd Hofmann are supported by the German Science Foundation (DFG) under grant HO 1454/13-1 (Project No. 453804957).

<sup>†</sup>Faculty of Mathematics, Chemnitz University of Technology, Reichenhainer Str. 39/41, 09107 Chemnitz, Germany (yu.deng@math.tu-chemnitz.de, hofmannb@mathematik.tu-chemnitz.de).

<sup>‡</sup>Institute for Mathematics, University of Würzburg, Emil-Fischer-Str. 30, 97074 Würzburg, Germany (frank.werner@mathematik.uni-wuerzburg.de).

$x*x$ . A new series of interdisciplinary autoconvolution studies was developed by a cooperation started in 2010 between a research group of the Max Born Institute for Nonlinear Optics and Short Pulse Spectroscopy (Berlin) led by Prof. Günter Steinmeyer and the Chemnitz research group on regularization. We refer to the publications [1, 9, 18, 19] for a presentation of the output of this cooperation. The goal of this cooperation between mathematics and laser optics was the extension of the one-dimensional deautoconvolution problem to complex-valued functions combining amplitude and phase functions for characterizing ultrashort laser pulses.

In this article, in an  $L^2$ -setting, we are considering a series of numerical case studies for the *nonlinear Volterra-type integral equation*

$$(1.1) \quad F(x) = y, \quad \text{with } F(x) := x * x,$$

the solution of which solves the *deautoconvolution problem in two dimensions*. Equation (1.1) is a special case of a nonlinear operator equation

$$F(x) = y, \quad F : \mathcal{D}(F) \subseteq X \rightarrow Y,$$

with *forward operator*  $F$  mapping between *real-valued Hilbert spaces*  $X$  and  $Y$  with norms  $\|\cdot\|_X$  and  $\|\cdot\|_Y$ , respectively, and domain  $\mathcal{D}(F)$ .

Depending on the data, we have to distinguish between the full and limited data cases. In the first case the forward operator  $F : X = L^2([0, 1]^2) \rightarrow Y = L^2([0, 2]^2)$  is defined by

$$(1.2) \quad [F(x)](s_1, s_2) := \int_{\max(s_2-1, 0)}^{\min(s_2, 1)} \int_{\max(s_1-1, 0)}^{\min(s_1, 1)} x(s_1-t_1, s_2-t_2) x(t_1, t_2) dt_1 dt_2 \quad (0 \leq s_1, s_2 \leq 2)$$

and in the second case the forward operator  $F : X = L^2([0, 1]^2) \rightarrow Y = L^2([0, 1]^2)$  is

$$(1.3) \quad [F(x)](s_1, s_2) := \int_0^{s_2} \int_0^{s_1} x(s_1-t_1, s_2-t_2) x(t_1, t_2) dt_1 dt_2 \quad (0 \leq s_1, s_2 \leq 1).$$

In general, we consider  $\mathcal{D}(F) = X = L^2([0, 1]^2)$ , but for the limited data case we partially focus on *non-negative solutions* expressed by the domain  $\mathcal{D}(F) = \mathcal{D}^+$  with

$$(1.4) \quad \mathcal{D}^+ := \{x \in X = L^2([0, 1]^2) : x \geq 0 \text{ a.e. on } [0, 1]^2\}.$$

For any function  $x \in L^2([0, 1]^2)$  the autoconvolution products  $F(x) = x * x$  and  $F(-x) = (-x) * (-x)$  coincide for both forward operator versions (1.2) and (1.3). However, it is of interest whether, for  $y = x * x$ , the elements  $x$  and  $-x$  are the only solutions of equation (1.1) or not. Moreover, it is of interest whether in the limited data case the restriction of the domain  $\mathcal{D}(F)$  to  $\mathcal{D}^+$  from (1.4) leads to unique solutions. Some answers to those questions will be given in the subsequent Section 2.

The remainder of the paper is organized as follows: Section 2 is devoted to assertions on twofoldness and uniqueness for the deautoconvolution problem in two dimensions, preceded by a subsection with relevant lemmas and definitions. As an inverse problem, deautoconvolution tends to be ill-posed in the setting of infinite dimensional  $L^2$ -spaces. After the presentation of two functions defined over the unit square as basis for later numerical case studies, in Section 3, the specific ill-posedness characteristics for the deautoconvolution of a real function of two real variables with compact support is analyzed and illustrated. To suppress ill-posedness phenomena, variants of variational and iterative regularization methods are used, which will be introduced in Section 4. The numerical treatment, including discretization approaches of forward operator and penalty functionals for the Tikhonov regularization as well as for the iterative regularization by using the Fourier transform, is outlined in Section 5. Section 6 completes the article with comprehensive numerical case studies.

**2. Assertions on twofoldness and uniqueness for the deautoconvolution problem in two dimensions.** In this section we discuss the uniqueness of the solution of the deautoconvolution problem in two dimensions.

**2.1. Preliminaries.** Assertions on twofoldness and uniqueness for the deautoconvolution problem in one dimension have been formulated in the articles [20] for the limited data case and in [19] for the full data case. The respective proofs are based on the Titchmarsh convolution theorem from [31], which was formulated as [20, Lemma 3], and will be recalled below, in a slightly reformulated form, as Lemma 2.1.

LEMMA 2.1. *Let the functions  $f, g \in L^2(\mathbb{R})$  have compact supports  $\text{supp}(f)$  and  $\text{supp}(g)$ , respectively. Then, we have for the convolution that  $f * g \in L^2(\mathbb{R})$  and that the equation*

$$\inf \text{supp}(f * g) = \inf \text{supp}(f) + \inf \text{supp}(g)$$

*holds. In particular, if  $\text{supp}(f)$  and  $\text{supp}(g)$  are covered by the unit interval  $[0, 1]$ , we conclude, from*

$$[f * g](s) = \int_{\max(s-1, 0)}^{\min(s, 1)} f(s-t)g(t) dt = 0 \quad \text{a.e. for } s \in [0, \gamma] \quad (\gamma \leq 2),$$

*that there are numbers  $\gamma_1, \gamma_2 \in [0, 1]$ , with  $\gamma_1 + \gamma_2 \geq \gamma$ , such that*

$$f(t) = 0 \quad \text{a.e. for } t \in [0, \gamma_1] \quad \text{and} \quad g(t) = 0 \quad \text{a.e. for } t \in [0, \gamma_2].$$

For an extension of the Titchmarsh convolution theorem to two dimensions, we mention the following Lemma 2.2; see [26, 27].

LEMMA 2.2. *Let the functions  $f, g \in L^2(\mathbb{R}^2)$  have compact supports  $\text{supp}(f)$  and  $\text{supp}(g)$ , respectively. Then, we have, for the convolution, that  $f * g \in L^2(\mathbb{R}^2)$  and that the equation*

$$(2.1) \quad \text{ch} \text{supp}(f * g) = \text{ch} \text{supp}(f) + \text{ch} \text{supp}(g)$$

*holds, where  $\text{ch } M$  denotes the convex hull of a set  $M \subseteq \mathbb{R}^2$ . In the special case that  $\text{supp}(f * g) = \emptyset$  we have that at least one of the supports  $\text{supp}(f)$  or  $\text{supp}(g)$  is the empty set.*

DEFINITION 2.3. *We call  $x^\dagger \in L^2([0, 1]^2)$ , with  $\text{supp}(x^\dagger) \subseteq [0, 1]^2$ , in the full data case, a solution of the operator equation (1.1), for a given  $y \in L^2([0, 2]^2)$ , if it satisfies the condition*

$$(2.2) \quad [x^\dagger * x^\dagger](s_1, s_2) = y(s_1, s_2) \quad \text{a.e. for } (s_1, s_2) \in [0, 2]^2.$$

DEFINITION 2.4. *We call  $x^\dagger \in L^2([0, 1]^2)$ , with  $\text{supp}(x^\dagger) \subseteq [0, 1]^2$ , in the limited data case, a solution to the operator equation (1.1), for a given  $y \in L^2([0, 1]^2)$ , if it satisfies the condition*

$$(2.3) \quad [x^\dagger * x^\dagger](s_1, s_2) = y(s_1, s_2) \quad \text{a.e. for } (s_1, s_2) \in [0, 1]^2.$$

*If  $x^\dagger \in \mathcal{D}^+$ , with  $\mathcal{D}^+$  in (1.4), we call it a non-negative solution in the limited data case.*

DEFINITION 2.5. *We call  $x \in L^2([0, 1]^2)$ , with  $\text{supp}(x) \subseteq [0, 1]^2$ , satisfying (2.2) or (2.3) a factored solution of equation (1.1) in the full data case or in the limited data case, respectively, if it has the structure*

$$x(t_1, t_2) = x_1(t_1)x_2(t_2) \quad (0 \leq t_1, t_2 \leq 1),$$

with  $x_i \in L^2([0, 1])$ ,  $\text{supp}(x_i) \subseteq [0, 1]$ , for  $i = 1$  and  $i = 2$ . If, moreover,  $x_i \geq 0$  a.e. on  $[0, 1]$  for  $i = 1$  and  $i = 2$ , then we call it a non-negative factored solution in the respective case.

**2.2. Results for the full data case.** Lemma 2.2 allows us to prove the following theorem for the forward autoconvolution operator  $F : L^2([0, 1]^2) \rightarrow L^2([0, 2]^2)$  in (1.2). This result is an extension of [19, Theorem 4.2] to the two-dimensional case of the deautoconvolution problem.

**THEOREM 2.6.** *If, for a given  $y \in L^2([0, 2]^2)$ , the function  $x^\dagger \in L^2([0, 1]^2)$ , with  $\text{supp}(x^\dagger) \subseteq [0, 1]^2$  is a solution of (1.1) with  $F$  in (1.2), then  $x^\dagger$  and  $-x^\dagger$  are the only solutions of this equation in the full data case.*

*Proof.* Let  $x^\dagger \in L^2([0, 1]^2)$ , with  $\text{supp}(x^\dagger) \subseteq [0, 1]^2$ , and  $h \in L^2([0, 1]^2)$ , with  $\text{supp}(h) \subseteq [0, 1]^2$ . Assume that  $x^\dagger$  and  $x^\dagger + h$  solve the equation (1.1), i.e.,

$$[x^\dagger * x^\dagger](s_1, s_2) = [(x^\dagger + h) * (x^\dagger + h)](s_1, s_2) \text{ a.e. for } (s_1, s_2) \in [0, 2]^2.$$

Therefore, we have

$$[(x^\dagger + h) * (x^\dagger + h) - x^\dagger * x^\dagger](s_1, s_2) = [h * (2x^\dagger + h)](s_1, s_2) = 0 \text{ a.e. for } (s_1, s_2) \in [0, 2]^2.$$

We can apply Lemma 2.2 setting  $f := h$  and  $g := 2x^\dagger + h$ . Taking into account that  $\text{supp}(h * (2x^\dagger + h)) \subseteq [0, 2]^2$ , we have  $\text{supp}(h * (2x^\dagger + h)) = \emptyset$  and, consequently,  $\text{ch supp}(h * (2x^\dagger + h)) = \emptyset$ . This implies, due to equation (2.1), that either  $\text{supp}(h) = \emptyset$  or  $\text{supp}(2x^\dagger + h) = \emptyset$ . On the one hand,  $\text{supp}(h) = \emptyset$  leads to the solution  $x^\dagger$  itself. Whereas, on the other hand,  $\text{supp}(2x^\dagger + h) = \emptyset$  leads to  $[2x^\dagger + h](t_1, t_2) = 0$  a.e. for  $(t_1, t_2) \in [0, 1]^2$  and, consequently, with  $h = -2x^\dagger$ , to the second solution  $-x^\dagger$ . Alternative solutions are thus excluded, which proves the theorem.  $\square$

**2.3. Results for the limited data case.** For solutions  $x^\dagger \in L^2([0, 1]^2)$  of equation (1.1), with  $\text{supp}(x^\dagger) \subseteq [0, 1]^2$ , the condition  $0 \in \text{supp}(x^\dagger)$  plays a prominent role in the limited data case. This condition means that, for any ball  $B_r(0)$  around the origin with arbitrary small radius  $r > 0$ , there exists a set  $M_r \subset B_r(0) \cap [0, 1]^2$  with Lebesgue measure  $\lambda(M_r) > 0$  such that  $x^\dagger(t_1, t_2) \neq 0$  a.e. for  $(t_1, t_2) \in M_r$ . Vice versa, for  $0 \notin \text{supp}(x^\dagger)$ , we have some sufficiently small radius  $r > 0$  such that  $x^\dagger(t_1, t_2) = 0$  a.e. for  $(t_1, t_2) \in B_r(0) \cap [0, 1]^2$ .

First, we generalize in Theorem 2.7 those aspects of [20, Theorem 1] that concern the strong non-injectivity of the autoconvolution operator in the limited data case.

**THEOREM 2.7.** *If, for a given  $y \in L^2([0, 1]^2)$ , the function  $x^\dagger \in L^2([0, 1]^2)$ , with  $\text{supp}(x^\dagger) \subseteq [0, 1]^2$ , is a solution of (1.1), with  $F$  in (1.3), that fulfills the condition*

$$(2.4) \quad 0 \notin \text{supp}(x^\dagger),$$

then, in the limited data case, there exists infinitely many other solutions  $\hat{x}^\dagger \in L^2([0, 1]^2)$  of (1.1), with  $\text{supp}(\hat{x}^\dagger) \subseteq [0, 1]^2$ .

*Proof.* If (2.4) holds, there exists some  $0 < \varepsilon < 1/2$  such that  $x^\dagger(t_1, t_2) = 0$  a.e. for  $(t_1, t_2) \in [0, \varepsilon]^2$ . Then, for all elements  $h \in L^2([0, 1]^2)$ , with  $\text{supp}(h) \subseteq [0, 1]^2$ , satisfying the condition

$$h(t_1, t_2) = 0 \text{ a.e. for } (t_1, t_2) \in [0, 1]^2 \setminus [1 - \varepsilon, 1]^2,$$

we have that  $\hat{x}^\dagger = x^\dagger + h$  obeys the condition

$$[\hat{x}^\dagger * \hat{x}^\dagger](s_1, s_2) = y(s_1, s_2) \text{ a.e. for } (s_1, s_2) \in [0, 1]^2.$$

This is a consequence of the fact that  $[h * (2x^\dagger + h)](s_1, s_2) = 0$  a.e. for  $(s_1, s_2) \in [0, 1]^2$  is true for each such element  $h$ .  $\square$

To formulate uniqueness assertions for solutions  $x^\dagger$  of equation (1.1) in the limited data case, we restrict our considerations now to non-negative solutions and the domain  $\mathcal{D}(F) = \mathcal{D}^+$  in (1.4) for the forward operator  $F$  in (1.3). We present, in Theorem 2.8, a result that extends to the two-dimensional autoconvolution operator  $F : \mathcal{D}^+ \subset L^2([0, 1]^2) \rightarrow L^2([0, 1]^2)$  in (1.3) those aspects of [20, Theorem 1] that concern the solution uniqueness. More in details, we are able to handle the special case of *factored non-negative solutions* in the sense of Definition 2.5. This occurs, for instance, when  $x^\dagger$  is a density function for the two-dimensional random variable  $\mathfrak{X} = (\mathfrak{X}_1, \mathfrak{X}_2)$ , where  $\mathfrak{X}_1$  and  $\mathfrak{X}_2$  are uncorrelated one-dimensional random variables.

**THEOREM 2.8.** *Let, for a given  $y \in L^2([0, 1]^2)$ ,  $x^\dagger$  be a non-negative factored solution of equation (1.1) in the limited data case, which satisfies the condition*

$$(2.5) \quad 0 \in \text{supp}(x^\dagger).$$

*Then, there are no other non-negative factored solutions.*

*Proof.* For the factored situation, we have that the right-hand side  $y$  is also factored as

$$y(s_1, s_2) = y_1(s_1) y_2(s_2) \quad (0 \leq s_1, s_2 \leq 1) \quad \text{and} \quad y_1 = x_1^\dagger * x_1^\dagger, \quad y_2 = x_2^\dagger * x_2^\dagger.$$

Moreover, condition (2.5) implies that

$$(2.6) \quad \inf \text{supp}(x_1^\dagger) = \inf \text{supp}(x_2^\dagger) = 0.$$

Otherwise, there would be a square  $[0, \varepsilon]^2$ , with  $\varepsilon = \max\{\inf \text{supp}(x_1^\dagger), \inf \text{supp}(x_2^\dagger)\} > 0$ , on which  $x^\dagger$  vanishes almost everywhere such that  $0 \notin \text{supp}(x^\dagger)$ . Now we suppose that, for  $i = 1$  and  $i = 2$ , quadratically integrable perturbations  $h_i(t_i)$  ( $0 \leq t_i \leq 1$ ) exist such that  $x_i^\dagger + h_i \geq 0$  a.e. on  $[0, 1]$  and

$$(2.7) \quad (x_1^\dagger + h_1) * (x_1^\dagger + h_1) = y_1 \quad \text{and} \quad (x_2^\dagger + h_2) * (x_2^\dagger + h_2) = y_2.$$

To complete the proof of the theorem we show that  $h_1$  and  $h_2$  have to vanish almost everywhere on  $[0, 1]$ . This can be done with the help of Titchmarsh's convolution theorem in the one-dimensional case; see Lemma 2.1. From (2.7) we derive, for  $i = 1$  and  $i = 2$ , that

$$[h_i * (2x_i^\dagger + h_i)](s_i) = 0 \quad \text{a.e. for } s_i \in [0, 1],$$

where  $x_i^\dagger + h_i \geq 0$  implies that  $2x_i^\dagger + h_i \geq x_i^\dagger$  and  $\inf \text{supp}(2x_i^\dagger + h_i) = 0$  as a consequence of (2.6). Then, it follows, from Lemma 2.1, that

$$\inf \text{supp}(h_i) + \inf \text{supp}(2x_i^\dagger + h_i) = \inf \text{supp}(h_i * (2x_i^\dagger + h_i)) \geq 1$$

and, hence,  $\inf \text{supp}(h_i) \geq 1$  for both  $i = 1, 2$ . This implies  $h_i = 0$  a.e. on  $[0, 1]$  and completes the proof.  $\square$

### 3. Examples and ill-posedness phenomena of deautoconvolution in two dimensions.

We now present two examples to illustrate the ill-posedness phenomena of deautoconvolution in 2D.

**3.1. Two examples.** For the numerical case studies of deautoconvolution in 2D, we will present two examples of solutions  $x^\dagger$  to the autoconvolution equation in 2D. The first one refers to the function

$$(3.1) \quad x^\dagger(t_1, t_2) = \left(-3t_1^2 + 3t_1 + \frac{1}{4}\right) (\sin(1.5\pi t_2) + 1) \quad (0 \leq t_1, t_2 \leq 1)$$

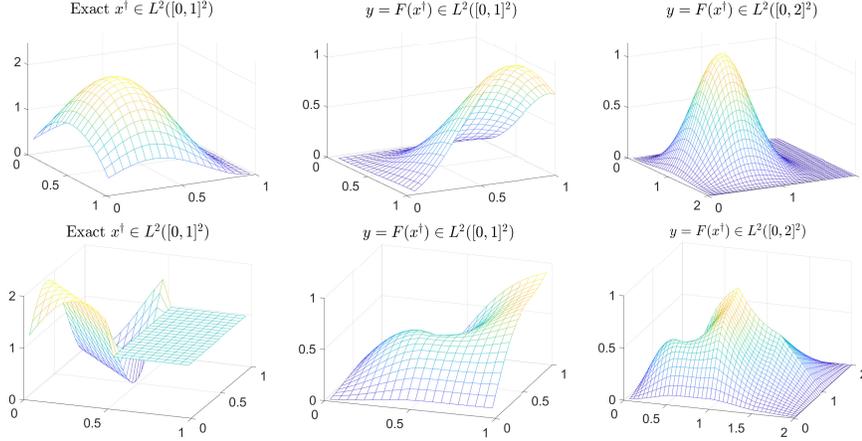


FIGURE 3.1. Functions in the two considered example. The first line reports, from left to right, the smooth and factored function  $x^\dagger(t_1, t_2)$  in (3.1),  $F(x^\dagger)$  with limited data, and  $F(x^\dagger)$  with full data. The second line reports, from left to right, the non-smooth and non-factored function  $x^\dagger(t_1, t_2)$  in (3.2),  $F(x^\dagger)$  with limited data, and  $F(x^\dagger)$  with full data.

to be reconstructed from its own autoconvolution  $F(x^\dagger) = x^\dagger * x^\dagger$ . This smooth and non-negative factored function  $x^\dagger$  is illustrated in the first line of Figure 3.1, alongside with the  $F(x^\dagger)$ -images for the limited and full data case, respectively.

The second example refers to the non-smooth, non-factored, and non-negative solution

$$(3.2) \quad x^\dagger(t_1, t_2) = \begin{cases} \sin(1.5\pi(t_1 + t_2)) + 1 & (0 \leq t_1 \leq 0.5, 0 \leq t_2 \leq 1), \\ 1 & (0.5 < t_1 \leq 1, 0 \leq t_2 \leq 1), \end{cases}$$

which is illustrated in the second line of Figure 3.1, together with  $F(x^\dagger)$ -images for the limited and full data case, respectively.

**3.2. Ill-posedness phenomenon.** As we will see in the numerical case studies presented below, especially in the limited data case, least-squares solutions of the discretized two-dimensional problem of deautoconvolution tend to become strongly oscillating even if the noise level  $\delta > 0$ , in the observed right-hand side  $y^\delta$ , is small. This indicates instability as ill-conditioning phenomenon for a discretized variant of deautoconvolution and ill-posedness for the underlying operator equation (1.1) in infinite dimensional  $L^2$ -spaces. For a theoretical verification we adopt the concept of local ill-posedness along the lines of [21, Def. 1.1] for nonlinear operator equations, and we recall this concept in the following definition.

**DEFINITION 3.1.** An operator equation  $F(x) = y$  with nonlinear forward operator  $F : \mathcal{D}(F) \subseteq X \rightarrow Y$  between the Hilbert spaces  $X$  and  $Y$  with domain  $\mathcal{D}(F)$  is called locally ill-posed at a solution point  $x^\dagger \in \mathcal{D}(F)$ , if there exist, for all closed balls  $\overline{\mathcal{B}}_r(x^\dagger)$  with radius  $r > 0$  and center  $x^\dagger$ , sequences  $\{x_n\} \subset \overline{\mathcal{B}}_r(x^\dagger) \cap \mathcal{D}(F)$  satisfying the condition

$$\|F(x_n) - F(x^\dagger)\|_Y \rightarrow 0, \quad \text{but} \quad \|x_n - x^\dagger\|_X \not\rightarrow 0, \quad \text{as} \quad n \rightarrow \infty.$$

Otherwise, the operator equation is called locally well-posed at  $x^\dagger$ .

Local ill-posedness everywhere on the non-negativity domain

$$\mathcal{D}(F) = \{x \in X = L^2([0, 1]) : x \geq 0 \text{ a.e. on } [0, 1]\}$$

was proven for the one-dimensional deautoconvolution problem in the limited data case in [20, Lemma 6]. With the following proposition we extend, by using similar proof ideas, this

assertion to the two-dimensional case and  $\mathcal{D}(F) = \mathcal{D}^+$ , with  $\mathcal{D}^+$  in (1.4). We should mention, as an overall consequence of the observed ill-posedness, that the *stable approximate solution* of the two-dimensional deautoconvolution problem *requires* the use of variational or iterative regularization methods.

**PROPOSITION 3.2.** *Let  $X = Y = L^2([0, 1]^2)$ . For the limited data case, the operator equation (1.1) with forward operator  $F$  in (1.3) restricted to the non-negativity domain  $\mathcal{D}(F) = \mathcal{D}^+$  in (1.4), is locally ill-posed everywhere on  $\mathcal{D}^+$ .*

*Proof.* Let  $x^\dagger \in \mathcal{D}^+$  be a solution of the operator equation under consideration here. To show local ill-posedness at  $x^\dagger$  we introduce, for a fixed  $r > 0$ , a sequence  $\{h_n\}_{n=3}^\infty$  of perturbations of the form

$$h_n(t_1, t_2) := \begin{cases} nr & \text{for } (t_1, t_2) \in [1 - \frac{1}{n}, 1]^2, \\ 0 & \text{for } (t_1, t_2) \in [0, 1]^2 \setminus [1 - \frac{1}{n}, 1]^2. \end{cases}$$

Let  $x_n := x^\dagger + h_n \in \mathcal{D}^+$ ,  $\|h_n\|_{L^2([0,1]^2)} = r$ , and, consequently,  $x_n \in \overline{\mathcal{B}_r(x^\dagger)} \cap \mathcal{D}^+$  for all  $n \geq 3$ . To complete the proof of the proposition we need to show that the norm  $\|F(x_n) - F(x^\dagger)\|_{L^2([0,1]^2)}$  tends to zero as  $n$  tends to infinity. Due to the facts that

$$F(x_n) - F(x^\dagger) = 2x^\dagger * h_n + h_n * h_n \quad \text{and} \quad \|h_n * h_n\|_{L^2([0,1]^2)} = 0,$$

it is sufficient to show the limit condition

$$\|x^\dagger * h_n\|_{L^2([0,1]^2)} \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Evidently, the non-negative values

$$[x^\dagger * h_n](s_1, s_2) = \int_0^{s_2} \int_0^{s_1} h_n(s_1 - t_1, s_2 - t_2) x^\dagger(t_1, t_2) dt_1 dt_2$$

can be different from zero only for the pairs  $(s_1, s_2) \in [1 - \frac{1}{n}, 1]^2$ . Using the Cauchy-Schwarz inequality and taking into account that  $x^\dagger \in \mathcal{D}^+$ , we have for those pairs the estimate

$$[x^\dagger * h_n](s_1, s_2) = nr \int_0^{s_2 - (1 - \frac{1}{n})} \int_0^{s_1 - (1 - \frac{1}{n})} x^\dagger(t_1, t_2) dt_1 dt_2 \leq r \|x^\dagger\|_{L^2([0,1]^2)}.$$

This, however, yields

$$\|x^\dagger * h_n\|_{L^2([0,1]^2)} \leq r \|x^\dagger\|_{L^2([0,1]^2)} \left( \int_{1 - \frac{1}{n}}^1 \int_{1 - \frac{1}{n}}^1 ds_1 ds_2 \right)^{1/2} = \frac{r \|x^\dagger\|_{L^2([0,1]^2)}}{n}$$

tending to zero as  $n$  tends to infinity, which completes the proof.  $\square$

In the full data case of one-dimensional deautoconvolution, local ill-posedness everywhere has been shown in [17, Proposition 2.3]. The used counterexample, however, is much more sophisticated and requires perturbations with weak poles at the origin. This seems to indicate the significantly lower strength of ill-posedness for the full data case compared to the limited data case. For factored solutions, the counterexample from [17] can also be exploited to prove local ill-posedness for the two-dimensional deautoconvolution problem in the full data case. Numerical case studies confirm the lower level of instability in the full data case of the 2D deautoconvolution compared to the limited data case; see Figure 3.2.

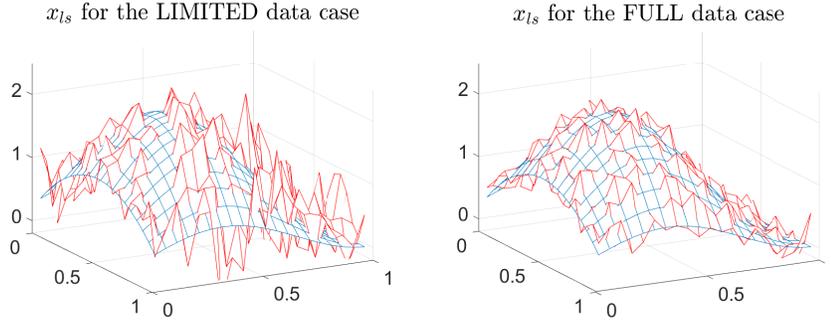


FIGURE 3.2. *Ill-posedness phenomenon of deautoconvolution for the example (1.2).*

Considering the *deterministic noise model*

$$\|y^\delta - y\|_Y \leq \delta,$$

we have calculated *discretized least-square solutions*  $x_{ls}$  of the deautoconvolution problem for the first example  $x^\dagger$  given in (3.1). A discretization with  $21 \times 21$  grid points over the unit square has been exploited for minimizing the Euclidean norm squares of the discretized residuals  $F(x^\dagger) - y^\delta$ . Setting a noise level  $\delta > 0$  that corresponds to a relative data error of 0.8%, Figure 3.2 shows a comparison between the least squares solutions for  $Y := L^2([0, 2]^2)$  of the full data case with  $F$  defined in (1.3) and for  $Y := L^2([0, 1]^2)$  of the limited data case with  $F$  defined in (1.2).

As a consequence of the *ill-posedness phenomenon* of the deautoconvolution problem in 2D, we observe in Figure 3.2 the occurrence of *strong oscillations* in both cases. The oscillations, however, are much heavier for the limited data case (left-hand graph) than for the full data case (right-hand graph). The difference is particularly pronounced for function values on the rear square half of the underlying unit square. The discretized  $L^2$ -norms of the deviation  $x_{ls} - x^\dagger$  correspond to relative errors of 34.54% (left) and 13.92% (right).

**4. Regularization methods.** This section presents two regularization methods for the solution of the deautoconvolution problem in two dimensions.

**4.1. Tikhonov regularization and regularization parameter choices.** As a first approach to overcome the ill-posedness of the two-dimensional deautoconvolution problem, we adopt the variational (Tikhonov-type) regularization, which is well-developed for solving ill-posed nonlinear operator equations. The stable approximate solutions (*regularized solutions*)  $x_\alpha^\delta$  are the global minimizers of the optimization problem

$$(4.1) \quad T_\alpha^\delta(x) := \|F(x) - y^\delta\|_Y^2 + \alpha \mathcal{R}(x) \rightarrow \min_{x \in \mathcal{D}(F) \subseteq X},$$

with *regularization parameter*  $\alpha > 0$  and some *penalty functional*  $\mathcal{R} : X \rightarrow [0, +\infty]$  with domain  $\mathcal{D}(\mathcal{R}) := \{x \in X = L^2([0, 1]^2) : \mathcal{R}(x) < \infty\}$ . The penalty functional is assumed to be *stabilizing*, *convex*, and *weakly sequentially continuous*. For the autoconvolution operator  $F$ , which is *weakly sequentially closed*, the general theory of variational regularization (see, e.g. [30, Section 4.1] and [14, 28]) with respect to *existence*, *stability*, and *convergence* of the Tikhonov-regularized solutions  $x_\alpha^\delta \in \mathcal{D}(F) \cap \mathcal{D}(\mathcal{R})$  applies. The following three penalty functionals are under consideration in this study.

- Classical norm square penalty

$$\mathcal{R}_1(x) := \|x - \bar{x}\|_X^2,$$

with prescribed reference element  $\bar{x} \in X$  and  $\mathcal{D}(\mathcal{R}_1) = X$ . Notably,  $\bar{x} \in X$  expresses some a priori knowledge about the potential solution.

- Gradient norm square penalty

$$\mathcal{R}_2(x) := \int_{[0,1]^2} \|\nabla x\|_2^2 dt_1 dt_2,$$

where  $\nabla x = \left( \frac{\partial x}{\partial t_1}, \frac{\partial x}{\partial t_2} \right)$  denotes the gradient with respect to both variables  $t_1, t_2$  and  $\|\cdot\|_2$  is the Euclidean norm. Here, we have  $\mathcal{D}(\mathcal{R}_2) = H^1([0,1]^2)$ . For this setting, the solution is assumed to have a certain smoothness.

- Total variation penalty

$$\mathcal{R}_3(x) := \|x\|_{TV([0,1]^2)} = \int_{[0,1]^2} \|\nabla x\|_2 dt_1 dt_2,$$

where  $\mathcal{D}(\mathcal{R}_3) = BV([0,1]^2) := \{x \in L^1([0,1]^2) : \|x\|_{TV([0,1]^2)} < \infty\}$  is the space of bounded variation over the unit square  $[0,1]^2$ . This approach was originally introduced for image restoration with the special aim of retaining the information on edges in an image, i.e., the penalty should work for solutions possessing jumps. A detailed analysis of TV-regularization can be found, for example, in [7, 32].

In a first step, we want to choose the optimal regularization parameter  $\alpha_{opt}$  for each input noise level  $\delta$  and corresponding  $y^\delta$  according to

$$\alpha_{opt}(\delta) = \operatorname{argmin}_{\alpha > 0} \|x_\alpha^\delta - x^\dagger\|_X.$$

It is well-known that for the practical use of *a priori choices* for finding the regularization parameter  $\alpha = \alpha(\delta)$ , some smoothness information about the exact solution  $x^\dagger$  is required, which is normally not available. Therefore, *a posteriori choices*  $\alpha = \alpha(\delta, y^\delta)$ , exploiting the measured noisy data  $y^\delta$  in combination with knowledge of the noise level  $\delta > 0$ , are an appropriate alternative; see [14, Sect. 3.1]. Under the limit conditions

$$(4.2) \quad \alpha(\delta, y^\delta) \rightarrow 0 \quad \text{and} \quad \frac{\delta^2}{\alpha(\delta, y^\delta)} \rightarrow 0 \quad \text{as} \quad \delta \rightarrow 0,$$

the regularized solutions  $x_\alpha^\delta$  solving the optimization problem (4.1) may possess a subsequence which converges to an exact solution  $x^\dagger$  as  $\delta \rightarrow 0$ . Due to the Fréchet differentiability of the autoconvolution operator  $F$  and convexity of the penalty functionals  $\mathcal{R}(x)$ , for the numerical experiment below, we implement, as a second step, the *sequential discrepancy principle (SDP)* which was analyzed, for example, in [2].

**DEFINITION 4.1.** *For given  $\tau > 1$ ,  $\alpha_0 > 0$ , and  $0 < q < 1$ , a parameter  $\alpha_{SDP}$  is chosen from the set  $\Delta_q := \{\alpha_l : \alpha_l = q^l \alpha_0, l \in \mathbb{Z}\}$  according to the sequential discrepancy principle (SDP), if*

$$\|F(x_{\alpha_{SDP}}^\delta) - y^\delta\|_Y \leq \tau\delta < \|F(x_{\alpha_{SDP/q}}^\delta) - y^\delta\|_Y$$

*holds true.*

We can directly apply [2, Theorem 1] to our autoconvolution problem and conclude that, with some  $\bar{\delta} > 0$ , the regularization parameters  $\alpha_{SDP} = \alpha_{SDP}(\delta, y^\delta)$ , chosen according to SDP, exist for  $0 < \delta < \bar{\delta}$  and satisfy the limit conditions (4.2). Then, the associated

regularized solutions  $x_{\alpha_{SDP}(\delta, y^\delta)}^\delta$  converge (at least in the sense of subsequences) to exact solutions  $x^\dagger$  as  $\delta \rightarrow 0$  and, moreover,  $\lim_{\delta \rightarrow 0} \mathcal{R}(x_{\alpha_{SDP}(\delta, y^\delta)}^\delta) = \mathcal{R}(x^\dagger)$ .

In a third step we search for *heuristic choices* of the regularization parameter  $\alpha > 0$  assuming that the noise level  $\delta > 0$  is not available or reliable. In our numerical case studies we focus only on the quasi-optimality criterion to find  $\alpha_{qo} = \alpha_{qo}(y^\delta)$ ; see [3, 4, 25] and references therein.

**DEFINITION 4.2.** *For sufficiently large  $\alpha_0 > 0$  and for some  $0 < q < 1$ , we call the parameter  $\alpha_{qo}$  chosen from the set  $\Delta_q := \{\alpha_l : \alpha_l = q^l \alpha_0, l \in \mathbb{N}\}$  according to*

$$\alpha_{qo}(y^\delta) = \operatorname{argmin}_{\alpha_l \in \Delta_q} \left\| x_{\alpha_l}^\delta - x_{\alpha_{l+1}}^\delta \right\|_X$$

*quasi-optimal regularization parameter.*

**4.2. An iteratively regularized Gauss–Newton method.** As an alternative to Tikhonov regularization we can consider the iteratively regularized Gauss–Newton method (IRGNM) and find the minimizers of the functional

$$(4.3) \quad J_{\alpha_n}^\delta(x) := \|F(x_n^\delta) + F'(x_n^\delta)(x - x_n^\delta) - y^\delta\|_Y^2 + \alpha_n \mathcal{R}(x) \rightarrow \min_{x \in \mathcal{D}(F) \subseteq X},$$

with some initial guess  $x_0^\delta \in X$  for a fixed noise level  $\delta > 0$ . Here  $F'(x) : X \rightarrow Y$  is the Fréchet derivative of  $F$  at  $x \in X$  and the sequence  $\{\alpha_n\}_{n=1}^\infty$  of regularization parameters satisfies

$$1 \leq \frac{\alpha_n}{\alpha_{n+1}} \leq C,$$

for some constant  $C > 0$ . The central advantage of (4.3) over (4.1) is that  $x_{n+1}^\delta$  is defined as the solution of (due to the linearity of  $F'(x_n^\delta)$ ) a *convex* optimization problem, which can efficiently be tackled by algorithms such as Chambolle-Pock [12] or FISTA [6]. When the norm square penalty  $\mathcal{R}_1(x)$  is considered, (4.3) can be solved explicitly as

$$(4.4) \quad (F'(x_n^\delta)^* [F'(x_n^\delta)] + \alpha_n \mathbb{I}) x = F'(x_n^\delta)^* [F'(x_n^\delta) x_n^\delta + y^\delta - F(x_n^\delta)] + \alpha_n \bar{x}.$$

Since the gradient operator is also linear, we can solve the linear equation

$$(4.5) \quad (F'(x_n^\delta)^* [F'(x_n^\delta)] + \alpha_n \nabla^* \nabla) x = F'(x_n^\delta)^* [F'(x_n^\delta) x_n^\delta + y^\delta - F(x_n^\delta)],$$

for the gradient norm square penalty  $\mathcal{R}_2(x)$ .

As a computational drawback, however, a full sequence of minimization problems (or linear equations) has to be solved. A convergence analysis for the IRGNM as depicted in (4.3) can, e.g., be found in [22, 33].

Note that a similar approach has been proposed in [29], where the least-squares residuum  $\|F(x) - y^\delta\|_Y^2$  is minimized over a finite dimensional ansatz space  $x \in \operatorname{span}\{\mu_1, \dots, \mu_n\}$ , e.g., consisting of splines, by linearization and iterative updating. In contrast to (4.1) and (4.3), regularization is there obtained by restriction to a finite-dimensional space, but the computational procedure is comparable to our update formula in (4.3).

Instead of choosing a regularization parameter  $\alpha$  in Tikhonov regularization, here we have to select an appropriate stopping index  $n \in \mathbb{N}_0$ . This can, in principle, be done by the same rules as discussed in Section 4.1. The running index  $n \in \mathbb{N}_0$  can be selected by this two approaches.

- In the best case as  $n_{opt}$  with

$$n_{opt}(\delta) = \operatorname{argmin}_{n \in \mathbb{N}_0} \|x_n^\delta - x^\dagger\|_X.$$

- According to a posteriori sequential discrepancy principle for a given constant parameter  $\tau > 1$  as  $n_{SDP}$ , if

$$\|F(x_{n_{SDP}}^\delta) - y^\delta\|_Y \leq \tau\delta < \|F(x_{n_{SDP}-1}^\delta) - y^\delta\|_Y$$

holds true; see [24] and the references therein.

**5. Numerical treatment.** We now discuss how we discretize problem (1.1) and how we numerically solve it.

**5.1. Discretization via the composite midpoint rule.** To discretize the continuous problem, we consider two different approaches. The first option is to divide each direction of the unit square equidistantly in  $n$  partitions with the uniform length  $h := 1/n$ . To discretize the nonlinear convolution equation or deduce the discretized forward operator, it is reasonable to replace the function values  $x(t_1, t_2)$  and  $y(s_1, s_2)$  by countable values  $x_{i,j}$  and  $y_{k,l}$  with

$$x_{i,j} := x\left(\frac{1}{2}(i + (i - 1))h, \frac{1}{2}(j + (j - 1))h\right), \quad y_{k,l} = y(kh, lh)$$

for all  $i, j = 1, \dots, n$  and  $k, l = 1, \dots, n$ , respectively.

**5.1.1. Discretization of forward operator for the limited data case.** The autoconvolution equation in the limited data case can be approximated, by means of the composite midpoint rule, by the discrete equations

$$\sum_{j=1}^l \sum_{i=1}^k h^2 x_{k-i+1, l-j+1} x_{i,j} = y_{k,l}.$$

Since the function  $x$  does not vanish only for  $0 \leq t_1 \leq s_1$  and  $0 \leq t_2 \leq s_2$ , only the indices  $i \leq k$  and  $j \leq l$  need to be taken into account in the discretized version. Collecting the values  $x_{i,j}$  and  $y_{k,l}$  into vectors  $\mathbf{x} := (x_1, \dots, x_p, \dots, x_{n^2})^T$  and  $\mathbf{y} := (y_1, \dots, y_q, \dots, y_{n^2})^T$ , respectively, with  $p := (i - 1) \cdot n + j$  and  $q := (k - 1) \cdot n + l$  for all  $p, q = 1, \dots, n^2$ , we can rewrite the weakly nonlinear forward operator of autoconvolution as

$$F_1(\mathbf{x}) := h^2 M_1(\mathbf{x})\mathbf{x},$$

where  $M_1(\mathbf{x}) \in \mathbb{R}^{n^2 \times n^2}$  is a lower triangular block matrix and has the structure

$$M_1(\mathbf{x}) := \begin{bmatrix} B_1 & 0 & \cdots & 0 \\ B_2 & B_1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ B_n & B_{n-1} & \cdots & B_1 \end{bmatrix},$$

with

$$B_m = \begin{bmatrix} x_{(m-1)n+1} & 0 & \cdots & 0 \\ x_{(m-1)n+2} & x_{(m-1)n+1} & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ x_{mn} & x_{mn-1} & \cdots & x_{(m-1)n+1} \end{bmatrix},$$

for  $1 \leq m \leq n$ . The first derivative of  $F_1(\mathbf{x})$  can be easily obtained as

$$F_1'(\mathbf{x}) = 2h^2 M_1(\mathbf{x}).$$

**5.1.2. Discretization of forward operator for the full data case.** In the full data case, the discrete forward operator of the autoconvolution equation can be derived in a similar manner. On four subareas  $[0, 1]^2$ ,  $[0, 1] \times (1, 2)$ ,  $(1, 2) \times [0, 1]$ , and  $(1, 2)^2$  we write the discrete equations

$$\begin{aligned}
 \sum_{j=1}^l \sum_{i=1}^k h^2 x_{k-i+1, l-j+1} x_{i, j} &= y_{k, l}, & \text{for } 1 \leq k, l \leq n, \\
 \sum_{j=l-n}^{n-1} \sum_{i=1}^k h^2 x_{k-i+1, l-j} x_{i, j+1} &= y_{k, l}, & \text{for } 1 \leq k \leq n \text{ and } n+1 \leq l \leq 2n-1, \\
 \sum_{j=1}^l \sum_{i=k-n}^{n-1} h^2 x_{k-i, l-j+1} x_{i+1, j} &= y_{k, l}, & \text{for } n+1 \leq k \leq 2n-1 \text{ and } 1 \leq l \leq n, \\
 \sum_{j=l-n}^{n-1} \sum_{i=k-n}^{n-1} h^2 x_{k-i, l-j} x_{i+1, j+1} &= y_{k, l}, & \text{for } n+1 \leq k, l \leq 2n-1.
 \end{aligned}$$

On the boundary of  $[0, 2] \times [0, 2]$ , i.e., for either  $k = 2n$  or  $l = 2n$ ,  $y$  vanishes. Note that the grid width remains  $h = 1/n$ . The discretized forward operator can be written as

$$F_2(\mathbf{x}) := h^2 M_2(\mathbf{x}) \mathbf{x}$$

with an extended block matrix  $M_2(\mathbf{x}) \in \mathbb{R}^{4n^2 \times n^2}$ , where

$$M_2(\mathbf{x}) := \begin{bmatrix}
 B_1 & 0 & \cdots & \cdots & 0 \\
 C_1 & 0 & \cdots & \cdots & 0 \\
 B_2 & B_1 & \cdots & \cdots & 0 \\
 C_2 & C_1 & \cdots & \cdots & 0 \\
 \vdots & \ddots & \ddots & \ddots & \vdots \\
 B_n & B_{n-1} & \cdots & \cdots & B_1 \\
 C_n & C_{n-1} & \cdots & \cdots & C_1 \\
 0 & B_n & B_{n-1} & \cdots & B_2 \\
 0 & C_n & C_{n-1} & \cdots & C_2 \\
 \vdots & \ddots & \ddots & \ddots & \vdots \\
 0 & 0 & 0 & \cdots & B_n \\
 0 & 0 & 0 & \cdots & C_n \\
 0 & 0 & 0 & \cdots & 0 \\
 0 & 0 & 0 & \cdots & 0
 \end{bmatrix},$$

with

$$C_m = \begin{bmatrix}
 0 & x_{mn} & x_{mn-1} & \cdots & x_{(m-1)n+2} \\
 0 & 0 & x_{mn} & \ddots & x_{(m-1)n+3} \\
 \vdots & \ddots & \ddots & \vdots & \\
 0 & 0 & 0 & \cdots & x_{mn} \\
 0 & 0 & 0 & \cdots & 0
 \end{bmatrix}$$

and  $B_m$  as above, for  $1 \leq m \leq n$ .

The first derivative of the extended forward operator is given by

$$F_2'(\mathbf{x}) = 2h^2 M_2(\mathbf{x}).$$

**5.1.3. Discretization of the penalty functionals.** The penalty functionals  $\mathcal{R}_1$  and  $\mathcal{R}_2$  can be discretized in a straightforward manner, while the derivatives of  $\nabla x$  are approximated by finite differences. However, to ensure differentiability of the discretization of  $\mathcal{R}_3$  for the Euclidean norm  $\|w\|_2 = \sqrt{|w_1|^2 + |w_2|^2}$  of an arbitrary vector  $w = (w_1, w_2)$ , we take, similarly to [32, Chapt. 8], the approximation  $\|w\|_{2,\beta} := \sqrt{|w_1|^2 + |w_2|^2 + \beta^2}$ , with a proper small positive parameter  $\beta \in (0, 1)$ . This leads to the discretization

$$(5.1) \quad \mathcal{R}_3(x) \approx h \left( \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} \sqrt{(x_{i+1,j} - x_{i,j})^2 + (x_{i,j+1} - x_{i,j})^2 + h^2 \beta^2} \right. \\ \left. + \sum_{i=1}^{n-1} \sqrt{(x_{i+1,n} - x_{i,n})^2 + h^2 \beta^2} + \sum_{j=1}^{n-1} \sqrt{(x_{n,j+1} - x_{n,j})^2 + h^2 \beta^2} \right).$$

**5.2. Discretization of forward operator via Fourier transform.** If we assume that the function  $x \in L^2((0, 1)^2)$  can be extended on the whole domain  $\mathbb{R}^2$ , with the support of  $x$  given by  $\text{supp}(x) \subset [0, 1] \times [0, 1]$ , i.e.,  $x(t) = 0$  for  $t = (t_1, t_2) \notin [0, 1]^2$ , then, the Fourier transform of  $x$  is

$$\mathcal{F}(x)(\omega) := \frac{1}{\sqrt{2\pi}} \int_{[0,1]^2} x(t) e^{i\omega t} dt, \quad \omega \in \mathbb{R}^2.$$

According to the convolution theorem, the autoconvolution operator can be represented by

$$(5.2) \quad F(x) = \mathcal{F}^{-1}(\mathcal{F}(x)^2).$$

Since the Fourier transform operator  $\mathcal{F}$  is linear w.r.t.  $x$ , we can easily obtain the Fréchet derivative and the adjoint operator of the autoconvolution forward operator as

$$(5.3) \quad F'(x)(u) = \mathcal{F}^{-1}(2\mathcal{F}(x)\mathcal{F}(u)), \\ F'(x)^*(v) = \mathcal{F}^{-1}(2\overline{\mathcal{F}(x)}\mathcal{F}(v)),$$

where  $u \in X$ ,  $v \in Y$ , and  $\bar{z}$  denotes the conjugate complex value of  $z$ .

A discretization of the above formulas is directly available by means of the Fast Fourier Transform (FFT) and its inverse (IFFT). However, we should take into account that they consider *periodic* functions, and, hence, the corresponding discretization will be inaccurate especially close to the boundary. Moreover, it is not able to distinguish between the limited and the full data case.

For the limited data case we perform a zero-padding, i.e., we replace the discretization  $\mathbf{x} \in \mathbb{R}^{n \times n}$  by an extended matrix  $\mathbf{x}^z \in \mathbb{R}^{2n \times 2n}$  of the form

$$\mathbf{x}^z = \begin{bmatrix} \mathbf{x} & 0 \\ 0 & 0 \end{bmatrix}.$$

More precisely, if we denote the zero-padding operator by  $Z : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{2n \times 2n}$  and the corresponding left-inverse (restriction) by  $R : \mathbb{R}^{2n \times 2n} \rightarrow \mathbb{R}^{n \times n}$ , then we obtain the discretization of (5.2) as

$$F(x) \approx R(\text{IFFT}(\text{FFT}(Z(\mathbf{x}))^2)).$$

and the corresponding discretizations of (5.3) as

$$F'(x)(u) \approx R(\text{IFFT}(2\text{FFT}(Z(\mathbf{x}))\text{FFT}(Z(\mathbf{u})))) \\ F'(x)^*(v) \approx R(\text{IFFT}(2\overline{\text{FFT}(Z(\mathbf{x}))}\text{FFT}(Z(\mathbf{v}))))).$$

The only difference between the limited data and the full data case is whether the restriction  $R$  as to be applied as a very last step or not.

**5.3. Computational implementation.** To tackle the autoconvolution problem in a stable way, we can either solve the Tikhonov regularized problem (4.1) with different penalty functionals or solve the iteratively regularized problem (4.3) equipped with norm square penalty  $\mathcal{R}_1(x)$  or gradient norm square penalty  $\mathcal{R}_2(x)$ . For both approaches, we need to initialize the regularization parameter  $\alpha_0$  and set the iteration step  $q \in (0, 1)$ , with  $\alpha_l = q^l \alpha_0$ , for  $l = 1, 2, 3, \dots$ .

To solve the Tikhonov regularized problem (4.1) with  $\alpha := \alpha_l$ , we use the discretization via the composite midpoint rule and consider its first-order optimality condition

$$(5.4) \quad 2h^2(M(\mathbf{x}))^T(h^2 M(\mathbf{x})\mathbf{x} - y^\delta) + \alpha \mathcal{R}'(\mathbf{x}) = 0,$$

which will be solved using a damped Newton method. In this nonlinear equation, either  $M := M_1$ , for the limited data case, or  $M := M_2$ , in the full data case. Note that the Newton-type method for solving (5.4) is also an iterative procedure and needs initialization of the solution as well.

For our academical experiments ( $x^\dagger$  is known), we carry out the process according to Algorithm 1.

---

**Algorithm 1** Conceptual algorithm for solving (4.1)

---

**S0:** Let  $\{\delta_k\}_{k \in \mathbb{N}}$  be a finite sequence of positive noise levels tending to 0 as  $k \rightarrow \infty$ . Fix  $\alpha_0 > 0$  and  $0 < q < 1$ . Set  $x_0$  be a starting point,  $l := 0$ ,  $k := 0$  and  $x_{\alpha_l}^{\delta_k} := x_0$ .

**S1:** Compute the error  $E_l^k = \|x_{\alpha_l}^{\delta_k} - x^\dagger\|_X$ .

**S2:** Solve the discretized problem (5.4) for fixed  $\delta := \delta_k$  and  $\alpha := \alpha_l$  using a damped Newton method with starting point  $x_{\alpha_l}^{\delta_k}$ . Let  $x_{\alpha_{l+1}}^{\delta_k}$  be the associated solution and compute the error  $E_{l+1}^k = \|x_{\alpha_{l+1}}^{\delta_k} - x^\dagger\|_X$ .

**S3:** If  $E_{l+1}^k < E_l^k$ , then  $\alpha_{l+1} := q\alpha_l$ ,  $l := l + 1$  and go to **S2**. Otherwise save  $x_{\alpha_l}^{\delta_k}$  and go to **S4**.

**S4:** Set  $x_{\alpha_0}^{\delta_{k+1}} := x_{\alpha_l}^{\delta_k}$ ,  $l := 0$ ,  $k := k + 1$  and go to **S2**.

---

Note that this algorithm can be efficiently implemented due to the square rate of convergence of Newton-type method. However, the derivation of the first and second derivative of the autoconvolution operator and penalty functionals are computationally expensive.

To solve the iteratively regularized problem (4.3), with either  $\mathcal{R} := \mathcal{R}_1$  or  $\mathcal{R} := \mathcal{R}_2$ , more precisely, to solve the linear equations (4.4) and (4.5), we use the Fourier transform technique to discretize the forward operator and apply the Conjugate Gradient (CG) method. Besides benefiting from the fast computational time, we can also avoid the computation of the second derivative matrices of all functionals. Actually, the associated algorithm is similar to Algorithm 1. The notation  $x_{\alpha_l}^{\delta_k}$  in Algorithm 1 corresponds to the solution of (4.4) and (4.5) on the step  $n := l$  for  $\delta := \delta_k$ .

**6. Numerical experiments.** In this section, we present numerical experiments based on both the examples that have been introduced in Section 3.1. For all three penalties  $\mathcal{R}_i(x)$  ( $i = 1, 2, 3$ ) introduced above, properties and accuracy of regularized solutions are compared and illustrated for various regularization parameter choice rules. We define a sequence of finite discrete noise levels  $\{\delta_k\}_{k \in \mathbb{N}}$  with corresponding relative noise levels between 10% and 0.05%. Let input data  $y^{\delta_k} \in Y$  satisfy  $\|y^{\delta_k} - y\|_Y = \delta_k$ . To find appropriate

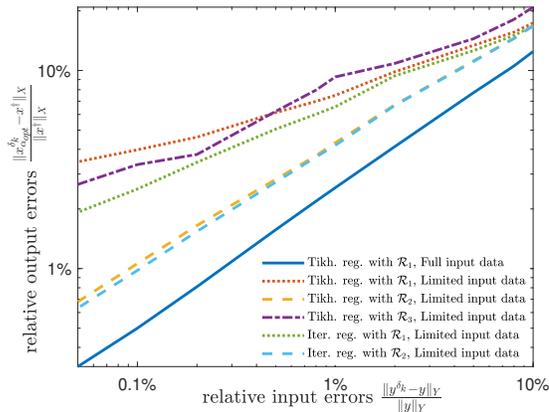


FIGURE 6.1. Comparison of relative error norms of regularized solutions  $x_{\alpha_{opt}}^{\delta_k}$  with optimal regularization parameter for example 1.

regularization parameters  $\alpha$  for each fixed  $\delta_k$ , we set for the sequence  $(\alpha_l)_{l \in \mathbb{N}}$  the starting value  $\alpha_0 = 1$ , step size  $q = 0.5$ , and, thus,  $\alpha_{l+1} = \alpha_l/2$ ; see Algorithm 1. On the discretized domain, with discretization level  $n = 20$ , we solve the nonlinear equation (5.4), discretized counterpart of the linear equations (4.4) and (4.5), for each  $\delta_k$  and regularization parameter  $\alpha_l$ . Additionally, the constant function  $x_0 \equiv 1$  is set as initialization for all computations. For the penalty term  $\mathcal{R}_1(x)$ , the reference element is set as a constant function  $\bar{x} = 0.5$ . If we set  $\bar{x} = -0.5$ , the solutions  $x_{\alpha_l}^{\delta_k}$  will converge to  $-x^\dagger$  for both examples. For the SDP choosing rule to determine the regularization parameters  $\alpha_{SDP}^\delta$  in Tikhonov regularized problem or the running index  $n_{SDP}$  in iteratively regularized problem, the constant parameter  $\tau = 1.2$  is fixed in our experiments. The smoothing parameter  $\beta$  is fixed to 0.1 in the discretized TV penalty (5.1).

**6.1. Results for Example 1.** First, we compare in Figure 6.1 the relative output errors of the regularized solutions  $x_{\alpha_{opt}}^{\delta_k}$  obtained using different penalties and by distinguishing the full data case and the limited data case. Since, for all penalties, the accuracies obtained with the full input data are uniformly better than those obtained with limited input data, we have illustrated in Figure 6.1, for the full data case, only the accuracies related to the  $\mathcal{R}_1(x)$  penalty with Tikhonov regularization. On the other hand, results for all penalties and both regularization methods are displayed for the limited data case.

Obviously, the reconstruction with the full input data achieves the smallest and best output error. For the limited data case, the gradient norm square penalty is the most suitable for this example. A perspicuous reason is the high level of smoothness of the exact solution  $x^\dagger$ . The quality of Tikhonov-regularized solutions based on the classical norm square penalty  $\mathcal{R}_1$  and of the TV-penalty  $\mathcal{R}_3$  is almost indiscernible. Moreover, when using this gradient penalty  $\mathcal{R}_2$ , the quality of results for Tikhonov regularization and iterative regularization are nearly indistinguishable.

In Table 6.1 we present Hölder exponents  $\kappa \in (0, 1)$ , estimated by regression from a series of  $\delta$ -values, which emulate numerically convergence rates results for regularized solutions with the best possible regularization parameter  $\alpha_{opt}$  for each  $\delta$ . More in details, we have listed the exponents  $\kappa$  such that approximately  $\|x_{\alpha_{opt}}^\delta - x^\dagger\|_X \sim \delta^\kappa$  as  $\delta \rightarrow 0$ .

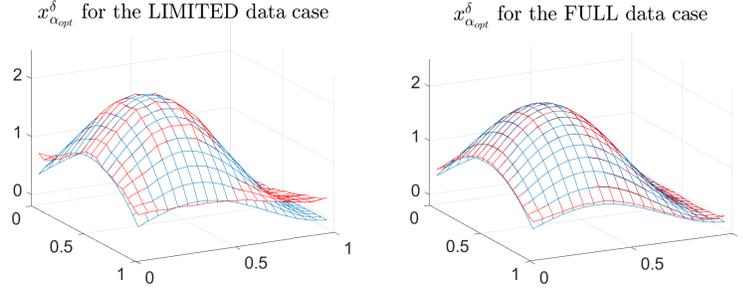


FIGURE 6.2. Regularized solutions with optimal regularization parameter for the limited and the full data case for example 1.

TABLE 6.1

Estimated Hölder exponents  $\kappa \in (0, 1)$  for Hölder convergence rates  $\|x_{\alpha_{opt}}^{\delta} - x^{\dagger}\|_X \sim \delta^{\kappa}$  as  $\delta \rightarrow 0$  in example 1.

Data	Penalty	Tikhonov reg.	Iterative reg.
$Y = L^2([0, 2]^2)$	$\mathcal{R}_1(x)$	0.6946	0.7184
	$\mathcal{R}_2(x)$	0.6638	0.6936
	$\mathcal{R}_3(x)$	0.4685	–
$Y = L^2([0, 1]^2)$	$\mathcal{R}_1(x)$	0.3118	0.4088
	$\mathcal{R}_2(x)$	0.6015	0.6183
	$\mathcal{R}_3(x)$	0.3919	–

For the limited data case, the gradient norm square penalty also delivers the largest Hölder rate exponent for the regularized solutions with best possible regularization parameter, which is consistent with the insights from Figure 6.1. Partially, the rate exponents for the iterative regularization method seems to be higher than in case of Tikhonov regularization.

Next, we compare in Table 6.2 the regularized solutions with respect to different regularization parameter choice rules, for a fixed noise level  $\delta$  of a relative input error  $\frac{\|y^{\delta} - y\|_Y}{\|y\|_Y} = 1\%$ .

TABLE 6.2

Comparison of relative error norms of regularized solutions for a fixed relative input error 1% for example 1.

Method	Data	Penalty	$\frac{\ x_{\alpha_{opt}}^{\delta} - x^{\dagger}\ _X}{\ x^{\dagger}\ _X}$	$\frac{\ x_{\alpha_{SDP}}^{\delta} - x^{\dagger}\ _X}{\ x^{\dagger}\ _X}$	$\frac{\ x_{\alpha_{qo}}^{\delta} - x^{\dagger}\ _X}{\ x^{\dagger}\ _X}$
Tikhonov reg.	$Y = L^2([0, 2]^2)$	$\mathcal{R}_1(x)$	2.56%	2.71%	2.71%
		$\mathcal{R}_1(x)$	7.49%	9.12%	9.12%
	$Y = L^2([0, 1]^2)$	$\mathcal{R}_2(x)$	4.32%	9.91%	4.39%
		$\mathcal{R}_3(x)$	9.28%	13.77%	10.30%
Iterative reg.	$Y = L^2([0, 1]^2)$	$\mathcal{R}_1(x)$	6.56%	9.07%	–
		$\mathcal{R}_2(x)$	4.19%	9.88%	–

The regularized solutions with the regularization parameters via discrepancy principle can be obtained in a more stable way than the use of quasi-optimality criterion, especially in the case of iteratively regularized Gauss–Newton method.

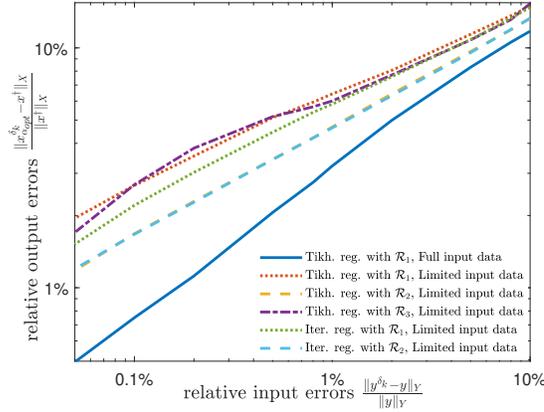


FIGURE 6.3. Comparison of relative error norms of regularized solutions  $x_{\alpha_{opt}}^\delta$  with optimal regularization parameter for example 2.

For a visual comparison of regularized solutions with least-square solutions, which have been presented in Figure 3.2, we report the Tikhonov regularized solutions with the best possible regularization parameter, obtained with classical norm square penalty  $\mathcal{R}_1(x)$  for the same noise level of 0.8%, in Figure 6.2 in the limited and in the full data case, respectively.

We can observe that there are almost no longer oscillations in the solutions. However, we can still notice the deviation of  $x_{\alpha_{opt}}^\delta$  from the exact  $x^\dagger$  on the far back corner of the underlying unit square for the limited data case. These deviations can be eliminated with the use of full input data.

**6.2. Results for Example 2.** In Figure 6.3 we present analog results for Example 2 with  $x^\dagger$  from (3.2), which represents a non-smooth and non-factored function.

From the results in Table 6.3 we can observe that the Total Variation penalty  $\mathcal{R}_3(x)$  yields the best results among all penalties. Intuitively this is due to the fact that the function  $x^\dagger$  of Example 2 presents jumps and is less smooth than the one in Example 1.

Figure 6.4 shows the Tikhonov-regularized solutions with best possible regularization parameters, calculated for different penalties and limited input data with a relative noise level of 0.8%. The improvement obtained using the TV penalty  $\mathcal{R}_3(x)$  can be especially observed on the rear area, where the function values tend to be constant. By contrast, the gradient norm square penalty  $\mathcal{R}_2(x)$  makes the associated regularized solution smoother on the front area.

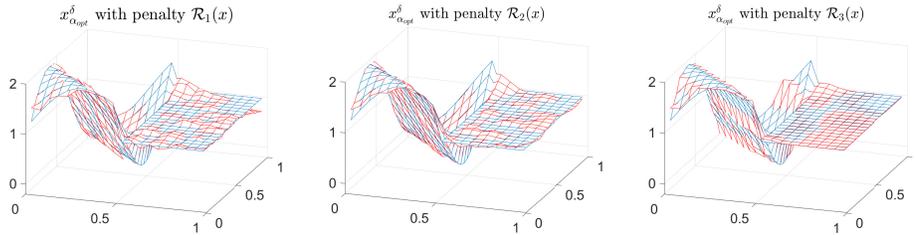


FIGURE 6.4. Regularized solutions for example 2 obtained with optimal regularization parameter for different penalties.

Similarly to the previous example we present the estimated Hölder exponents  $\kappa$  for regularized solutions  $x_{\alpha_{opt}}^{\delta_k}$  with optimal regularization parameter in Table 6.3.

TABLE 6.3  
*Estimated Hölder exponents  $\kappa \in (0, 1)$  for Hölder convergence rates  $\|x_{\alpha_{opt}}^{\delta} - x^{\dagger}\|_X \sim \delta^{\kappa}$  as  $\delta \rightarrow 0$ .*

Data	Penalty	Tikhonov reg.	Iterative reg.
$Y = L^2([0, 2]^2)$	$\mathcal{R}_1(x)$	0.6059	0.6699
	$\mathcal{R}_2(x)$	0.6320	0.6705
	$\mathcal{R}_3(x)$	0.5083	–
$Y = L^2([0, 1]^2)$	$\mathcal{R}_1(x)$	0.3753	0.4164
	$\mathcal{R}_2(x)$	0.4522	0.4505
	$\mathcal{R}_3(x)$	0.3787	–

Regarding the convergence rates of  $x_{\alpha_{opt}}^{\delta}$  converging to  $x^{\dagger}$  as  $\delta \rightarrow 0$ , for both penalties  $\mathcal{R}_1$  and  $\mathcal{R}_2$ , the iterative regularization approach proves to be advantageous.

We list in Table 6.4 the relative output errors of regularized solutions  $x_{\alpha_{opt}}^{\delta}$ ,  $x_{\alpha_{SDP}}^{\delta}$  and  $x_{\alpha_{qo}}^{\delta}$  for the fixed noise level  $\delta$  of relative input error 1%.

TABLE 6.4  
*Comparison of relative error norms of regularized solutions for a fixed relative input error 1%.*

Method	Data	Penalty	$\frac{\ x_{\alpha_{opt}}^{\delta} - x^{\dagger}\ _X}{\ x^{\dagger}\ _X}$	$\frac{\ x_{\alpha_{SDP}}^{\delta} - x^{\dagger}\ _X}{\ x^{\dagger}\ _X}$	$\frac{\ x_{\alpha_{qo}}^{\delta} - x^{\dagger}\ _X}{\ x^{\dagger}\ _X}$
Tikhonov reg.	$Y = L^2([0, 2]^2)$	$\mathcal{R}_1(x)$	3.22%	5.75%	3.47%
	$Y = L^2([0, 1]^2)$	$\mathcal{R}_1(x)$	6.43%	8.59%	7.23%
		$\mathcal{R}_2(x)$	4.68%	7.75%	5.83%
		$\mathcal{R}_3(x)$	6.01%	10.60%	6.01%
Iterative reg.	$Y = L^2([0, 1]^2)$	$\mathcal{R}_1(x)$	5.83%	9.06%	–
		$\mathcal{R}_2(x)$	4.64%	8.92%	–

Summarizing, we can state that it is always possible to solve the two-dimensional deautoconvolution problem in a rather stable way by either Tikhonov or iterative regularization. In this context, we also obtain a reasonable accuracy for an appropriate choice of the regularization parameter.

**Acknowledgment.** The authors are very grateful to two anonymous reviewers for their valuable suggestions, which improved the final version of the paper.

REFERENCES

- [1] S. W. ANZENGRUBER, S. BÜRGER, B. HOFMANN, AND G. STEINMEYER, *Variational regularization of complex deautoconvolution and phase retrieval in ultrashort laser pulse characterization*, Inverse Problems, 32 (2016), 035002 (27 pages).
- [2] S. W. ANZENGRUBER, B. HOFMANN, AND P. MATHÉ, *Regularization properties of the sequential discrepancy principle for Tikhonov regularization in Banach spaces*, Appl. Anal., 93 (2014), pp. 1382–1400.
- [3] A. BAKUSHINSKII, *Remarks on choosing a regularization parameter using the quasi-optimality and ratio criterion*, U.S.S.R. Comput. Math. and Math. Phys., 24 (1984), pp. 181–182.
- [4] F. BAUER AND S. KINDERMANN, *The quasi-optimality criterion for classical inverse problems*, Inverse Problems, 24 (2008), 035002 (20 pages).

- [5] J. BAUMEISTER, *Deconvolution of appearance potential spectra*, in Direct and Inverse Boundary Value Problems, R. Kleinman, R. Kress, and E. Martensen, eds., Peter Lang, Frankfurt am Main, 1991, pp. 1–13.
- [6] A. BECK AND M. TEBoulLE, *A fast dual proximal gradient algorithm for convex minimization and applications*, Oper. Res. Lett., 42 (2014), pp. 1–6.
- [7] M. BÜRGER, A. C. MENNUCCI, S. OSHER, AND M. RUMPF, *Level Set and PDE Based Reconstruction Methods in Imaging: Cetraro, Italy 2008*, Editors: Martin Burger, Stanley Osher, Springer, 2013.
- [8] S. BÜRGER AND J. FLEMMING, *Deautoconvolution: a new decomposition approach versus TIGRA and local regularization*, J. Inverse Ill-Posed Probl., 23 (2015), pp. 231–243.
- [9] S. BÜRGER, J. FLEMMING, AND B. HOFMANN, *On complex-valued deautoconvolution of compactly supported functions with sparse fourier representation*, Inverse Problems, 32 (2016), 104006 (12 pages).
- [10] S. BÜRGER AND B. HOFMANN, *About a deficit in low-order convergence rates on the example of autoconvolution*, Appl. Anal., 94 (2015), pp. 477–493.
- [11] S. BÜRGER AND P. MATHÉ, *Discretized Lavrent'ev regularization for the autoconvolution equation*, Appl. Anal., 96 (2017), pp. 1618–1637.
- [12] A. CHAMBOLLE AND T. POCK, *A first-order primal-dual algorithm for convex problems with applications to imaging*, J. Math. Imaging Vision, 40 (2011), pp. 120–145.
- [13] K. CHOI AND A. D. LANTERMAN, *An iterative deautoconvolution algorithm for nonnegative functions*, Inverse Problems, 21 (2005), pp. 981–995.
- [14] H. ENGL, M. HANKE, AND A. NEUBAUER, *Regularization of Inverse Problems*, Springer, Kluwer, Dordrecht, 1996.
- [15] L. FINESSO AND P. SPREIJ, *The inverse problem of positive autoconvolution*, IEEE Trans. Inform. Theory, 2023, to appear, (12 pages).
- [16] G. FLEISCHER, R. GORENFLO, AND B. HOFMANN, *On the autoconvolution equation and total variation constraints*, ZAMM Z. Angew. Math. Mech., 79 (1999), pp. 149–159.
- [17] G. FLEISCHER AND B. HOFMANN, *On inversion rates for the autoconvolution equation*, Inverse Problems, 12 (1996), pp. 419–435.
- [18] J. FLEMMING, *Variational Source Conditions, Quadratic Inverse Problems, Sparsity Promoting Regularization: New Results in Modern Theory of Inverse Problems and an Application in Laser Optics*, Springer, Cham, 2018.
- [19] D. GERTH, B. HOFMANN, S. BIRKHOLZ, S. KOKE, AND G. STEINMEYER, *Regularization of an autoconvolution problem in ultrashort laser pulse characterization*, Inverse Probl. Sci. Eng., 22 (2014), pp. 245–266.
- [20] R. GORENFLO AND B. HOFMANN, *On autoconvolution and regularization*, Inverse Problems, 10 (1994), pp. 353–373.
- [21] B. HOFMANN AND O. SCHERZER, *Local ill-posedness and source conditions of operator equations in Hilbert spaces*, Inverse Problems, 14 (1998), pp. 1189–1206.
- [22] T. HOHAGE AND F. WERNER, *Iteratively regularized Newton-type methods for general data misfit functionals and applications to Poisson data*, Numer. Math., 123 (2013), pp. 745–779.
- [23] J. JANNO, *Lavrent'ev regularization of ill-posed problems containing nonlinear near-to-monotone operators with application to autoconvolution equation*, Inverse Problems, 16 (2000), pp. 333–348.
- [24] B. KALTENBACHER, A. NEUBAUER, AND O. SCHERZER, *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*, de Gruyter, Berlin, 2008.
- [25] S. KINDERMANN AND A. NEUBAUER, *On the convergence of the quasioptimality criterion for (iterated) Tikhonov regularization*, Inverse Probl. Imaging, 2 (2008), pp. 291–299.
- [26] J. LIONS, *Supports dans la transformation de Laplace (in French)*, J. Analyze Math., 2 (1953), pp. 369–380.
- [27] J.-L. LIONS, *Supports de produits de composition. i (in French)*, C. R. Acad. Sci. Paris, 232 (1951), pp. 1530–1532.
- [28] O. SCHERZER, M. GRASMAIR, H. GROSSAUER, M. HALTMEIER, AND F. LENZEN, *Variational Methods in Imaging*, Springer, New York, 2009.
- [29] K.-T. SCHLEICHER, S. SCHULZ, R. GMEINER, AND H.-U. CHUN, *A computational method for the evaluation of highly resolved dos functions from aps measurements*, Journal of Electron Spectroscopy and Related Phenomena, 31 (1983), pp. 33–56.
- [30] T. SCHUSTER, B. KALTENBACHER, B. HOFMANN, AND K. S. KAZIMIERSKI, *Regularization Methods in Banach Spaces*, de Gruyter, Berlin, 2012.
- [31] E. C. TITCHMARSH, *The zeros of certain integral functions*, Proc. Lond. Math. Soc. (2), 25 (1926), pp. 283–302.
- [32] C. R. VOGEL, *Computational Methods for Inverse Problems*, SIAM, Philadelphia, 2002.
- [33] F. WERNER, *On convergence rates for iteratively regularized Newton-type methods under a Lipschitz-type nonlinearity condition*, J. Inverse Ill-Posed Probl., 23 (2015), pp. 75–84.