Numerische
Mathematik

Check for
updates

On the iterative regularization of non-linear illposed problems in L^∞

Lukas Pieronek[1] · **Andreas Rieder[1]**

Abstract

Parameter identification tasks for partial differential equations are non-linear illposed problems where the parameters are typically assumed to be in L^∞. This Banach space is non-smooth, non-reflexive and non-separable and requires therefore a more sophisticated regularization treatment than the more regular L^p-spaces with $1 < p < \infty$. We propose a novel inexact Newton-like iterative solver where the Newton update is an approximate minimizer of a smoothed Tikhonov functional over a finite-dimensional space whose dimension increases as the iteration progresses. In this way, all iterates stay bounded in L^∞ and the regularizer, delivered by a discrepancy principle, converges weakly-⋆ to a solution when the noise level decreases to zero. Our theoretical results are demonstrated by numerical experiments based on the acoustic wave equation in one spatial dimension. This model problem satisfies all assumptions from our theoretical analysis.

Mathematics Subject Classification 65J20 · 65J22 · 65N20

1 Introduction

We consider the numerical solution of non-linear illposed and inverse problems where the underlying non-linearity F maps from a possibly multi-component version of L^∞ into a normed space Y. This scenario appears quite naturally in many parameter identification tasks for partial differential equations.

✉ Lukas Pieronek
  lukas.pieronek@kit.edu

✉ Andreas Rieder
  andreas.rieder@kit.edu

[1] Department of Mathematics, Karlsruhe Institute of Technology (KIT), Englerstr. 2, 76131 Karlsruhe, Germany

For instance, in electrical impedance tomography (EIT) one wants to identify the conductivity distribution in a body. To this end, electric currents are applied through electrodes and the resulting voltages are measured. In the quasi-static regime, an adapted version of the Laplace equation connects the currents with the voltages, see, e.g., [1]. Another application is full waveform inversion (FWI), the most advanced inversion technique in seismic imaging, see, e.g., [2, 3]. Depending on the used mathematical model for wave propagation (acoustic, elastic, or visco-elastic regime) the searched-for parameters include bulk density, pressure and shear wave velocities, and corresponding relaxation times. Here, a wave field is initiated by a source (explosion or earthquake) and the parts of the wave field that are reflected from the Earth's internal structure are then picked up by receivers on the Earth's surface or by hydrophones in the ocean.

From an abstract point of view, in both examples we have to solve an equation

$$\text{find } u \in \mathcal{D}(F) \text{ such that } F(u) \approx y^\delta \tag{1}$$

where the non-linear operator

$$F : \mathcal{D}(F) \subset L^\infty(D)^\ell \to Y \tag{2}$$

maps ℓ parameter functions u located on some domain of interest D to the respective measurements. Further, y^δ are the (noisy) measurements satisfying $\|y^\delta - F(u^+)\|_Y \leq \delta$ for one $u^+ \in \mathcal{D}(F)$.

Newton-like regularization schemes are well-established iterations for getting a meaningful approximate solution of non-linear inverse problems. In this work we explore the Newton-like solver REGINN$^\infty$ which extends REGINN of [4, 5] to a non-linear inverse problem with generic operators F as in (2). Here, F is required to fulfill a few specific properties which are satisfied by EIT and, except for one, also by FWI in all regimes. This not yet verified property of FWI is a structural assumption known as tangential cone condition (TCC), see (3) below and consult [6] for a first promising result. Apart from establishing the non-linearity constraint, the main challenge about regularization in L^∞ is its non-reflexive nature. As this space is further non-separable, convergence of discretization schemes in the strong topology, which is desirable for practical implementations, cannot be achieved, see [7].

In this work we will exploit the weak-$\star$ topology and semi-discrete approximations to F based on a family $\{X^n\}_n$ of finite-dimensional nested subspaces of $L^\infty(D)^\ell$. Such an ansatz is very close to an implementation and addresses discretization errors directly in the theory. Apart from minor mathematical restrictions which will be specified below, the subspaces will also be held quite generic and may reflect features of the reconstructions one wants to focus on, e.g. using piecewise constant functions to model sharp interfaces. By the well-known limit

$$\|u\|_{L^\infty(D)^\ell} = \lim_{q \to \infty} \|u\|_{L^q(D)^\ell} \quad \text{for all } u \in L^\infty(D)^\ell,$$

for bounded D, X^n can be equivalently furnished with the $L^{q_n}(D)^\ell$-topology for properly chosen $q_n < \infty$ such that $q_n \to \infty$ as $n \to \infty$. This approach then implies that the Newton update for the n-th iterate can be obtained as an approximate minimizer of a smooth (provided Y is smooth) and convex Tikhonov functional over X^n. As a result, the underlying minimizing procedure can be implemented for general $\{X^n\}_n$ with classical techniques from smooth and unconstrained optimization.

There is a wealth of literature on the analysis of Newton-like regularization schemes, mainly in a Hilbert space but meanwhile also in a Banach space setting; we refer only to the monographs [8, 9] for a first reading. However, most of the Banach space methods are formulated in abstract spaces requiring smoothness and reflexivity at least as they rely heavily on duality mappings to mimic the Riesz isomorphism. To the best of our knowledge, regularization schemes applicable to non-reflexive spaces are only considered in [7, 10–15]. The first six of these publications consider iterative schemes on the basis of proximal point methods, Morozov, Ivanov, or Tikhonov regularization, respectively. We will compare our scheme with the somewhat similar IRGNM-Tikhonov method of [12] in Remark 2.6 below.

Our material is organized as follows: In Sect. 2 we introduce and analyze two versions of REGINN$^\infty$ which differ in what information about the smoothness of the ground truth is available a priori. Under reasonable assumptions both algorithms are well defined and terminate with a regularized solution $u_{M_\delta} \in \mathcal{D}(F)$ of (1). Further, we prove the regularization property, that is, weak-$\star$ convergence of $\{u_{M_\delta}\}_{\delta>0}$ to an exact solution of $F(\cdot) = F(u^+)$ as the noise level δ tends to zero. Our hypotheses are reasonable in fact as they are met by EIT as well as FWI with exception of the TCC (Sect. 3) as mentioned above. For FWI in the acoustic regime we are even able to validate the stronger assumption about the ground truth which enters the second version of REGINN$^\infty$. Finally, Sect. 4 contains some experiments to identify two parameters (density and pressure wave velocity) in the acoustic wave equation. Here, all our assumptions including TCC in a semi-discrete setting actually hold. Some technical details that would otherwise interrupt the flow of reading have been moved to three appendices.

2 REGINN$^\infty$

For some bounded domain $D \subset \mathbb{R}^d$ let $F \colon \mathcal{D}(F) \subset L^\infty(D)^\ell \to Y$ be Fréchet-differentiable and satisfy the tangential cone condition (TCC) at u^+, i.e., there are a positive constant $\omega < 1$ and a ball $B_r(u^+) \subset \mathcal{D}(F)$ with radius $r > 0$ such that

$$\|F(u) - F(\widetilde{u}) - F'(\widetilde{u})(u - \widetilde{u})\|_Y \le \omega \|F(u) - F(\widetilde{u})\|_Y$$
$$\text{for all } u, \widetilde{u} \in B_r(u^+). \tag{3}$$

Here, $F' \colon \mathcal{D}(F) \subset L^\infty(D)^\ell \to \mathcal{L}(L^\infty(D)^\ell, Y)$ denotes the Fréchet-derivative of F and $L^\infty(D)^\ell$ is endowed with $\|\cdot\|_{L^\infty(D)^\ell}^2$ where

$$\|u\|_{L^q(D)^\ell}^2 = \|(u_1, \ldots, u_\ell)\|_{L^q(D)^\ell}^2 := \sum_{j=1}^{\ell} \|u_i\|_{L^q(D)}^2, \quad q \in [1, \infty].$$

The TCC, introduced in [16], is a well-established and widely used assumption in the convergence analysis of a variety of iterative regularization schemes for non-linear illposed problems; we only refer to the monographs [8, 9]. Nevertheless, only a few academic examples are reported in the infinite-dimensional setting where it holds, see, e.g., [8, 17, 18].

For $\omega < 1/2$ we can equivalently restate the TCC as

$$\|F(u) - F(\widetilde{u}) - F'(\widetilde{u})(u - \widetilde{u})\|_Y \le L \|F'(\widetilde{u})(u - \widetilde{u})\|_Y, \tag{4}$$

with $L = \omega/(1 - \omega) < 1$. Since our method will be explicitly based on discretizing $L^\infty(D)^\ell$, we impose the following assumptions on corresponding spaces X^n:

(S1) $\{X^n\}_{n \in \mathbb{N}}$ is a sequence of nested subspaces of $L^\infty(D)^\ell$, i.e., $X^n \subset X^{n+1} \subset L^\infty(D)^\ell$ for all $n \in \mathbb{N}$.

(S2) For each X^n there exists a linear projection operator $\mathcal{P}^n \colon L^\infty(D)^\ell \to X^n$, that is, $\mathcal{P}^n u = u$ for all $u \in X^n$, satisfying $\|\mathcal{P}^n u\|_{L^\infty(D)^\ell} \le C_\mathcal{P} \|u\|_{L^\infty(D)^\ell}$ where the constant $C_\mathcal{P} \ge 1$ is independent of n.

(S3) For any fixed $C_\infty > 1$ we can find a positive increasing sequence $\{q_n\}_{n \in \mathbb{N}}$ such that

$$\|u\|_{L^\infty(D)^\ell} \le C_\infty \|u\|_{L^{q_n}(D)^\ell} \quad \text{for all } u \in X^n.$$

Note that the $L^\infty(D)^\ell$-norm is always stronger than the $L^{q_n}(D)^\ell$-norm, hence the magnitude of $C_\infty > 1$ in (S3) determines how tight the norm equivalence of $\|\cdot\|_{L^\infty(D)^\ell}$ and $\|\cdot\|_{L^{q_n}(D)^\ell}$ is, restricted to X^n. Thus, having set $\{q_n\}_{n \in \mathbb{N}}$ once for a fixed feasible subspace sequence $\{X^n\}_{n \in \mathbb{N}}$ and some $C_\infty > 1$, we can easily switch between non-smooth and smooth norms. We emphasize that C_∞ is independent of n. A family $\{X^n\}_{n \in \mathbb{N}}$ enjoying (S1)–(S3) is constructed in Appendix A.1 on the basis of tensor product B-splines (see Appendix A.2 for a possible other construction, which is, however, not considered further in this work).

To ensure that our discretizations do indeed approximate the actual inverse problem in $L^\infty(D)^\ell$ sufficiently well, we require a compatibility condition for F of the form

$$\liminf_{n \to \infty} \|F'(u)(\widehat{u} - \mathcal{P}^n \widehat{u})\|_Y = 0 \tag{5}$$

for all $u \in \mathcal{D}(F)$ and all $\widehat{u} \in L^\infty(D)^\ell$. This extra condition is necessary since, in general, one cannot expect $\lim_{n \to \infty} \mathcal{P}^n \widehat{u} = \widehat{u}$ in $L^\infty(D)^\ell$ strongly (which would yield (5) by continuity of F') because the union of X^n is countable while $L^\infty(D)^\ell$ is not separable. However, as a consequence, the closure of the range of $F'(u)$ must be separable, in contrast to Y itself, for all $u \in \mathcal{D}(F)$.

Additionally, to prove the regularization property of our scheme (Algorithm 1), we will require either the weak-$\star$ closedness of F or the range inclusion

$$\mathcal{R}(F'(u^+)^*) \subset L^1(D)^\ell \tag{6}$$

where $F'(u^+)^* \colon Y^* \to \left(L^\infty(D)^*\right)^\ell$ denotes the adjoint operator. In Sects. 3.1 and 3.2 below we will verify (5) as well as (6) for the forward maps belonging to EIT and FWI, respectively.

As motivated in the introduction, the guideline for designing our regularization algorithm is to generate uniformly bounded iterates u_m in $L^\infty(D)^\ell$ which give sufficiently small residuals

$$b_m^\delta := y^\delta - F(u_m), \tag{7}$$

where $y^\delta \in Y$ is the perturbed datum at hand. We build on an inexact-Newton framework and find the updates from the linearization approximately via Tikhonov regularization linked to X^n as follows: Given an initial guess $u_0 \in \mathcal{D}(F)$ and an initial space X^{n_0}, we iterate

$$u_{m+1} = u_m + s_m, \quad m = 0, 1, 2, \ldots, \tag{8}$$

where $n_m \in \mathbb{N}$ with $n_m \geq n_{m-1}$ and $s_m \in X^{n_m}$ are chosen such that

$$J_{n_m, m}(s_m) \leq \mu_m^2 \|b_m^\delta\|_Y^2. \tag{9}$$

Here, the Tikhonov functional $J_{n,m} \colon X^n \to [0, \infty)$ is

$$J_{n,m}(s) := \|F'(u_m)s - b_m^\delta\|_Y^2 + \alpha_m \|s + (u_m - u_0)\|_{L^{q_n}(D)^\ell}^2 \tag{10}$$

with penalty parameters set a posteriori by

$$\alpha_m = \frac{\|b_m^\delta\|_Y^2}{\gamma^2}. \tag{11}$$

We recall from assumption (S3) that an approximate minimizer of (10) is also one of the corresponding Tikhonov functional with $\| \cdot \|_{L^\infty(D)^\ell}$-penalty term and vice versa. However, given the sequence $\{q_n\}_{n \in \mathbb{N}}$ for some $C_\infty > 1$ and a subspace sequence $\{X^n\}_{n \in \mathbb{N}}$, we have more flexibility in choosing a numerical method to find s_m in (9) via the smoothed version.

So far, the parameters γ, $\{\mu_m\}_{m \in \mathbb{N}}$ are restricted to fulfill $0 < \mu_m < 1$ and $\gamma \neq 0$. While μ_m serves as a stopping criterion in the spirit of an inexact Newton condition to set the m-th update s_m, γ will be responsible for keeping the resulting iterate u_{m+1} sufficiently close to u^+: the larger γ is chosen, the better the initial guess has to be. The effect of the initial discretization level n_0 on the iterates is demonstrated in Section 4 by numerical experiments. We stop the Newton iteration (8) by a discrepancy principle

---

Algorithm 1 REGINN$^\infty$

Input: F; u_0; y^δ; δ; $\{\mu_m\}_m$; τ; γ; C_∞; n_0
Output: u_M with $\|y^\delta - F(u_M)\|_Y \le \tau\delta$
    $m := 0$
    $b_m^\delta := y^\delta - F(u_m)$
    while $\|b_m^\delta\|_Y > \tau\delta$ **do**
       $\alpha_m := \|b_m^\delta\|_Y^2 / \gamma^2$
       find (smallest) $n \ge n_m \in \mathbb{N}$ such that $J_{n,m}(s_m) \le \mu_m^2 \|b_m^\delta\|_Y^2$ for some $s_m \in X^n$
       $u_{m+1} := u_m + s_m$
       $m := m + 1$
       $n_m := n$
       $b_m^\delta := y^\delta - F(u_m)$
    end while
    $M := m$

---

with constant $\tau > 1$. The resulting inversion scheme REGINN$^\infty$ is summarized in Algorithm 1. It is well defined under reasonable assumptions according to the following theorem. Its regularization property is then specified in Theorem 2.3.

Theorem 2.1 (Well-definedness and termination of REGINN$^\infty$) *Let* $F \colon \mathcal{D}(F) \subset L^\infty(D)^\ell \to Y$ *be as above satisfying (3) with* $\omega < 1/3$ *in* $B_r(u^+) \subset \text{int}(\mathcal{D}(F))$ *and (5), where* $\{X^n\}_{n\in\mathbb{N}}$ *and* $\{\mathcal{P}^n\}_{n\in\mathbb{N}}$ *fulfill assumptions* (S1)–(S3). *Let* y^δ *be given such that* $\|F(u^+) - y^\delta\|_Y \le \delta$ *for one* $\delta > 0$. *Let* $\Lambda \in (\frac{2\omega}{1-\omega}, 1)$ *and set*

$$\mu_{\max} := (1 - \omega)\Lambda - \omega.$$

For

$$\gamma \in \left(0, \frac{r}{\sqrt{C_\infty}\mu_{\max}}\right)$$

and

$$r_0 \in \left(0, \min\left\{r - \sqrt{C_\infty}\mu_{\max}\gamma, \frac{\gamma}{C_\mathcal{P}}\sqrt{\mu_{\max}^2 - \omega^2}\right\}\right)$$

choose

$$\tau > \frac{1 + \omega}{\sqrt{\mu_{\max}^2 - C_\mathcal{P}^2 r_0^2/\gamma^2} - \omega}.$$

Further, define

$$\mu_{\min} := \sqrt{\left(\omega + \frac{1 + \omega}{\tau}\right)^2 + C_\mathcal{P}^2 \frac{r_0^2}{\gamma^2}}.$$

Restrict all tolerances $\{\mu_m\}$ *to* $(\mu_{\min}, \mu_{\max})$ *and start with some arbitrary* $n_0 \in \mathbb{N}$ *and* $u_0 \in B_{r_0}(u^+)$. *Then, there exists an* $M_\delta \in \mathbb{N}$ *such that all iterates* $\{u_1, \ldots, u_{M_\delta}\}$

of REGINN$^\infty$ *are well-defined and stay in* $B_r(u^+)$. *Moreover,* $\|b_{m+1}^\delta\|_Y \leq \Lambda \|b_m^\delta\|_Y$
for $m = 0, \ldots, M_\delta - 1$, $\|b_{M_\delta}^\delta\|_Y \leq \tau\delta$, *and* $M_\delta = \mathcal{O}(|\log\delta|)$ *as* $\delta \searrow 0$.

Proof Before we begin with the proof we discuss our assumptions on the parameters. First, observe that the open interval for choosing Λ is non-empty by $\omega < 1/3$. The lower bound for Λ guarantees that $\mu_{\max} > \omega$. Together with the upper bound on γ this yields a positive upper bound for r_0. Further, the radicand and the denominator of the lower bound for τ are positive. Finally, $\mu_{\min} < \mu_{\max}$.

We use an inductive argument and assume therefore that $\|b_m^\delta\|_Y \leq \Lambda^m \|b_0^\delta\|_Y$ as well as $\|u_i - u^+\|_{L^\infty(D)^\ell} < r$ for $i \leq m$, which holds in particular for $m = 0$ because of $\|u_0 - u^+\|_{L^\infty(D)^\ell} < r_0 < r$. If $\|b_m^\delta\|_Y \leq \tau\delta$, REGINN$^\infty$ stops with $u_{M_\delta} := u_m$ and nothing else needs to be shown. Otherwise, $\|b_m^\delta\|_Y > \tau\delta$ and we next show that a Newton update is well defined by (9). Let $s_{n,m} := \arg\min_{s \in X^n} J_{n,m}(s)$ which exists as the unique minimizer of a strictly convex functional over a finite dimensional space. Then,

$$
\begin{aligned}
J_{n,m}(s_{n,m}) &\leq J_{n,m}\big(\mathcal{P}^n(u^+ - u_m)\big) \\
&= \|F'(u_m)\mathcal{P}^n(u^+ - u_m) - b_m^\delta\|_Y^2 \\
&\quad + \alpha_m \|\mathcal{P}^n(u^+ - u_m) + (u_m - u_0)\|_{L^{q_n}(D)^\ell}^2.
\end{aligned}
$$

Recursively, we get $u_m - u_0 = s_0 + s_1 + \cdots + s_{m-1}$ from which we deduce that $u_m - u_0$ is in $X^{n_{m-1}}$ as the spaces are nested by (S1). Hence, by (S2),

$$
\mathcal{P}^n(u_m - u_0) = u_m - u_0 \tag{12}
$$

for $n \geq n_{m-1}$ and by linearity of $\mathcal{P}^n$ we may simplify

$$
J_{n,m}(s_{n,m}) \leq \|F'(u_m)\mathcal{P}^n(u^+ - u_m) - b_m^\delta\|_Y^2 + \alpha_m \|\mathcal{P}^n(u^+ - u_0)\|_{L^{q_n}(D)^\ell}^2
$$

$$
\leq \|F'(u_m)\mathcal{P}^n(u^+ - u_m) - b_m^\delta\|_Y^2 + \text{vol}_{\text{d}}(D)^{2/q_n} C_\mathcal{P}^2 \frac{r_0^2}{\gamma^2} \|b_m^\delta\|_Y^2.
$$

In the last step we additionally used (11), Hölder's inequality and $\|\mathcal{P}^n(u^+ - u_0)\|_{L^\infty(D)^\ell} \leq C_\mathcal{P}\|u^+ - u_0\|_{L^\infty(D)^\ell} < C_\mathcal{P}r_0$, see (S2). We continue by splitting the residual term according to

$$
\begin{aligned}
\|F'(u_m)\mathcal{P}^n(u^+ - u_m) - b_m^\delta\|_Y &\leq \|F'(u_m)(u^+ - u_m) - F(u^+) - F(u_m)\|_Y \\
&\quad + \|F(u^+) - y^\delta\|_Y \\
&\quad + \|F'(u_m)\big(\mathcal{P}^n(u^+ - u_m) - (u^+ - u_m)\big)\|_Y \\
&\leq \|F'(u_m)(u^+ - u_m) - F(u^+) - F(u_m)\|_Y + \delta \\
&\quad + \|F'(u_m)\big(\mathcal{P}^n(u^+ - u_0) - (u^+ - u_0)\big)\|_Y,
\end{aligned}
$$

employing again (12) to get the bottom line. Since $\|u_m - u^+\|_{L^\infty(D)^\ell} < r$ by induction, TCC (3) yields

$$\|F'(u_m)\mathcal{P}^n(u^+ - u_m) - (y^\delta - F(u_m))\|_Y$$
$$\leq \omega\|F(u^+) - F(u_m)\|_Y + \delta + \|F'(u_m)(\mathcal{P}^n(u^+ - u_0) - (u^+ - u_0))\|_Y$$

and with $\|F(u^+) - F(u_m)\|_Y \leq \|F(u^+) - y^\delta\|_Y + \|b_m^\delta\|_Y \leq \delta + \|b_m^\delta\|_Y$ we deduce further

$$\|F'(u_m)\mathcal{P}^n(u^+ - u_m) - (y^\delta - F(u_m))\|_Y$$
$$\leq \omega(\delta + \|b_m^\delta\|_Y) + \delta + \|F'(u_m)(\mathcal{P}^n(u^+ - u_0) - (u^+ - u_0))\|_Y.$$

Taking into account that $\|b_m^\delta\|_Y > \tau\delta$, we get

$$\|F'(u_m)\mathcal{P}^n(u^+ - u_m) - (y^\delta - F(u_m))\|_Y$$
$$\leq \|b_m^\delta\|_Y\left(\omega + \frac{1+\omega}{\tau}\right) + \|F'(u_m)(\mathcal{P}^n(u^+ - u_0) - (u^+ - u_0))\|_Y$$

and finally

$$J_{n,m}(s_{n,m}) \leq \left(\|b_m^\delta\|_Y\left(\omega + \frac{1+\omega}{\tau}\right) + \|F'(u_m)(\mathcal{P}^n(u^+ - u_0) - (u^+ - u_0))\|_Y\right)^2$$
$$+ \mathrm{vol}_d(D)^{2/q_n} C_\mathcal{P}^2 \frac{r_0^2}{\gamma^2}\|b_m^\delta\|_{L^2(D)}^2. \tag{13}$$

Since $\mathrm{vol}_d(D)^{2/q_n} \to 1$ as $n \to \infty$ and in view of (5), we find that

$$\liminf_{n\to\infty} J_{n,m}(s_{n,m}) \leq \|b_m^\delta\|_Y^2\left(\left(\omega + \frac{1+\omega}{\tau}\right)^2 + C_\mathcal{P}^2\frac{r_0^2}{\gamma^2}\right) = \mu_{\min}^2\|b_m^\delta\|_Y^2. \tag{14}$$

Consequently, condition (9) with $\mu_m > \mu_{\min}$ is feasible for n_m large and appropriate $s_m \in X^{n_m}\backslash\{0\}$, where $s_m \neq 0$ follows by $J_{n,m}(0) \geq \|b_m^\delta\|_Y^2$. Hence, $u_{m+1} = u_m + s_m$ is well defined and, relying on (S3) as well as (11), we proceed with

$$\|u_{m+1} - u_0\|_{L^\infty(D)^\ell}^2 = \|s_m + (u_m - u_0)\|_{L^\infty(D)^\ell}^2 \leq C_\infty \frac{J_{n_m,m}(s_m)}{\alpha_m}$$
$$\leq C_\infty \frac{\mu_m^2\|b_m^\delta\|_Y^2}{\alpha_m} < C_\infty\mu_m^2\gamma^2.$$

Hence,

$$\|u_{m+1} - u^+\|_{L^\infty(D)^\ell} \leq \|u_{m+1} - u_0\|_{L^\infty(D)^\ell} + \|u^+ - u_0\|_{L^\infty(D)^\ell}$$
$$< \sqrt{C_\infty}\mu_{\max}\gamma + r_0 < r.$$

by the upper bound of r_0, yielding $u_{m+1} \in B_r(u^+) \subset \mathrm{int}(\mathcal{D}(F))$. Finally, we estimate on the basis of (4) and (9)

$$
\begin{aligned}
\|b^\delta_{m+1}\|_Y &= \|(b^\delta_m - F'(u_m)s_m) - (F(u_{m+1}) - F(u_m) - F'(u_m)s_m)\|_Y \\
&\leq \|b^\delta_m - F'(u_m)s_m\|_Y + \frac{\omega}{1-\omega}\|F'(u_m)s_m\|_Y \\
&\leq \sqrt{J_{n,m}(s_m)} + \frac{\omega}{1-\omega}\left(\|b^\delta_m\|_Y + \|F'(u_m)s_m - b^\delta_m\|_Y\right) \\
&\leq \mu_m\|b^\delta_m\|_Y + \frac{\omega}{1-\omega}(1+\mu_m)\|b^\delta_m\|_Y \\
&= \left(\mu_m + \frac{\omega}{1-\omega}(1+\mu_m)\right)\|b^\delta_m\|_Y \\
&< \left(\mu_{\max} + \frac{\omega}{1-\omega}(1+\mu_{\max})\right)\|b^\delta_m\|_Y < \Lambda\|b^\delta_m\|_Y.
\end{aligned}
\tag{15}
$$

Having thus proven the induction part, REGINN$^\infty$ is forced to terminate for any $\delta > 0$ due to $\|b^\delta_m\|_Y \leq \Lambda^m\|b^\delta_0\|_Y \leq \tau\delta$ for m sufficiently large. From this estimate, we may even deduce $M_\delta = \mathcal{O}(|\log\delta|)$ as $\delta \searrow 0$. $\qquad\square$

Remark 2.2 (a) The name REGINN$^\infty$ for Algorithm 1 is justified by the stopping condition (9) for determining the Newton update which is, in view of (10) and (11), equivalent to

$$
\frac{\|F'(u_m)s_m - b^\delta_m\|_Y^2}{\|b^\delta_m\|_Y^2} + \frac{\|s_m + (u_m - u_0)\|_{L^{q_n}(D)^\ell}^2}{\gamma^2} \leq \mu_m^2.
$$

In particular, s_m satisfies the stopping condition of REGINN [4], i.e., the above condition without penalty term.

(b) Recall that REGINN fulfills in the Hilbert space setting (and likewise for smooth reflexive Banach spaces) the error reducing property for the iterates of many inner linear solvers, keeping thus $u_m \in B_r(u^+)$ if the initial guess was chosen so. However, this does not hold any longer for our L^∞-tailored REGINN$^\infty$ in general. Therefore our parameters need to be controlled in terms of both ω and r, whereas standard REGINN only requires the knowledge of ω for defining admissible tolerances μ and stopping constants τ, see Theorem 3.1 in [19].

In case that F is linear, i.e., $F(u) = Au$ for some $A \in \mathcal{L}(L^\infty(D)^\ell, Y)$, the TCC holds with $\omega = 0$ and $r = \infty$. Some observations are in order:

- Although r_0 can be arbitrarily large now, to still ensure a finite and uniform L^∞-bound on the iterates, $r_0 < \gamma < \infty$ needs to be chosen compatibly.
- Because of

$$
\|F'(u_m)s_m - b^\delta_m\|_Y^2 = \|As_m - (y^\delta - Au_m)\|_Y^2 = \|Au_{m+1} - y^\delta\|_Y^2
$$

the iterate u_{m+1} satisfies

$$\|Au_{m+1} - y^\delta\|_Y^2 + \alpha_m \|u_{m+1} - u_0\|_{L^{q_{n_m}}(D)}^2 \le \mu_m^2 \|b_m^\delta\|_Y^2.$$

Hence, u_{m+1} can be considered an approximate minimizer of the Tikhonov functional $u \mapsto \|Au - y^\delta\|_Y^2 + \alpha_m \|u - u_0\|_{L^{q_{n_m}}(D)}^2$ in the set $u_0 + X^{n_m}$. Put differently: in the linear case, REGINN$^\infty$ can be viewed as a cascading Tikhonov regularization iterating over nested finite-dimensional spaces where the penalty term is determined a posteriori by the previous iterate.

Theorem 2.3 (Regularization property of REGINN$^\infty$) *Adopt all assumptions and notations from Theorem 2.1 with $\overline{B_r(u^+)} \subset \mathcal{D}(F)$ and set $F(u^+) = y$. Additionally, assume that $F'(u^+)$ fulfills (6) or that F is weakly-$\star$ sequentially closed, that is, $w_n \overset{\star}{\rightharpoonup} w$ in $L^\infty(D)^\ell$ and $F(w_n) \rightharpoonup z$ imply that $F(w) = z$. Let $\{\delta_i\}_{i\in\mathbb{N}}$ be a positive zero sequence and let $u_{M_{\delta_i}}$ be the output of Algorithm 1 with respect to perturbed data y^{δ_i}. Then the set of weak-$\star$ accumulation points of the sequence $\{u_{M_{\delta_i}}\}_{i\in\mathbb{N}}$ is non-empty and consists of solutions to $F(\cdot) = y$. If u^+ is the only solution in $\overline{B_r(u^+)}$, then even the whole sequence $\{u_{M_{\delta_i}}\}_{i\in\mathbb{N}}$ converges weakly-$\star$ to u^+ in $L^\infty(D)^\ell$.*

Proof By construction in Theorem 2.1 we know that $\{u_{M_{\delta_i}}\}_{i\in\mathbb{N}}$ yields

$$\|y - F(u_{M_{\delta_i}})\|_Y \le \|y - y^{\delta_i}\|_Y + \|b_{M_{\delta_i}}^{\delta_i}\|_Y \le (1+\tau)\delta_i \to 0 \qquad (16)$$

and is uniformly bounded in $L^\infty(D)^\ell$, so there exists weak-$\star$ accumulation points in $\overline{B_r(u^+)}$ by sequential weak-$\star$-compactness. Take representatively $u_{M_{\delta_{i_k}}} \overset{\star}{\rightharpoonup} \tilde{u}$. In case that F is weakly-$\star$ sequentially closed, we can directly deduce $F(\tilde{u}) = y$, hence any weak-$\star$ accumulation point solves the equation. In case that (6) holds, we first note that the TCC (3) implies by the reverse triangle inequality for any $u \in B_r(u^+)$ that

$$(1-\omega)\|F(u^+) - F(u)\|_Y \le \|F'(u^+)(u^+ - u)\|_Y$$
$$\le (1+\omega)\|F(u^+) - F(u)\|_Y. \qquad (17)$$

With (6) we then obtain for any $g \in Y^*$

$$\begin{aligned}\langle F'(u^+)\tilde{u}, g\rangle_{Y,Y^*} &= \langle \tilde{u}, F'(u^+)^* g\rangle_{L^\infty(D)^\ell, L^1(D)^\ell} \\ &= \lim_{k\to\infty} \langle u_{M_{\delta_{i_k}}}, F'(u^+)^* g\rangle_{L^\infty(D)^\ell, L^1(D)^\ell} \\ &= \lim_{k\to\infty} \langle F'(u^+)u_{M_{\delta_{i_k}}}, g\rangle_{Y,Y^*} \\ &= \langle F'(u^+)u^+, g\rangle_{Y,Y^*},\end{aligned}$$

where the last equality above follows by the second inequality in (17) with $u = u_{M_{\delta_{i_k}}}$ and $F(u_{M_{\delta_{i_k}}}) \to y$ in Y. We deduce that $F'(u^+)(u^+ - \tilde{u}) = 0$ and combining this relation now with the first inequality in (17) using $\tilde{u} = u$, we may again conclude

$F(\widetilde{u}) = y$. Finally, if u^+ is the only solution to $F(\cdot) = y$ in $\overline{B_r(u^+)}$, the weak-$\star$ convergence of the whole sequence $\{u_{M_{\delta_i}}\}_{i \in \mathbb{N}}$ follows by a standard subsequence argument, see [20, Prop. 10.13(2)]. □

Remark 2.4 If $F'(u^+)$ is injective, $\|u\|_\star := \|F'(u^+)u\|$ constitutes a norm on $L^\infty(D)^\ell$ with respect to which $\{u_{M_\delta}\}_{\delta > 0}$ then converges strongly to u^+ according to (17) and (16) at the rate $\|u^+ - u_{M_\delta}\|_\star = \mathcal{O}(\delta)$ as $\delta \to 0$. However, this norm is generally weaker than $\|\cdot\|_{L^\infty(D)^\ell}$ with equivalence if and only if $F'(u^+)$ is boundedly invertible. However, for locally illposed problems $F(\cdot) = y$ we expect its linearization to be illposed as well.

Remark 2.5 We discuss how the statements of the theorems from above carry over to a semi-discrete situation as it appears under an implementation of Algorithm 1. Typically, one $X^{n_{\max}}$ represents the finest possible or finest chosen resolution for the sought-for quantity $u^+ \in X^{n_{\max}}$ and models the implementation from a mathematical point of view.[1] Here, $X^{n_{\max}}$ is equipped with the L^∞-topology. Now, Theorems 2.1 and 2.3 apply to

$$F_{n_{\max}} : \mathcal{D}(F) \cap X^{n_{\max}} \subset L^\infty(D)^\ell \to Y, \quad u \mapsto F(u),$$

where (5) can be omitted due to

$$F'_{n_{\max}}(u)\big((u^+ - u_0) - \mathcal{P}^{n_{\max}}(u^+ - u_0)\big) = 0$$

since both u_0 and u^+ are assumed to be in $X^{n_{\max}}$. Further, (14) then reads

$$J_{n_{\max},m}(s_{n_{\max},m}) \leq \|b_m^\delta\|_Y^2 \left(\left(\omega + \frac{1+\omega}{\tau} \right)^2 + \mathrm{vol}_d(D) C_{\mathcal{P}}^2 \frac{r_0^2}{\gamma^2} \right)$$

and as the only consequence the constant $C_{\mathcal{P}}$ needs to be replaced by $\mathrm{vol}_d(D)^{1/2} C_{\mathcal{P}}$ in the definition of corresponding REGINN^∞ parameters.

We emphasize that the underlying semi-discrete inverse problem is: given $y^\delta \in Y$ find $u^\delta \in X^{n_{\max}}$ such that $F_{n_{\max}}(u^\delta) \approx y^\delta$ where y^δ now incorporates measurement noise and discretization error.

Remark 2.6 At first glance the IRGNM-Tikhonov method of [12] and REGINN^∞ seem to be quite similar, but they are separated by significant structural differences: In IRGNM-Tikhonov the penalty parameter is determined a priori and is assigend to a fixed regularization term, whereas for REGINN^∞ the regularization term is explicitly n-dependent and relies on an the posteriori parameter choice (11). Further, the Newton update for the IRGNM-Tikhonov method has to be an exact minimizer of the Tikhonov functional. In contrast, the Newton update for REGINN^∞ only requires

---

[1] Recall that Theorem 2.1 in its original version requires an initial guess $u_0 \in B_{r_0}(u^+)$. Since $L^\infty(D)^\ell$ is not separable, however, there might be no element in X^n for any $n \in \mathbb{N}$ which satisfies this closeness condition. As remedy we may even enlarge the parameter space for the semi-discrete problem to $U_0 + X^{n_{\max}}$ where $U_0 \subset L^\infty(D)^\ell$ is a proper finite dimensional subspace such that $u_0 \in U_0$. In this case, we assume $u^+ \in U_0 + X^{n_{\max}}$.

an approximate minimizer, see (9), which is especially convenient when it comes to practical realizations. Note also that the discrete spaces X^n are an integral part of REGINN$^\infty$ and its theory, so numerical effects are included in a natural way. As part of our results, the discrete regularizers $u_{M_{\delta_i}} \in X^{n_{M_{\delta_i}}}$ of Theorem 2.3 converge to the continuous limit $u^+ \in L^\infty(D)^\ell$. A similar connection of the discrete with the continuous world is (to our knowledge) still pending for the method in [12].

The previous version of REGINN$^\infty$ requires the determination of successive discretization levels $n_m \geq n_{m-1}$ for possibly strongly increasing n_m so that many intermediate (almost) minimizers $s \in X^n$ of (10) for $n_{m-1} \leq n \leq n_m$ need to be computed before meeting the given μ_m-criterion in (9). As this adaptive discretization strategy can become numerically expensive, we want to present an alternative version which directly links n to m. A closer look at the proof of Theorem 2.1 reveals that n_m actually depends on the decay of $\|F'(u_m)(\mathcal{P}^n(u^+ - u_0) - (u^+ - u_0))\|_Y$. Hence, if we have a concrete upper bound for this discretization residual in terms of n, feasible choices of n_m can be found by simple algebraic manipulation. Such upper bounds can be deduced on the basis of better initial guesses which are governed by some stronger norm. For this purpose we state the following refined version of assumption (5): If $X \subset L^\infty(D)^\ell$ is an embedded normed space such that

$$\|\widehat{u}\|_{L^\infty(D)^\ell} \leq C_X \|\widehat{u}\|_X \quad \text{for all } \widehat{u} \in X, \tag{18}$$

then for any u^+ such that $B_{\widetilde{r}}(u^+) \subset \text{int}(\mathcal{D}(F))$, $u \in B_{\widetilde{r}}(u^+)$ and $\widehat{u} \in X$ we assume that

$$\|F'(u)(\widehat{u} - \mathcal{P}^n \widehat{u})\|_Y \leq C^+ \|\widehat{u}\|_X \beta(n), \tag{19}$$

where $C^+ > 0$ and β fulfills $\beta(n) \searrow 0$ with $\beta(0) = 1$. We think of β as being rather independent of $u \in \mathcal{D}(F)$ once X and $\{X^n\}_{n\in\mathbb{N}}$ are set while the magnitude of C^+ is strongly $B_{\widetilde{r}}(u^+)$-dependent. We will verify in Sect. 3.3 below that (19) is satisfied, for instance, by the modeling assumptions of FWI in the acoustic regime. The next theorem shows that we can indeed determine n_m conveniently for successive Newton steps of REGINN$^\infty$ when replacing condition (5) by (19).

Theorem 2.7 *Adopt all assumptions, notations and parameters from Theorems 2.1 and 2.3, and assume that (19) holds—without loss of generality with $\widetilde{r} = r$ by shrinking one of the radii otherwise. Start with $u_0 \in L^\infty(D)^\ell$ such that $\|u^+ - u_0\|_X < \min\{r_0/C_X, 1/C^+\}$ and restrict $\{\mu_m\}$ to $(\mu_{\min}^\varepsilon, \mu_{\max})$, where*

$$\mu_{\min}^\varepsilon := \sqrt{\left(\omega + \frac{1+\omega}{\tau} + \varepsilon\right)^2 + \max\left\{vol_d(D)^{2/q_{n_0}}, 1\right\} C_\mathcal{P}^2 \frac{r_0^2}{\gamma^2}} < \mu_{\max} \tag{20}$$

for some $\varepsilon > 0$ sufficiently small and q_{n_0} large with $n_0 \in \mathbb{N}$. Further, let n_m be successively defined by

$$n_m := \min\left\{n \geq n_0 : \beta(n) \leq \varepsilon \|b_m^\delta\|_Y\right\}. \tag{21}$$

Algorithm 2 REGINN$^\infty$ for improved initial guesses

Input: F; u_0; y^δ; δ; $\{\mu_m\}_m$; τ; γ; C_∞; n_0 ; ε; β
Output: u_M with $\|y^\delta - F(u_M)\|_Y \leq \tau\delta$
   $m := 0$
   $b_m^\delta := y^\delta - F(u_m)$
   while $\|b_m^\delta\|_Y > \tau\delta$ **do**
      $\alpha_m := \|b_m^\delta\|_Y^2/\gamma^2$
      $n_m := \min\left\{n \geq n_0 : \beta(n) \leq \varepsilon\|b_m^\delta\|_Y\right\}$
      determine $s_m \in X^{n_m}$: $J_{n_m,m}(s_m) \leq \mu_m^2\|b_m^\delta\|_Y^2$
      $u_{m+1} := u_m + s_m$
      $m := m + 1$
      $b_m^\delta := y^\delta - F(u_m)$
   end while
   $M := m$

Then we can find $s_m \in X^{n_m}$ satisfying (9)*. In particular,* REGINN$^\infty$ *also terminates in this case and the regularization property still holds.*

Proof First, n_m according to (21) is well defined since $\lim_{n\to\infty} \beta(n) = 0$. Besides, since $\|b_m^\delta\|_Y$ is monotonously decreasing in m, we get that n_m is non-decreasing, too. Using $s_m := \arg\min_{s \in X^{n_m}} J_{n_m,m}(s)$, we may compute with (13) and by (19)

$$J_{n_m,m}(s_m) \leq \left(\|b_m^\delta\|_Y\left(\omega + \frac{1+\omega}{\tau}\right) + \|F'(u_m)\left(\mathcal{P}^n(u^+ - u_0) - (u^+ - u_0)\right)\|_Y\right)^2$$

$$+ \mathrm{vol}_d(D)^{2/q_{n_m}}\frac{r_0^2}{\gamma^2}C_\mathcal{P}^2\|b_m^\delta\|_{L^2(D)}^2$$

$$\leq \left(\|b_m^\delta\|_Y\left(\omega + \frac{1+\omega}{\tau}\right) + \underbrace{C^+\|u^+ - u_0\|_X}_{<1}\underbrace{\beta(n_m)}_{\leq\varepsilon\|b_m^\delta\|_Y}\right)^2$$

$$+ \max\{\mathrm{vol}_d(D)^{2/q_{n_0}}, 1\}C_\mathcal{P}^2\frac{r_0^2}{\gamma^2}\|b_m^\delta\|_{L^2(D)}^2$$

$$\leq (\mu_{\min}^\varepsilon)^2\|b_m^\delta\|_{L^2(D)}^2.$$

The fact that REGINN$^\infty$ still terminates and also admits the regularization property follows by Theorems 2.1, 2.3 and $\|u_0 - u^+\|_{L^\infty(D)^\ell} \leq C_X\|u_0 - u^+\|_X < r_0$. $\qquad\square$

For convenience, we restate REGINN$^\infty$ in Algorithm 2 subject to $u^+ - u_0 \in X$ for which we need to provide ε and β as additional input. This version is especially of interest if the regularity $u^+ \in X$ is known a priori so that $u_0 \in X$ ensures $u^+ - u_0 \in X$, as desired.

3 Applications: parameter identification tasks in PDEs

In this section we will verify our abstract assumptions (5) and (6) in the concrete settings of electrical impedance tomography and time domain full waveform inversion

(FWI) in the visco-elastic regime which is the state-of-the-art imaging modality in exploration geophysics, see, e.g., [2, 3]. Both inverse problems are locally illposed. We will even show that (19) is satisfied in the acoustic regime. Moreover, our results for FWI carry over to parameter identification tasks for other hyperbolic evolution equations, such as an inverse problem in electromagnetic scattering.

For all applications we rely on the discrete B-spline spaces constructed in Appendix A.1.

3.1 Electrical impedance tomography under the complete electrode model

Electrical impedance tomography (EIT) entails the determination of the electric conductivity distribution of an object by applying electric currents at the boundary through electrodes and measuring the resulting voltages at the boundary as well. Potential applications are, for instance, medical imaging and non-destructive testing.

Let $\sigma \colon D \to [\sigma_{\min}, \infty)$, $\sigma_{\min} > 0$, be the searched-for conductivity distribution in the simply connected Lipschitz-domain $D \subset \mathbb{R}^2$. Further, the p electrodes are denoted by $E_1, \ldots, E_p$ and are assumed to be open subsets of ∂D, the boundary of D, having positive surface measure: $|E_j| > 0$, $j = 1, \ldots, p$. Moreover, let the electrodes be connected and separated: $\overline{E}_i \cap \overline{E}_j = \emptyset$, $i \neq j$. To this electrode configuration we associate the electrode space

$$\mathscr{E}_p := \mathrm{span}\{\chi_{E_1}, \ldots, \chi_{E_p}\} \cap L^2_\diamond(\partial D) \subset L^2_\diamond(\partial D)$$

where χ_{E_i} is the indicator function of the ith electrode and $L^2_\diamond(\partial D)$ is the space of L^2-functions on the boundary of D having vanishing mean.

The forward problem of impedance tomography under the *complete electrode model* (CEM) in the weak formulation is based on the bilinear form $a \colon Y_p \times Y_p \to \mathbb{R}$, $Y_p := H^1(D) \oplus \mathscr{E}_p$,

$$a\big((v, V), (w, W)\big) := \int_D \sigma \nabla v \cdot \nabla w \, \mathrm{d}x + \frac{1}{z} \int_{\partial D} (v - V)(w - W) \, \mathrm{d}s$$

Given an electrode (mean) current $I \in \mathscr{E}_p$ and a contact impedance $z > 0$, find a voltage potential $u \in H^1(D)$ and an electrode voltage $U \in \mathscr{E}_p$ such that

$$a\big((u, U), (w, W)\big) = \int_{\partial D} I \, W \mathrm{d}s \quad \text{for all } (w, W) \in Y_p. \tag{22}$$

The vanishing mean of I and U can be interpreted as conservation of charge and grounding the potential, respectively. CEM (22) is the most accurate mathematical model for EIT currently in use and has a unique solution for $\sigma \in L^\infty_+(D)$, where

$$L^\infty_+(D) := \{\sigma \in L^\infty(D) \colon \sigma \geq \sigma_{\min} > 0 \text{ a.e. in } D\}.$$

The latter follows from the Lax-Milgram theorem, see [21].

The nonlinear forward operator F describing CEM is

$$F\colon L^\infty_+(D) \subset L^\infty(D) \to \mathcal{L}(\mathscr{E}_p), \quad \sigma \mapsto \{I \mapsto U\}.$$

In other words: $F(\sigma)$ maps the current I to the voltage U of the solution of (22). Its Frechét derivative $F'(\sigma) \in \mathcal{L}\big(L^\infty(D), \mathcal{L}(\mathscr{E}_p)\big)$ is given by

$$F'(\sigma)[h]I = U'$$

where $(u', U') \in Y_p$ uniquely solves

$$a\big((u', U'), (w, W)\big) = -\int_D h\nabla u(I) \cdot \nabla w \mathrm{d}x \quad \text{for all } (w, W) \in Y_p \qquad (23)$$

with $u = u(I)$ being the first component of the solution of (22) with respect to the current I, see, e.g., [22]. We will similarly use $U(I)$ for the second component of the solution of (22) and $u'(I)$ for the first component of (23).

We equip $\mathcal{L}(\mathscr{E}_p)$ with the Hilbert-Schmidt inner product: $\langle \Lambda, \Gamma \rangle_{\mathrm{HS}} = \sum_{j=1}^{p-1} \langle \Lambda I_j, \Gamma I_j \rangle_{L^2(\partial D)}$ where $\{I_1, \dots, I_{p-1}\}$ is an orthonormal basis of $\mathscr{E}_p$. This inner product is independent of the chosen orthonormal basis.

Lemma 3.1 *The adjoint operator* $F'(\sigma)^*\colon \mathcal{L}(\mathscr{E}_p)^* \to L^\infty(D)^*$ *is given by*

$$F'(\sigma)^*\Gamma = -\sum_{j=1}^{p-1} \nabla u(\Gamma I_j) \cdot \nabla u(I_j)$$

where $u(\Gamma I_j)$ *and* $u(I_j)$ *denote the solutions of* (22) *with respect to* $I = \Gamma I_j$ *and* $I = I_j$, *respectively. As a sum of products of two L^2-functions, $F'(\sigma)^*\Gamma$ is even in $L^1(D) \subset L^\infty(D)^*$, so that (6) holds.*

Proof Set $\Lambda = F'(\sigma)[h]$, $\Lambda_j = \Lambda I_j$, and $\Gamma_j = \Gamma I_j$ for one $\Gamma \in \mathcal{L}(\mathscr{E}_p)$. Then,

$$\langle \Lambda, \Gamma \rangle_{\mathrm{HS}} = \sum_{j=1}^{p-1} \langle \Lambda_j, \Gamma_j \rangle_{L^2(\partial D)} \overset{(22)}{=} \sum_{j=1}^{p-1} a\big((u(\Gamma_j), U(\Gamma_j)), (u'(I_j), \Lambda_j)\big)$$

$$\overset{(23)}{=} -\int_{\partial D} h \sum_{j=1}^{p-1} \nabla u(\Gamma_j) \cdot \nabla u(I_j)\, \mathrm{d}s = \langle h, F'(\sigma)^*\Gamma \rangle_{L^\infty(D) \times L^\infty(D)^*}$$

which settles the argument. $\qquad\qquad\square$

Lemma 3.2 *Let* $X^n = X_N^n(D)$, $n \in \mathbb{N}$, *be the tensor product spline space of Appendix A.1 with projector* $\mathcal{P}^n$ *defined in* (A5). *Then, for any* $\sigma \in L^\infty_+(D)$,

$$\lim_{n \to \infty} \big\| F'(\sigma)[h - \mathcal{P}^n h] \big\|_{\mathcal{L}(\mathscr{E}_p)} = 0 \quad \text{for any } h \in L^\infty(D).$$

Proof We start with

$$\left\| F'(\sigma)[h - \mathcal{P}^n h] \right\|^2_{\mathcal{L}(\mathscr{E}_p)} = \sum_{j=1}^{p-1} \left\| F'(\sigma)[h - \mathcal{P}^n h] I_j \right\|^2_{L^2(\partial D)}.$$

The bilinear form a is bounded and elliptic on Y_p. Hence, it follows from (23) that

$$\left\| F'(\sigma)[h - \mathcal{P}^n h] I_j \right\|_{L^2(\partial D)} \lesssim \left\| (h - \mathcal{P}^n h) |\nabla u(I_j)| \right\|_{L^2(D)}.$$

Since $\mathcal{P}^n h \xrightarrow{n \to \infty} h$ pointwise a.e. (Appendix B), the norm on the right tends to zero as $n \to \infty$ by the dominated convergence theorem.   □

So we have validated (5) and (6) for the forward operator of EIT. Moreover, TCC (3) holds for the semi-discrete version $F_{n_{\max}}$ of F (Remark 2.5) provided the number of electrodes is sufficiently large [22, Theorem 4.5].

3.2 Full waveform inversion

Wave propagation in realistic media can be modeled by a visco-elastic wave equation which accounts for dispersion and attenuation [2, Chapter 5]. Here, we consider the formulation introduced in [23]: Let $D \subset \mathbb{R}^3$ be a Lipschitz domain. Using $L \in \mathbb{N}$ damping tensors $\boldsymbol{\sigma}_l \colon [0, T] \times D \to \mathbb{R}^{3 \times 3}_{\text{sym}}$, $l = 1, \ldots, L$, the evolution of the particle velocity field $\mathbf{v} \colon [0, T] \times D \to \mathbb{R}^3$ and stress tensor $\boldsymbol{\sigma}_0 \colon [0, T] \times D \to \mathbb{R}^{3 \times 3}_{\text{sym}}$ reads

$$\rho \, \partial_t \mathbf{v} = \operatorname{div} \left(\sum_{l=0}^{L} \boldsymbol{\sigma}_l \right) + \boldsymbol{f} \qquad \text{in } (0, T) \times D, \qquad (24\text{a})$$

$$\partial_t \boldsymbol{\sigma}_0 = C\big(\mu, \pi\big) \, \varepsilon(\mathbf{v}) \qquad \text{in } (0, T) \times D, \qquad (24\text{b})$$

$$\tau_{\sigma, l} \, \partial_t \boldsymbol{\sigma}_l = C\big(\tau_{\text{s}} \mu, \tau_{\text{p}} \pi\big) \, \varepsilon(\mathbf{v}) - \boldsymbol{\sigma}_l, \qquad l = 1, \ldots, L, \qquad \text{in } (0, T) \times D, \qquad (24\text{c})$$

with zero initial conditions[2]

$$\mathbf{v}(0) = \mathbf{0} \ \text{ and } \ \boldsymbol{\sigma}_l(0) = \mathbf{0}, \ l = 0, \ldots, L. \qquad (24\text{d})$$

Above, $\boldsymbol{f} \colon (0, T) \times D \to \mathbb{R}^3$ is the volume force, which initiates the wave, and the functions $\tau_{\text{p}}, \tau_{\text{s}} \colon D \to \mathbb{R}$ are scaling factors for the unrelaxed bulk modulus $\pi \colon D \to \mathbb{R}$ and shear modulus $\mu \colon D \to \mathbb{R}$, respectively. Further, $\rho \colon D \to \mathbb{R}$ is the mass density. The linear map $C(m, p)$, $m, p \in \mathbb{R}$, is Hooke's tensor

$$C(m, p) \colon \mathbb{R}^{3 \times 3} \to \mathbb{R}^{3 \times 3}, \qquad C(m, p)\mathbf{M} = 2m\mathbf{M} + (p - 2m) \operatorname{tr}(\mathbf{M})\mathbf{I},$$

---

[2] Zero initial conditions are not a general restriction. We use them here only to ease the presentation.

where $\mathbf{I} \in \mathbb{R}^{3\times3}$ is the identity matrix and tr($\mathbf{M}$) denotes the trace of $\mathbf{M} \in \mathbb{R}^{3\times3}$. Finally,

$$\boldsymbol{\varepsilon}(\mathbf{v}) = \frac{1}{2}\big[(\nabla_x \mathbf{v})^\top + \nabla_x \mathbf{v}\big]$$

is the linearized strain rate.

Wave propagation is frequency-dependent and the numbers $\tau_{\boldsymbol{\sigma},l} > 0, l = 1, \ldots, L$, model this dependency about the center frequency ω_0, see [24, 25]. Introducing the frequency-dependent phase velocities of P- and S-waves,

$$v_{\mathrm{P}}^2 = \frac{\pi}{\rho}(1 + \tau_{\mathrm{P}}\alpha) \text{ and } v_{\mathrm{S}}^2 = \frac{\mu}{\rho}(1 + \tau_{\mathrm{S}}\alpha) \text{ with } \alpha = \sum_{l=1}^{L} \frac{\omega_0^2 \tau_{\boldsymbol{\sigma},l}^2}{1 + \omega_0^2 \tau_{\boldsymbol{\sigma},l}^2},$$

full waveform inversion entails the identification of the five spatially dependent parameters

$$u = (\rho, v_{\mathrm{S}}, \tau_{\mathrm{S}}, v_{\mathrm{P}}, \tau_{\mathrm{P}})$$

from partial wave field measurements. For a physically meaningful open subset[3] $\mathcal{D}(F) \subset L^\infty(D)^5$ the *full waveform forward operator*

$$F \colon \mathcal{D}(F) \subset L^\infty(D)^5 \to L^2((0,T), H), \quad u \mapsto y := (\mathbf{v}, \boldsymbol{\sigma}_0, \ldots, \boldsymbol{\sigma}_L),$$

is well defined with

$$H = L^2(D, \mathbb{R}^3) \times L^2(D, \mathbb{R}_{\mathrm{sym}}^{3\times3})^{1+L}$$

where $y \in C^1([0,T], H)$ is the unique classical solution of (24) with source $f \in W^{1,1}((0,T), L^2(D, \mathbb{R}^3))$, see [26] (of course, y satisfies boundary conditions, whose concrete formulations we have omitted for simplicity).

Remark 3.3 Actually, the operator modeling seismic imaging is $\Phi = M \circ F$ where $M \colon L^2((0,T), H) \to S$ denotes the measurement operator (S is the space of seismograms). We can reasonably assume that M is linear and bounded. If F satisfies (5) and (6), so does Φ.

---

[3] That is, each of the 5 parameters is restricted from below and from above by physically meaningful bounds. Consult (35) for a concrete example.

For suitable $w = (\mathbf{w}, \boldsymbol{\psi}_0, \ldots, \boldsymbol{\psi}_L) \in H$ we define operators A, B, and Q mapping into H by

$$Aw = -\begin{pmatrix} \mathrm{div}\left(\sum_{l=0}^{L} \boldsymbol{\psi}_l\right) \\ \varepsilon(\mathbf{w}) \\ \vdots \\ \varepsilon(\mathbf{w}) \end{pmatrix}, \quad B^{-1}w = \begin{pmatrix} \frac{1}{\rho}\mathbf{w} \\ C(\mu, \pi)\boldsymbol{\psi}_0 \\ C(\tau_\mathrm{s}\mu, \tau_\mathrm{p}\pi)\boldsymbol{\psi}_1 \\ \vdots \\ C(\tau_\mathrm{s}\mu, \tau_\mathrm{p}\pi)\boldsymbol{\psi}_L \end{pmatrix}, \quad Qw = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \frac{1}{\tau_{\sigma,1}}\boldsymbol{\psi}_1 \\ \vdots \\ \frac{1}{\tau_{\sigma,L}}\boldsymbol{\psi}_L \end{pmatrix},$$

so that (24) collapses to

$$\begin{aligned} B\partial_t y(t) + Ay(t) + BQy(t) &= f(t) \quad \text{in } (0, T), \\ y(0) &= 0. \end{aligned} \tag{25}$$

Note that only $B \colon \mathcal{D}(F) \subset L^\infty(D)^5 \to \mathcal{L}(H)$ depends on the parameters to be identified: $B = B(u)$. It is Fréchet-differentiable with

$$B'(u)[\widehat{u}]\begin{pmatrix} \mathbf{w} \\ \boldsymbol{\psi}_0 \\ \vdots \\ \boldsymbol{\psi}_L \end{pmatrix} = \begin{pmatrix} \widehat{\rho}\,\mathbf{w} \\ -\dfrac{\widehat{\rho}}{\rho}\widetilde{C}(\mu, \pi)\boldsymbol{\psi}_0 + \rho\,\widetilde{C}'(\mu, \pi)\begin{bmatrix} \widetilde{\mu} \\ \widetilde{\pi} \end{bmatrix}\boldsymbol{\psi}_0 \\ -\dfrac{\widehat{\rho}}{\rho}\widetilde{C}(\tau_\mathrm{s}\mu, \tau_\mathrm{p}\pi)\boldsymbol{\psi}_1 + \rho\,\widetilde{C}'(\tau_\mathrm{s}\mu, \tau_\mathrm{p}\pi)\begin{bmatrix} \widehat{\mu} \\ \widehat{\pi} \end{bmatrix}\boldsymbol{\psi}_1 \\ \vdots \\ -\dfrac{\widehat{\rho}}{\rho}\widetilde{C}(\tau_\mathrm{s}\mu, \tau_\mathrm{p}\pi)\boldsymbol{\psi}_L + \rho\,\widetilde{C}'(\tau_\mathrm{s}\mu, \tau_\mathrm{p}\pi)\begin{bmatrix} \widehat{\mu} \\ \widehat{\pi} \end{bmatrix}\boldsymbol{\psi}_L \end{pmatrix} \tag{26}$$

where $\widehat{u} = (\widehat{\rho}, \widehat{v}_\mathrm{s}, \widehat{\tau}_\mathrm{s}, \widehat{v}_\mathrm{p}, \widehat{\tau}_\mathrm{p})$ and

$$\widetilde{\mu} = \frac{2v_\mathrm{s}}{1 + \tau_\mathrm{s}\alpha}\widehat{v}_\mathrm{s} - \frac{\alpha v_\mathrm{s}^2}{(1 + \tau_\mathrm{s}\alpha)^2}\widehat{\tau}_\mathrm{s}, \qquad \widetilde{\pi} = \frac{2v_\mathrm{p}}{1 + \tau_\mathrm{p}\alpha}\widehat{v}_\mathrm{p} - \frac{\alpha v_\mathrm{p}^2}{(1 + \tau_\mathrm{p}\alpha)^2}\widehat{\tau}_\mathrm{p},$$

$$\widehat{\mu} = \frac{2\tau_\mathrm{s} v_\mathrm{s}}{1 + \tau_\mathrm{s}\alpha}\widehat{v}_\mathrm{s} + \frac{v_\mathrm{s}^2}{(1 + \tau_\mathrm{s}\alpha)^2}\widehat{\tau}_\mathrm{s}, \qquad \widehat{\pi} = \frac{2\tau_\mathrm{p} v_\mathrm{p}}{1 + \tau_\mathrm{p}\alpha}\widehat{v}_\mathrm{p} + \frac{v_\mathrm{p}^2}{(1 + \tau_\mathrm{p}\alpha)^2}\widehat{\tau}_\mathrm{p}.$$

Further,

$$\widetilde{C}(m, p) := C(m, p)^{-1} = C\left(\frac{1}{4m}, \frac{p - m}{m(3p - 4m)}\right).$$

and

$$\widetilde{C}'(m, p) \begin{bmatrix} \widehat{m} \\ \widehat{p} \end{bmatrix} = -\widetilde{C}(m, p) \circ C(\widehat{m}, \widehat{p}) \circ \widetilde{C}(m, p). \tag{27}$$

Both, $\widetilde{C}(\mu, \pi)$ and $\widetilde{C}(\tau_s\mu, \tau_p\pi)$ are well defined for $u \in \mathcal{D}(F)$, see [26, Section 4.1] for the involved definition of $\mathcal{D}(F)$.

Lemma 3.4 *Let $\{\widehat{u}_n\}_{n\in\mathbb{N}} \subset L^\infty(D)^5$ be bounded and be convergent to $\widehat{u} \in L^\infty(D)^5$ pointwise a.e. Then, for any $u \in \mathcal{D}(F)$,*

$$\lim_{n\to\infty} \left\| B'(u)[\widehat{u}_n - \widehat{u}]h \right\|_H = 0 \ \ \text{for any } h \in H. \tag{28}$$

Proof The L^2-space H has inner product

$$\big\langle (\mathbf{v}, \boldsymbol{\sigma}_0, \dots, \boldsymbol{\sigma}_L), (\mathbf{w}, \boldsymbol{\psi}_0, \dots, \boldsymbol{\psi}_l) \big\rangle_H = \int_D \left(\mathbf{v} \cdot \mathbf{w} + \sum_{l=0}^{L} \boldsymbol{\sigma}_l : \boldsymbol{\psi}_l \right) \mathrm{d}x$$

where the colon indicates the Frobenius inner product on $\mathbb{R}^{3\times3}$. Now, in view of (26) and (27) we see that the integrand of $\|B'(u)[\widehat{u}_n - u]h\|_H^2$ converges to zero pointwise a.e. Furthermore, as $\{\widehat{u}_n\}_{n\in\mathbb{N}}$ is bounded in $L^\infty(D)^5$, the absolute value of the integrand is bounded by an integrable function (as a matter of fact, the integrand is the sum of products of two L^2-functions with an L^∞-function). The assertions follows from the dominated convergence theorem. $\square$

For completion, we quote a result from [26, Theorem 4.4] which we will need below.

Lemma 3.5 *Under the assumptions from above, F is Fréchet-differentiable at any $u \in \mathcal{D}(F)$. For $\widehat{u} \in L^\infty(D)^5$ we have $\overline{y} := F'(u)\widehat{u} \in C([0, T], H)$ given as the unique weak solution of*

$$B\partial_t \overline{y}(t) + A\overline{y}(t) + BQ\overline{y}(t) = -B'(u)[\widehat{u}] \left(\partial_t y(t) + Qy(t)\right),$$
$$\overline{y}(0) = 0,$$

with $y = F(u)$. Further, for any $t \in [0, T]$,

$$\|\overline{y}\|_{C([0,t],H)} \le C \|B'(u)[\widehat{u}](\partial_t y + Qy)\|_{L^1((0,t),H)}, \tag{29}$$

where C depends continuously on the operator norms of B, B^{-1}, Q, and on T.

We define

$$X^n = X_N^n \times \cdots \times X_N^n \quad \text{(5 factors)} \tag{30}$$

as well as

$$\mathcal{P}^n \colon L^\infty(D)^5 \to X^n, \quad \mathcal{P}^n u = (\mathcal{P}^n_N u_1, \ldots, \mathcal{P}^n_N u_5),$$

where X^n_N and $\mathcal{P}^n_N$ are given by the cardinal B-spline spaces in (A3) and (A5), respectively. We point out that also heterogeneous choices for the factors in X^n are possible as long as X^n keeps nested with respect to n. The convergence properties of $\{\mathcal{P}^n\}_n$, see Appendix B, then guarantee (5) according to the next lemma.

Lemma 3.6 *Let X^n and $\mathcal{P}^n$ be as above. For the full waveform forward operator F with $f \in W^{1,1}((0,T), L^2(D, \mathbb{R}^3))$ we have that*

$$\lim_{n \to \infty} \|F'(u)(\widehat{u} - \mathcal{P}^n \widehat{u})\|_{L^2((0,T),H)} = 0$$

for any $u \in \mathcal{D}(F)$ and all $\widehat{u} \in L^\infty(D)^5$.

Proof The function $\overline{y}_n := F'(u)(\widehat{u} - \mathcal{P}^n \widehat{u})$ is the unique weak solution of

$$B\partial_t \overline{y}_n(t) + A\overline{y}_n(t) + BQ\overline{y}_n(t) = -B'(u)[\widehat{u} - \mathcal{P}^n \widehat{u}](\partial_t y(t) + Qy(t)),$$
$$\overline{y}_n(0) = 0,$$

which according to (29) satisfies

$$\|F'(u)(\widehat{u} - \mathcal{P}^n \widehat{u})\|_{C([0,T],H)} \leq C\|[B'(u)[\widehat{u} - \mathcal{P}^n \widehat{u}](\partial_t y + Qy)\|_{L^1((0,T),H)}.$$

Since $C([0,T], H) \hookrightarrow L^2((0,T), H)$ is continuous, the assertion of the theorem follows if we can show that the right-hand side of the above stability estimate goes to zero as $n \to \infty$. Applying Proposition B.1 componentwise, we deduce that $\mathcal{P}^{n_k} \widehat{u} \to \widehat{u}$ pointwise a.e. for a subsequence $\{n_k\}_{k \in \mathbb{N}}$. Using $\|\mathcal{P}^{n_k} \widehat{u}\|_{L^\infty(D)^5} \leq C_\mathcal{P} \|\widehat{u}\|_{L^\infty(D)^5}$ and $\partial_t y + Qy \in C([0,T], H)$, we see that (26) corresponds to linear combinations of uniformly bounded functions in t for all k. In particular, we can find $g \in L^1(0,T)$ such that

$$\sup_k \|[B'(u)[\widehat{u} - \mathcal{P}^{n_k} \widehat{u}](\partial_t y + Qy)(t)\|_H \leq g(t).$$

By (28) we can then apply the dominated convergence theorem for integration in time which yields $\|B'(u)[\widehat{u} - \mathcal{P}^{n_k} \widehat{u}](\partial_t y + Qy)\|_{L^1((0,T),H)} \to 0$ as $k \to \infty$. By uniqueness of the pointwise limit $\widehat{u}$ the latter convergence even holds for the whole sequence, see [20, Prop. 10.13(2)]. □

Condition (6), which guarantees the regularization property (Theorem 2.3), was proven for the visco-elastic case (24) in [26, Theorem 4.5 and Remark 4.6]. Unfortunately, the TCC, which is the remaining condition for the rigorous applicability of REGINN$^\infty$, is subject of current research in the context of FWI and only special cases are known to hold. For example, in [6] the TCC has recently been shown for a semi-discrete setting in the acoustic regime, cf. (31) below.

Remark 3.7 The findings for the full waveform forward operator in this subsection, that is, (5) and (6) hold, carry over to parameter identification tasks for other hyperbolic evolution equations of the form (25). Examples are the parameter-to-state maps with respect to the acoustic and elastic wave equations (which are in fact simplifications of the visco-elastic model). A further example is the Maxwell system where the components of $y = (\mathbf{E}, \mathbf{H})$ are the electric and magnetic fields, respectively. The involved operators are

$$A = \begin{pmatrix} \sigma \, \mathrm{Id} & -\mathrm{curl} \\ \mathrm{curl} & 0 \end{pmatrix}, \quad B = \begin{pmatrix} \varepsilon \, \mathrm{Id} & 0 \\ 0 & \mu \, \mathrm{Id} \end{pmatrix}, \quad Q = 0, \text{ and } f = \begin{pmatrix} -\mathbf{J}_{\mathrm{e}} \\ \mathbf{J}_{\mathrm{m}} \end{pmatrix}$$

where $\varepsilon, \mu, \sigma \colon D \to \mathbb{R}$ are the (electric) permittivity, the (magnetic) permeability, and the conductivity. Further, $\mathbf{J}_{\mathrm{e}}, \mathbf{J}_{\mathrm{m}} \colon [0, \infty) \times D \to \mathbb{R}^3$ are the current and magnetic densities. See [27, Section 5] for the details to verify (5) and (6) for the map $F \colon (\varepsilon, \mu) \mapsto y$.

3.3 A stronger result for FWI in the acoustic regime

Setting $\mu = 0$, $\tau_{\sigma, l} = 0$, $\tau_{\mathrm{P}} = \tau_{\mathrm{S}} = 0$ in (24) and introducing the hydrostatic pressure $p = \mathrm{tr}(\boldsymbol{\sigma}_0)/3$, we can formally derive the acoustic wave equation as first order system,

$$\rho \, \partial_t \mathbf{v} - \nabla p = \boldsymbol{f}_1 \quad \text{in } (0, T) \times D, \tag{31a}$$

$$\frac{1}{\rho v^2} \, \partial_t p - \mathrm{div}\, \mathbf{v} = f_2 \quad \text{in } (0, T) \times D, \tag{31b}$$

where we allow two source terms and where $v = v_{\mathrm{P}}$ is the compression wave speed. In our abstract formulation (25), it is represented by $y = (\mathbf{v}, p) \in L^2((0, T), H)$, $H = L^2(D, \mathbb{R}^d) \times L^2(D)$, and the operators

$$A = - \begin{pmatrix} 0 & \nabla \\ \mathrm{div} & 0 \end{pmatrix}, \quad B = \begin{pmatrix} \rho \, \mathrm{Id} & 0 \\ 0 & \frac{1}{\rho v^2} \mathrm{Id} \end{pmatrix}, \quad Q = 0. \tag{32}$$

Further, A is defined on

$$\mathcal{D}(A) = \left\{ (\mathbf{v}, p) \in L^2(D, \mathbb{R}^d) \times H_0^1(D) : \mathrm{div}\, \mathbf{v} \in L^2(D) \right\}. \tag{33}$$

If $\boldsymbol{f} = (f_1, f_2) \in W^{1,1}((0, T), H)$ then (31) has a unique classical solution in $C^1([0, T], H) \cap C([0, T], \mathcal{D}(A))$, see [27] (recall that we have zero initial conditions (24d)).

The parameter-to-state map here is

$$F \colon \mathcal{D}(F) \subset L^\infty(D)^2 \to L^2((0, T), H), \quad u = (\rho, v) \mapsto y = (\mathbf{v}, p), \tag{34}$$

with domain of definition

$$\mathcal{D}(F) = \big\{ (\rho, \nu) \in L^\infty(D)^2 : 0 < \rho_{\min} < \rho < \rho_{\max} < \infty \text{ and}$$
$$0 < \nu_{\min} < \nu < \nu_{\max} < \infty \text{ a.e. in } D \big\}. \tag{35}$$

As explained in Remark 3.7, F satisfies (5) and (6). In this part we even verify the stronger compatibility condition (19), that is, for a special choice of X^n and X, see (18), we will specify the decay function β in Theorem 3.8 below. For this purpose, we restrict ourselves to the discretization space $X^n = X_1^n \times X_1^n$ with $X_1^n = X_1^n(D)$ from (A3) with $N = 1$. Thus, X^n consists of piecewise constant functions in the sequel. The associated projector onto X^n then reads

$$\mathcal{P}^n \colon L^\infty(D)^2 \to X^n, \quad \mathcal{P}^n u = (\mathcal{P}_1^n u_1, \mathcal{P}_1^n u_2), \tag{36}$$

whose components are given by

$$\mathcal{P}_1^n u_j = \sum_{k \in \mathcal{I}_n} 2^{nd} \left(\int_{\square_k^n} u_j(x) \, dx \right) \mathbb{1}_{\square_k^n}$$

for $j \in \{1, 2\}$ according to (A5). Here,

$$\square_k^n := 2^{-n}([0, 1]^d + k) \tag{37}$$

is the translated and dilated unit cube. Further, $\mathcal{I}_n = \{ k \in \mathbb{Z}^d : \square_k^n \subset \overline{D} \}$, see (A4).

We are left to determine $X \subset L^\infty(D)^2$ where the subspace X is governed by a stronger topology measuring some kind of smoothness. Intuitively, X should be large enough to still contain a wide class of discontinuous profiles, on the other hand we need some minimal a priori regularity such that its X^n-projections facilitate a common decay rate in (19). For $s > 0$ fixed we set

$$X := L_s^\infty(D)^2$$

whose component spaces are characterized by

$$L_s^\infty(D) := \left\{ w \in L^\infty(D) : \sup_{h \neq 0} \frac{\|w(\cdot - h) - w\|_{L^2(D^h)}}{|h|^s} < \infty \right\} \tag{38}$$

with $D^h := \{ x \in D : x - h \in D \}$ for any $h \in \mathbb{R}^d$. We assign the norm

$$\| \cdot \|_{L_s^\infty(D)} := \| \cdot \|_{L^\infty(D)} + [\cdot]_{B_{2,\infty}^s(D)},$$

where $[\cdot]_{B_{2,\infty}^s(D)}$ is a semi-norm given by the numerical value of the supremum in (38). Originally, $[\cdot]_{B_{2,\infty}^s(D)}$ emerges from the definition of Hilbertian Besov-Nikolskii

spaces

$$B_{2,\infty}^s(D) = \left\{ w \in L^2(D) : \sup_{h \neq 0} \frac{\|w(\cdot - h) - w\|_{L^2(D^h)}}{|h|^s} < \infty \right\}$$

with $\|\cdot\|_{B_{2,\infty}^s(D)} := \|\cdot\|_{L^2(D)} + [\cdot]_{B_{2,\infty}^s(D)}$, see [28]. We set

$$\|u\|_X^2 := \|\rho\|_{L_s^\infty(D)}^2 + \|v\|_{L_s^\infty(D)}^2 \tag{39}$$

and obviously have that $\|u\|_{L^\infty(D)^2} \leq \|u\|_X$ for all $u \in X$.

Theorem 3.8 *Let $D \subset \mathbb{R}^d$, $d \in \{1, 2, 3\}$, be a domain with C^1 boundary and assume that $f = (f_1, f_2) \in W^{2,1}((0, T), H)$ with $f(0) = \partial_t f(0) = \mathbf{0}$ and $f_1 \in C([0, T], L^q(D, \mathbb{R}^d))$ for some $q \in (2, q_{\max})$, where $q_{\max}$ only depends on D and the ratio $\rho_{\max}/\rho_{\min} < \infty$ of the parameters from the definition of $\mathcal{D}(F)$. Let $\mathcal{P}^n$ be as in (36) and let $X = L_s^\infty(D)^2$ be as above for some $0 < s \leq 1/2$. Then, there is a neighborhood U of $u^+ \in \mathcal{D}(F)$ such that*

$$\|F'(u)(\widehat{u} - \mathcal{P}^n \widehat{u})\|_{L^2((0,T),H)} \leq C^+ \|\widehat{u}\|_X \left(2^{-s(q-2)/(4d(q-1))} \right)^n \tag{40}$$

for all $u \in U$ and any $\widehat{u} \in X$.

To prove the theorem, some preparation is required. We start with the observation that $L_s^\infty(D) \subset B_{2,\infty}^s(D)$ and $X_1^n \subset L_s^\infty(D)$ if $s \leq 1/2$. The latter inclusion can be seen as follows: Let $h = (h_1, \dots, h_d)$ with $|h| < 2^{-n}$. Without loss of generality we may assume that $h_i \geq 0$. Now, thanks to the symmetry of the cubes, we estimate

$$\begin{aligned}
\|\mathbb{1}_{-h+\square_k^n} &- \mathbb{1}_{\square_k^n}\|_{L^2(D^h)}^2 \\
&\leq 2\mathrm{vol}_d\left([0, 2^{-n}]^d \setminus [0, 2^{-n} - h_1] \times \cdots \times [0, 2^{-n} - h_d] \right) \\
&= 2\left(2^{-nd} - \prod_{i=1}^d (2^{-n} - h_i) \right) \\
&\leq 2\left(2^{-nd} - (2^{-n} - |h|)^d \right) \leq d\, 2^{-n(d-1)+1} |h|
\end{aligned}$$

for all $k \in \mathcal{I}_n$, where we used the mean value theorem in the last step.

Lemma 3.9 *Let D be a bounded Lipschitz domain and let $\mathcal{P}^n$ be as in (36). Then, for any $u \in X$ and $0 < s \leq 1/2$,*

$$\|u - \mathcal{P}^n u\|_{L^2(D)^2} \leq C \|u\|_X 2^{-ns/d},$$

where C only depends on D and the dimension d.

Proof In view of (39) it suffices to prove the assertion for each component $\mathcal{P}_1^n$ of $\mathcal{P}^n$ and for any $w \in L_s^\infty(D)$. Let $\square \subset \mathbb{R}^d$ be a sufficiently large rectangle containing D. According to [29] there exists $\widetilde{w} \in B_{2,\infty}^s(\square)$ such that $\widetilde{w}|_D = w$ and $\|\widetilde{w}\|_{B_{2,\infty}^s(\square)} \leq \widetilde{C}\|w\|_{B_{2,\infty}^s(D)}$, where $\|w\|_{B_{2,\infty}^s(D)}$ can be replaced by the stronger norm $\|w\|_{L_s^\infty(D)}$. Using a dyadic partition of $\square$ at level n based on the cubes $\{\square_k^n\}_k$ (for which we restrict $\square$ to have integer side length), we have

$$\|\widetilde{w} - \widetilde{\mathcal{P}}_1^n \widetilde{w}\|_{L^2(\square)} \leq [\widetilde{w}]_{B_{2,\infty}^s(\square)} 2^{-ns/d},$$

see [30], where $\widetilde{\mathcal{P}}_1^n$ is defined as in (36) but with respect to the larger index set $\mathcal{I}_n(\square) = \{k \in \mathbb{Z}^d : \square_k^n \subset \overline{\square}\}$. Setting

$$D_n := \bigcup_{k \in \mathcal{I}_n(D)} \square_k^n,$$

we conclude that $\mathcal{P}_1^n w|_{D_n} = \widetilde{\mathcal{P}}_1^n \widetilde{w}|_{D_n}$ and $\mathcal{P}_1^n w|_{D \setminus D_n} = 0$. The latter implies

$$\begin{aligned}\|w - \mathcal{P}_1^n w\|_{L^2(D \setminus D_n)} = \|w\|_{L^2(D \setminus D_n)} &\leq \sqrt{\mathrm{vol}_d(D \setminus D_n)} \|w\|_{L^\infty(D)} \\ &\leq \sqrt{\mathrm{vol}_d(D \setminus D_n)} \|w\|_{L_s^\infty(D)}\end{aligned}$$

for which we can estimate for some $C_D < \infty$

$$\mathrm{vol}(D \setminus D_n) \leq C_D \mathcal{H}^{d-1}(\partial D) 2^{-n}$$

according to the inclusion $D \setminus D_n \subset \cup_{x \in \partial D}(x + 2^{-n}[-1,1]^d)$, where $\mathcal{H}^{d-1}(\partial D)$ denotes the $d-1$-dimensional Hausdorff measure of ∂D. Altogether, we obtain for $s \leq 1/2$ that

$$\begin{aligned}\|w - \mathcal{P}_1^n w\|_{L^2(D)} &\leq \|w - \mathcal{P}_1^n w\|_{L^2(D_n)} + \|w - \mathcal{P}_1^n w\|_{L^2(D \setminus D_n)} \\ &\leq \left(\widetilde{C} 2^{-ns/d} + \sqrt{C_D \mathcal{H}^{d-1}(\partial D)} 2^{-n/2}\right)\|w\|_{L_s^\infty(D)} \\ &\leq C\|w\|_{L_s^\infty(D)} 2^{-ns/d}\end{aligned}$$

which proves the lemma. $\square$

Lemma 3.10 *Under the assumptions of Theorem 3.8, there is a $q_{\max} > 2$ such that for any $u^+ = (\rho^+, v^+) \in \mathcal{D}(F)$ there exists a neighborhood U of u^+ such that*

$$\sup\left\{\|\partial_t y\|_{L^1((0,T),H^q)} : y = (p, \mathbf{v}) = F(u),\ u = (\rho, v) \in U\right\} < \infty$$

for any $q \in (2, q_{\max})$ where

$$H^q := L^q(D) \times L^q(D, \mathbb{R}^d). \tag{41}$$

Here, $q_{\max}$ only depends on D and the ratio $\rho_{\max}/\rho_{\min} < \infty$ of the parameters from the definition of $\mathcal{D}(F)$.

Proof The proof makes use of converting higher time regularity to higher spatial integrability. By Theorem 2.6 of [27], we know that for $u^+ \in \mathcal{D}(F)$ we have $\partial_t y = (\partial_t p, \partial_t \mathbf{v}) \in C([0, T], \mathcal{D}(A)) \cap C^1([0, T], H)$, in particular $\partial_t p \in L^1((0, T), H_0^1(D))$. By Sobolev embedding, see [31], we obtain at least $\partial_t p \in L^1((0, T), L^6(D))$ for $d \in \{1, 2, 3\}$. Again by Theorem 2.6 of [27], since the constant of the stability estimate depends continuously on B and thus on u, we can actually conclude $\|\partial_t p\|_{L^1((0,T),L^6(D))} < \infty$ uniformly in a neighborhood of u^+. Concerning $\partial_t \mathbf{v}$, we only have integrability information about its divergence. Therefore, we first note that by (31) we have

$$-\operatorname{div}\left(\frac{1}{\rho}\nabla p\right) = \operatorname{div}\left(\frac{\boldsymbol{f}_1}{\rho}\right) - \operatorname{div}\partial_t \mathbf{v} = \operatorname{div}\left(\frac{\boldsymbol{f}_1}{\rho}\right) + \left(\partial_t f_2 - \frac{\partial_t^2 p}{\rho \nu^2}\right)$$

in the sense of distributions. As $(\rho, \nu) \in \mathcal{D}(F)$ is bounded away from zero uniformly, the parameters do not affect the integrability of the right-hand side terms. Now Meyers' estimate, see [32, Thm. 1], implies for fixed $t \in [0, T]$ and any $q \in (2, q_{\max})$ that

$$\|\nabla p\|_{L^q(D,\mathbb{R}^d)} \leq C\left(\frac{\|\boldsymbol{f}_1\|_{L^q(D,\mathbb{R}^d)}}{\rho_{\min}} + \|\partial_t f_2\|_{L^2(D)} + \frac{\|\partial_t^2 p\|_{L^2(D)}}{\rho_{\min}\nu_{\min}^2}\right),$$

where $q_{\max}$ depends on D and $\rho_{\max}/\rho_{\min} < \infty$, while C additionally depends on q. Since $\|\partial_t^2 p\|_{L^2(D)}$ can be bounded in terms of $\|\partial_t^2 \boldsymbol{f}\|_{L^1((0,T),H)}$ and a constant which depends continuously on B, that is, on u, see Theorem 2.6 of [27], we concude that $\|\nabla p\|_{L^1((0,T),L^q(D,\mathbb{R}^d))} < \infty$ uniformly in a neighborhood of u^+. Comparing with (31), we finally get that also $\|\partial_t \mathbf{v}\|_{L^1((0,T),L^q(D,\mathbb{R}^d))} < \infty$ locally uniform in u. This completes the proof. □

Finally, we can prove the main result.

Proof of Theorem 3.8 The proof uses a more elaborate analysis of the stability estimate

$$\|F'(u)(\widehat{u} - \mathcal{P}^n\widehat{u})(t)\|_{L^2((0,T),H)} \leq C\|B'(u)(\widehat{u} - \mathcal{P}^n\widehat{u})\partial_t y\|_{L^1((0,T),H)},$$

compared to Lemma 3.6. Recall that the constant C depends continuously on the operator norms of $B(u)$, $B^{-1}(u)$ and on T according to (29). Since also $B \mapsto B^{-1}$ is (locally) continuous, we can assume without loss of generality that the above inequality holds for some fixed $C = C_{r',u^+} > 0$ for all $\|u - u^+\|_{L^\infty(D)^2} < r'$ by shrinking the original $r' > 0$ otherwise. Then, for any $\delta > 0$,

$$\|B'(u)[\widehat{u} - \mathcal{P}^n\widehat{u}]\partial_t y\|_{L^1((0,T),H)}$$
$$= \left\|B'(u)\left[\frac{\widehat{u} - \mathcal{P}^n\widehat{u}}{\|\widehat{u}\|_X}\right]\left(\partial_t y(\mathbb{1}_{\{|\widehat{u}-\mathcal{P}^n\widehat{u}|\geq\delta\|\widehat{u}\|_X\}} + \mathbb{1}_{\{|\widehat{u}-\mathcal{P}^n\widehat{u}|<\delta\|\widehat{u}\|_X\}})\right)\right\|_{L^1((0,T),H)}$$
$$\times \|\widehat{u}\|_X$$

$$\leq \|B'(u)\|_{\mathcal{L}(L^\infty(D)^2,\mathcal{L}(H))} \frac{\|\widehat{u} - \mathcal{P}^n\widehat{u}\|_{L^\infty(D)^2}}{\|\widehat{u}\|_X} \left\| \partial_t y \mathbb{1}_{\{|\widehat{u} - \mathcal{P}^n\widehat{u}| \geq \delta \|\widehat{u}\|_X\}} \right\|_{L^1((0,T),H)}$$

$$\times \|\widehat{u}\|_X$$

$$+ \|B'(u)\|_{\mathcal{L}(L^\infty(D)^2,\mathcal{L}(H))} \delta \left\| \partial_t y \mathbb{1}_{\{|\widehat{u} - \mathcal{P}^n\widehat{u}| < \delta \|\widehat{u}\|_X\}} \right\|_{L^1((0,T),H)} \|\widehat{u}\|_X.$$

Due to $\|\widehat{u} - \mathcal{P}^n\widehat{u}\|_{L^\infty(D)^2} \leq (1 + C_\mathcal{P})\|\widehat{u}\|_{L^\infty(D)^2} \leq (1 + C_\mathcal{P})\|\widehat{u}\|_X$ by (S2) and $C_X = 1$ in (18), we obtain, by dropping the complementary indicator function in the bottom line above, that

$$\|B'(u)[\widehat{u} - \mathcal{P}^n\widehat{u}]\partial_t y\|_{L^1((0,T),H)}$$

$$\leq (1 + C_\mathcal{P})\|B'(u)\|_{\mathcal{L}(L^\infty(D)^2,\mathcal{L}(H))} \left\| \partial_t y \mathbb{1}_{\{|\widehat{u} - \mathcal{P}^n\widehat{u}| \geq \delta \|\widehat{u}\|_X\}} \right\|_{L^1((0,T),H)} \|\widehat{u}\|_X$$

$$+ \|B'(u)\|_{\mathcal{L}(L^\infty(D)^2,\mathcal{L}(H))} \delta \|\partial_t y\|_{L^1((0,T),H)} \|\widehat{u}\|_X.$$

The middle line here can also be directly expressed in terms of δ. To this end, we apply Hölder's inequality with exponent $q/2 > 1$ in view of (41) to get

$$\left\| \partial_t y (\mathbb{1}_{\{|\widehat{u} - \mathcal{P}^n\widehat{u}| \geq \delta \|\widehat{u}\|_X\}} \right\|_{L^1((0,T),H)}$$

$$= \int_0^T \left(\int_D |\partial_t y(t)|^2(x) \mathbb{1}_{\{|\widehat{u} - \mathcal{P}^n\widehat{u}| \geq \delta \|\widehat{u}\|_X\}}(x) \, dx \right)^{1/2} dt$$

$$\leq \int_0^T \left(\|\partial_t y(t)\|_{H^q}^2 \mathrm{vol}(\{|\widehat{u} - \mathcal{P}^n\widehat{u}| \geq \delta \|\widehat{u}\|_X\})^{(q-2)/q} \right)^{1/2} dt$$

$$= \|\partial_t y\|_{L^1((0,T),H^q)} \mathrm{vol}(\{|\widehat{u} - \mathcal{P}^n\widehat{u}| \geq \delta \|\widehat{u}\|_X\})^{(q-2)/(2q)}$$

$$\leq \|\partial_t y\|_{L^1((0,T),H^q)} \left(\frac{\|\widehat{u} - \mathcal{P}^n\widehat{u}\|_{L^2(D)^2}^2}{\delta^2 \|\widehat{u}\|_X^2} \right)^{(q-2)/(2q)},$$

where we additionally employed Tschebyscheff's inequality in the bottom line, see e.g., [33]. With Lemma 3.9 we can further estimate

$$\frac{\|\widehat{u} - \mathcal{P}^n\widehat{u}\|_{L^2(D)^2}^2}{\delta^2 \|\widehat{u}\|_X^2} \leq \frac{C \|\widehat{u}\|_X^2 (2^{-(s/d)})^n}{\delta^2 \|\widehat{u}\|_X^2} = \frac{C (2^{-(s/d)})^n}{\delta^2}.$$

Altogether, we get with a similar Hölder-inequality argument for

$$\|\partial_t y\|_{L^1((0,T),H)} \leq \|\partial_t y\|_{L^1((0,T),H^q)} \mathrm{vol}_d(D)^{(q-2)/(2q)}$$

that

$$\|B'(u)[\widehat{u} - \mathcal{P}^n\widehat{u}]\partial_t y\|_{L^1((0,T),H)} \leq \|\widehat{u}\|_X \|B'(u)\|_{\mathcal{L}(L^\infty(D)^2,\mathcal{L}(H))} \|\partial_t y\|_{L^1((0,T),H^q)}$$

$$\times \left((1 + C_\mathcal{P}) \left(\frac{C(2^{-(s/d)})^n}{\delta^2} \right)^{(q-2)/(2q)} + \delta \mathrm{vol}_d(D)^{(q-2)/(2q)} \right)$$

which holds for all $\delta > 0$. Optimization in δ then yields $\delta \propto \left(2^{-s(q-2)/(4d(q-1))}\right)^n$. By Lemma 3.10 and the continuity of B', we can indeed find $C^+ < \infty$ such that (40) holds. □

4 Numerical Results

We present numerical experiments[4] on a two-parameter reconstruction to demonstrate the operation of $\texttt{REGINN}^\infty$ in a test scenario where all assumptions required for our analysis in the previous sections are satisfied.

Recall from Theorem 2.3 that, in general, the regularization property holds only in the weak-$\star$ topology permitting a kind of strange convergence behavior. Therefore, we test Algorithm 1 as the noise level approaches zero and also how it behaves under different initial spaces X^{n_0}. We will start with a rather low dimensional X^{n_0} such that n_m increases successively in the course of the Newton iteration (`while`-loop) and in contrast also with some large dimensional X^{n_0} which corresponds to a more static use of Tikhonov regularization throughout all iterations.

Our experiments rely on the acoustic wave equation in one spatial dimension, $d = 1$, where $D = (0, 1)$ and $T = 1$:

$$
\begin{aligned}
\frac{1}{\rho v^2} \, \partial_t p - \partial_x v &= f_1 &&\text{in } (0, 1) \times (0, 1), \\
\rho \, \partial_t v - \partial_x p &= f_2 &&\text{in } (0, 1) \times (0, 1), \\
v(0, \cdot) = p(0, \cdot) &= 0 &&\text{on } (0, 1), \\
p(\cdot, 0) = p(\cdot, 1) &= 0 &&\text{on } (0, 1).
\end{aligned}
\tag{42}
$$

The source components $f_1, f_2 \colon [0, 1] \times [0, 1] \to \mathbb{R}$ are

$$
\begin{aligned}
f_1(t, x) = f_1(x) &= 100\Big(x(x - 1)\frac{1}{\rho(x)v(x)^2} - \frac{\pi}{2} \cos\left(\frac{\pi}{2}x\right)\Big), \\
f_2(t, x) &= 100\Big(-t(2x - 1) + \sin\left(\frac{\pi}{2}x\right)\rho(x)\Big),
\end{aligned}
\tag{43}
$$

where

$$
\rho(x) = 1 + \frac{1}{5}\, \mathbb{1}_{[7/30, 17/30]}(x) \quad \text{and} \quad v(x) = 1 - \frac{1}{10}\, \mathbb{1}_{[13/30, 23/30]}(x).
\tag{44}
$$

The corresponding exact data, that is, the solution of (42) and (43), are given by

$$
p(t, x) = 100tx(x - 1) \quad \text{and} \quad v(t, x) = 100t \sin\left(\frac{\pi}{2}x\right).
\tag{45}
$$

---

[4] For the reader's own experiments we provide our MATLAB code on http://www.math.kit.edu/ianm3/~rieder/media/reginn_infty_fig2.m. Executed in MATLAB (R2021a) on an Intel(R) Core(TM) i5-1035G4 CPU under Windows 10, the code produces the output shown in Fig. 2. In our program we use a routine by John D'Errico (2021): Piecewise functions (https://www.mathworks.com/matlabcentral/fileexchange/9394-piecewise-functions), MATLAB Central File Exchange. Retrieved November 29, 2021.

We solve the appearing wave equations during inversion for the parameters by the FEM-based MATLAB (R2021a) command `pdepe` with 300 spatial and 100 temporal grid points. Both sets of points are distributed equidistantly in [0, 1].

Our discrete parameter spaces $X^n = X_1^n \times X_1^n$ are generated by the piecewise constant cardinal B-spline as explained in Appendix A.1. So, conditions (S1)–(S3) are fulfilled. Note that the dimension of X_1^n is 2^n. In view of Remark 2.5 we set $n_{\max} = 8$ yielding the semi-discrete parameter-to-state map

$$F_{n_{\max}} : \mathcal{D}(F_{n_{\max}}) \subset L^\infty(D)^2 \to L^2([0,1], H), \quad (\rho, \nu) \mapsto (p, v), \qquad (46)$$

where $(p, v) \in L^2([0,1], H)$, $H = L^2(0,1)^2$, solves (42) with (43) and $(\rho, \nu) \in \mathcal{D}(F_{n_{\max}}) = X^{n_{\max}} \cap \mathcal{D}(F)$, see (35) for $\mathcal{D}(F)$. Within our computations, $\|\cdot\|_{L^2([0,1], H)}$ is discretized by the corresponding space-time Euclidean norm and denoted by $\|\cdot\|$. Since $F_{n_{\max}}$ satisfies the TCC (3), see Appendix C, Theorem 2.1 guarantees termination of REGINN$^\infty$ applied to the inverse problem

$$\text{find } (\rho, \nu) \in X^{n_{\max}} : \quad F_{n_{\max}}(\rho, \nu) \approx (p^\delta, v^\delta). \qquad (47)$$

To simplify notation we use the same symbols for the continuous and the discrete versions of functions such as p, v, ρ, ν, etc.

We apply REGINN$^\infty$ (Algorithm 1) to (47) where we choose $q_n = n / \log_2 C_\infty$ for $J_{n,m}$ from (10) in accordance with the lower bound in (A6) below. For each m the computation of Newton update candidates $s_m \in X^n$ is realized—benefiting greatly from the smoothness of the Tikhonov functionals—by a steepest descent routine with Armijo stepsize rule in a loop over n until (9) is met. We adapt μ_m during iteration according to the rule proposed in [4]: we start with $\mu_1 = \mu_0$ and set

$$\mu_m = \begin{cases} \min\{1 - \frac{j_{m-2}}{j_{m-1}}(1 - \mu_{m-1}), 0.999\}, & j_{m-1} \geq j_{m-2}, \\ 0.9\mu_{m-1}, & \text{otherwise}, \end{cases} \quad m \geq 2,$$

where $\mu_0 \in (0, 1)$ is user-supplied and j_m denotes the number of gradient decent steps needed to compute the update s_m. Complementary, the underlying discretization level n will be increased if the gradient descent loop stagnates on X^n, which we consider to occur if the ratio of two successive gradient step evaluations does not exceed a fixed threshold close to 1, say 0.99999. We stop the algorithm either by the discrepancy principle or if $n \geq n_{\max}$ happens, that is, if the discretization of X^n would become finer than the computational grid used in the `pdepe`-routine for solving the wave equation. In the latter case, we still perform $u_{m+1} = u_m + \widetilde{s}_m$ with the last update candidate $\widetilde{s}_m \in X^{n_{\max}}$ before abortion. We emphasize that $\widetilde{s}_m$ is not a Newton update in the sense of (9), but the corresponding u_{m+1} might still fulfill the discrepancy principle unlike u_m.

In our experiments we especially want to detect the jump regions [7/30, 17/30] and [13/30, 23/30], where the parameters differ from the homogeneous background material $(\rho_0, \nu_0) = (1, 1) \in X^{n_{\max}}$, respectively, that we take as initial guess. Note that no grid point of X^n coincides with either of the jump discontinuity points for all n so that
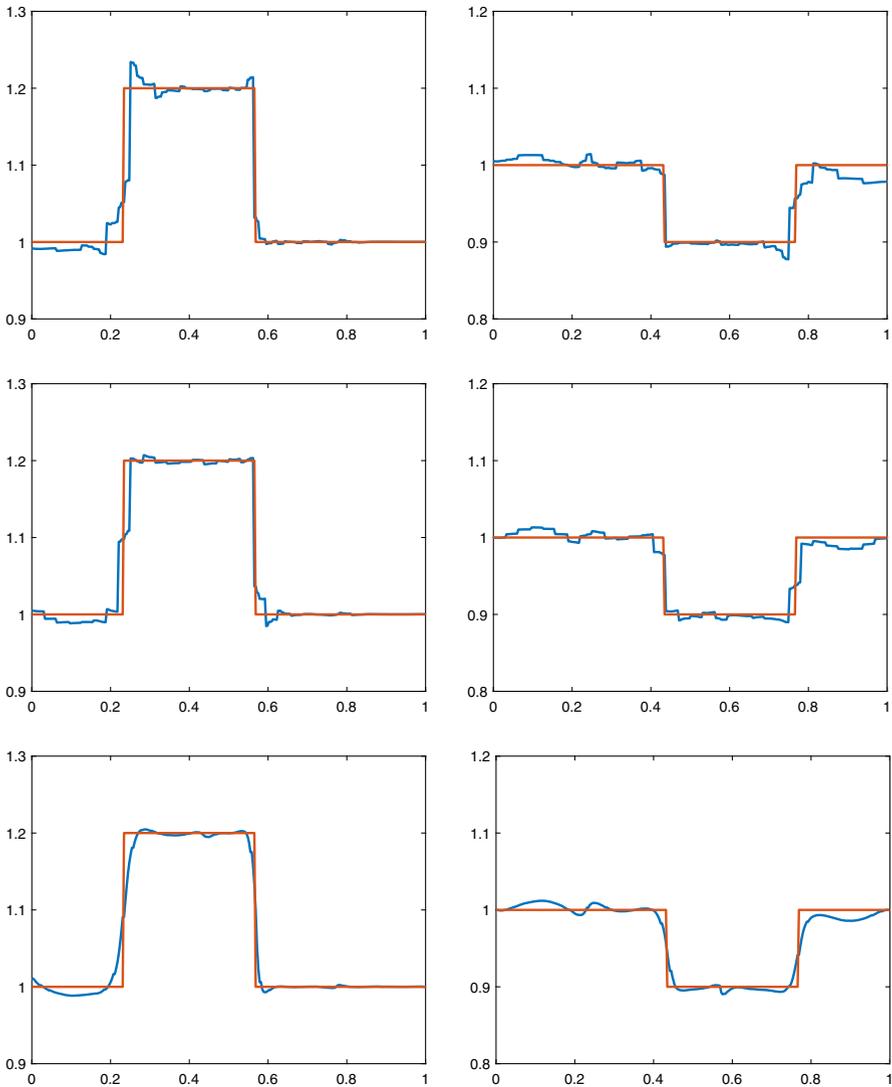
Fig. 1 Approximate solutions ρ_M (blue, left column) and v_M (blue, right column) by Algorithm 1 with initial spaces X^2, X^5, and X^8 (top to bottom) (color figure online)

the error of any reconstruction of ρ and v will always be at least $(\max \rho - \min \rho)/2 = 1/10$ and $(\max v - \min v)/2 = 1/20$ with respect to the L^∞-norm, respectively.

First, we investigate the case of 'exact' data, that is, $(p^\delta, v^\delta) = (p, v)$ with (p, v) from (45). Despite of $\delta = 0$ our data might still be contaminated by some discretization error with respect to $F_{n_{\max}}$ since the corresponding analytical solution (ρ, v) given by (44) is not contained in $X^{n_{\max}}$. Choosing $\mu_0 = 0.7$, $\gamma = 0.8$, $C_\infty = 1.1$, we run Algorithm 1 for different n_0 to observe how its choice affects the outcome. Note that setting τ is redundant here because termination is solely forced by n exceeding $n_{\max}$.

m	n_m	j_m	μ_m	$\|b_m^0\|$
0	2	0	0.7	2.56207
1	2	2	0.7	1.31590
2	2	3	0.8	0.67130
3	2	2	0.72	0.43376
4	2	5	0.888	0.29172
5	2	17	0.96706	0.24329
6	3	39	0.98564	0.20439
7	3	1	0.88708	0.17912
8	3	10	0.98871	0.14127
9	3	8	0.88984	0.13206
10	4	70	0.98741	0.10649
11	4	7	0.88867	0.09944
12	4	10	0.92207	0.08374
13	4	9	0.82986	0.07341
14	4	20	0.92344	0.05562
15	4	16	0.83109	0.04895
16	4	36	0.92493	0.03842
17	4	33	0.83244	0.03383
18	6	327	0.98309	0.02641
19	7	190	0.88478	0.02493
20	8	57	–	0.02442
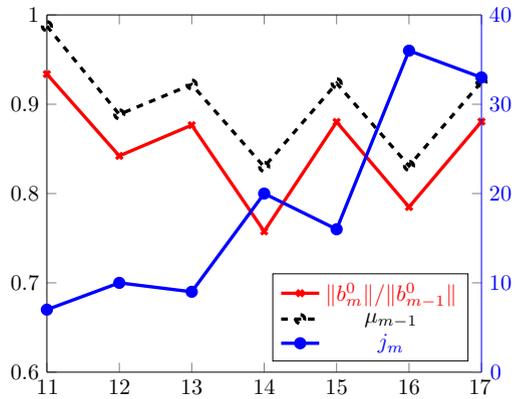
Fig. 2 Left: Convergence history for exact data case $n_0 = 2$ from Fig. 1. Peaks for j_m arise whenever the discretization level n_m is increased as cumulative contribution. Right: Graphical presentation of the values j_m (blue) and μ_{m-1} (black dashed) as functions of $m \in \{11, \ldots, 17\}$ where $n_m = 4$. Moreover, we have included the quotient $\|b_m^0\|/\|b_{m-1}^0\|$ (red) which is always below μ_{m-1}. This illustrates that (15) holds for a tiny ω (color figure online)

Figure 1 displays the exact parameter functions ρ and ν (red) and the corresponding outputs ρ_M and ν_M (blue) of Algorithm 1 when starting with $n_0 \in \{2, 5, 8\}$, respectively. We see that the larger n_0 is, the smoother the output becomes, while the points of discontinuity are more sharply located for smaller n_0. Hence, for the reconstruction of jump discontinuities, n_0 shall be chosen large enough to locate discontinuity points sufficiently precise while at the same time it should not be too large to prevent over-smoothing. Figure 2 shows a more detailed convergence history in the case $n_0 = 2$ and confirms that the majority of Newton steps is indeed undertaken with $n_m \leq 4$.

Next, we study the case of noisy data. For this purpose we generate noise vectors ζ as random samples from a centered Gaussian distribution and scale it such that $\|\zeta\| = \delta \|(p, \nu)\|$. Since δ is a relative perturbation here, the discrepancy principle must be adjusted accordingly. As before, we employ Algorithm 1 with $\mu_0 = 0.7$, $\gamma = 0.8$, $C_\infty = 1.1$, $\tau = 1.1$. Using the insights from our exact data case we set $n_0 = 5$ as initial value to balance the aforementioned effects of globally smooth and locally oscillating reconstructions. The corresponding results for ρ_M and ν_M are shown in Fig. 3 for $\delta = 5\%$, $\delta = 2\%$, and 1%. We see that the reconstructions' profiles approach the correct jump height of the exact solution as δ becomes smaller. In all three cases, termination occurs by reaching the discretization limit, however, each
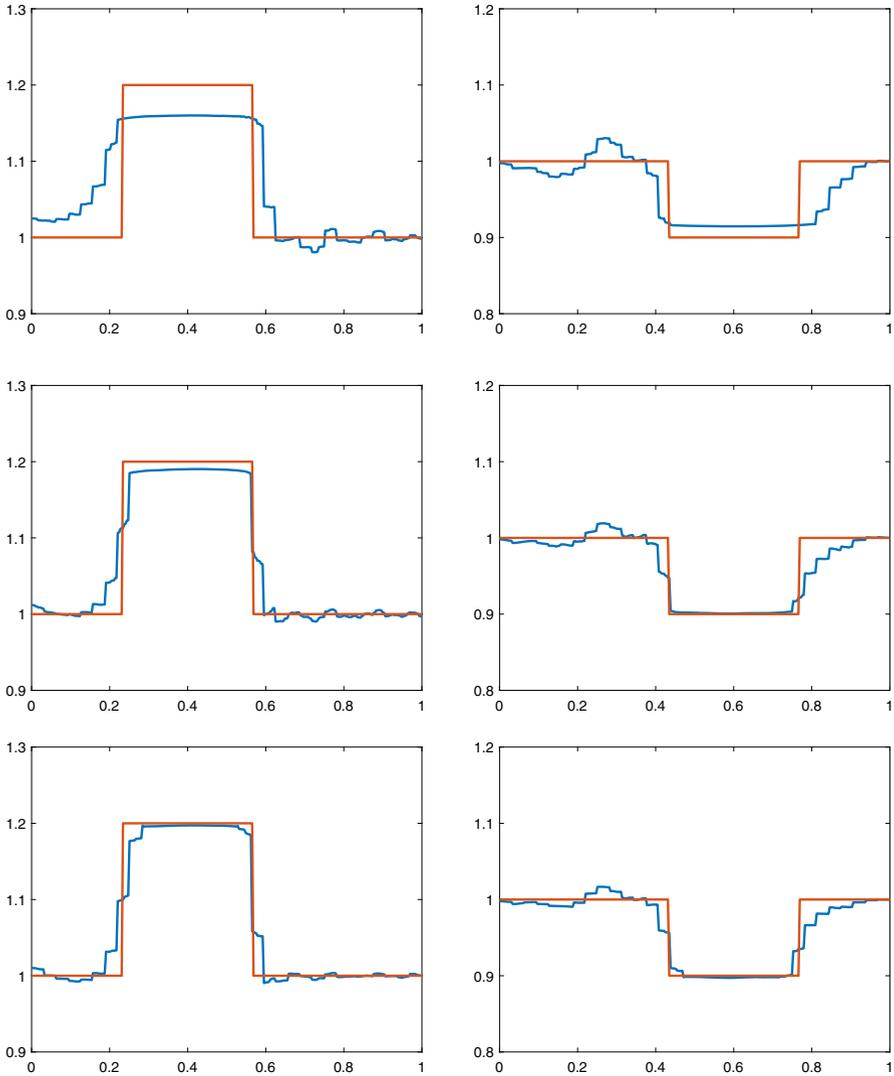
Fig. 3 Regularized solutions ρ_M (blue, left column) and ρ_M (blue, right column) by Algorithm 1 for relative noise levels of 5%, 2% and 1% (top to bottom) (color figure online)

last update fulfills the discrepancy principle afterwards. Altogether, the plots are in agreement with the weak-$\star$ regularization property of REGINN$^\infty$ (Theorem 2.3).

5 Conclusion

We have investigated a novel iterative regularization algorithm tailored for non-linear illposed problems between $L^\infty(D)^\ell$ and a normed space Y. The main focus was

on generating uniformly bounded iterates relying on a Tikhonov-like regularization term. Due to the non-smooth structure of $L^\infty(D)^\ell$, a straightforward implementation would require non-smooth or box-constrained calculus which we could circumvent by using discretization in combination with equivalent $L^p(D)^\ell$-norms for $p < \infty$. Under reasonable assumptions on the input parameters, our algorithm REGINN$^\infty$ terminates after finitely many steps. Further, it fulfills the regularization property in the weak-$\star$ topology as the noise level of the Y-data tends to zero. Depending on the non-linearity, this convergence can be reformulated as convergence with respect to a norm. Numerical experiments with a model problem illustrate the theoretical findings.

Future research may include a convergence rate analysis under higher regularity assumptions as in (19) or under more general variational source conditions with respect to a Bregman distance [13]. We could even incorporate a metric to overcome that $L^\infty(D)^\ell$ is non-separable; an approach proposed in [7]. Concerning the data space, especially the task of finding proper measures for the misfit in seismograms, the Kantorovich-Rubinstein (KR) norm has recently proven advantageous in exploration geophysics, see, e.g., [34, 35]. This fact suggests an implementation of our method under the KR-norm on Y. In fact, our theory also allows more general distance functions on Y. For instance, any distance concept is admissible which is convex in one of its two arguments (e.g. Bregman distances). Such a concept can in particular be learned within a predefined set of admissible distance functions in practice and is subject of ongoing research, see [36].

Declarations

Conflict of interest We hereby confirm that we do not have any conflicts of interest.

Appendix A A family of admissible subspaces

In this appendix we give a concrete construction for a family $\{X^n\}_{n \in \mathbb{N}}$ of subspaces of $L^\infty(D)^\ell$ which satisfies our assumptions (S1)–(S3) of Sect. 2 for D an open and bounded subset of $\mathbb{R}^d$. Using Cartesian products in case of $\ell > 1$, cf. (30), we restrict our attention to $\ell = 1$ here.

A.1 The discrete subspaces used in this work

We will rely on the cardinal B-spline $\varphi_N : \mathbb{R} \to \mathbb{R}$ of order $N \in \mathbb{N}$ which is recursively defined by

$$\varphi_N(t) := \varphi_{N-1} \star \varphi_1(t) = \int_0^1 \varphi_{N-1}(t-s) \, ds, \quad \varphi_1 = \mathbb{1}_{[0,1]}.$$

It obeys the scaling relation

$$\varphi_N(t) = 2^{1-N} \sum_{k=0}^N \binom{N}{k} \varphi_N(2t - k). \tag{A1}$$

Further properties are

- $\operatorname{supp}\varphi_N = [0, N], \quad \varphi_N|_{]0,N[} > 0, \quad \varphi_N \in C^{N-2}$,
- for each $k \in \mathbb{Z}$, $\varphi_N|_{[k,k+1]}$ is a polynomial of degree $N - 1$,
- for all $t \in \mathbb{R}$,

$$1 = \sum_{m \in \mathbb{Z}} \varphi_N(t - m), \tag{A2}$$

see, e.g., [37].

Using the tensor product B-Spline $\Phi(x) := \prod_{i=1}^d \varphi_N(x_i)$, $x = (x_1, \ldots, x_d)^\top \in \mathbb{R}^d$, and the notation $\Phi_{n,k}(x) = 2^{nd/2} \Phi(2^n x - k)$, $n \in \mathbb{N}$, $k \in \mathbb{Z}^d$, we define

$$X^n = X_N^n(D) := \operatorname{span}\{\Phi_{n,k}|_D : k \in \mathcal{I}_n(D)\} \tag{A3}$$

with

$$\mathcal{I}_n = \mathcal{I}_n(D) := \{k \in \mathbb{Z}^d : \operatorname{supp}\Phi_{n,k} \subset \overline{D}\}. \tag{A4}$$

These spaces are nested due to (A1), so that (S1) holds. Note that $\cup_{k \in \mathcal{I}_n} \operatorname{supp}\Phi_{n,k} \subset \overline{D}$ which is a proper inclusion in general.

Next we demonstrate (S2). To this end we set

$$\mathcal{P}^n u = \mathcal{P}_N^n u := \sum_{k \in \mathcal{I}_n} \langle u, \widetilde{\Phi}_{n,k} \rangle_{L^2(D)} \Phi_{n,k} \quad \text{for } u \in L^\infty(D) \tag{A5}$$

where $\widetilde{\Phi}$ is a compactly supported dual function to Φ satisfying the biorthogonality

$$\langle \widetilde{\Phi}(\cdot), \Phi(\cdot - k) \rangle_{L^2(\mathbb{R}^d)} = \delta_{0,k}.$$

The existence of such functions has been shown in [38]. The biorthogonality yields $\mathcal{P}_N^n \Phi_{n,k} = \Phi_{n,k}$, for all $k \in \mathcal{I}_n$. Hence, the required projection property holds:

$\mathcal{P}_N^n u = u$ for all $u \in X_N^n(D)$. We proceed with

$$\|\mathcal{P}_N^n u\|_{L^\infty(D)} = \sup_{x \in D} \Big| \sum_{k \in \mathcal{I}_n} \langle u, \widetilde{\Phi}_{n,k} \rangle_{L^2(D)} \Phi_{n,k}(x) \Big|$$

$$= \Big| \sum_{k \in \mathcal{I}_n} \langle u, \widetilde{\Phi}_{n,k} \rangle_{L^2(D)} \Phi_{n,k}(x^*) \Big|$$

$$\leq \sum_{k \in \mathcal{I}_n} |\langle u, \widetilde{\Phi}_{n,k} \rangle_{L^2(D)}| \, \Phi_{n,k}(x^*)$$

$$\leq \|u\|_{L^\infty(D)} \sum_{k \in \mathcal{I}_n} \|\widetilde{\Phi}_{n,k}\|_{L^1(D)} \, \Phi_{n,k}(x^*).$$

Since

$$\|\widetilde{\Phi}_{n,k}\|_{L^1(D)} \leq 2^{-nd/2} \int_{\mathbb{R}^d} |\widetilde{\Phi}(x)| \, \mathrm{d}x$$

$$\text{and} \quad \sum_{k \in \mathcal{I}_n} \Phi_{n,k}(x^*) \leq \sum_{k \in \mathbb{Z}^d} \Phi_{n,k}(x^*) \overset{(A2)}{=} 2^{nd/2}$$

we have established (S2) with $C_{\mathcal{P}} \leq \|\widetilde{\Phi}\|_{L^1(\mathbb{R}^d)}$. Observe that $\|\widetilde{\Phi}\|_{L^1(\mathbb{R}^d)} \geq 1$ as $\widetilde{\Phi}$ has mean value 1 just as Φ.

It remains to validate (S3). Let $u \in X_N^n(D)$ with $\|u\|_{L^\infty(D)} = 1$. Then,

$$\|u\|_{L^q(D)} \geq \delta_q > 0$$

for

$$\delta_q = \min_{u \in M} \|u\|_{L^q(D)} \quad \text{where} \quad M = \big\{ u \in X_N^n(D) : \|u\|_{L^\infty(D)} = 1 \big\}.$$

This minimum is non-zero and exists as M is compact in the finite dimensional space $X_N^n(D)$. Since $\delta_q = \|u_q\|_{L^q(D)}$ for one $u_q \in M$ and as $\delta_q \to 1$ for $q \to \infty$ (see below), we find a q with $\delta_q \geq 1/C_\infty$ for any fixed $C_\infty > 1$. Hence, $1 \leq C_\infty \|u\|_{L^q(D)}$ for all $u \in M$ and (S3) follows by the homogeneity of norms.

We finish with proving $\lim_{q \to \infty} \delta_q = 1$. Obviously, $\delta_q = \|u_q\|_{L^q(D)} \leq \mathrm{vol}_d(D)^{1/q} \to 1$ as $q \to \infty$. Therefore $\{\delta_q\}$ is bounded and admits a convergent subsequence, say, $\lim_{i \to \infty} \delta_{q_i} = \delta^* \leq 1$. For each q let $x_q^* \in D$ with $|u_q(x_q^*)| = 1$. If $N = 1$, u_q must be equal to unity in a whole cube of length 2^{-n} as a subset of D containing x_q^*. So we can estimate

$$\|u_q\|_{L^q(D)} \geq \left(2^{-nd} \right)^{1/q} \to 1 \tag{A6}$$

as $q \to \infty$ which proves the assertion in this case. If $N > 1$, we can still find for any $\varepsilon > 0$ sufficiently small a $\delta > 0$ such that $|u_q|_{B_\delta(x_q^*) \cap D}| \geq 1 - \varepsilon$ for all q. This follows

by uniform equicontinuity ensured by the Arzelà-Ascoli theorem since M is compact in $\mathcal{C}(\overline{D})$ as a bounded, closed, and finite dimensional set. Further, we have that

$$\text{vol}_d(B_\delta(x_q^*) \cap D) > c > 0$$

for all q. This follows by the more general observation that the union over $m \in \mathbb{N}$ of

$$V_m := \left\{ x \in \mathbb{R}^d : \text{vol}_d(B_\delta(x) \cap D) > \frac{1}{m} \right\}$$

is an open cover for $\overline{D}$, so we can find a minimal m such that $\overline{D} \subset V_m$ by compactness of $\overline{D}$. Altogether, we can again deduce a lower bound of the form

$$\|u_q\|_{L^q(D)} \geq (1-\varepsilon)\text{vol}_d(B_\delta(x_q^*) \cap D)^{1/q} \geq (1-\varepsilon)c^{1/q} \to 1 - \varepsilon$$

as $q \to \infty$. We conclude $\delta^* = 1$ since $\varepsilon > 0$ can be chosen as small as we wish. Finally, any subsequence of $\{\delta_q\}$ contains a subsequence which converges to 1. So, the whole sequence must converge to 1, see, e.g., [20, Prop. 10.13(2)].

A.2 A different approach

The functions of $X_N^n(D)$ from the above construction vanish on the set $\Delta = D \setminus \cup_{k \in \mathcal{I}_n}$ supp$\Phi_{n,k}$ which is non-empty in general. Thus, $\mathcal{P}_N^n u$ is a poor approximation of u near to the boundary of D in general: $\mathcal{P}_N^n u|_\Delta = 0$. Here we sketch an alternative approach to overcome this drawback. We emphasize that this approach has not been relied on in the previous sections as it requires the verification of an additional assumption on F', see text below following (A7).

Basically, we extend the preimage-space of the map F of Sect. 2 while keeping all its necessary properties to carry over Theorems 2.1 and 2.3 to the extension. Let $\widetilde{D}$ be an open superset of $\overline{D}$: $\overline{D} \subset \widetilde{D}$. We will need two operators: the extension operator $E \colon L^\infty(D) \to L^\infty(\widetilde{D})$, which extends a function by zero, and the restriction operator $R \colon L^\infty(\widetilde{D}) \to L^\infty(D)$, which multiplies a function by the indicator $\mathbb{1}_D$.

We define $\widetilde{F} \colon \mathcal{D}(\widetilde{F}) \subset L^\infty(\widetilde{D}) \to Y$ by $\mathcal{D}(\widetilde{F}) = E\mathcal{D}(F)$ and $\widetilde{F}(\widetilde{u}) = F(R\widetilde{u})$. This $\widetilde{F}$ is Fréchet-differentiable just like F. Moreover, the TCC holds in $B_r(Eu^+) \subset \mathcal{D}(\widetilde{F})$. Indeed, let $\widetilde{u} \in B_r(Eu^+)$ then $\|R\widetilde{u} - u^+\|_{L^\infty(D)} = \|R\widetilde{u} - REu^+\|_{L^\infty(D)} \leq \|\widetilde{u} - Eu^+\|_{L^\infty(\widetilde{D})} \leq r$, that is, $R\widetilde{u} \in B_r(u^+) \subset \mathcal{D}(F)$. Thus, for $u, \widetilde{u} \in B_r(Eu^+)$, we get

$$\|\widetilde{F}(u) - \widetilde{F}(\widetilde{u}) - \widetilde{F}'(\widetilde{u})(u - \widetilde{u})\|_Y = \|F(Ru) - F(R\widetilde{u}) - F'(R\widetilde{u})(Ru - R\widetilde{u})\|_Y$$
$$\leq \omega \|F(Ru) - F(R\widetilde{u})\|_Y = \omega \|\widetilde{F}(u) - \widetilde{F}(\widetilde{u})\|_Y.$$

Further, (6) is also satisfied by $\widetilde{F}$ as $R^* = E$.

For this $\widetilde{F}$ we can define spaces $X^n = X_N^n(\widetilde{D})$ as in Appendix A.1 but with respect to $\widetilde{D}$ rather than D. Now, the union of the supports of $\Phi_{n,k}$, for $k \in \mathcal{I}_n(\widetilde{D})$ covers D when n is large enough. Unfortunately, condition (5) does not transfer immediately to

the new construction. One must prove it for the concrete case, for example, as follows: for $\widetilde{u} \in \mathcal{D}(\widetilde{F})$ and $\widehat{u} \in L^\infty(\widetilde{D})$ we have that

$$
\begin{aligned}
\|\widetilde{F}'(\widetilde{u})(\widehat{u} - \widetilde{\mathcal{P}}_N^n \widehat{u})\|_Y &\leq \|F'(R\widetilde{u})(R\widehat{u} - \mathcal{P}_N^n R\widehat{u})\|_Y \\
&+ \|F'(R\widetilde{u})(\mathcal{P}_N^n R\widehat{u} - R\widetilde{\mathcal{P}}_N^n \widehat{u})\|_Y
\end{aligned}
\tag{A7}
$$

where $\mathcal{P}_N^n \colon L^\infty(D) \to X_N^n(D)$ and $\widetilde{\mathcal{P}}_N^n \colon L^\infty(\widetilde{D}) \to X_N^n(\widetilde{D})$ are the corresponding projection operators in accordance with (S2). The left norm on the right hand side of (A7) tends to 0 for $n \to \infty$ by (5). The right norm converges to 0, for instance, if $F'(u) \colon L^\infty(D) \to Y$ is weak-$\star$ continuous for all $u \in \mathcal{D}(F)$: both sequences $\{\mathcal{P}_N^n R\widehat{u}\}_n$ and $\{R\widetilde{\mathcal{P}}_N^n \widehat{u}\}_n$ converge to $R\widehat{u}$ pointwise a.e. This convergence can be verified by standard arguments, see e.g., [37, Chap. 12.3] and [39, Chap. 2]. Further, both sequences are uniformly bounded due to (S2). Hence, $\mathcal{P}_N^n R\widehat{u} - R\widetilde{\mathcal{P}}_N^n \widehat{u} \to 0$ weakly-$\star$.

Appendix B An approximation result

Proposition B.1 *For $u \in L^\infty(D)$ and $\{\mathcal{P}^n\}_n$ as in (A5) we have that $\mathcal{P}^n u \to u$ in $L^q(D)$ for all $q < \infty$.*

Proof Let $\square$ be a rectangular superset of D and $\widetilde{\square}$ a superset of $\square$. Further, extend u by zero outside of D. The convergence results of Section 12.3 from [37] yield that $\widetilde{\mathcal{P}}^n u \to u$ in $L^q(\square)$ for any $q < \infty$ where $\widetilde{\mathcal{P}}^n$ is defined as in (A5), however, with respect to $\mathcal{I}_n(\widetilde{\square})$ and $X^n(\widetilde{\square})$. Hence, for any $x \in D$ we have that $\widetilde{\mathcal{P}}^n u(x) = \mathcal{P}^n u(x)$ for n large enough such that $x \in D_n$, where we set

$$
D_n := \left\{ x \in D : \sum_{\mathbf{k} \in \mathcal{I}_n} \Phi_{n,\mathbf{k}}(x) = 1 \right\}
$$

in particular $\operatorname{vol}_d(D \backslash D_n) \to 0$. Because of $\|\mathcal{P}^n u\|_{L^\infty(D)} \leq C_{\mathcal{P}} \|u\|_{L^\infty(D)}$ by (S2), we can estimate

$$
\begin{aligned}
\|u - \mathcal{P}^n u\|_{L^q(D)} &= \|u - \mathcal{P}^n u\|_{L^q(D \backslash D_n)} + \|u - \mathcal{P}^n u\|_{L^q(D_n)} \\
&\leq \operatorname{vol}_d(D \backslash D_n)^{1/q} (C_{\mathcal{P}} + 1)\|u\|_{L^\infty(D)} + \|u - \widetilde{\mathcal{P}}^n u\|_{L^q(D_n)} \\
&\leq \operatorname{vol}_d(D \backslash D_n)^{1/q} (C_{\mathcal{P}} + 1)\|u\|_{L^\infty(D)} + \|u - \widetilde{\mathcal{P}}^n u\|_{L^q(\square)}
\end{aligned}
$$

and the assertion follows. $\qquad\square$

Appendix C On the tangential cone condition for the operator in (46)

Here we argue that the semi-discrete non-linear operator $F_{n_{\max}}$ defined in (46) satisfies the TCC (3).

The underlying abstract system is (25) with the concrete settings (32), $u_0 = 0$, and (33) where $D = (0, 1)$, $T = 1$, that is, $d = 1$ and $u = (p, v) \in L^2([0, 1], H)$,

$H = L^2(0, 1)^2$. Thus, we obtain the acoustic system (42) which has a unique classical solution under $(\rho, v) \in \mathcal{D}(F_{n_{\max}})$ and for the sources (43). In view of Lemma 3.5, $F_{n_{\max}}$ is Fréchet-differentiable and we have $F'_{n_{\max}}(\rho, v)(\widehat{\rho}, \widehat{v}) = (\overline{p}, \overline{v})$ where $(\overline{p}, \overline{v})$ weakly solves

$$
\begin{aligned}
\frac{1}{\rho v^2} \partial_t \overline{p} - \partial_x \overline{v} &= -\widehat{\rho} \partial_t p && \text{in } (0, 1) \times (0, 1), \\
\rho \partial_t \overline{v} - \partial_x \overline{p} &= -\widehat{v} \partial_t v && \text{in } (0, 1) \times (0, 1), \\
\overline{v}(0, \cdot) = \overline{p}(0, \cdot) &= 0 && \text{on } (0, 1) \\
\overline{v}(\cdot, 0) = \overline{p}(\cdot, 1) &= 0 && \text{on } (0, 1).
\end{aligned}
\tag{C8}
$$

In a first step we validate injectivity of $F'_{n_{\max}}(\rho, v)$ for any $(\rho, v) \in \mathcal{D}(F_{n_{\max}})$. To this end, assume $F'_{n_{\max}}(\rho, v)(\widehat{\rho}, \widehat{v}) = (0, 0)$. From (C8) we get

$$
0 = \widehat{\rho} \partial_t p \quad \text{and} \quad 0 = \widehat{v} \partial_t v \quad \text{in } (0, 1) \times (0, 1).
$$

Assume $0 \neq \widehat{\rho} \in X^{n_{\max}}$. Then, there is a non-empty interval $[a, b]$, $a = 2^{-n_{\max}} k$, $b = 2^{-n_{\max}}(k+1)$, $k \in \mathbb{N}_0$, where $\widehat{\rho}$ does not vanish. Hence, $\partial_t p = 0$ in $[0, 1] \times [a, b]$. By the first equation in (42), $-\partial_x v = f_1$ in $[0, 1] \times [a, b]$, that is,

$$
v(t, x) = v(t, a) - \int_a^x f_1(t, y) \mathrm{d}y, \quad (t, x) \in (0, 1) \times [a, b].
$$

Recalling the zero initial value $v(0, \cdot) = 0$ we get the contradiction $0 = \int_a^x f_1(0, y) \mathrm{d}y < 0$ for $x \in [a, b]$ according to (43). So, $\widehat{\rho} = 0$ in $(0, 1)$. One argues analogously to validate $\widehat{v} = 0$ in $(0, 1)$. Hence, $F'_{n_{\max}}(\rho, v)$ is one-to-one which implies the TCC at any interior point of $\mathcal{D}(F_{n_{\max}})$ due to Lemma C.1 of [6].

References

1. Mueller, J.L., Siltanen, S.: Linear and Nonlinear Inverse Problems with Practical Applications. Computational Science & Engineering, vol. 10. Society for Industrial and Applied Mathematics (SIAM), Philadelphia (2012). https://doi.org/10.1137/1.9781611972344
2. Fichtner, A.: Full Seismic Waveform Modelling and Inversion Advances in Geophysical and Environmental Mechanics and Mathematics. Springer, Berlin (2011). https://doi.org/10.1007/978-3-642-15807-0
3. Virieux, J., Operto, S.: An overview of full-waveform inversion in exploration geophysics. Geophysics **74**(6), 1–26 (2009). https://doi.org/10.1190/1.3238367
4. Rieder, A.: On the regularization of nonlinear ill-posed problems via inexact Newton iterations. Inverse Prob. **15**(1), 309–327 (1999). https://doi.org/10.1088/0266-5611/15/1/028
5. Margotti, F.J.: On inexact Newton methods for inverse problems in Banach spaces. Ph.D. thesis, Karlsruher Institut für Technologie (KIT) (2015). https://doi.org/10.5445/IR/1000048606
6. Eller, M., Rieder, A.: Tangential cone condition and Lipschitz stability for the full waveform forward operator in the acoustic regime. Inverse Prob. **37**, 085011 (2021). https://doi.org/10.1088/1361-6420/ac11c5
7. Pöschl, C., Resmerita, E., Scherzer, O.: Discretization of variational regularization in Banach spaces. Inverse Probl. **26**(10), 105017 (2010). https://doi.org/10.1088/0266-5611/26/10/105017

8. Kaltenbacher, B., Neubauer, A., Scherzer, O.: Iterative Regularization Methods for Nonlinear Ill-Posed Problems. Radon Series on Computational and Applied Mathematics, vol. 6. Walter de Gruyter GmbH & Co. KG, Berlin (2008). https://doi.org/10.1515/9783110208276

9. Schuster, T., Kaltenbacher, B., Hofmann, B., Kazimierski, K.S.: Regularization Methods in Banach Spaces. De Gruyter, Berlin (2012). https://doi.org/10.1515/9783110255720

10. Clason, C., Klassen, A.: Quasi-solution of linear inverse problems in non-reflexive Banach spaces. J. Inverse Ill-Posed Probl. **26**(5), 689–702 (2018). https://doi.org/10.1515/jiip-2018-0026

11. Iusem, A.N., Resmerita, E.: A proximal point method in nonreflexive Banach spaces. Set-Valued Var. Anal. **18**(1), 109–120 (2010). https://doi.org/10.1007/s11228-009-0126-z

12. Kaltenbacher, B., de Souza, M.L.P.: Convergence and adaptive discretization of the IRGNM Tikhonov and the IRGNM Ivanov method under a tangential cone condition in Banach space. Numer. Math. **140**(2), 449–478 (2018). https://doi.org/10.1007/s00211-018-0971-5

13. Kaltenbacher, B., Klassen, A.: On convergence and convergence rates for Ivanov and Morozov regularization and application to some parameter identification problems in elliptic PDEs. Inverse Probl. **34**(5), 055008 (2018). https://doi.org/10.1088/1361-6420/aab739

14. Kaltenbacher, B., Klassen, A., Previatti de Souza, M.L.: The Ivanov regularized Gauss-Newton method in Banach space with an a posteriori choice of the regularization radius. J. Inverse Ill-Posed Probl. **27**(4), 539–557 (2019). https://doi.org/10.1515/jiip-2018-0093

15. Schuster, T., Rieder, A., Schöpfer, F.: The approximate inverse in action: IV. Semi-discrete equations in a Banach space setting. Inverse Probl. **28**(10), 104001 (2012). https://doi.org/10.1088/0266-5611/28/10/104001

16. Scherzer, O.: The use of Morozov's discrepancy principle for Tikhonov regularization for solving nonlinear ill-posed problems. Computing **51**(1), 45–60 (1993). https://doi.org/10.1007/BF02243828

17. Hanke, M., Neubauer, A., Scherzer, O.: A convergence analysis of the Landweber iteration for nonlinear ill-posed problems. Numer. Math. **72**(1), 21–37 (1995). https://doi.org/10.1007/s002110050158

18. Kaltenbacher, B., Nguyen, T.T.N., Scherzer, O.: The tangential cone condition for some coefficient identification model problems in parabolic PDEs. In: Kaltenbacher, B., Schuster, T., Wald, A. (eds.) Time-dependent Problems in Imaging and Parameter Identification, pp. 121–163. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-57784-1_5

19. Lechleiter, A., Rieder, A.: Towards a general convergence theory for inexact Newton regularizations. Numer. Math. **114**(3), 521–548 (2010). https://doi.org/10.1007/s00211-009-0256-0

20. Zeidler, E.: Nonlinear Functional Analysis and Its Applications I: Fixed-Point Theorems. Springer, New York (1986)

21. Somersalo, E., Cheney, M., Isaacson, D.: Existence and uniqueness for electrode models for electric current computed tomography. SIAM J. Appl. Math. **52**(4), 1023–1040 (1992). https://doi.org/10.1137/0152060

22. Lechleiter, A., Rieder, A.: Newton regularizations for impedance tomography: convergence by local injectivity. Inverse Probl. **24**(6), 065009 (2008). https://doi.org/10.1088/0266-5611/24/6/065009

23. Zeltmann, U.: The viscoelastic seismic model: existence, uniqueness and differentiability with respect to parameters. Ph.D. thesis, Karlsruhe Institute of Technology (2018). https://doi.org/10.5445/IR/1000093989

24. Bohlen, T.: Viskoelastische FD-Modellierung seismischer Wellen zur Interpretation gemessener Seismogramme. Ph.D. thesis, Christian-Albrechts-Universität zu Kiel (1998). https://bit.ly/2LM0SWr

25. Bohlen, T.: Parallel 3-D viscoelastic finite difference seismic modelling. Comput. Geosci. **28**(8), 887–899 (2002)

26. Kirsch, A., Rieder, A.: Inverse problems for abstract evolution equations II: higher order differentiability for viscoelasticity. SIAM J. Appl. Math. **79**(6), 2639–2662 (2019). https://doi.org/10.1137/19M1269403. Corrigendum under https://doi.org/10.48550/arXiv.2203.01309

27. Kirsch, A., Rieder, A.: Inverse problems for abstract evolution equations with applications in electrodynamics and elasticity. Inverse Probl. **32**(8), 085001 (2016). https://doi.org/10.1088/0266-5611/32/8/085001

28. Triebel, H.: Theory of Function Spaces. II. Monographs in Mathematics, vol. 84. Birkhäuser Verlag, Basel (1992). https://doi.org/10.1007/978-3-0346-0419-2

29. Rychkov, V.S.: On restrictions and extensions of the Besov and Triebel–Lizorkin spaces with respect to Lipschitz domains. J. Lond. Math. Soc. **60**(1), 237–257 (1999). https://doi.org/10.1112/S0024610799007723

30. Binev, P., Cohen, A., Dahmen, W., DeVore, R., Temlyakov, V.: Universal algorithms for learning theory I. Piecewise constant functions. J. Mach. Learn. Res. **6**, 1297–1321 (2005)

31. Evans, L.C.: Partial Differential Equations. Graduate Studies in Mathematics, vol. 19, 2nd edn. American Mathematical Society, Providence (2010). https://doi.org/10.1090/gsm/019

32. Meyers, N.G.: An L^p-estimate for the gradient of solutions of second order elliptic divergence equations. Ann. Sc. Norm. Sup. Pisa Cl. Sci. (3) **17**, 189–206 (1963)

33. Ibe, O.C.: Markov Processes for Stochastic Modeling, 2nd edn. Elsevier, Amsterdam (2013). https://doi.org/10.1016/B978-0-12-407795-9.00001-3

34. Messud, J., Poncet, R., Lambaré, G.: Optimal transport in full-waveform inversion: analysis and practice of the multidimensional Kantorovich–Rubinstein norm. Inverse Prob. **37**(6), 065012 (2021). https://doi.org/10.1088/1361-6420/abfb4c

35. Métivier, L., Brossier, R., Mérigot, Q., Oudet, E., Virieux, J.: An optimal transport approach for seismic tomography: application to 3D full waveform inversion. Inverse Probl. **32**(11), 115008 (2016). https://doi.org/10.1088/0266-5611/32/11/115008

36. Sun, B., Alkhalifah, T.: Ml-misfit: learning a robust misfit function for full-waveform inversion using machine learning. In: Conference Proceedings, EAGE 2020 Annual Conference & Exhibition Online, 1–5 (2020). https://doi.org/10.3997/2214-4609.202010466

37. Schumaker, L.L.: Spline Functions: Basic Theory. Cambridge Mathematical Library, 3rd edn. Cambridge University Press, Cambridge (2007). https://doi.org/10.1017/CBO9780511618994

38. Cohen, A., Daubechies, I., Feauveau, J.-C.: Biorthogonal bases of compactly supported wavelets. Commun. Pure Appl. Math. **45**(5), 485–560 (1992). https://doi.org/10.1002/cpa.3160450502

39. Oswald, P.: Multilevel Finite Element Approximation. Teubner Skripten zur Numerik [Teubner Scripts on Numerical Mathematics]. B. G. Teubner, Stuttgart (1994). https://doi.org/10.1007/978-3-322-91215-2