

# **Variational regularization of complex deautoconvolution and phase retrieval in ultrashort laser pulse characterization**

**S. Anzengruber, S. Bürger, B. Hofmann,  
G. Steinmeyer**

**RICAM-Report 2015-31**

**October 2015**

# VARIATIONAL REGULARIZATION OF COMPLEX DEAUTOCONVOLUTION AND PHASE RETRIEVAL IN ULTRASHORT LASER PULSE CHARACTERIZATION

STEPHAN W. ANZENGRUBER, STEVEN BÜRGER, BERND HOFMANN,  
AND GÜNTER STEINMEYER

**ABSTRACT.** The SD-SPIDER method for the characterization of ultrashort laser pulses requires the solution of a nonlinear integral equation of autoconvolution type with a device-based kernel function. Taking into account the analytical background of a variational regularization approach for solving the corresponding ill-posed operator equation formulated in complex-valued  $L^2$ -spaces over finite real intervals, we suggest and evaluate numerical procedures using NURBS and the TIGRA method for calculating the regularized solutions in a stable manner. In this context, besides the complex deautoconvolution problem with noisy but full data, a phase retrieval problem is introduced which adapts to the experimental state of the art in laser optics. For the treatment of this problem facet, which is formulated as a tensor product operator equation, we derive well-posedness of variational regularization methods. Case studies with synthetic and real optical data show the capability of the implemented approach as well as its limitation due to measurement deficits.

**MSC2010 subject classification:** 47A52, 47J06, 78A60, 65R32, 45Q05, 65J15

**Keywords:** Deautoconvolution, phase retrieval, regularization, laser pulse characterization, ill-posed problem, SD-SPIDER, NURBS, TIGRA.

## 1. INTRODUCTION

About two decades ago and motivated by a problem from spectroscopy, namely the evaluation of highly resolved functions of the density of unoccupied states from appearance-potential spectra (cf, e.g., [29]), the inverse problem of deautoconvolution came into the focus of the mathematical literature for the first time (cf. [2, 16]). In particular, the stable approximate determination of real functions  $f$  over  $\mathbb{R}$  with compact support, say  $\text{supp}(f) \subseteq [0, 1]$ , from noisy data of its self-convolution  $g$  with  $\text{supp}(g) \subseteq [0, 2]$  became of interest, which is equivalent to the solution of the Volterra type nonlinear integral equation

$$(1) \quad \int_0^s f(s-\tau) f(\tau) d\tau = g(s), \quad 0 \leq s \leq 2,$$

---

*Date:* August 4, 2015.

occurring also in stochastics if the probability density function  $f(\tau)$  with support in  $[0, 1]$  of a random variable  $\mathfrak{X}$  is to be determined from data of the density function  $g(s)$ . In this context,  $g(s)$  corresponds to the random variable  $\mathfrak{Y} + \mathfrak{Z}$ , where  $\mathfrak{Y}$  and  $\mathfrak{Z}$  are stochastically independent random variables and  $\mathfrak{X}, \mathfrak{Y}, \mathfrak{Z}$  are identically distributed. It was shown in [16] resp. [11] that (1) written as an operator equation with the nonlinear forward operator of autoconvolution between the real spaces  $L^2(0, 1)$  and  $L^2(0, 1)$  resp.  $L^2(0, 2)$  of quadratically integrable functions is locally ill-posed everywhere in the sense of [19, Definition 2]. Furthermore, regularization approaches were developed and evaluated, where the astonishing fact appeared that in spite of the simple quadratic structure of the forward operator convergence rates results are difficult to obtain. Indeed, autoconvolution ‘scrambles’ the input function in such a way that qualified nonlinearity conditions like the tangential cone condition and classical source conditions are generally not satisfied. For details of the corresponding results we refer to [11, 16, 20, 21] and the recent paper [7]. Moreover, we refer to [8, 12] for alternative approaches to deautoconvolution.

A completely new approach to the autoconvolution problem from a mathematical point of view was started in 2011 at the TU Chemnitz in collaboration with the research group ‘Solid State Light Sources’ of the Max Born Institute for Non-linear Optics and Short Pulse Spectroscopy, Berlin. The new onset was motivated by the development of the *SD-SPIDER* (self-diffraction spectral phase interferometry for direct electric-field reconstruction) technique in ultrashort (femtosecond) laser pulse characterization at the Max Born institute. A mathematical model for this method can be formulated by generalizing (1) as a kernel-based autoconvolution equation

$$(2) \quad \int_0^s k(s, \tau) f(s - \tau) f(\tau) d\tau = g(s), \quad 0 \leq s \leq 2,$$

with complex-valued functions  $f$  over  $\mathbb{R}$  with  $\text{supp}(f) \subseteq [0, 1]$  and  $g$  over  $\mathbb{R}$  with  $\text{supp}(g) \subseteq [0, 2]$ , and a complex-valued kernel  $k(s, \tau)$  with  $(s, \tau) \in \mathbb{R}^2$  and  $\text{supp}(k) \subseteq [0, 2] \times [0, 1]$ . In the sequel we use the polar coordinate representations

$$(3) \quad f(\tau) = a(\tau) e^{i\varphi(\tau)} \quad \text{and} \quad g(s) = b(s) e^{i\psi(s)}$$

for the searched-for function  $f$  and the right-hand side function  $g$ , respectively. For a detailed explanation of the physical background, i.e., of the SD-SPIDER approach in light of (2) and concerning the availability of optical measurements for the amplitude functions  $a, b$  and the phase functions  $\varphi, \psi$  in (3), we refer to the subsequent Section 2.

To be precise, we consider (2) as a nonlinear operator equation in a Hilbert space setting. Our focus is on the spaces

$$(4) \quad X := L_{\mathbb{C}}^2(0, 1) \quad \text{and} \quad Y := L_{\mathbb{C}}^2(0, 2)$$

of complex-valued square integrable functions with associated norms  $\|\cdot\|_X, \|\cdot\|_Y$  and inner products  $\langle \cdot, \cdot \rangle_X, \langle \cdot, \cdot \rangle_Y$ , respectively. Then we write (2) in the concise form

$$(5) \quad F(f) = g, \quad f \in X, \quad g \in Y,$$

where  $f^\dagger \in X$  denotes exact solutions to (5) for given exact right-hand side  $g = g^\dagger \in Y$ , and where the forward operator  $F : X \rightarrow Y$ , taking into account the specific support intervals of pre-image and image functions, is defined as

$$(6) \quad \begin{aligned} [F(f)](s) &= [B(f)](s) e^{i[\Psi(f)](s)} \\ &:= \int_{\max(s-1,0)}^{\min(s,1)} k(s, \tau) f(s - \tau) f(\tau) d\tau, \quad 0 \leq s \leq 2, \end{aligned}$$

for preimage function  $f(\tau) = a(\tau) e^{i\varphi(\tau)}$ ,  $0 \leq \tau \leq 1$ . From (6) it is seen that the support of the kernel function  $k$  is contained in the parallelogram

$$(7) \quad \mathfrak{P} := \{(s, \tau) \in [0, 2] \times [0, 1] : 0 \leq \tau \leq 1, \tau \leq s \leq \tau + 1\}.$$

According to physical models (cf. [4, 14, 15]) the kernel functions occurring in laser pulse characterization are continuous complex-valued functions on  $\mathfrak{P}$ . We will thus restrict our considerations to kernel functions, such that

$$(8) \quad \text{supp}(k) \subset \mathfrak{P}, \quad k \in C(\mathfrak{P}), \quad \text{and} \quad k(s, \tau) = k(s, s - \tau) \text{ for } (s, \tau) \in \mathfrak{P}.$$

The last assumption in (8) indeed holds without loss of generality as we have, for arbitrary kernel functions  $k(s, \tau)$  and all  $s \in [0, 2]$ , the identity

$$\int_{\max(s-1,0)}^{\min(s,1)} k(s, \tau) f(s - \tau) f(\tau) d\tau = \int_{\max(s-1,0)}^{\min(s,1)} \frac{k(s, \tau) + k(s, s - \tau)}{2} f(s - \tau) f(\tau) d\tau.$$

We mention that the structure (8) is in particular satisfied if the kernel is generated by a complex-valued function  $\kappa$  with  $\text{supp}(\kappa) \subset [0, 1]$  and  $\kappa \in C[0, 1]$  such that

$$(9) \quad k(s, \tau) = \kappa(\tau) \kappa(s - \tau) \quad \text{if } (\tau, s) \in \mathfrak{P}, \quad \text{and} \quad k(s, \tau) = 0 \quad \text{otherwise.}$$

For some statements in Section 3, we will be forced to focus on the special case (9).

In this paper, we distinguish two different inverse problems in the context of the forward operator  $F$  from (6) under the assumption (8) for the kernel  $k$ , which are briefly presented in the following.

**1.1. The deautoconvolution problem.** We call the inverse problem of identifying the complex-valued function  $f$  solving the operator equation (5) deautoconvolution problem. In this context, it is assumed that only noisy data  $g^\delta$  of the complex-valued function  $g^\dagger$  with

$$(10) \quad \|g^\dagger - g^\delta\|_Y \leq \delta$$

are available for a reasonably small noise level  $\delta > 0$ .

This deautoconvolution problem was tackled in [15] in a direct manner, where theoretic consideration were only made for the trivial kernel  $k \equiv 1$ . For the stable approximate solution including a nontrivial kernel, an iterative regularization approach based on a variant of the Levenberg-Marquardt method was used there, which showed quite good results in numerical case studies with synthetic data. In

this paper, we extend assertions on properties of the forward operator  $F$  and on the solution of the operator equation (5) to kernels  $k$  from (8) or (9).

The real data situation in laser optics, see for details Section 2 below, has an advantage and a disadvantage. Fortunately, noisy measurements for the amplitude function  $a$  in the solution  $f$  (see (3)) can be provided. These data were exploited in [15] to control a regularization parameter. However, measurements for the amplitude function  $b$  in  $g$ , which were required in combination with data for the phase  $\psi$ , are unfortunately not sufficiently reliable in practice. As a result, the method in [15] failed for real data from optical experiments. In order to resolve this shortcoming, we clarify the objective and the data situation in the following manner.

**1.2. The phase retrieval problem.** We call the inverse problem of identifying the phase function  $\varphi$  in the solution  $f$  (see (3)) of equation (5) from noisy data of the phase function  $\psi$  in the right-hand side  $g$  phase retrieval problem when noisy data of the amplitude function  $a$  in  $f$ , but no data of the amplitude function  $b$  in  $g$  are available.

A first trial of a more adapted regularization approach with focus on the phase retrieval problem was included in the recent paper [4] taking into account the real data situation. However, the well-posedness of the non-standard variational regularization methods was not considered there. We are going to close this gap in the present paper, in addition to an essential improvement of the numerical implementation based on *NURBS* (Non-Uniform Rational B-splines; see, e.g., [24]) in combination with a *TIGRA*-type algorithm (the name is derived from Tikhonov-GRAdient method; cf. [25, 26, 32]).

The paper is organized as follows. After an explanation of the background from laser optics in Section 2 we summarize properties of the forward operator  $F$  as well as of the deautoconvolution problem solving equation (5) in Section 3. Well-posedness results for the phase retrieval problem are presented in Section 4. Turning to numerical considerations, we then briefly introduce planar *NURBS* and describe how they may be utilized in the context of the complex-valued deautoconvolution problems in Section 5. Finally, in Section 6 we show numerical results obtained with a *TIGRA*-type method for both synthetic and real data.

## 2. PHYSICAL BACKGROUND

The motivation for the complex-valued and kernel-based deconvolution problem of solving (5) in the introduction is the SD-SPIDER (Self-Diffraction Spectral Phase Interferometry for Direct Electric-field Reconstruction) method in laser optics. The aim of this method is the reconstruction of the electric field  $E(t)$  of ultrashort (femtosecond) laser pulses, which is a real-valued oscillatory function of time  $t$ . This function is usually decomposed into an amplitude part and an oscillating part by

$$(11) \quad E(t) = \sqrt{I(t)} \cos\{\omega_0 t + \eta(t)\},$$

where  $I(t)$  is the intensity (up to re-normalization),  $\omega_0$  the carrier frequency, and  $\eta(t)$  is called temporal phase. Since measurements are available for the spectral domain only, we have to consider the Fourier transform of  $E(t)$

$$(12) \quad \hat{E}(\omega) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} E(t) e^{-i\omega t} dt$$

as a complex-valued function of the frequency  $\omega$ . The Fourier transform can be written in polar coordinates as

$$(13) \quad \hat{E}(\omega) = \sqrt{S(\omega)} e^{i\phi(\omega)},$$

where  $S(\omega)$  is the spectral power density (up to re-normalization), and  $\phi(\omega)$  is called spectral phase. Fortunately, one is able to measure the spectrum directly, but contaminated with noise. Hence, the approximate identification of the desired physical quantity  $E(t)$  by inverse Fourier transform requires first the determination of  $\phi$ . Careful denoising, e.g., by Fourier filtering or adjacent averaging of the measurements for  $S$  may prove helpful to reduce a possible oscillatory behavior of the algorithms. Since the spectral phase  $\phi$  cannot be measured directly, nonlinear optical processes have to be employed to infer the spectral phase from an indirect measurement. To this end, one suitable process is self-diffraction (SD), which is a spectrally degenerate variant of the more general four-wave mixing process. The electric field of the generated pulses in the spectral domain  $\hat{E}_{\text{SD}}(\omega)$  is related to  $\hat{E}(\omega)$  through

$$(14) \quad \hat{E}_{\text{SD}}(\omega) = \int_0^{\omega+\omega_{\text{cw}}} K(\omega, \Omega) \hat{E}(\omega + \omega_{\text{cw}} - \Omega) \hat{E}(\Omega) d\Omega.$$

The SD process involves interaction with a continuous wave at a known frequency  $\omega_{\text{cw}}$ , which has been incorporated into the kernel function  $K(\omega, \Omega)$ . Moreover, the kernel function  $K(\omega, \Omega)$  as an apparatus function is in principle known from physical modeling and can be assumed to be smooth in both arguments  $\omega$  and  $\Omega$ . For a detailed description of the kernel see [14]. We also write  $\hat{E}_{\text{SD}}$  in polar coordinates as

$$(15) \quad \hat{E}_{\text{SD}}(\omega) = \sqrt{S_{\text{SD}}(\omega)} e^{i\phi_{\text{SD}}(\omega)}.$$

Perfect knowledge of all efficiency calibration factors in  $K$  provided, one can directly use the spectral power density  $S_{\text{SD}}$  for reconstruction of  $E(t)$ . Even a relatively small miscalibration, however, may have a dramatic influence on  $S_{\text{SD}}$ , which is why algorithms relying on the spectral power density may have trouble to produce reasonable results [14]. The phase  $\phi_{\text{SD}}$ , on the other hand, remains widely unaffected from such amplitude calibration problems. The measurement setup for the SD-interferogram is shown in Figure 1. For physical details we refer to [14, 15, 23].

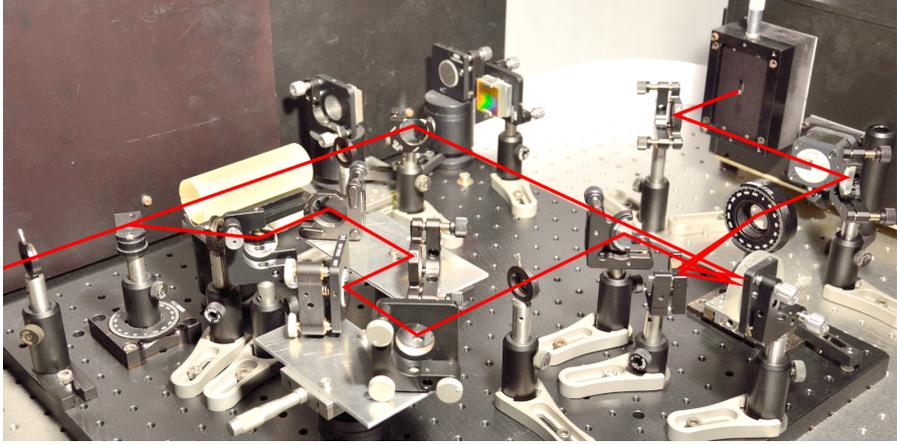


FIGURE 1. Measurement setup in self-diffraction spectral interferometry.

The measurements for the spectrum of  $\hat{E}$  indicate that the corresponding function can be neglected outside a compact interval  $[\omega_{\text{low}}, \omega_{\text{up}}]$ , i.e. we assume that

$$(16) \quad \text{supp } \hat{E} \subset [\omega_{\text{low}}, \omega_{\text{up}}].$$

Consequently,  $\hat{E}_{\text{SD}}$  can also be neglected outside the compact interval

$$(17) \quad \text{supp } \hat{E}_{\text{SD}} \subset [2\omega_{\text{low}} - \omega_{\text{cw}}, 2\omega_{\text{up}} - \omega_{\text{cw}}].$$

Thus, (14) can be written as

$$(18) \quad \hat{E}_{\text{SD}}(\omega) = \int_{\max\{\omega_{\text{low}}, \omega + \omega_{\text{cw}} - \omega_{\text{up}}\}}^{\min\{\omega_{\text{up}}, \omega + \omega_{\text{cw}} - \omega_{\text{low}}\}} K(\omega, \Omega) \hat{E}(\omega + \omega_{\text{cw}} - \Omega) \hat{E}(\Omega) d\Omega.$$

The substitutions

$$\begin{aligned} g(s) &:= \hat{E}_{\text{SD}}(-\omega_{\text{cw}} + 2\omega_{\text{low}} + s(\omega_{\text{up}} - \omega_{\text{low}})), \\ f(\tau) &:= \hat{E}(\omega_{\text{low}} + \tau(\omega_{\text{up}} - \omega_{\text{low}})), \\ k(s, \tau) &:= K(-\omega_{\text{cw}} + 2\omega_{\text{low}} + s(\omega_{\text{up}} - \omega_{\text{low}}), \omega_{\text{low}} + \tau(\omega_{\text{up}} - \omega_{\text{low}})) \end{aligned}$$

reformulate (18) as

$$(19) \quad \int_{\max\{0, s-1\}}^{\min\{1, s\}} k(s, \tau) f(s - \tau) f(\tau) ds = g(s), \quad 0 \leq s \leq 2,$$

which is precisely the mathematical model presented in the introduction. The physical background justifies the continuity of the kernel  $k$  in  $\mathfrak{P}$  as required in (8).

### 3. SOME PROPERTIES OF COMPLEX AUTOCONVOLUTION OPERATOR AND DEAUTOCONVOLUTION PROBLEM

In this section, we summarize properties of the complex autoconvolution operator and discuss the classical Tikhonov regularization for the deautoconvolution

problem aimed at solving the operator equation (5) based on noisy data  $g^\delta$  under the noise model (10).

The Fréchet differentiability of the autoconvolution operator mapping in the real Hilbert space  $L^2(0, 1)$  and the structure of its Fréchet derivative were outlined in [16]. Under the kernel assumption (8) imposed on the kernel  $k$ , such result can also be formulated for the complex case (6) with the Fréchet derivative  $F'(f) : X = L^2_{\mathbb{C}}(0, 1) \rightarrow Y = L^2_{\mathbb{C}}(0, 2)$  given for all  $f \in X$  by the formula

$$(20) \quad [F'(f)h](s) = 2 \int_{\max(s-1,0)}^{\min(s,1)} k(s, \tau) f(s - \tau) h(\tau) d\tau, \quad 0 \leq s \leq 2, \quad h \in X,$$

from which we easily derive

$$(21) \quad [F(f+h) - F(f) - F'(f)h](s) = [F(h)](s), \quad 0 \leq s \leq 2, \quad h \in X.$$

Hence, we obtain with  $\mathfrak{P}$  from (7) and using

$$(22) \quad \bar{k} := \max_{(s,\tau) \in \mathfrak{P}} |k(s, \tau)|,$$

the norm equation

$$\|F(f+h) - F(f) - F'(f)h\|_Y = \|F(h)\|_Y \quad \text{for all } f, h \in X$$

and the nonlinearity condition

$$(23) \quad \|F(f+h) - F(f) - F'(f)h\|_Y \leq \bar{k} \|h\|_X^2 \quad \text{for all } f, h \in X.$$

Such condition, which was already used in [9, Section 10.2] to obtain convergence rates for Tikhonov regularization of nonlinear operator equations, is by now the only available nonlinearity condition for the operator  $F$  from (6).

Note that the adjoint operator  $F'(f)^* : Y \rightarrow X$  can also be written down explicitly, namely as

$$(24) \quad [F'(f)^*r](\tau) = 2 \int_{\tau}^{\tau+1} k(s, \tau) f(s - \tau) r(s) ds, \quad 0 \leq \tau \leq 1, \quad f \in X, \quad r \in Y.$$

As usual we apply in the sequel the symbol  $\mathcal{B}(f, r)$  for the closed ball in  $X$  with center  $f$  and radius  $r > 0$  and the symbols  $\rightharpoonup$  and  $\rightarrow$  for weak and norm convergence, respectively, in the occurring Hilbert spaces. The proofs of the following two results have been postponed to the Appendix.

**Proposition 1.** *Suppose that the kernel  $k$  satisfies (8). Then the autoconvolution operator  $F : X \rightarrow Y$  from (6) is weakly continuous, i.e., for every sequence  $\{f_n\}_{n \in \mathbb{N}} \in X$  with  $f_n \rightharpoonup f_0$  in  $X$  as  $n \rightarrow \infty$  we have  $F(f_n) \rightharpoonup F(f_0)$  in  $Y$ .*

The following proposition outlines a very specific property of the autoconvolution operator, namely that the nonlinear  $F : L^2_{\mathbb{C}}(0, 1) \rightarrow L^2_{\mathbb{C}}(0, 2)$  from (6) is non-compact, whereas the linear Fréchet derivative operator  $F'(f)$  is compact in all points  $f \in L^2_{\mathbb{C}}(0, 1)$ . This is a remarkable property of the autoconvolution operator, which rarely occurs for nonlinear operators. Conversely, it is well-known

that the Fréchet derivative of a compact nonlinear operator is always compact. Again the proof can be found in the Appendix.

**Proposition 2.** *Suppose that the kernel  $k \not\equiv 0$  satisfies (8). Then the kernel-based nonlinear autoconvolution operator  $F : X \rightarrow Y$  from (6) is not compact. More precisely, there exists a sequence  $\{h_n\}_{n \in \mathbb{N}} \subset \mathcal{B}(0, 1)$  such that  $h_n \rightarrow 0$  but  $h_n \not\rightarrow 0$  in  $X$  as  $n \rightarrow \infty$ , which implies that we have, for all  $f \in X$  and all  $r > 0$ , a sequence  $\{f_n := f + r h_n\}_{n \in \mathbb{N}} \subset \mathcal{B}(f, r)$  with  $f_n \rightarrow f$  in  $X$ ,  $F(f_n) \rightarrow F(f)$  in  $Y$  but  $F(f_n) \not\rightarrow F(f)$  in  $Y$  as  $n \rightarrow \infty$ . On the other hand, the Fréchet derivative  $F'(f)$  given by (20) is a compact linear operator for all  $f \in X$ .*

**Remark 3.** The proof of Proposition 2 is based on finding sequences  $f_n \rightarrow f$  in  $X$  such that  $F(f_n) \not\rightarrow F(f)$  in  $Y$  as  $n \rightarrow \infty$ . A conversely related concept is the local ill-posedness of the operator equation (5) at  $f \in X$ , which requires sequences  $f_n$  in any neighborhood  $\mathcal{B}(f, r)$  of  $f$  such that  $f_n \not\rightarrow f$ , but  $F(f_n) \rightarrow F(f)$ . For the specific case  $k \equiv 1$  on  $\mathfrak{B}$ , an example of the form

$$f_n = f + r h_n, \quad \text{with } h_n(\tau) = e^{in^2 \tau^2},$$

was provided in [7, Example 3.2], showing that the operator equation (5) is locally ill-posed everywhere in this situation. It is worth noting that, while the freedom in choosing  $r > 0$  is exploited here for local arguments with arbitrarily small  $r$ , it may also be used for construction of elements  $f_n$  at arbitrarily large distances,  $\|f - f_n\|_X = r$ , whose images  $F(f)$  and  $F(f_n)$  are virtually indistinguishable, say  $\|F(f) - F(f_n)\|_Y \leq \frac{1}{n}$ . This means that highly oscillating perturbations imposed on  $f$  cannot be recovered by a simple least-squares approach from noisy data  $g^\delta$  of  $g = F(f)$  with noise model (10), even if the noise level  $\delta > 0$  is arbitrarily small. Since the same phenomenon has to be expected for more general kernels  $k$ , a regularization method seems to be always required in order to avoid oscillating numerical approximations.

These observations are particularly interesting in combination with the uniqueness assertion in Proposition 5 below which is based on the following well-known Titchmarsh convolution theorem; cf. [31].

**Lemma 4.** *Let  $\xi_1, \xi_2 \in L^1_{\mathbb{C}}(\mathbb{R})$  with  $\text{supp}(\xi_l) \subset [0, \infty)$ ,  $l = 1, 2$ , and let for some constant  $a > 0$*

$$\int_0^s \xi_1(s - \tau) \xi_2(\tau) d\tau = 0 \quad \text{for almost all } s \in [0, a].$$

*Then there are nonnegative constants  $a_1$  and  $a_2$  such that  $a_1 + a_2 \geq a$  and*

$$\xi_1(\tau) = 0 \quad \text{for almost all } \tau \in [0, a_1] \quad \text{and} \quad \xi_2(\tau) = 0 \quad \text{for almost all } \tau \in [0, a_2].$$

**Proposition 5.** *Let the kernel  $k$  from (8) be generated by a function  $\kappa \in C[0, 1]$  with  $\text{supp}(\kappa) = [0, 1]$  such that (9) is satisfied. If for given  $g \in Y = L^2_{\mathbb{C}}(0, 2)$  the function  $f^\dagger \in X = L^2_{\mathbb{C}}(0, 1)$  solves (5) then  $f^\dagger$  and  $-f^\dagger$  are the only solutions of this equation for the right-hand side  $g$ .*

*Proof.* Based on Lemma 4, the assertion of this proposition can be shown in analogy to the proof of Theorem 4.2 in [15], taking into account that under the stated assumptions on  $\kappa$  the equality  $\kappa(\tau) f^\dagger(\tau) = \kappa(\tau) \tilde{f}(\tau)$  with  $\tilde{f} \in X$  and for almost all  $\tau \in [0, 1]$  implies that  $f^\dagger$  and  $\tilde{f}$  are the same elements of  $L^2_{\mathbb{C}}(0, 1)$  when taking into account that

$$\text{supp}(\kappa) = [0, 1] \quad \text{holds if and only if} \quad \kappa \neq 0 \text{ almost everywhere on } [0, 1].$$

□

Note that Lemma 4 under the assumption (8) also provides us with a necessary and sufficient condition for the injectivity of the Fréchet derivative; see (20). Namely, the bounded linear operator  $F'(f) : X \rightarrow Y$  is non-injective if either  $f$  is the zero function almost everywhere on  $[0, 1]$  or if  $\kappa \equiv 0$ . Otherwise the Fréchet derivative is injective.

Variational (cf. [28]) and iterative (cf. [22]) regularization methods form two classes of standard methods for stabilizing ill-posed nonlinear operator equations (5) in Hilbert spaces. The most prominent representative for the first class is the Tikhonov regularization, where in the simplest form (cf., e.g., [9, Chapt.10]) regularized solutions  $f_\alpha^\delta$  are minimizers of the extremal problem

$$(25) \quad \|F(f) - g^\delta\|_Y^2 + \alpha \|f - \bar{f}\|_X^2 \rightarrow \min, \quad \text{subject to } f \in X,$$

with some initial guess  $\bar{f} \in X$ . For obtaining convergence rates of the regularized solutions, an appropriate interplay of solution smoothness and structural conditions expressing the nonlinearity of  $F$  in a neighborhood of the solution is required (see, e.g., [17] for an overview). For the Tikhonov regularization of the form (25) and  $F$  from (6), the condition (23) acts as nonlinearity condition sufficiently well and allows proving the convergence rate

$$\|f_{\alpha(\delta)}^\delta - f^\dagger\|_X = \mathcal{O}(\sqrt{\delta}) \quad \text{as } \delta \rightarrow 0 \quad \text{and} \quad \alpha(\delta) \sim \delta,$$

since the operator  $F$  is weakly continuous (cf. Proposition 1), which implies that  $F$  is weakly closed. The latter result additionally requires the existence of a source element  $w \in Y$  satisfying the smallness condition  $\|w\|_Y \leq 1$  such that the source condition

$$(26) \quad f^\dagger(\tau) - \bar{f}(\tau) = \int_{\tau}^{\tau+1} k(s, \tau) f^\dagger(s - \tau) w(s) ds, \quad 0 \leq \tau \leq 1,$$

is fulfilled (cf. [9, Theorem 10.4]). However, in [7, Proposition 2.6] it was shown that such source condition (26) is hardly possible to achieve even in the simplest cases of a kernel  $k \equiv 1$ .

For wide classes of iterative regularization methods, however, the tangential cone condition

$$(27) \quad \|F(f+h) - F(f) - F'(f)h\|_Y \leq \bar{c} \|F(f+h) - F(f)\|_Y \quad \text{for all } h \in \mathcal{B}(f, r)$$

is required, with a constant  $0 \leq \bar{c} < 1$ , some radius  $r > 0$ , and at least for  $f \in X$  in a neighborhood of the solution  $f^\dagger$  to (5) (see, e.g., [9, Chapt.11]). Slightly modified variants of the nonlinearity condition (27) for constants  $0 \leq \bar{c} < \infty$  and terms  $\theta(\|F(f+h) - F(f)\|_Y)$  instead of  $\|F(f+h) - F(f)\|_Y$ , with concave strictly increasing functions  $\theta : (0, \infty) \rightarrow (0, \infty)$  and  $\lim_{t \rightarrow +0} \theta(t) = 0$ , are also relevant for obtaining convergence rates in Tikhonov regularization, in particular if the solution smoothness is low and approximate source conditions apply (see [30, Section 4.2]).

**Proposition 6.** *For the autoconvolution operator  $F : X \rightarrow Y$  from (6), the tangential cone condition (27) cannot hold for any  $f \in X$  with a constant  $0 \leq \bar{c} < 1$  and a radius  $r > 0$ .*

*Proof.* The assertion is, in principle, a consequence of the noncompactness of  $F$  in contrast to the compactness of  $F'(f)$  for all  $f \in X$ . Using the triangle inequality

$$\|F(f+h) - F(f)\|_Y \leq \|F(f+h) - F(f) - F'(f)h\|_Y + \|F'(f)h\|_Y,$$

we would have from (27), for fixed  $f \in X$  and  $r > 0$ , the inequality

$$(1 - \bar{c})\|F(f+h) - F(f)\|_Y \leq \|F'(f)h\|_Y \quad \text{for all } h \in \mathcal{B}(f, r).$$

By substituting  $h := h_n$  in that inequality with  $\{h_n\}_{n \in \mathbb{N}} \subset \mathcal{B}(f, r)$  from Proposition 2 satisfying  $h_n \rightarrow 0$ ,  $h_n \not\rightarrow 0$  in  $X$  and  $F(f+h_n) \not\rightarrow F(f)$  in  $Y$  as  $n \rightarrow \infty$ , we arrive at a contradiction, because  $F'(f)$  is compact and hence satisfies the limit condition  $\lim_{n \rightarrow \infty} \|F'(f)h_n\|_Y = 0$ . This contradiction proves the proposition.  $\square$

**Remark 7.** As originally discussed in [19], the four norm terms associated with the Taylor remainder equation,

$$F(f+h) - F(f) = F'(f)h + \text{Rem}(f, h),$$

namely  $\|F(f+h) - F(f)\|_Y$ ,  $\|\text{Rem}(f, h)\|_Y = \|F(f+h) - F(f) - F'(f)h\|_Y$ ,  $\|F'(f)h\|_Y$ , and  $\|h\|_X$  show distinguished cross connections depending on whether the corresponding nonlinear operator equation (5) is well-posed or ill-posed. The convergence of well-posed problems usually results from the fact that the remainder  $\|\text{Rem}(f, h)\|_Y$  converges to zero faster than the term  $\|F(f+h) - F(f)\|_Y$  as  $\|h\|_X \rightarrow 0$ . In the ill-posed situation, however, where  $F$  is a ‘smoothing’ operator, the term  $\|F(f+h) - F(f)\|_Y$  may be significantly smaller than  $\|\text{Rem}(f, h)\|_Y$  even for arbitrarily small  $\|h\|_X$ . In the latter case, there exist operators  $F$  for which the tangential cone condition (27) may fail to hold even for large constants  $\bar{c} \geq 1$  as well as for modifications with terms  $\theta(\|F(f+h) - F(f)\|_Y)$  rather than  $\|F(f+h) - F(f)\|_Y$ . Then only Lipschitz continuity conditions for  $F'(f)$  such as (23) can be seen as nonlinearity conditions. In this particular situation, the linear operator  $F'(f)$  does not cover sufficient information about the nonlinear operator in a neighborhood of  $\mathcal{B}(f, r)$  for ensuring convergence rates of regularized solutions. This information deficit seems to be the case for the autoconvolution operator  $F$  from (6) and, unfortunately, prevents the establishment of convergence rates in Tikhonov regularization when the standard source condition (26) fails.

## 4. A REGULARIZATION APPROACH FOR THE PHASE RETRIEVAL PROBLEM

In this section, we collect some basic well-posedness results concerning the regularization approach by means of minimizing Tikhonov-type variational functionals for the different problems summarized in Subsections 1.1 and 1.2.

The theory of Tikhonov-regularization is well understood for ill-posed problems of the general form

$$(28) \quad \mathcal{F}(x) = y,$$

where  $x$  is a searched-for quantity of interest and  $y$  is approximately known from measurements in the form of data  $y^\delta$ . We refer to the monographs [9, 28, 30] as well as to the seminal works [10, 18] for detailed regularization results. For our purposes, let us recall the main sufficient condition on  $\mathcal{F} : \text{dom}(\mathcal{F}) \subset \mathcal{X} \rightarrow \mathcal{Y}$  [28, Section 3.2] for the well-posedness of minimizing a Tikhonov-type functional

$$\mathcal{T}_\alpha^\delta(x) = \frac{1}{2} \|\mathcal{F}(x) - y^\delta\|_Y^2 + \alpha \mathcal{R}(x)$$

in Hilbert spaces  $\mathcal{X}, \mathcal{Y}$ : namely, that the forward operator  $\mathcal{F}$  be weakly sequentially closed in the sense that

$$x_n \rightharpoonup \bar{x} \in \mathcal{X} \text{ and } \mathcal{F}(x_n) \rightharpoonup \bar{y} \in \mathcal{Y} \quad \Rightarrow \quad \bar{x} \in \text{dom}(\mathcal{F}) \text{ and } \mathcal{F}(\bar{x}) = \bar{y}$$

hold for all sequences  $\{x_n\} \subset \text{dom}(\mathcal{F})$ .

**Remark 8.** Note that we do not specify further the choice of the penalty term  $\mathcal{R}(x)$ . For the following results to hold true,  $\mathcal{R}(x)$  is required to be proper, weakly sequentially lower semicontinuous and to have weakly sequentially precompact sublevelsets. We will summarize these properties by saying that  $\mathcal{R}(x)$  is assumed to be *stabilizing*. The interested reader will easily verify that for stabilizing penalty terms  $\mathcal{R}(x)$  in combination with weakly sequentially closed forward operator  $\mathcal{F}$  with  $\text{dom}(\mathcal{F}) = \mathcal{X}$ , Assumption 3.13 in [28] is satisfied (with the exception of convexity of  $\mathcal{R}$  which is, however, not required here). Consequently, the well-posedness results in [28, Section 3.2] hold true. In particular, we have:

- Existence of minimizers (cf. Theorem 3.22);
- Weak, subsequential stability of the minimizers (cf. Theorem 3.23);
- Weak, subsequential convergence of the minimizers to an  $\mathcal{R}$ -minimizing solution as  $\delta \rightarrow 0$  under suitable parameter choice rules (cf. Theorem 3.26).

The deautoconvolution problem of Subsection 1.1 is naturally of the form (28) with  $\mathcal{F} = F$  from (6), and due to the properties in Section 3 we immediately obtain the following regularization result.

**Proposition 9.** *Suppose that the kernel  $k$  satisfies (8) and that the penalty term  $\mathcal{R}(f)$  is stabilizing in  $X = L^2_{\mathbb{C}}(0, 1)$ . Then for any  $g^\delta \in Y = L^2_{\mathbb{C}}(0, 2)$ , minimizing the Tikhonov functional*

$$\mathcal{T}_\alpha^\delta(f) = \frac{1}{2} \|F(f) - g^\delta\|_Y^2 + \alpha \mathcal{R}(f)$$

for  $f \in X$  is well-posed in the sense of Remark 8.

*Proof.* The weak sequential closedness of  $F : X \rightarrow Y$  follows readily from the weak sequential continuity in Proposition 1. The result then follows arguing as in Remark 8.  $\square$

In the particular situation of phase retrieval problems as in Subsection 1.2, the data consists of two separate parts. On the one hand  $a^\sigma$  contains information about the absolute value of the unknown function  $f$ , and, on the other hand,  $\psi^\delta \approx \arg(y^\dagger)$  is related to the image space  $Y$ . For formulating this problem in the context of Tikhonov regularization, we first define the Sign operator on  $L^2_{\mathbb{C}}(0, 2)$  pointwise almost everywhere (a.e.) by

$$[\text{Sign}(g)](s) = \begin{cases} \frac{g(s)}{|g(s)|} & \text{if } g(s) \neq 0 \\ 0 & \text{else.} \end{cases}$$

Then the forward operator in the sense of (28) also consists of two parts and maps as

$$\begin{aligned} \mathcal{F}_{\text{PR}} : X = L^2_{\mathbb{C}}(0, 1) &\rightarrow \mathcal{Y} = L^2_{\mathbb{C}}(0, 2) \times L^2(0, 1) \\ \mathcal{F}_{\text{PR}}(f) &= (\text{Sign}(F(f)), |f|), \end{aligned}$$

where the function  $|f| \in L^2(0, 1)$  is again defined pointwise a.e. by  $|f|(\tau) = |f(\tau)|$  for  $f \in X$ . Note that the product space  $\mathcal{Y} = Y \times L^2(0, 1)$  is a Hilbert space endowed with any of the equivalent norms  $\|(y, a)\|_{\mathcal{Y}, \beta}^2 := \|y\|_Y^2 + \beta \|a\|_X^2$  corresponding to discrepancy terms

$$\|\mathcal{F}_{\text{PR}}(f) - (e^{i\psi^\delta}, a^\sigma)\|_{\mathcal{Y}, \beta}^2 = \|\text{Sign}(F(f)) - e^{i\psi^\delta}\|_Y^2 + \beta \| |f| - a^\sigma \|_X^2.$$

Even though the latter functional by itself resembles a Tikhonov functional, we emphasize that here the factor  $\beta > 0$  acts as a balancing weight between two discrepancy terms and does not assume the role of a regularization parameter. This difference is reflected in our notation, where in the following we clearly distinguish between regularization parameters (denoted by  $\alpha$ ) and the discrepancy weight  $\beta$ . Nevertheless, it proved reasonable to consider well-established regularization parameter choice rules also for choosing  $\beta$ , and we present a suitable example in Section 6.

The operator  $\mathcal{F}_{\text{PR}} : X \rightarrow \mathcal{Y}$  is, however, *not* weakly sequentially closed as neither  $\text{Sign}(g)$  nor  $|f|$  have this property, which is easily seen by considering the sequences  $g_n \equiv \frac{1}{n}$  and  $f_n(t) = e^{int}$ , respectively. Thus, following the approach suggested in [13], we approximate the Sign operator for  $\varepsilon \rightarrow +0$  by

$$[\text{Sign}_\varepsilon(g)](s) = \frac{g(s)}{\max(\varepsilon, |g(s)|)}.$$

While the latter operators are Lipschitz-continuous with constant  $\frac{1}{\varepsilon}$ , they still do not satisfy the sufficient condition for well-posedness of Tikhonov regularization.

**Lemma 10.** *For any  $\varepsilon > 0$ , the operator  $\text{Sign}_\varepsilon(g) : L^2_{\mathbb{C}}(0, 2) \rightarrow L^2_{\mathbb{C}}(0, 2)$  is not weakly sequentially closed.*

*Proof.* To prove the assertion we construct a counterexample. Let  $K > \varepsilon$  be fixed and define the sequence  $\{g_n\}$  recursively by

$$g_0(t) = \begin{cases} 2K & \text{if } 0 < t < 2/3 \\ -K & \text{if } 2/3 \leq t < 2 \end{cases} \quad \text{and} \quad g_{n+1}(t) = \begin{cases} g_n(2t) & \text{if } 0 < t < 1 \\ g_n(2t-1) & \text{if } 1 \leq t < 2. \end{cases}$$

In fact, for these choices  $g_n \rightharpoonup 0$  holds, yet  $\text{Sign}_\varepsilon(g_n) \rightharpoonup -\frac{1}{3} \neq \text{Sign}_\varepsilon(0)$ , which shows that  $\text{Sign}_\varepsilon$  is not weakly sequentially closed.  $\square$

In a sense, the weak topology in  $L^2_{\mathbb{C}}(0, 2)$  is insufficient to derive continuity results even for  $\text{Sign}_\varepsilon(g)$ . A common remedy for obtaining regularization properties for operators that are continuous is to restrict ourselves to some subspace with compact embedding into  $L^2_{\mathbb{C}}(0, 2)$ . For the phase retrieval problem, we may thus formulate the following well-posedness result.

**Proposition 11.** *Suppose that the kernel  $k$  satisfies (8) and that the penalty term  $\mathcal{R}(f)$  is stabilizing in  $X_0 = H^1_{\mathbb{C}}(0, 1)$ . Then for any  $\varepsilon > 0$ ,  $a^\sigma \in L^2(0, 1)$  and  $\psi^\delta \in L^2(0, 2)$ , minimizing the Tikhonov functional*

$$(29) \quad \mathcal{T}_{\alpha, \beta}^{\sigma, \delta}(f) = \frac{1}{2} \left\| \text{Sign}_\varepsilon(F(f)) - e^{i\psi^\delta} \right\|_Y^2 + \frac{\beta}{2} \| |f| - a^\sigma \|_X^2 + \alpha \mathcal{R}(f)$$

for  $f \in X_0$  is well-posed in the sense of Remark 8.

*Proof.* Due to the compact embedding of  $X_0 \hookrightarrow X$ , both operators

$$\text{Sign}_\varepsilon(F(f)) : X_0 \rightarrow Y \quad \text{and} \quad |f| : X_0 \rightarrow X$$

are strongly sequentially continuous. Hence they are in particular weakly sequentially closed. The result then follows from the identical argument as in Remark 8.  $\square$

**Remark 12.** As our notation  $\psi^\delta$  suggests, the measurement errors affect the phase function data additively, i.e.,

$$\psi^\delta = \psi^\dagger + \vartheta,$$

where  $\vartheta \in L^2(0, 2)$  denotes the noise. Due to discontinuities in the principal part of the complex argument, however, we evaluate data discrepancy on the complex unit sphere  $S^1_{\mathbb{C}}$ , where we have a multiplicative noise model,

$$e^{i\psi^\delta} = e^{i\psi^\dagger} \cdot e^{i\vartheta}.$$

Finding a discrepancy term that suitably addresses this particular situation is certainly future work, but motivated by the regularization results in Proposition 11 and the successful numerical experiments in Section 6, we use the subspace topology from  $Y = L^2_{\mathbb{C}}(0, 2)$  on  $\{e^{i\psi} : \psi \in L^2(0, 2)\} \subset Y$  instead, and hence introduce the pseudo-metric  $d(\psi_1, \psi_2) := \|e^{i\psi_1} - e^{i\psi_2}\|_Y$  for  $\psi_1, \psi_2 \in L^2(0, 2)$ . Note that this distance penalizes phase differences modulo  $2\pi$  and does not increase the noise level, as

$$d(\psi^\delta, \psi^\dagger) = \left\| 2 \sin \left( \frac{\psi^\delta - \psi^\dagger}{2} \right) \right\|_{L^2(0, 2)} \leq \| \psi^\delta - \psi^\dagger \|_{L^2(0, 2)}.$$

As mentioned in Section 2, the laser pulses generated in optical experiments exhibit a limited bandwidth; cf. (16). In the context of our current notation this requires that the searched-for function  $f^\dagger$  (which corresponds to the Fourier transform  $\hat{E}$  of the laser pulse) can be neglected outside a certain parameter range  $[\tau_1, \tau_2] \subset (0, 1)$ . Due to the structure of the autoconvolution operator, the same is to be expected for the image  $y^\dagger = F(f^\dagger)$ . Therefore, the phase data  $\psi^\delta \approx \arg(y^\dagger)$  carries little or no useful information in regions where  $|y^\dagger|$  is close to zero. As reliable measurements for  $|y^\dagger|$  are, however, not at hand, this fact is not accounted for in the Tikhonov-functional (29). To overcome this issue, a different data fidelity term has been proposed in [4], where the phase data discrepancy is weighted by  $|F(f)|$  or, more precisely, by the normalized function  $|F(f)|/\|F(f)\|_Y$ . While this normalization is necessary to avoid an artificial bias towards reconstructions for which  $|F(f)|$  is small, it also introduces a singularity for  $\|F(f)\|_Y = 0$ . Therefore, we again introduce an approximation level  $\varepsilon > 0$  and consider the variational functional

$$(30) \quad \mathcal{T}_{\varepsilon, \alpha, \beta}^{\sigma, \delta}(f) = \begin{cases} \frac{\|F(f) - |F(f)|e^{i\psi^\delta}\|_Y^2}{\|F(f)\|_Y^2} + \frac{\beta}{2} \| |f| - a^\sigma \|_X^2 + \alpha \mathcal{R}(f), & \text{if } \|F(f)\|_Y \geq \varepsilon, \\ +\infty, & \text{else,} \end{cases}$$

with  $Y = L^2_{\mathbb{C}}(0, 2)$ . The proof that this functional also admits a minimizer in  $H^1_{\mathbb{C}}(0, 1)$  is included in the Appendix.

**Proposition 13.** *Suppose that the kernel  $k$  satisfies (8) and that the penalty term  $\mathcal{R}(f)$  is stabilizing in  $X_0 = H^1_{\mathbb{C}}(0, 1)$  in the sense of Remark 8. Then for any  $a^\sigma \in L^2(0, 1)$ ,  $\psi^\delta \in L^2(0, 2)$ , and  $\varepsilon > 0$  such that*

$$\text{dom } \mathcal{R}(f) \cap \{f \in X_0 : \|F(f)\|_Y \geq \varepsilon\} \neq \emptyset,$$

*a minimizer of  $\mathcal{T}_{\varepsilon, \alpha, \beta}^{\sigma, \delta}(f)$  defined by (30) exists in  $X_0$ .*

**Remark 14.** Another phenomenon of ill-posedness in solving the autoconvolution equation (cf. [11, 15] and [7, Example 3.2]) is due to amplitudes  $a(\tau)$  in  $f(\tau) = a(\tau)e^{i\varphi(\tau)}$ . It occurs when  $a(\tau)$  blows up to infinity locally near some  $\tau_0 \in [0, 1]$  in a way that hardly leaves effects on  $F(f)$ . This phenomenon, however, is suppressed in case of the phase retrieval problem because the amplitude function  $a^\dagger$  is known up to measurement errors. It was, indeed, proven in [7, Proposition 3] that even locally well-posed situations arise under the assumption that the amplitude function  $a$  be fixed and essentially bounded.

## 5. NON-UNIFORM RATIONAL B-SPLINES (NURBS)

The numerical solution of the deautoconvolution problem 1.1 as well as of phase retrieval problems 1.2 requires discretization of the complex-valued autoconvolution equation with kernel function (6). The most natural choice are discretizations using piecewise constant functions, either in terms of step functions  $\{\chi_{[i/n, (i+1)/n)}\}_{i=0, \dots, n-1}$  (see, e.g., [11, 21, 6, 5]) or by means of Haar wavelets (e.g.,

[1, 32]). While step functions yield simple (and computationally efficient) formulae, essentially reducing the continuous autoconvolution to its discrete counterpart, Haar wavelets are particularly suitable for the reconstruction of functions in  $L^2(0, 1)$  as they yield an orthonormal basis both of the infinite dimensional Lebesgue space as well as of its truncated, finite-dimensional approximations. Motivated by the results in Section 4 concerning the existence of minimizers of the Tikhonov functional in the smoother space  $H_{\mathbb{C}}^1(0, 1)$  and by the underlying physical problem of ultrashort laser pulse characterization (see Section 2), we focus on a different representation in the sequel, which is tailored towards the reconstruction of smooth functions. Namely, we will model the curve  $(\operatorname{Re} f, \operatorname{Im} f)$  in the complex plane (and hence  $f$  itself) as a rational B-spline curve.

The shape of a non-uniform rational B-spline (short: NURBS) is determined by a set of  $n$  control points  $\{P_1, \dots, P_n\} \subset \mathbb{R}^2$ , corresponding positive weights  $w = \{w_1, \dots, w_n\}$  and a non-decreasing knot vector  $\eta = \{\eta_1, \dots, \eta_{n+p+1}\}$ , where  $p$  is the polynomial degree of the spline. From the knot vector  $\eta$ , the B-spline basis functions  $N_{j,p}$  are determined via the *Cox-de Boor* recursion

$$N_{j,0}(\tau) = \begin{cases} 1 & \text{if } \eta_j \leq \tau < \eta_{j+1} \\ 0 & \text{otherwise,} \end{cases}$$

$$N_{j,p}(\tau) = \frac{\tau - \eta_j}{\eta_{j+p} - \eta_j} N_{j,p-1}(\tau) + \frac{\eta_{j+p+1} - \tau}{\eta_{j+p+1} - \eta_{j+1}} N_{j+1,p-1}(\tau).$$

Notice that throughout this section we adopt the convention  $\frac{0}{0} := 0$  as is customary in this context. The NURBS curve is then given by

$$\gamma[P, w](\tau) = \frac{\sum_{j=1}^n P_j w_j N_{j,p}(\tau)}{\sum_{j=1}^n w_j N_{j,p}(\tau)}, \quad \tau_{\min} \leq \tau \leq \tau_{\max},$$

where  $\tau_{\min} := \eta_{p+1}$ ,  $\tau_{\max} := \eta_{n+1}$ . We refer the interested reader to [24, 27] for a comprehensive introduction to NURBS. Defining *rational basis functions* as

$$R_{j,p}(\tau) = \frac{w_j N_{j,p}(\tau)}{\sum_{l=1}^n w_l N_{l,p}(\tau)}$$

the curve can be equivalently written as

$$\gamma[P, w](\tau) = \sum_{j=1}^n P_j R_{j,p}(\tau).$$

It can be easily seen from these definitions that NURBS curves are invariant (up to re-parametrization) under rescalings and shifts of the knot vector. Without loss of generality, we may thus assume that the knot vector satisfies  $\eta_{p+1} = 0$  and  $\eta_{n+1} = 1$ , so that the NURBS curve is parametrized by  $\tau \in [0, 1]$ . A common choice are *open* knot vectors of the form

$$\eta = (0, \dots, 0, \eta_{p+2}, \dots, \eta_n, 1, \dots, 1),$$

which have the additional property that the resulting NURBS curve  $\gamma(\tau)$  begins and ends in the first and last control point, respectively, i.e., it satisfies  $\gamma(0) = P_1$  and  $\gamma(1) = P_n$ .

**Autoconvolution and NURBS.** For planar curves the NURBS control points  $P_j$  are typically assumed to be in  $\mathbb{R}^2$ . Nevertheless, the formulae and results in Section 5 remain valid, if we model them as  $P_j \in \mathbb{C}$  which allows to conveniently view the parametrized spline curves as complex-valued functions. In addition, we shall refer to the real and imaginary parts of the control points by  $u_j$  and  $v_j$ , respectively, so that  $P_j = u_j + iv_j$ . The main motivation for working with real parameters is the fact that real-valued functionals, such as the variational Tikhonov-type objective functionals defined in Section 4, are not holomorphic. When using gradient based optimization methods, it is therefore necessary to consider partial derivatives with respect to  $u$  and  $v$ .

Now let the spline degree  $p$ , the number  $n \geq p + 1$  of control points, as well as the knot vector  $\eta$  be fixed. Then, we denote the finite dimensional space of NURBS *design parameters*  $x = (u, v, w)$  by

$$X_n := \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}_+^n \subset \mathbb{R}^{3n}$$

and by  $\gamma : X_n \rightarrow X = L_{\mathbb{C}}^2(0, 1)$  the synthesis operator

$$\gamma[x](\tau) = \sum_{j=1}^n (u_j + iv_j) R_{j,p}(\tau), \quad \text{with } R_{j,p}(\tau) = \frac{w_j N_{j,p}(\tau)}{\sum_{l=1}^n w_l N_{l,p}(\tau)}.$$

In terms of the complex-valued NURBS curves, the autoconvolution operator maps as

$$(31) \quad F_n = F \circ \gamma : X_n \rightarrow Y = L_{\mathbb{C}}^2(0, 2),$$

where  $F : X \rightarrow Y$  is given by (6) with  $f = \gamma[x]$ . In this way, we aim to reconstruct approximations  $f$  in the discretized space

$$X_{n,p}^{\eta} := \{\gamma[x] : x \in X_n\} \subset H_{\mathbb{C}}^1(0, 1).$$

**Regularization for NURBS design parameters.** As illustrated in Section 4, the regularity of the forward operator for the phase retrieval problem 1.2 in its continuous form guarantees existence of minimizers of Tikhonov-type functionals in spaces that embed compactly into  $L_{\mathbb{C}}^2(0, 1)$ . In choosing a discretization space that consists of sufficiently regular spline curves, we readily ensure that the reconstructed solutions belong to  $H_{\mathbb{C}}^1(0, 1)$ . As penalty term for the NURBS design parameters  $x = (u, v, w)$ , defined in the previous subsection, we propose

$$(32) \quad \mathcal{R}(x) = \beta_P \mathcal{R}_P(u, v) + \beta_w \mathcal{R}_{w_0}(w),$$

where we penalize the distance between control points  $P_j = u_j + iv_j$ , and hence to some extent the length of the curve (corresponding to the  $H_{\mathbb{C}}^1(0, 1)$  seminorm), by

$$\mathcal{R}_P(u, v) = \frac{1}{2n} \left[ \sum_{j=1}^{n-1} (u_{j+1} - u_j)^2 + (v_{j+1} - v_j)^2 \right].$$

To account for the constraint that the NURBS weights are required to be positive we penalize them by

$$\mathcal{R}_{w_0}(w) = \frac{1}{2n} \sum_{j=1}^n f_{w_0}(w_j),$$

where

$$f_{w_0}(w) = \begin{cases} +\infty & \text{if } w \leq 0 \\ w^{-2} & \text{if } 0 < w < w_0/2 \\ (w_0/2)^{-4}(w - w_0)^2 & \text{if } w \geq w_0/2. \end{cases}$$

Clearly, this functional, which is shown in Figure 2, acts as a (quadratic) barrier for  $w \rightarrow +0$  as well as for  $w \rightarrow +\infty$ . In addition, it gives preference to values near a reference weight  $w_0 > 0$ . It is worth noting, however, that  $w_0$  does not represent *a priori* knowledge on the average size of the weights, because NURBS weights are only determined up to a constant factor. Instead, the choice of  $w_0$  in combination with the size of the parameter  $\beta_w$  typically impacts the relative difference  $\max(w)/\min(w)$  in the approximate solutions.

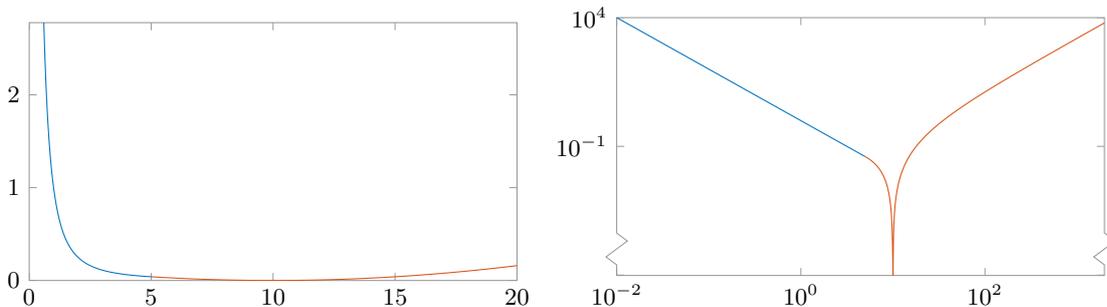


FIGURE 2. Regularizing functional  $f_{w_0}(w)$  for  $w_0 = 10$  (left) and its log-log plot (right).

## 6. NUMERICAL RESULTS

We have tested our method both on synthetic data as well as on real data obtained from an SD-SPIDER apparatus at the Max-Born-Institute for Nonlinear Optics and Short Pulse Spectroscopy in Berlin, Germany. The regularized solutions are NURBS design matrices  $x_{\alpha,\beta}^{\sigma,\delta} = (u_{\alpha,\beta}^{\sigma,\delta}, v_{\alpha,\beta}^{\sigma,\delta}, w_{\alpha,\beta}^{\sigma,\delta})$ , minimizing a variational Tikhonov-type functional

$$\mathcal{T}_{\alpha,\beta}^{\sigma,\delta}(x) = \frac{\|F_n(x) - |F_n(x)|e^{i\psi^\delta}\|_Y^2}{\|F_n(x)\|_Y^2} + \frac{\beta}{2} \|\gamma[x] - a^\sigma\|_X^2 + \alpha \mathcal{R}(x),$$

with penalty term given by (32). The corresponding functions in  $X = L^2_{\mathbb{C}}(0, 1)$  are obtained as  $f_{\alpha,\beta}^{\sigma,\delta} = \gamma[x_{\alpha,\beta}^{\sigma,\delta}]$ . To be precise, we used the functional (30) with approximation level  $\varepsilon = 10^{-10}$ , but as situations with  $\|F_n(x)\|_Y < \varepsilon$  never occurred during the computations, we shall omit this additional parameter for the sake of brevity.

For finding an approximation of  $x_{\alpha,\beta}^{\sigma,\delta}$  we use a Quasi-Newton method,

$$(33) \quad x_{j+1} = x_j + \rho_j H_j^{-1} \nabla T_{\alpha,\beta}^{\sigma,\delta}(x_j),$$

with step-sizes  $\rho_k$  satisfying the Wolfe conditions and BFGS-updates of the approximation  $H_k$  of the Hessian. It is well-known, that gradient based optimization methods for nonlinear problems generally suffer from local minima [3]. To obtain a better approximation of the global optimizer, we employ a strategy known as TIGRA (the name being derived from Tikhonov-GRAdient method), which was introduced by Ramlau [25]. This method was proven to converge globally for suitable problems and with suitably chosen parameters in [25, 26, 32].

Aiming at a well-balanced data fit both for  $\psi^\delta \approx \arg(y^\dagger)$  and for  $a^\sigma \approx |f^\dagger| = |\gamma[x^\dagger]|$ , we used a similar approach to TIGRA, but with respect to the weight  $\beta$  balancing the discrepancy terms (cmp. Section 4). Starting with large  $\beta_0$  and  $q < 1$  (here:  $\beta_0 = 100$ ,  $q = 0.25$ ), the resulting method is described as follows.

- Find  $x_{\alpha,\beta_k}^{\sigma,\delta} \in \arg \min T_{\alpha,\beta_k}^{\sigma,\delta}(x)$  using (33) starting from  $x_{k,0} = x_{\alpha,\beta_{k-1}}^{\sigma,\delta}$ ,

where  $\beta_k = q \beta_{k-1}$ . Initially, for large  $\beta_0$  only  $a^\sigma$  is emphasized while reconstructions will not typically provide a good match for  $\psi^\delta$ . Then the weight is gradually shifted giving more and more importance to  $\psi^\delta$  while the initial good match for  $a^\sigma$  declines only mildly. A globalization approach (such as the proposed reweighing of the discrepancy terms, for example) is certainly required to reach an approximate solution of the original problem, but comes at the cost of a higher computational effort. In our case the additional effort is due to the repeated optimization with different values of the weight  $\beta_k$ . To keep the number of iterations to a minimum, however, we may solve the earlier optimization problems inexactly and increase the required precision while the iteration proceeds. In our experiments we have employed this technique using  $\text{tol}_k = \max(\frac{\beta_k}{2000}, 10^{-9})$  as stopping tolerance for  $\nabla T_{\alpha,\beta_k}^{\sigma,\delta}(x_{k,j})$ , limiting the number of iterations during each optimization procedure to  $\text{maxIt} = 10000$ .

The opposing trends of the two data fidelity terms (cmp. Figure 3) during the iteration can be exploited to obtain a stopping rule for the procedure. Considering that our main objective is to optimize the *overall* data fit, we define a weighted relative least-squares functional

$$(34) \quad e(x)^2 := 2d(x)^2 + r(x)^2,$$

in terms of the auxilliary quantities

$$d(x)^2 := \frac{\|F_n(x) - |F_n(x)|e^{i\psi^\delta}\|_Y^2}{\|F_n(x)\|_Y^2} \quad \text{and} \quad r(x)^2 := \frac{\| |\gamma[x]| - a^\sigma \|_X^2}{\|a^\sigma\|_X^2},$$

and stop the iteration when  $e(x_{\alpha,\beta_k}^{\sigma,\delta})$  reaches its lowest value. Putting more emphasis on the discrepancy term  $d(x)$  in  $Y$  proved beneficial in the numerical experiments. We remark that this procedure may be regarded as a bilevel optimization approach for choosing the discrepancy weight. Namely, for fixed  $\alpha > 0$ , we would

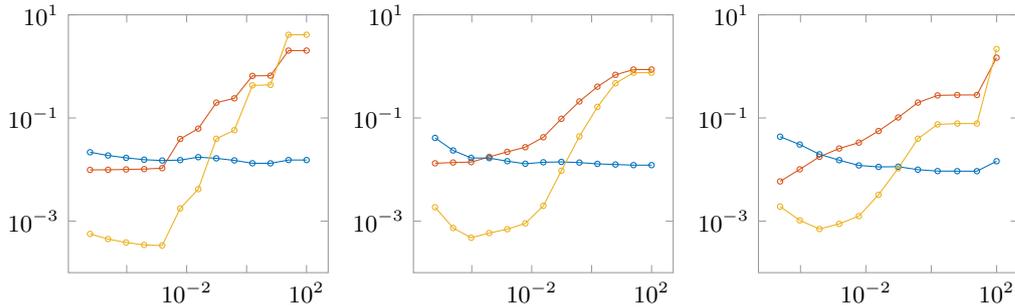


FIGURE 3. Relative data error quantities  $d(x_{\alpha,\beta}^{\sigma,\delta})$  (red) and  $r(x_{\alpha,\beta}^{\sigma,\delta})$  (blue) from the following subsections in dependence on  $\beta$ . The final value  $\beta^*$  is chosen as minimizer of  $e(x_{\alpha,\beta}^{\sigma,\delta})^2$  (yellow).

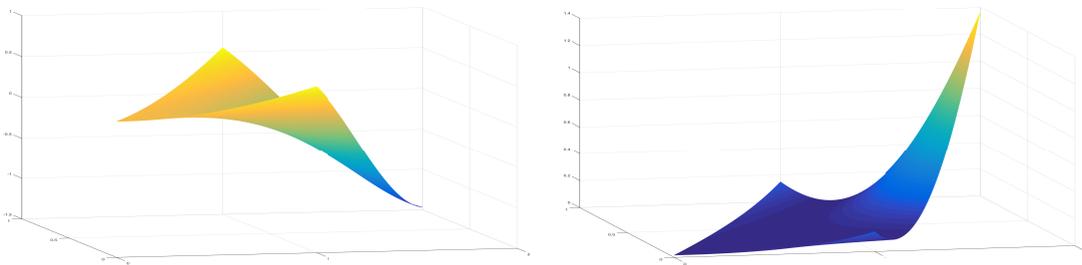


FIGURE 4. The real (left) and imaginary (right) part of the kernel function.

choose  $\beta^*$  as an approximate solution of

$$e(x_{\alpha,\beta}^{\sigma,\delta})^2 \rightarrow \min_{\beta > 0}$$

such that  $x_{\alpha,\beta}^{\sigma,\delta} \in \arg \min_{x \in X_n} \mathcal{T}_{\alpha,\beta}^{\sigma,\delta}(x)$ .

For our purposes, however, the primary motivation is to improve the global convergence properties of the Quasi-Newton method rather than to solve the latter problem. The final values from the experiments in the following subsections have been collected in Table 1.

We first present some results obtained with synthetic data. In order to stay as close to the realistic data situations as possible, we use the same kernel function throughout the following subsections. The kernel shown in Figure 4 was obtained from physical modelling of the nonlinear optical processes that result in the measurements (cf. [23, 15, 4]). The target phase function  $\varphi^\dagger = \arg(f^\dagger)$  was chosen identical to [14, Section 3.2] (see also [15]), and the noisy data were generated by adding 1% relative noise.

The approximate solutions were represented as complex-valued NURBS curves  $f = \gamma[x]$  as introduced in Section 5. In our experiments we used quadratic splines, i.e.,  $p = 2$ , with  $n = 150$  control points and corresponding positive weights. The knot vector  $\eta$ , which determines where and how the control points impact the

parametrized curve, was chosen as *open uniform*. This is to say that

$$\eta_1 = \cdots = \eta_{p+1} = 0, \quad \eta_{n+1} = \cdots = \eta_{n+p+1} = 1,$$

and

$$\eta_j = \frac{j - p - 1}{n - p} \quad \text{for } p + 2 \leq j \leq n.$$

In  $X = L^2_{\mathbb{C}}(0, 1)$  we have discretized the resulting curves by choosing  $N = 1000$  equidistant sampling points  $\tau_k = \frac{k-1}{N-1}$  and, correspondingly, in the image space  $Y = L^2_{\mathbb{C}}(0, 2)$  by  $s_m = \frac{m-1}{N-1}$  for  $m = 1, \dots, 2N - 1$ . Computations were started with an initial guess consisting of constant weights  $w_j = w_0 = 10$  and control points  $P_j = u_j + iv_j$  interpolated from  $a^\sigma$  such that  $u_j = v_j \approx \sqrt{2} \cdot a^\sigma(\frac{j-1}{n-1})$ . In the penalty term (32) we also used  $w_0 = 10$  and balanced both terms equally by  $\beta_P = \beta_w = 1$ . The regularization parameter was chosen as  $\alpha = 10^{-6}$ .

Data	CPU-time	#Iter	$\beta^*$	$d(x_{\alpha, \beta^*}^{\sigma, \delta})$	$r(x_{\alpha, \beta^*}^{\sigma, \delta})$	$e(x_{\alpha, \beta^*}^{\sigma, \delta})^2$
$(a^\sigma, y^\delta)$	1585 s	19942	$3.81 \cdot 10^{-4}$	$1.02 \cdot 10^{-2}$	$1.55 \cdot 10^{-2}$	$4.47 \cdot 10^{-4}$
$(a^\sigma, \psi^\delta)$	1663 s	19391	$9.54 \cdot 10^{-5}$	$1.39 \cdot 10^{-2}$	$1.68 \cdot 10^{-2}$	$6.69 \cdot 10^{-4}$
measured	1848 s	24403	$3.81 \cdot 10^{-4}$	$1.77 \cdot 10^{-2}$	$1.97 \cdot 10^{-2}$	$1.01 \cdot 10^{-3}$

TABLE 1. Comparison of results for various data situations from the following subsections with  $\beta_0 = 100$ ,  $\alpha = 10^{-6}$ .

**Phase retrieval with synthetic data for  $y$ .** We first consider the ideal data situation, assuming that measurements are available for both the modulus and the argument of  $y^\delta \approx y$  as well as for the modulus  $a^\sigma \approx |f^\dagger| = |\gamma[x^\dagger]|$ . To simulate the real data situation we used the same measurements for  $a^\sigma$  as in the final subsection. These measurements are shown in Figure 5 together with the reconstructed solution.

In order to be able to work with the same parameter values for  $\beta_0$ ,  $\text{tol}_k$ , and  $\alpha$  as in the other test cases, we consider here the Tikhonov functional given by

$$\mathcal{T}_{\alpha, \beta}^{\sigma, \delta}(x) = \frac{\|F_n(x) - y^\delta\|_Y^2}{\|y^\delta\|_Y^2} + \frac{\beta}{2} \|\gamma[x] - a^\sigma\|_X^2 + \alpha \mathcal{R}(x),$$

where the NURBS penalty term is as defined by (32) and  $\alpha = 10^{-6}$ . Similarly, the relative data misfit term

$$d(x) = \frac{\|F_n(x) - y^\delta\|_Y}{\|y^\delta\|_Y}$$

was used in the least-squares error functional  $e(x)$  defined by (34).

The results are shown in Figures 5 and 6. Note that the real and imaginary parts of the reconstruction  $f_{\alpha, \beta^*}^{\sigma, \delta} = \gamma[x_{\alpha, \beta^*}^{\sigma, \delta}]$  (right column of Figure 5) evidently provide a good approximation of  $f^\dagger$ . Small oscillations in regions where the modulus  $|f^\dagger|$  is close to zero, however, result in quite large deviations of their arguments in these

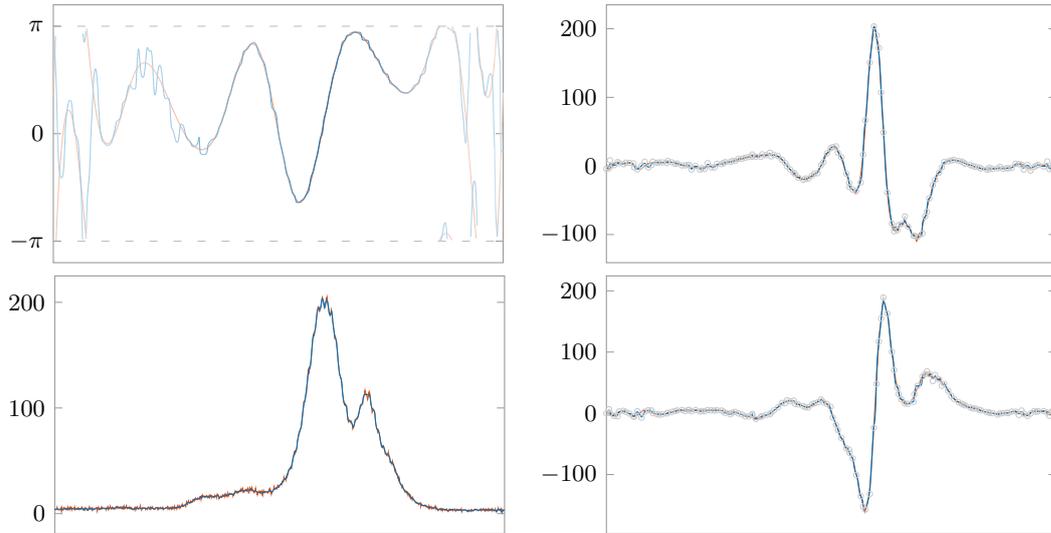


FIGURE 5. Top-left: Target phase  $\varphi^\dagger$  (red) and reconstruction  $\arg(f_{\alpha,\beta^*}^{\sigma,\delta})$  (blue). Bottom-left: Data  $a^\sigma \approx |f^\dagger|$  (red) and reconstruction  $|f_{\alpha,\beta^*}^{\sigma,\delta}|$  (blue). Right column: Real (top) and imaginary (bottom) part of  $f^\dagger$  (red) and  $f_{\alpha,\beta^*}^{\sigma,\delta}$  (blue) together with NURBS control points.

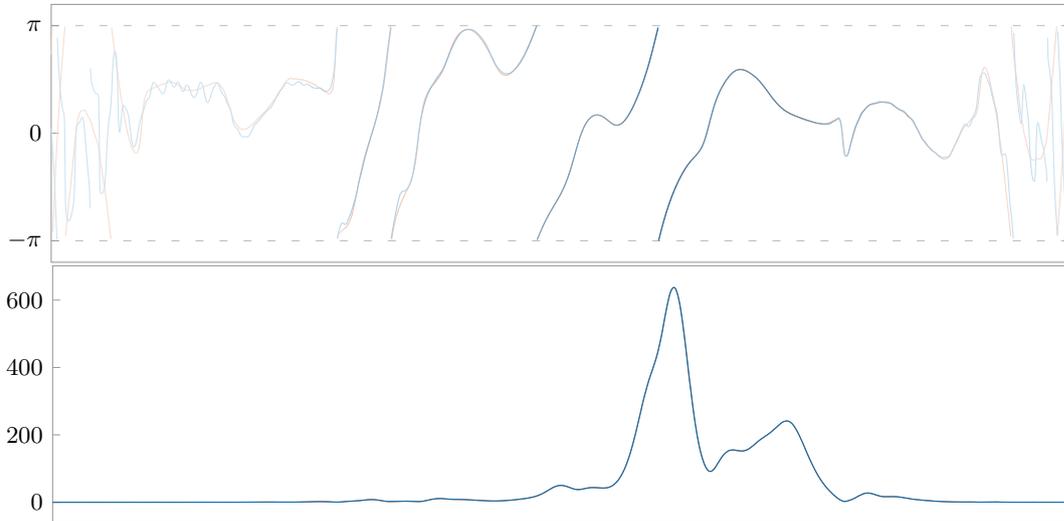


FIGURE 6. Top: True  $\arg(y^\dagger)$  (red) and reconstructed  $\arg F_n(x_{\alpha,\beta^*}^{\sigma,\delta})$  (blue). The brightness is chosen in (logarithmic) dependence on  $|F_n(x_{\alpha,\beta^*}^{\sigma,\delta})|$ . Bottom: True  $|y^\dagger|$  (red) and  $|F_n(x_{\alpha,\beta^*}^{\sigma,\delta})|$  (blue)

areas (left column of Figure 5). As one might expect reconstructions can therefore only be reliable in those regions where  $|f_{\alpha,\beta^*}^{\sigma,\delta}| \gg 0$ . To emphasize this observation

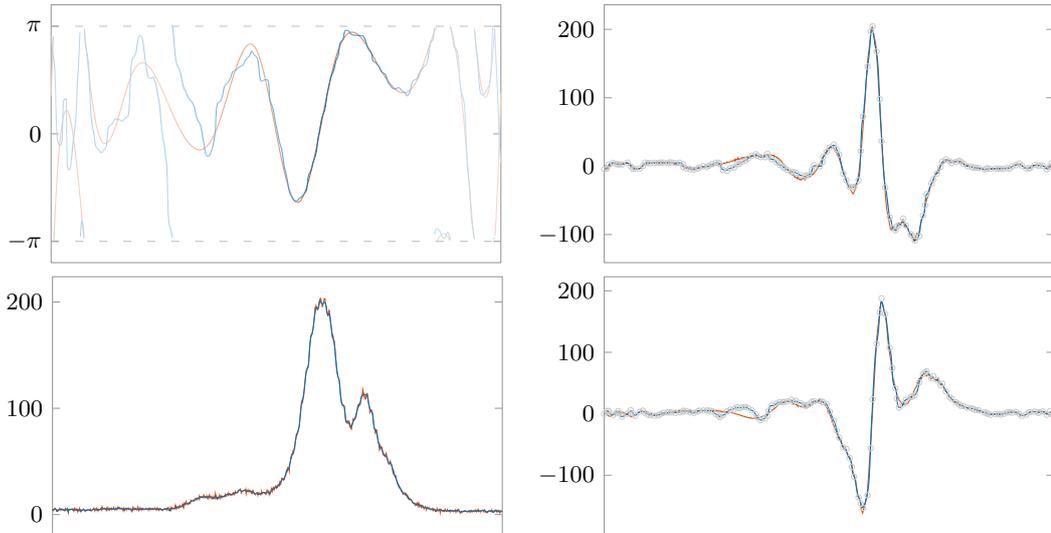


FIGURE 7. Top-left: Target phase  $\varphi^\dagger$  (red) and reconstruction  $\arg(f_{\alpha,\beta^*}^{\sigma,\delta})$  (blue). Bottom-left: Data  $a^\sigma \approx |f^\dagger|$  (red) and reconstruction  $|f_{\alpha,\beta^*}^{\sigma,\delta}|$  (blue). Right column: Real (top) and imaginary (bottom) part of  $f_{\alpha,\beta^*}^{\sigma,\delta}$  together with NURBS control points.

we have weighted the brightness of the plot of  $\arg(f_{\alpha,\beta^*}^{\sigma,\delta})$  and  $\arg(f^\dagger)$  depending on their absolute values.

**Phase retrieval with synthetic data for  $\arg(\mathbf{y})$ .** As another academic example, we consider the phase retrieval problem corresponding to the real data situation in the following subsection. We assume that measurements are available for the argument of  $y^\delta \approx y$  and for the modulus  $a^\sigma \approx |f^\dagger| = |\gamma[x^\dagger]|$ . Again, we used the available experimental data for  $a^\sigma$  which is shown in Figure 7.

The results are shown in Figures 7 and 8. When comparing to the ideal data situation where both  $|y^\dagger|$  and  $|f^\dagger|$  are approximately known, a certain fall-off in quality is evident. However, especially in those regions where  $|f^\dagger| \gg 0$ , the reconstruction still provides a good approximation of the target.

**Phase retrieval with real data.** In the real data from optical measurements, the frequency band containing the support of the  $a^\sigma \approx |f^\dagger|$  is located in between  $\omega_{\text{low}} = 3.5 \times 10^{15}$  Hz and  $\omega_{\text{up}} = 4.1 \times 10^{15}$  Hz, and the frequency of the continuous wave (see Section 2) at  $\omega_{\text{cw}} \approx 3.86 \times 10^{15}$  Hz.

As above, the Tikhonov functional is given by

$$\mathcal{T}_{\alpha,\beta}^{\sigma,\delta}(x) = \frac{\|F_n(x) - |F_n(x)|e^{i\psi^\delta}\|_Y^2}{\|F_n(x)\|_Y^2} + \frac{\beta}{2} \|\gamma[x] - a^\sigma\|_X^2 + \alpha \mathcal{R}(x),$$

with penalty term  $\mathcal{R}(x)$  defined by (32) and  $\alpha = 10^{-6}$ . For the TIGRA-type approach with respect to the weight  $\beta$ , we used  $\beta_0 = 10^2$  and  $q = 0.25$ . Keeping

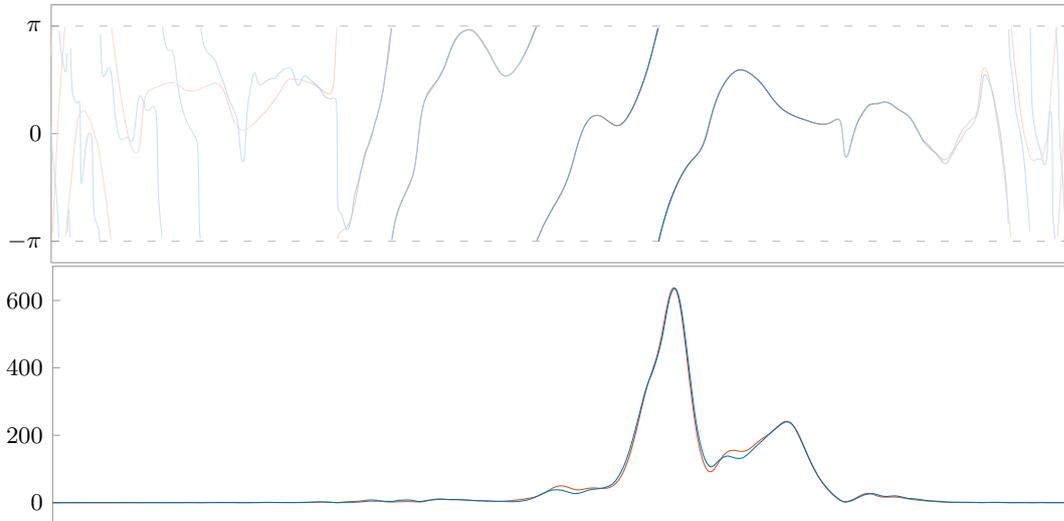


FIGURE 8. Top: Generated data  $\psi^\delta$  (red) and reconstruction (blue) for  $\arg F_n(x_{\alpha,\beta^*}^{\sigma,\delta})$ . The brightness is chosen in (logarithmic) dependence on  $|F_n(x_{\alpha,\beta^*}^{\sigma,\delta})|$ . Bottom: Target  $|y^\dagger|$  and reconstructed  $|F_n(x_{\alpha,\beta^*}^{\sigma,\delta})|$ .

in mind that phase data is reliable only in regions where the absolute value of the function is sufficiently large, we again observe a good data fit in Figures 9 and 10. To highlight this observation, we compare several reconstructions, which were obtained using different parameters in Figure 11.

## CONCLUSION

In this paper, we have studied complex-valued autoconvolution problems with continuous kernel-functions in different data situations, arising, for example, in the characterization of ultrashort laser pulses by means of the SD-SPIDER method. We have derived fundamental analytical properties, in particular, weak-to-weak continuity in  $L^2_{\mathbb{C}}$ -spaces of the autoconvolution operator, which ensure well-posedness of regularization approaches by minimizing Tikhonov-type functionals either in  $L^2_{\mathbb{C}}$  or in compactly embedded subspaces, depending on the data at hand. Inspired by the TIGRA method and using discretizations in terms of NURBS curves, we have also proposed a novel globalized numerical method for phase retrieval problems corresponding to the real-world data situation, which is able to find solutions that appropriately reproduce the given data.

Nevertheless, a number of open questions remain. Convergence rates results are, to be best of our knowledge, completely unavailable except for very special cases, and we have proven that classical nonlinearity conditions such as the tangential cone condition are not suitable to tackle autoconvolution problems. Concerning the noise model, further improvements might be possible with discrepancy terms that capture more adequately the multiplicative noise structure on the complex unit sphere for measurements of the complex phase function. Finally, we expect

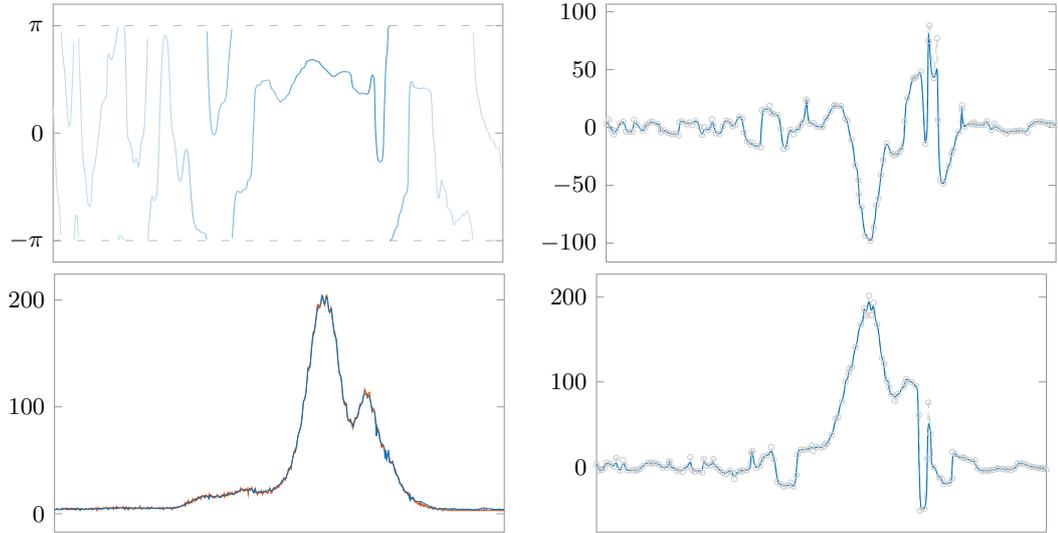


FIGURE 9. Top-left: Reconstructed phase  $\arg(f_{\alpha, \beta^*}^{\sigma, \delta})$ . Bottom-left: Data  $a^\sigma \approx |f^\dagger|$  (red) and reconstruction  $|f_{\alpha, \beta^*}^{\sigma, \delta}|$  (blue). Right column: Real (top) and imaginary (bottom) part of  $f_{\alpha, \beta^*}^{\sigma, \delta}$  together with NURBS control points.

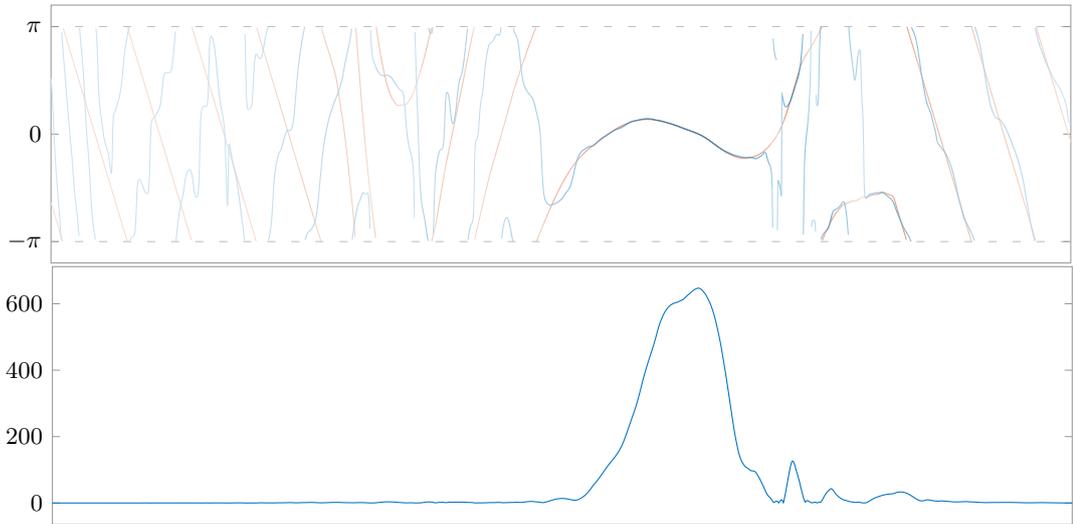


FIGURE 10. SD-phase data  $\psi^\delta$  (red) and reconstruction (blue) for  $\arg F_n(x_{\alpha^*}^\delta)$ . The brightness is chosen in (logarithmic) dependence on  $|F_n(x_{\alpha^*}^\delta)|$ .

that an even better data fit could be achieved numerically by devising a method for blind deautoconvolution, i.e., by including the kernel function as a free or parameter-dependent variable in the optimization method.

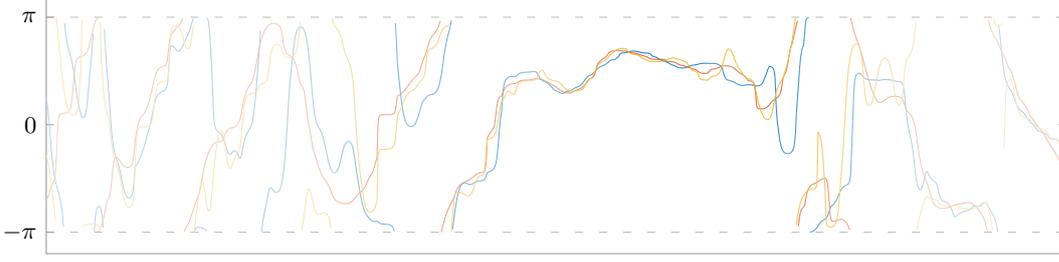


FIGURE 11. Comparison of various phase reconstructions  $\arg(f_{\alpha,\beta^*}^{\sigma,\delta})$  for  $\alpha = 10^{-6}$  (blue),  $5 \cdot 10^{-6}$  (yellow), and  $10^{-5}$  (red), showing a close match in the region where  $|f_{\alpha,\beta^*}^{\sigma,\delta}|$  is large.

#### ACKNOWLEDGEMENT

The first three authors greatly appreciate financial support by the German Research Foundation (DFG) under grants HO 1454/9-1 (SA and BH) and FL 832/1-2 (SB), respectively.

#### APPENDIX

**Proof of Proposition 1.** We consider a sequence  $\{f_n\}_{n \in \mathbb{N}} \in X$  with  $f_n \rightharpoonup f_0$  in  $X$  as  $n \rightarrow \infty$  and show that  $\lim_{n \rightarrow \infty} \langle F(f_n) - F(f_0), \eta \rangle_Y = 0$  for all  $\eta \in Y$ , which proves the proposition. Indeed, by Fubini's theorem we have

$$\begin{aligned}
 & \langle F(f_n) - F(f_0), \eta \rangle_Y \\
 &= \int_0^2 \int_{\max(s-1,0)}^{\min(s,1)} k(s,\tau) (f_n(s-\tau)f_n(\tau) - f_0(s-\tau)f_0(\tau)) d\tau \eta(s) ds \\
 &= \int_0^2 \left[ \int_{\max(s-1,0)}^{\min(s,1)} k(s,\tau) (f_n(s-\tau) - f_0(s-\tau))(f_n(\tau) + f_0(\tau)) d\tau \right] \eta(s) ds \\
 &= \int_0^1 (f_n(\tau) + f_0(\tau)) \left[ \int_{\tau}^{\tau+1} k(s,\tau) (f_n(s-\tau) - f_0(s-\tau)) \eta(s) ds \right] d\tau \\
 &= \int_0^1 (f_n(\tau) + f_0(\tau)) \left[ \int_0^1 k(\xi + \tau, \tau) (f_n(\xi) - f_0(\xi)) \eta(\xi + \tau) d\xi \right] d\tau.
 \end{aligned}$$

If we use the settings  $\Delta_n(\xi) := f_n(\xi) - f_0(\xi)$ ,  $\Theta_\tau(\xi) := k(\xi + \tau, \tau)\eta(\xi + \tau)$  and  $\Xi_n(\tau) := \int_0^1 \Theta_\tau(\xi)\Delta_n(\xi) d\xi$ , where  $\Theta_\tau \in X$  for all  $\tau \in [0, 1]$  due to the continuity

of the kernel  $k$  and  $\Delta_n, \Xi_n \in X$  for all  $n \in \mathbb{N}$ , we arrive at

$$\langle F(f_n) - F(f_0), \eta \rangle_Y = \int_0^1 (f_n(\tau) + f_0(\tau)) \Xi_n(\tau) \, d\tau = \langle f_n + f_0, \Xi_n \rangle_X.$$

By some calculations it can be shown that the family of functions  $\Xi_n$ ,  $n \in \mathbb{N}$ , is equicontinuous on the interval  $[0, 1]$ . On that interval, the sequence  $\{\Xi_n\}_{n \in \mathbb{N}}$  converges pointwise to zero, because  $\Delta_n \rightarrow 0$  in  $X$  implies

$$\Xi_n(\tau) = \langle \Theta_\tau, \Delta_n \rangle_X \rightarrow 0 \quad \text{as } n \rightarrow \infty \quad \text{for all } \tau \in [0, 1].$$

However, an equicontinuous and pointwise convergent sequence of functions is even uniformly convergent, which yields  $\lim_{n \rightarrow \infty} \|\Xi_n\|_X = 0$ . Then the limit condition

$$|\langle F(f_n) - F(f_0), \eta \rangle_Y| = |\langle f_n + f_0, \Xi_n \rangle_X| \leq \|f_n + f_0\|_X \|\Xi_n\|_X \rightarrow 0 \quad \text{as } n \rightarrow \infty,$$

which is based on the fact that  $\|f_n + f_0\|_X$  is bounded, completes the proof.  $\square$

**Proof of Proposition 2.** In the special case  $k \equiv 1$  on  $\mathfrak{P}$  from (7), the proof would be based on the fact that we have  $F(h_n) \not\rightarrow 0$  for  $h_n(t) := e^{int} \rightarrow 0$ . For a general kernel function  $k$ , however, we have to take into account some more details. Since  $k$  is not identically zero, there is some  $(s_0, \tau_0) \in \text{int}(\mathfrak{P})$  with  $\underline{k} := k(s_0, \tau_0) \neq 0$ . Due to the continuity of the kernel function  $k$  on  $\mathfrak{P}$  there exists an open neighborhood  $\mathcal{U}_0$  of  $(s_0, \tau_0)$  with  $|k(s, \tau) - \underline{k}| \leq \frac{|\underline{k}|}{2}$  for all  $(s, \tau) \in \mathcal{U}_0 \cap \text{int}(\mathfrak{P})$ . If we define  $D := \{(2\tau, \tau) : 0 \leq \tau \leq 1\}$ , then there exists  $(s_1, \tau_1)$  and  $\varepsilon > 0$  such that with

$$\mathcal{U}_1 := \{(s, \tau) \in \text{int}(\mathfrak{P}) : \tau_1 \leq \tau \leq \tau_1 + \varepsilon, s_1 - \tau_1 \leq s - \tau \leq s_1 - \tau_1 + \varepsilon\}$$

we have  $\mathcal{U}_1 \subset \mathcal{U}_0 \cap \text{int}(\mathfrak{P})$  and  $\mathcal{U}_1 \cap D = \emptyset$ . It is not hard to show that

$$(35) \quad \left| \int k(s, \tau) h(s - \tau) h(\tau) \, d\tau \right| \geq |\underline{k}| \left| \int h(s - \tau) h(\tau) \, d\tau \right| - \frac{|\underline{k}|}{2} \int |f(s - \tau) f(\tau)| \, d\tau$$

for  $h \in L^2_{\mathbb{C}}(0, 1)$  and  $(s, \tau) \in \mathcal{U}_1$ . Now we define sequences  $\{h_n\}_{n \in \mathbb{N}}$  and  $\{f_n\}_{n \in \mathbb{N}}$  by

$$h_n(\tau) := \begin{cases} e^{in\left(\frac{\tau - \tau_1}{\varepsilon}\right)} & \text{for } \tau_1 \leq \tau \leq \tau_1 + \varepsilon \\ e^{in\left(\frac{\tau - s_1 + \tau_1}{\varepsilon}\right)} & \text{for } s_1 - \tau_1 \leq \tau \leq s_1 - \tau_1 + \varepsilon \\ 0 & \text{else} \end{cases} \quad f_n := f + r h_n.$$

Obviously  $h_n \rightarrow 0$ ,  $f_n \rightarrow f$  in  $X$ , and due to the weak continuity of  $F$  (cf. Proposition 1)  $F(f_n) \rightarrow F(f)$  in  $Y$ . With (35) at hand it is easy to show that

$$\liminf_{n \rightarrow \infty} \|F(h_n)\|_Y > 0,$$

which means that  $F(h_n) \not\rightarrow 0$  in  $Y$ . Taking into account that the Fréchet derivative is compact and therefore completely continuous we have  $F'(f) h_n \rightarrow 0$  in  $Y$ . Together with formula (21), we then obtain

$$F(f_n) - F(f) = r^2 F(h_n) + r F'(f) h_n \quad \text{and thus} \quad F(f_n) \not\rightarrow F(f).$$

The compactness of  $F'(f)$  is an immediate consequence of the fact that (20) is a linear Fredholm integral operator with square integrable kernel. Such operators are always Hilbert-Schmidt operators and thus compact. The nonlinear operator  $F$ , however, is not compact, since the sequence  $\{f_n\}_{n \in \mathbb{N}} \subset X$  is bounded and weakly convergent to  $f$  but the associated sequence  $\{F(f_n)\}_{n \in \mathbb{N}} \subset Y$  cannot have a convergent subsequence. This completes the proof.  $\square$

**Proof of Proposition 13.** Let  $\{f_n\} \subset X_0$  be a sequence such that

$$\mathcal{T}_{\varepsilon, \alpha, \beta}^{\sigma, \delta}(f_n) \rightarrow \inf_{f \in X_0} \mathcal{T}_{\varepsilon, \alpha, \beta}^{\sigma, \delta}(f) < \infty.$$

Passing to a subsequence if necessary, we may assume that  $\|F(f_n)\|_Y \geq \varepsilon$  holds for all  $n$ . Due to the continuity of  $F(f)$ , the pre-image of the closed set  $\{\|y\|_Y \geq \varepsilon\}$ ,

$$A_\varepsilon := \{f \in X : \|F(f)\|_Y \geq \varepsilon\}$$

is a closed subset of  $X = L^2_{\mathbb{C}}(0, 1)$ . Now any accumulation point  $\bar{f}$  of  $\{f_n\}$  with respect to the weak topology in  $X_0$  ( $\bar{f}$  exists due to the stabilizing properties of  $\mathcal{R}(f)$ ) is a strong accumulation point of  $\{f_n\}$  in  $X$  and hence contained in  $A_\varepsilon$ . Taking a subsequence, again denoted by  $\{f_n\}$ , such that  $f_n \rightharpoonup \bar{f}$  in  $X_0$  we thus have  $f_n \rightarrow \bar{f}$  in  $X$  and  $F(f_n) \rightarrow F(\bar{f})$  in  $Y$  with  $\|F(\bar{f})\|_Y \geq \varepsilon$ . This yields

$$\frac{F(f_n) - |F(f_n)|e^{i\psi^\delta}}{\|F(f_n)\|_Y} \rightarrow \frac{F(\bar{f}) - |F(\bar{f})|e^{i\psi^\delta}}{\|F(\bar{f})\|_Y} \text{ in } Y \quad \text{and} \quad |f_n| \rightarrow |\bar{f}| \text{ in } X.$$

In combination with the weak lower semicontinuity of  $\mathcal{R}(f)$  in  $X_0$  we therefore obtain

$$\mathcal{T}_{\varepsilon, \alpha, \beta}^{\sigma, \delta}(\bar{f}) \leq \liminf_{n \rightarrow \infty} \mathcal{T}_{\varepsilon, \alpha, \beta}^{\sigma, \delta}(f_n) = \inf_{f \in X_0} \mathcal{T}_{\varepsilon, \alpha, \beta}^{\sigma, \delta}(f)$$

and the proof is complete.  $\square$

## REFERENCES

- [1] S. W. Anzengruber and R. Ramlau. Morozov's discrepancy principle for Tikhonov-type functionals with nonlinear operators. *Inverse Problems* **26** (2010), 025001.
- [2] J. Baumeister. Deconvolution of appearance potential spectra. In: *Direct and Inverse Boundary Value Problems* (eds Kleinman R, Kress R, Martensen E), vol. 37 of *Methoden und Verfahren der mathematischen Physik*. Peter Lang, Frankfurt am Main 1991, 1–13.
- [3] L. Beilina and M. V. Klibanov. *Approximate Global Convergence and Adaptivity for Coefficient Inverse Problems* Springer, New York, 2012.
- [4] S. Birkholz, G. Steinmeyer, S. Koke, D. Gerth, S. Bürger, and B. Hofmann. Phase retrieval via regularization in self-diffraction based spectral interferometry. *J. Opt. Soc. Am. B* **32** (2015), 983–992.
- [5] S. Bürger. About an autoconvolution problem arising in ultrashort laser pulse characterization. *Preprintreihe der Fakultät für Mathematik*, TU Chemnitz, Preprint 2014-16.
- [6] S. Bürger and J. Flemming. Deautoconvolution: A new decomposition approach versus TIGRA and local regularization. *J. Inverse Ill-Posed Probl.* **23** (2015), 231–243.
- [7] S. Bürger and B. Hofmann. About a deficit in low order convergence rates on the example of autoconvolution. *Applicable Analysis* **94** (2015), 477–493.

- [8] K. Choi and A. D. Lanterman. An iterative deautoconvolution algorithm for nonnegative functions. *Inverse Problems* **21** (2005), 981–995.
- [9] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Kluwer Academic Publishers, Dordrecht, 1996, 2nd Edition 2000.
- [10] H. W. Engl, K. Kunisch, and A. Neubauer. Convergence rates for Tikhonov regularisation of non-linear ill-posed problems. *Inverse Problems* **5** (1989), 523–540.
- [11] G. Fleischer and B. Hofmann. On inversion rates for the autoconvolution equation. *Inverse Problems* **12** (1996), 419–435.
- [12] J. Flemming. Regularization of autoconvolution and other ill-posed quadratic equations by decomposition. *J. Inverse Ill-Posed Probl.* **22** (2014), 551–567.
- [13] F. Frühauf, O. Scherzer, and A. Leitão. Analysis of regularization methods for the solution of ill-posed problems involving discontinuous operators. *SIAM J. Numer. Anal.* **43** (2005), 767–786.
- [14] D. Gerth. *Regularization of an autoconvolution problem occurring in measurements of ultrashort laser pulses*. Diploma thesis, TU Chemnitz 2011, <http://nbn-resolving.de/urn:nbn:de:bsz:ch1-qucosa-85485>.
- [15] D. Gerth, B. Hofmann, S. Birkholz, S. Koke, and G. Steinmeyer. Regularization of an autoconvolution problem in ultrashort laser pulse characterization. *Inverse Problems in Science and Engineering* **22** (2014), 245–266.
- [16] R. Gorenflo and B. Hofmann. On autoconvolution and regularization. *Inverse Problems* **10**, 353–373, 1994.
- [17] B. Hofmann. On smoothness concepts in regularization for nonlinear inverse problems in Banach spaces. Chapter 8 in *Mathematical and Computational Modeling: With Applications in the the Natural and Social Sciences, Engineering, and the Arts* (Ed.: R. Melnik). John Wiley, New Jersey 2015, pp. 192–221.
- [18] B. Hofmann, B. Kaltenbacher, C. Pöschl, and O. Scherzer. A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators. *Inverse Problems* **23** (2007), 987–1010.
- [19] B. Hofmann and O. Scherzer. Factors influencing the ill-posedness of nonlinear problems. *Inverse Problems* **10** (1994), 1277–1297.
- [20] J. Janno. On a regularization method for the autoconvolution equation. *Z. Angew. Math. Mech.* **77** (1997), 393–394.
- [21] J. Janno. Lavrent’ev regularization of ill-posed problems containing nonlinear near-to-monotone operators with application to autoconvolution equation. *Inverse Problems* **16** (2000), 333–348.
- [22] B. Kaltenbacher, A. Neubauer, and O. Scherzer. *Iterative Regularization Methods for Non-linear Ill-posed Problems*. Walter de Gruyter, Berlin, 2008.
- [23] S. Koke S. Birkholz, J. Bethge, C. Grebing C, and G. Steinmeyer. Self-diffraction SPIDER. *Conference on Lasers and Electro-Optics, OSA Technical Digest (CD) (Optical Society of America)* 2010, Paper CMK3, DOI 10.1364/CLEO.2010.CMK3.
- [24] L. Piegl and W. Tiller. *The NURBS book*. 2nd ed. Springer, Berlin, Heidelberg, 1997.
- [25] R. Ramlau. A steepest descent algorithm for the global minimization of the Tikhonov functional *Inverse Problems* **18** (2002), 381–405.
- [26] R. Ramlau. TIGRA —an iterative algorithm for regularizing nonlinear ill-posed problems. *Inverse Problems* **19** (2003), 433–465.
- [27] D. F. Rodgers. *An introduction to NURBS: with historical perspective*. Academic Press, San Diego, CA, USA, 2001.
- [28] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen. *Variational Methods in Imaging*, volume 167 of *Applied Mathematical Sciences*. Springer, New York, 2009.
- [29] K.-Th. Schleicher, S. W. 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), 33–56.

- [30] T. Schuster, B. Kaltenbacher, B. Hofmann, and K. S. Kazimierski. *Regularization Methods in Banach Spaces*, volume 10 of *Radon Ser. Comput. Appl. Math.*. Walter de Gruyter, Berlin/Boston, 2012.
- [31] E. C. Titchmarsh. The zeros of certain integral functions. *Proc. London Math. Society* **25** (1926), 283–302.
- [32] W. Wang, S.W. Anzengruber, R. Ramlau, and B. Han. A global minimization algorithm for Tikhonov functionals with sparsity constraints. *Applicable Analysis* **94** (2015), 580–611.

(S.W. Anzengruber) J. RADON INSTITUTE FOR COMPUTATIONAL AND APPLIED MATHEMATICS (RICAM), 4040 LINZ, AUSTRIA, and FACULTY OF MATHEMATICS, TU CHEMNITZ, 09107 CHEMNITZ, GERMANY

*E-mail address:* `stephan.anzengruber@ricam.oeaw.ac.at`

(S. Bürger) FACULTY OF MATHEMATICS, TU CHEMNITZ, 09107 CHEMNITZ, GERMANY

*E-mail address:* `steven.buerger@mathematik.tu-chemnitz.de`

(B. Hofmann) FACULTY OF MATHEMATICS, TU CHEMNITZ, 09107 CHEMNITZ, GERMANY

*E-mail address:* `bernd.hofmann@mathematik.tu-chemnitz.de`

(G. Steinmeyer) MAX-BORN-INSTITUTE, MAX-BORN-STRASSE 2A, 12489 BERLIN, GERMANY

*E-mail address:* `steinmey@mbi-berlin.de`