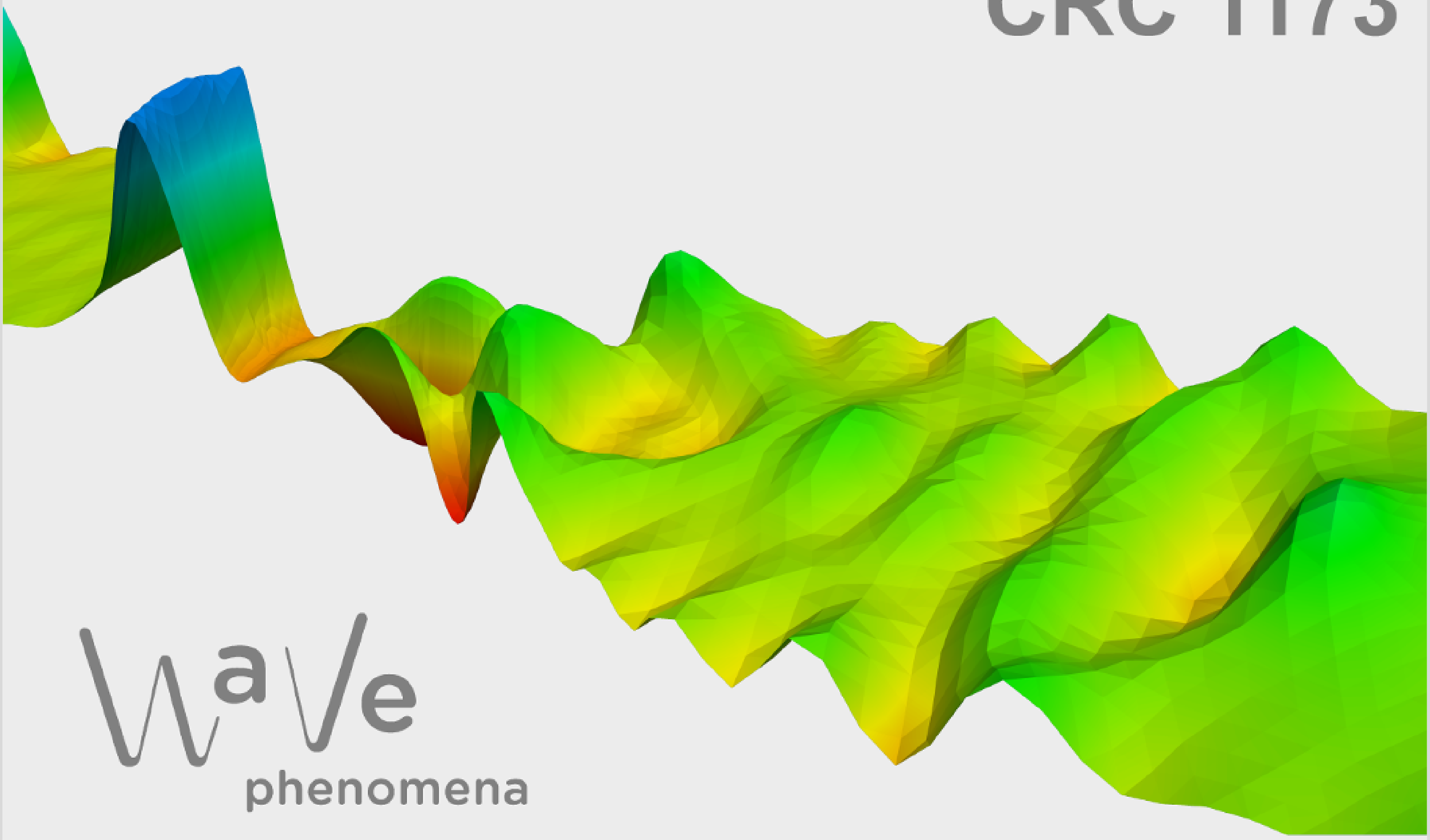
On the iterative regularization of non-linear illposed problems in L^∞

Lukas Pieronek, Andreas Rieder

CRC Preprint 2021/46, December 2021

KARLSRUHE INSTITUTE OF TECHNOLOGY

CRC 1173



Participating universities

Universität Stuttgart

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

TU WIEN

Funded by

DFG

ISSN 2365-662X

ON THE ITERATIVE REGULARIZATION OF NON-LINEAR ILLPOSED PROBLEMS IN L^∞

LUKAS PIERONEK AND ANDREAS RIEDER

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 smooth Tikhonov functional over a finite-dimensional space whose dimension increases as the iteration progresses. In this way, all iterates stay bounded and the regularizer, delivered by a discrepancy principle, converges weakly-$\star$ 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.

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. The application we have in mind, and which has triggered our research, is full waveform inversion (FWI), the most advanced inversion technique in seismic imaging, see, e.g., [6, 27]. 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. Let

$$F\colon \mathcal{D}(F) \subset L^\infty(D)^\ell \to Y \tag{1}$$

map these ℓ parameter functions located on some domain of interest D to the wave field initiated by a source (explosion or earthquake). The wave field can be recorded at receivers on the earth's surface or by hydrophones in the sea. From these measurements we then try to recover the parameters. Mathematically, we have to

$$\text{find } u \in \mathcal{D}(F) \text{ such that } F(u) \approx y^\delta \tag{2}$$

where y^δ are the (noisy) measurements satisfying $\|y^\delta - F(u^+)\|_Y \le \delta$ for one $u^+ \in \mathcal{D}(F)$.

For this purpose, Newton-like regularization schemes are well-established iterations for getting a meaningful approximate solution of non-linear inverse problems. There is a wealth of literature on the analysis of those methods, mainly in a Hilbert space but meanwhile also in a Banach space setting; we refer only to the monographs [12, 24] for

Date: December 1, 2021.

2020 *Mathematics Subject Classification.* 65J20, 65J22, 65N20.

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project-ID 258734477 - SFB 1173. The authors thank Christian Rheinbay for insightful discussions.

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 [2, 8, 9, 10, 11, 20, 25]. The first six of these publications consider iterative schemes on the basis of proximal point methods, Morozov, Ivanov, or Tikhonov regularization, respectively. However, all of them require norm-related minimization whose implementation in the context of L^∞ calls for non-smooth or constrained optimization techniques.

In this work we explore the Newton-like solver $\texttt{REGINN}^\infty$ which extends $\texttt{REGINN}$ of [21, 15] to a non-linear inverse problem with generic operators F as in (1). F is required to fulfill a few specific properties which are, except for one, satisfied by FWI in all regimes. This not yet verified property is a structural assumption known as tangential cone condition (TCC), see (3) below and consult [4] for a first promising result. Apart from establishing the non-linearity constraint, the main challenge about regularization in L^∞ is its non-reflexive and non-smooth nature. As this space is further non-separable, convergence of a discretization scheme in the strong topology cannot be expected, see [20]. Instead, using 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, our idea is to make use of semi-discrete approximations to F based on a family $\{X^n\}_n$ of finite dimensional nested subspaces of $L^\infty(D)^\ell$ which are 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 then implies that the Newton update for the n-th iterate is obtained as an approximate minimizer of a smooth (provided Y is smooth) and convex Tikhonov functional over X^n with just the L^{q_n}-norm as penalty term. In particular, the underlying minimizing procedure can be easily implemented numerically, which is one of the major advantage of our algorithm, while all iterates are still kept bounded in L^∞ uniformly.

At first glance the IRGNM-Tikhonov method of [9] and $\texttt{REGINN}^\infty$ seem to be quite similar, but they are separated by significant structural differences: In IRGNM-Tikhonov the penalty parameter is determined a priori and the Newton update has to be a minimizer of the Tikhonov functional, whereas for $\texttt{REGINN}^\infty$ the penalty parameter as well as the penalty functional itself depend on the current iterate and the Newton update is only an approximate minimizer. Further, the discrete spaces X^n are an intrinsic part of $\texttt{REGINN}^\infty$, so that we can get rid of the L^∞-norm with its geometrical and numerical difficulties. Standard adaptive discretization, in contrast, mainly aims at reducing computational expenses.

We present our material as follows: In Section 2 we introduce and analyze two versions of $\texttt{REGINN}^\infty$ 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 (2). 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 FWI with exception of TCC (Section 3) as mentioned above. For the acoustic regime we are even able to provide the information about the smoothness of the ground truth which enters the second version of $\texttt{REGINN}^\infty$. Finally, Section 4 contains some experiments concerning a toy model in the acoustic regime for which the TCC actually holds. 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

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

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

$$\|u\|^2_{L^\infty(D)^\ell} = \|(u_1, \dots, u_\ell)\|^2_{L^\infty(D)^\ell} := \sum_{j=1}^{\ell} \|u_i\|^2_{L^\infty(D)}.$$

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 some $C_\infty > 1$ 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 with respect to X^n is. A family $\{X^n\}_{n\in\mathbb{N}}$ enjoying (S1)-(S3) is constructed in Appendix A on the basis of tensor product B-splines. Finally, we require a compatibility condition of the form

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

for all $u \in \mathcal{D}(F)$ and all $\widehat{u} \in L^\infty(D)^\ell$. This relation ensures that $\mathcal{P}^n \widehat{u}$ converges to $\widehat{u}$ as $n \to \infty$ in a sense which still yields strong convergence of the images under $F'(u)$. In general, one cannot expect $\lim_{n\to\infty} \mathcal{P}^n \widehat{u} = \widehat{u}$ in $L^\infty(D)^\ell$ because the union of X^n is countable while $L^\infty(D)^\ell$ is not separable.

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

$$b_m^\delta := y^\delta - F(u_m). \tag{6}$$

To address all three aspects at once, we build on an inexact-Newton framework and find the updates from the linearization approximately via Tikhonov regularization with special adaptive discretization. The latter refers to minimizing proper Tikhonov functionals on X^n which are linked in a pre-defined manner. Note that thanks to assumption (S3) the penalty term therein can be reduced to the $L^{q_n}(D)^\ell$-norm while still controlling the

Algorithm 1 `REGINN`$^\infty$

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

corresponding $L^\infty(D)^\ell$-norm of actual interest which is difficult to cope with numerically. Hence, given iterates $\{u_0, \dots, u_m\} \subset \mathcal{D}(F)$, the next Newton step reads

$$u_{m+1} = u_m + s_m, \tag{7}$$

where $n_m \ge n_{m-1}$ and $s_m \in X^{n_m}$ are chosen such that

$$J_{n_m,m}(s_m) \le \mu_m^2 \|b^\delta_m\|^2_Y. \tag{8}$$

Here,

$$J_{n,m}(s) := \|F'(u_m)s - b^\delta_m\|^2_Y + \alpha_m \|s + (u_m - u_0)\|^2_{L^{q_n}(D)^\ell} \tag{9}$$

is a Tikhonov functional for fixed $y^\delta \in Y$ with domain X^n and

$$\alpha_m = \frac{\|b^\delta_m\|^2_Y}{\gamma^2} \tag{10}$$

is to set successively during run-time. So far, the parameters $\gamma, \{\mu_m\}_{m\in\mathbb{N}}$ 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. We stop the Newton iteration by a discrepancy principle 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.

Theorem 2.1 (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 $B_r(u^+) \subset \operatorname{int}(\mathcal{D}(F))$ 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* $n_0 \in \mathbb{N}$ *and* $u_0 \in B_{r_0}(u^+)$. *Then, there exists an* $M_\delta \in \mathbb{N}$ *such that all iterates* $\{u_1, \dots, u_{M_\delta}\}$ *of* $\texttt{REGINN}^\infty$ *are well-defined and stay in* $B_r(u^+)$. *Moreover,* $\|b_{m+1}^\delta\|_Y \leq \Lambda \|b_m^\delta\|_Y$ *for* $m = 0, \dots, M_\delta - 1$, $\|b_{M_\delta}^\delta\|_Y \leq \tau\delta$, *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 $\|b_m^\delta\|_Y \leq \tau\delta$, $\texttt{REGINN}^\infty$ 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 (8). 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,

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

Recursively, we get $u_m - u_0 = s_0 + s_1 + \ldots + s_{m-1}$ 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), $\mathcal{P}^n(u_m - u_0) = u_m - u_0$ for $n \geq n_{m-1}$ and by linearity of $\mathcal{P}^n$ we may simplify

$$\begin{aligned}
J_{n,m}(s_{n,m}) &\leq \|F'(u_m)\mathcal{P}^n(u^+ - u_m) - b_m^\delta\|_Y^2 + \alpha_m \|\mathcal{P}^n(u^+ - u_0)\|_{L^{q_n}(D)^\ell}^2 \\
&\leq \|F'(u_m)\mathcal{P}^n(u^+ - u_m) - b_m^\delta\|_Y^2 + \mathrm{vol_d}(D)^{2/q_n} C_{\mathcal{P}}^2 \frac{r_0^2}{\gamma^2} \|b_m^\delta\|_Y^2.
\end{aligned}$$

In the last step we additionally used (10), Hölder's inequality and $\|\mathcal{P}^h(u^+ - u_0)\|_{L^\infty(D)^\ell} \leq C_{\mathcal{P}} \|u^+ - u_0\|_{L^\infty(D)^\ell} < C_{\mathcal{P}} r_0$, see (S2). We continue by splitting the residual term according to

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

employing again $\mathcal{P}^n(u^+ - u_m) = u_m - u_0$ to get the bottom line. Since $\|u_m - u^+\|_{L^\infty(D)^\ell} < r$ by induction, TCC (3) yields

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

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

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

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

$$\begin{aligned}\|F'(u_m)\mathcal{P}^n(u^+ - u_m) - \big(y^\delta - F(u_m)\big)\|_Y \\ \leq \|b_m^\delta\|_Y\Big(\omega + \frac{1+\omega}{\tau}\Big) + \|F'(u_m)\big(\mathcal{P}^n(u^+ - u_0) - (u^+ - u_0)\big)\|_Y\end{aligned}$$

and finally

$$\begin{aligned}J_{n,m}(s_{n,m}) \leq \bigg(\|b_m^\delta\|_Y\Big(\omega + \frac{1+\omega}{\tau}\Big) + \|F'(u_m)\big(\mathcal{P}^n(u^+ - u_0) - (u^+ - u_0)\big)\|_Y\bigg)^2 \\ + \operatorname{vol}_d(D)^{2/q_n} C_{\mathcal{P}}^2 \frac{r_0^2}{\gamma^2}\|b_m^\delta\|_{L^2(D)}^2.\end{aligned} \tag{11}$$

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

$$\limsup_{n\to\infty} J_{n,m}(s_{n,m}) \leq \|b_m^\delta\|_Y^2\bigg(\Big(\omega + \frac{1+\omega}{\tau}\Big)^2 + C_{\mathcal{P}}^2\frac{r_0^2}{\gamma^2}\bigg) = \mu_{\min}^2\|b_m^\delta\|_Y^2. \tag{12}$$

Consequently, condition (8) with $\mu_m > \mu_{\min}$ is feasible for n_m large and appropriate $s_m \in X^{n_m}\setminus\{0\}$, 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 (10), we proceed with

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

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 $u_{m+1} \in B_r(u^+) \subset \operatorname{int}(\mathcal{D}(F))$. Finally, we estimate on the basis of (4) and (8)

$$\begin{aligned}\|b_{m+1}^\delta\|_Y &= \|\big(b_m^\delta - F'(u_m)s_m\big) - \big(F(u_{m+1}) - F(u_m) - F'(u_m)s_m\big)\|_Y \\ &\leq \|b_m^\delta - F'(u_m)s_m\|_Y + \frac{\omega}{1-\omega}\|F'(u_m)s_m\|_Y \\ &\leq \sqrt{J_{n,m}(s_m)} + \frac{\omega}{1-\omega}\big(\|b_m^\delta\|_Y + \|F'(u_m)s_m - b_m^\delta\|_Y\big) \\ &\leq \mu_m\|b_m^\delta\|_Y + \frac{\omega}{1-\omega}(1+\mu_m)\|b_m^\delta\|_Y \qquad (13)\end{aligned}$$

$$= \left(\mu_m + \frac{\omega}{1-\omega}(1+\mu_m)\right)\|b_m^\delta\|_Y$$
$$< \left(\mu_{\max} + \frac{\omega}{1-\omega}(1+\mu_{\max})\right)\|b_m^\delta\|_Y < \Lambda\|b_m^\delta\|_Y.$$

Having thus proven the induction part, $\texttt{REGINN}^\infty$ is forced to terminate for any $\delta > 0$ due to $\|b_m^\delta\|_Y \le \Lambda^m\|b_0^\delta\|_Y \le \tau\delta$ for m sufficiently large. From this estimate, we may even deduce $M_\delta = \mathcal{O}(|\log\delta|)$ as $\delta \searrow 0$. □

Remark 2.2. a) *The name* $\texttt{REGINN}^\infty$ *for Algorithm 1 is justified by the stopping condition (8) for determining the Newton update which is, in view of (9) and (10), equivalent to*

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

In particular, s_m satisfies the stopping condition of $\texttt{REGINN}$ *[21], i.e., the above condition without penalty term.*

b) *Recall that* $\texttt{REGINN}$ *admits in the Hilbert space setting (and likewise for smooth reflexive Banach spaces) a so-called 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* $\texttt{REGINN}^\infty$ *in general. Therefore our parameters need to be controlled in terms of both ω and r, whereas standard* $\texttt{REGINN}$ *only requires the knowledge of ω for defining admissible tolerances μ and stopping constants τ, see Theorem* 3.1 *in* [14].

Remark 2.3. *We discuss how the statement of the theorem from above carries 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, Theorem 2.1 applies to $F_{n_{\max}}$ 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, (12) then reads

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

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 $\texttt{REGINN}^\infty$ *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.

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:

[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 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 $u \in U_0$. In this case, we assume $u^+ \in U_0 + X^{n_{\max}}$.

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

Corollary 2.4 (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*

$$\mathcal{R}(F'(u^+)^*) \subset L^1(D)^\ell \tag{14}$$

or that F is weakly-$\star$ sequentially closed, that is, $u_n \overset{*}{\rightharpoonup} x$ in $L^\infty(D)^\ell$ and $F(u_n) \rightharpoonup y$ imply that $F(u) = y$. 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

$$\|y - F(u_{M_{\delta_i}})\|_Y \le \|y - y^{\delta_i}\|_Y + \|b_{M_{\delta_i}}^{\delta_i}\|_Y \le (1+\tau)\delta_i \to 0. \tag{15}$$

and is uniformly bounded in $L^\infty(D)^\ell$, so there exists weak-$\star$ accumulation points in $\overline{B_r(u^+)}$ by weak-$\star$-compactness. Take representatively $u_{M_{\delta_{i_k}}} \overset{*}{\rightharpoonup} \widetilde{u}$. In case that F is weak-$\star$ sequentially closed, we can directly deduce $F(\widetilde{u}) = y$, hence any weak-$\star$ accumulation point solves the equation. In case that (14) holds, we first note that the TCC (3) implies by the reverse triangle inequality for any $u \in B_r(u^+)$ that

$$(1-\omega)\|F(u^+) - F(u)\|_Y \le \|F'(u^+)(u^+ - u)\|_Y \le (1+\omega)\|F(u^+) - F(u)\|_Y. \tag{16}$$

With (14) we then obtain for any $g \in Y^*$

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

where the last equality above follows by the second inequality in (16) with $x = u_{M_{\delta_{i_k}}}$ and $F(u_{M_{\delta_{i_k}}}) \to y$ in Y. We deduce that $F'(u^+)(u^+ - \widetilde{u}) = 0$ and combining this relation now with the first inequality in (16) using $\widetilde{u} = u$, 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 [28, Prop. 10.13(2)]. ∎

Remark 2.5. *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 (16) and (15) 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.*

The previous version of $\texttt{REGINN}^\infty$ requires the determination of successive discretization levels $n_m \geq n_{m-1}$ for possibly many n and corresponding (almost) minimizers $s \in X^n$ of (9) need to be computed before meeting the given μ_m-criterion in (8). As this can be numerically expensive, we want to present an alternative version which directly links n to m. A closer look on 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 a subspace such that

$$\|\widehat{u}\|_{L^\infty(D)^\ell} \leq C_X \|\widehat{u}\|_X \quad \text{for all } \widehat{u} \in X, \tag{17}$$

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

$$\|F'(u)\big(\widehat{u} - \mathcal{P}^n \widehat{u}\big)\|_Y \leq C^+ \|\widehat{u}\|_X \beta(n) \ , \tag{18}$$

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. The next corollary shows that on this basis we can indeed determine n_m conveniently for successive Newton steps of $\texttt{REGINN}^\infty$.

Corollary 2.6. *Adopt all assumptions, notations and parameters from Theorem 2.1 and assume that (18) 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*

$$\mu_{\min}^\varepsilon := \left(\omega + \frac{1+\omega}{\tau} + \varepsilon\right)^2 + \max\left\{\operatorname{vol}_d(D)^{2/q_{n_0}}, 1\right\} C_{\mathcal{P}}^2 \frac{r_0^2}{\gamma^2} < \mu_{\max} \tag{19}$$

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

$$n_m := \min\left\{n \geq n_0 : \ \beta(n) \leq \varepsilon \|b_m^\delta\|_Y\right\}. \tag{20}$$

Then we can find $s_m \in X^{n_m}$ satisfying (8) for all m. In particular, $\texttt{REGINN}^\infty$ also terminates in this case and the regularization property still holds.

Proof. First, n_m according to (20) 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 (11) and by (18)

$$J_{n_m,m}(s_m) \leq \left(\|b_m^\delta\|_Y\left(\omega + \frac{1+\omega}{\tau}\right) + \|F'(u_m)\big(\mathcal{P}^n(u^+ - u_0) - (u^+ - u_0)\big)\|_Y\right)^2$$
$$+ \operatorname{vol}_d(D)^{2/q_{n_m}} \frac{r_0^2}{\gamma^2} C_{\mathcal{P}}^2 \|b_m^\delta\|_{L^2(D)}^2$$

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 $\|y^\delta - F(u_M)\|_Y \le \tau\delta$
$m := 0$
$b_m^\delta := y^\delta - F(u_m)$
while $\|b_m^\delta\|_Y > \tau\delta$ **do**
$\alpha_m := \|b_m^\delta\|_Y^2/\gamma^2$
$n_m := \min\left\{n \ge n_0 : \ \beta(n) \le \varepsilon\|b_m^\delta\|_Y\right\}$
determine $s_m \in X^{n_m}$: $J_{n_m,m}(s_m) \le \mu_m^2\|b_m^\delta\|_Y^2$
$u_{m+1} := u_m + s_m$
$m := m+1$
$b_m^\delta := y^\delta - F(u_m)$
end while
$M := m$

$$\le \left(\|b_m^\delta\|_Y\left(\omega + \frac{1+\omega}{\tau}\right) + \underbrace{C^+\|u^+ - u_0\|_X}_{<1}\underbrace{\beta(n_m)}_{\le\varepsilon\|b_m^\delta\|_Y}\right)^2$$
$$+ \max\{\mathrm{vol}_d(D)^{2/q_{n_0}}, 1\}C_{\mathcal{P}}^2\frac{r_0^2}{\gamma^2}\|b_m^\delta\|^2_{L^2(D)}$$
$$\le (\mu_{\min}^\varepsilon)^2\|b_m^\delta\|^2_{L^2(D)}.$$

The fact that `REGINN`$^\infty$ still terminates and also admits the regularization property follows by Theorem 2.1, Corollary 2.4 and $\|u_0 - u^+\|_{L^\infty(D)^\ell} \le C_X\|u_0 - u^+\|_X < r_0$. □

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: Full Waveform Inversion

Full waveform inversion (FWI) is the state-of-the-art imaging modality in exploration geophysics, see, e.g., [6, 27]. Basically, it consists of a parameter identification problem for the governing wave equation in time domain. In this section we will first verify the compatibility condition (5) for the underlying parameter-to-solution map in quite a general fashion, including the visco-elastic regime, and the B-spline subspaces from Appendix A. Confining to the acoustic regime and linear B-spline spaces then, we can even show (18).

We follow the abstract approach from [13] and consider, for some bounded domain $D \subset \mathbb{R}^d$, a (multi-)parameter-to-solution map

$$F\colon \mathcal{D}(F) \subset L^\infty(D)^\ell \to L^2([0,T],H) =: Y, \quad u = (u_1,\dots,u_\ell) \mapsto y,$$

which relates certain parameter functions $\{u_l\}_{1\le l\le\ell} \subset L^\infty(D)$ to corresponding solutions $y\colon [0,T] \to H$ given by

$$\begin{aligned} By'(t) + Ay(t) + BQy(t) &= f(t) \quad \text{in } (0,T), \\ y(0) &= y_0. \end{aligned} \tag{21}$$

Here, H is a Hilbert space and A, B, Q are time-independent operator-valued coefficients on H for which we impose that $A \in \mathcal{L}(\mathcal{D}(A), H)$ is unbounded and maximal monotone, $Q \in \mathcal{L}(H)$ is such that $Q\mathcal{D}(A) \subset \mathcal{D}(A)$ and $B \in \mathcal{L}(H)$ is invertible and self-adjoint. Further, we assume only B to depend on the ℓ parameter functions $\{u_l\}_{1\le l\le \ell}$. This suggests introducing an auxiliary operator $V\colon L^\infty(D)^\ell \to \mathcal{L}(H)$ given by $(u_1,\dots,u_\ell) =: u \mapsto V(u) := B$ which we require to be differentiable and such that any uniformly bounded $\{\widehat{u}_n\}_{n\in\mathbb{N}} \subset L^\infty(D)^\ell$ with $\widehat{u}_n \to \widehat{u}$ pointwise a.e. implies for all $h \in H$ and any $u \in L^\infty(D)^\ell$ that

$$[V'(u)\widehat{u}_n]h \to [V'(u)\widehat{u}]h. \tag{22}$$

For a concrete definition of the involved spaces and operators for the visco-elastic wave equation we refer to [13] where it was shown that (21) admits a unique strong solution $y \in C([0,T], \mathcal{D}(A)) \cap C^1([0,T], H) \subset L^2([0,T], H)$ for $y_0 \in \mathcal{D}(A)$ and $f \in W^{1,1}([0,T], H)$. However, the inverse problem is locally illposed at any interior point $u \in \mathcal{D}(F)$.

We will need the following two results of [13] for our further considerations.

Lemma 3.1. *For $f \in W^{1,1}([0,T], H)$ and $y_0 \in \mathcal{D}(A)$, F is Fréchet-differentiable at any interior point $u \in \mathcal{D}(F)$. Setting $y = F(u)$, $\overline{y} := F'(u)\widehat{u} \in C([0,T], H)$ for $\widehat{u} \in L^\infty(D)^\ell$ is given as the unique weak solution of*

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

that is

$$\frac{d}{dt}(B\overline{y}, v)_H + (\overline{y}, A^* v)_H + (BQ\overline{y}, v)_H = -\big([V'(u)\widehat{u}]\left(y'(t) + Qy(t)\right), v\big)_H$$

for a.e. $t \in (0,T)$ and all $v \in \mathcal{D}(A^*)$. Further, we have the stability estimate

$$\|\overline{y}(t)\|_H \le C\|[V'(u)\widehat{u}](y' + Qy)\|_{L^1((0,t),H)}, \tag{23}$$

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

Lemma 3.2. *If $f \in W^{k,1}([0,T], H)$ for $k \ge 1$ and the compatibility conditions*

$$y_{0,l} := (B^{-1}A + Q)^l y_0 + \sum_{j=0}^{l-1} (B^{-1}A + Q)^j B^{-1} f^{(l-1+j)} \in \mathcal{D}(A), \quad l = 0, \dots, k-1,$$

hold, then the solution u of (21) is in $C^{k-1}([0,T], \mathcal{D}(A)) \cap C^k([0,T], H)$. Further, we also have the stability estimate

$$\|y^{(l)}(t)\|_H \le C\left(\|y_{0,l}\|_H + \|f^{(l)}\|_{L^1((0,t),H)}\right), \tag{24}$$

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

For $\ell \in \mathbb{N}$ we define

$$X^n = X_N^n \times \cdots \times X_N^n \quad (\ell \text{ factors}) \tag{25}$$

as well as

$$\mathcal{P}^n\colon L^\infty(D)^\ell \to X^n, \quad \mathcal{P}^n u = (\mathcal{P}_N^n u_1, \dots, \mathcal{P}_N^n u_\ell), \tag{26}$$

where X_N^n and $\mathcal{P}_N^n$ are given by the cardinal B-spline spaces in (44) and (46), respectively. The convergence properties of $\{\mathcal{P}^n\}_n$, see Appendix B, then guarantee (5) according to the next lemma.

Lemma 3.3. *Let X^n and $\mathcal{P}^n$ be as above and assume that F is defined for $f \in W^{1,1}([0,T],H)$ and $y_0 \in \mathcal{D}(A)$. Then we have that*

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

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

Proof. Let $\overline{y}_n := F'(u)\big(\widehat{u} - \mathcal{P}^n\widehat{u}\big)$ be the unique weak solution of

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

which satisfies according to (23) for all $0 \le t \le T$

$$\|F'(u)\big(\widehat{u} - \mathcal{P}^n\widehat{u}\big)(t)\|_H \le C\|[V'(u)\big(\widehat{u} - \mathcal{P}^n\widehat{u}\big)](y' + Qy)\|_{L^1((0,T),H)}.$$

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 there exists a subsequence such that $\mathcal{P}^{n_k}\widehat{u} \to \widehat{u}$ pointwise a.e.. Using $\|\mathcal{P}^{n_k}\widehat{u}\|_{L^\infty(D)^\ell} \le C_{\mathcal{P}}\|\widehat{u}\|_{L^\infty(D)^\ell}$ and (22), we can apply the dominated convergence theorem for integration in time which yields $\|[V'(u)\big(\widehat{u}-\mathcal{P}^{n_k}\widehat{u}\big)](y'+Qy)\|_{L^1([0,T],H)} \to 0$ as $k \to \infty$. By uniqueness of the pointwise limit $\widehat{u}$ the latter convergence even holds for the whole sequence, see [28, Prop. 10.13(2)]. □

Unfortunately, the TCC, which is the remaining condition for the rigorous applicability of $\texttt{REGINN}^\infty$, is subject of current research in the context of FWI and only special cases are known to hold. For example, the TCC has recently been shown for a semi-discrete setting in the acoustic regime which does, however, not meet our requirements, see [4]. To conclude our discussion about the rigorous scope of $\texttt{REGINN}^\infty$, we mention that condition (14) from Corollary 2.4, which guarantees the regularization property, was proven in [13] for the visco-elastic case.

A stronger result for the acoustic regime. In the remainder of this section we focus on the stronger compatibility condition (18). For the acoustic wave equation and a special choice of X^n and X, see (17), we will specify the decay function β.

In our abstract formulation (21), the acoustic wave equation is represented when setting $u = (\rho, \nu)$, $y = (p, v) \in L^2([0,T],H)$, $H = L^2(D) \times L^2(D,\mathbb{R}^d)$ and

$$Ay = \begin{pmatrix} \operatorname{div}(v) \\ \nabla p \end{pmatrix}, \quad B^{-1}y = \begin{pmatrix} \frac{1}{\rho\nu^2}p \\ \rho v \end{pmatrix}, \quad Q = 0, \tag{27}$$

subject to

$$\mathcal{D}(A) = \big\{(p,v) \in H_0^1(D) \times L^2(D,\mathbb{R}^d) :\ \operatorname{div}(v) \in L^2(D,\mathbb{R}^d)\big\}. \tag{28}$$

The two parameter functions are the bulk density ρ and the pressure wave speed ν taken from the set

$$\begin{aligned} \mathcal{D}(F) = \big\{(\rho,\nu) \in L^\infty(D)^2 : 0 < \rho_{\min} \le \rho \le \rho_{\max} < \infty \ \text{ and} \\ 0 < \nu_{\min} \le \nu \le \nu_{\max} < \infty \ \text{ a.e. in } D\big\} \end{aligned} \tag{29}$$

which is the domain of definition of the corresponding parameter-to-solution map.

Our discretization space $X^n = X_1^n \times X_1^n$ from (25) is now specified by $\ell = 2$ and $N = 1$ and thus consists of piecewise constant functions. The associate projector (46) reads

$$\mathcal{P}^n u = \sum_{k\in\mathcal{I}_n} 2^{-nd} \left(\int_{\square_k^n} u(x)\,\mathrm{d}x \right) \mathbb{1}_{\square_k^n} \tag{30}$$

where

$$\square_k^n := 2^{-n}([0,1]^d + k) \tag{31}$$

is the translated and dilated unit cube. Further, $\mathcal{I}_n = \{k \in \mathbb{Z}^d : \square_k^n \subset \overline{D}\}$, see (45).

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 (18). For $s > 0$ fixed we set

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

whose component spaces are characterized by

$$L_s^\infty(D) := \left\{ u \in L^\infty(D) :\ \sup_{h\neq 0} \frac{\|u(\cdot - h) - u\|_{L^2(D^h)}}{|h|^s} < \infty \right\} \tag{32}$$

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 magnitude of the sup within (32). Originally, $[\cdot]_{B_{2,\infty}^s(D)}$ emerges from the definition of Hilbertian Besov-Nikolskii spaces

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

with

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

cf. [26]. We obviously have $\|\cdot\|_{L^\infty(D)^2} \le \|\cdot\|_X$ for

$$\|u\|_X^2 := \|\rho\|_{L_s^\infty(D)}^2 + \|\nu\|_{L_s^\infty(D)}^2,$$

$L_s^\infty(D) \subset B_{2,\infty}^s(D)$ and $X_1^n \subset L_s^\infty(D)$ if $s \le 1/2$. The latter can be seen for $|h| < 2^{-n}$ with $h = (h_1, \dots, h_d)$ and $h_i \ge 0$ without loss of generality thanks to symmetry of the cube by the estimate

$$\begin{aligned}
\|\mathbb{1}_{-h+\square_k^n} - \mathbb{1}_{\square_k^n}\|_{L^2(D^h)}^2 &\le 2\,\mathrm{vol}_d \left(\big[0, 2^{-n}\big]^d \backslash \big[0, 2^{-n} - h_1\big] \times \dots \times \big[0, 2^{-n} - h_d\big] \right) \\
&= 2\Big(2^{-nd} - \prod_{i=1}^d \big(2^{-n} - h_i\big)\Big) \\
&\le 2\Big(2^{-nd} - \big(2^{-n} - |h|\big)^d\Big) \\
&\le d2^{-n(d-1)+1}|h|
\end{aligned}$$

for all $k \in \mathcal{I}_n$, where we used the mean value theorem in the last step. We have a first approximation result.

Lemma 3.4. *Let D be a bounded Lipschitz domain and let $\mathcal{P}^n$ be as in (30) for general $\ell \in \mathbb{N}$. Then, for any $u \in L_s^\infty(D)^\ell$ and $0 < s \leq 1/2$,*

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

where C only depends on D and the dimension d.

Proof. It suffices to prove the assertion for $\ell = 1$ since the case $\ell > 1$ follows by a componentwise consideration. Let $\square \subset \mathbb{R}^d$ be a sufficiently large rectangle containing D. By [22] there exists $\widetilde{u} \in B_{2,\infty}^s(\square)$ such that $\widetilde{u}|_D = u$ and $\|\widetilde{u}\|_{B_{2,\infty}^s(\square)} \leq \widetilde{C}\|u\|_{B_{2,\infty}^s(D)}$, where $\|u\|_{B_{2,\infty}^s(D)}$ can be replaced by the stronger norm $\|u\|_{L_s^\infty(D)}$. Using a dyadic partition of $\square$ at level n based on our cubes $\{\square_k^n\}_k$ (for which we restrict $\square$ to have integer side length), we have

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

see [1], where $\widetilde{\mathcal{P}}^n$ is defined as in (30) 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 $\mathcal{P}^n u|_{D_n} = \widetilde{\mathcal{P}}^n\widetilde{u}|_{D_n}$ and $\mathcal{P}^n u|_{D\setminus D_n} = 0$. The latter implies

$$\|u - \mathcal{P}^n u\|_{L^2(D\setminus D_n)} = \|u\|_{L^2(D\setminus D_n)} \leq \sqrt{\mathrm{vol}_d(D\setminus D_n)}\|u\|_{L^\infty(D)} \leq \sqrt{\mathrm{vol}_d(D\setminus D_n)}\|u\|_{L_s^\infty(D)}$$

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 $D\setminus D_n \subset \cup_{x\in\partial D}\big(x + 2^{-n}[-1,1]^d]\big)$, 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

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

which proves the lemma. $\square$

If $F(u)$ admits higher integrability in space for all u in a neighborhood of the exact solution u^+, we can use the latter approximation result to verify (18) in quite a general fashion.

Lemma 3.5. *Let $X = L_s^\infty(D)^\ell$ for some $0 < s \leq 1/2$ and let $u^+ \in \mathcal{D}(F)$ be some interior point such that $\|y' + Qy\|_{L^1([0,T],H^q)} < C'$ for all $\|u - u^+\|_{L^\infty(D)^\ell} < r'$, where $y = F(u)$ is defined by (21) and H^q is the L^q-version of our L^2-based Hilbert space H for some $q > 2$, that is in the acoustic case*

$$H^q := L^q(D) \times L^q(D, \mathbb{R}^d). \tag{33}$$

Then, on the basis of (26) and (30), we have for all $\widehat{u} \in X$ and u in a neighborhood of u^+ that

$$\|F'(u)\big(\widehat{u} - \mathcal{P}^n\widehat{u}\big)\|_{L^2([0,T],H)} \leq C^+\|\widehat{u}\|_X\big(2^{-s(q-2)/(3qd-4d)}\big)^n. \tag{34}$$

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

$$\|F'(u)\big(\widehat{u}-\mathcal{P}^n\widehat{u}\big)(t)\|_{L^2([0,T],H)} \le C\|V'(u)\big(\widehat{u}-\mathcal{P}^n\widehat{u}\big)(y'+Qy)\|_{L^1((0,T),H)},$$

compared to Lemma 3.3. Recall that the constant C depends continuously on the operator norm of $B=V(u)$, B^{-1}, Q and on T according to (23). Since V and $B\mapsto B^{-1}$ are (locally) continuous, we can assume without loss of generality that the above inequality holds for some fixed $C=C_{r',u^+}>0$ for any $\|u-u^+\|_{L^\infty(D)^\ell}<r'$ by shrinking the original $r'>0$ otherwise. Then, for any $\delta>0$,

$$\begin{aligned}
&\|V'(u)\big(\widehat{u}-\mathcal{P}^n\widehat{u}\big)(y'+Qy)\|_{L^1((0,T),H)}\\
&=\left\|V'(u)\left(\frac{\widehat{u}-\mathcal{P}^n\widehat{u}}{\|\widehat{u}\|_X}\right)\big((y'+Qy)(\mathbb{1}_{\{|\widehat{u}-\mathcal{P}^n\widehat{u}|\ge\delta\|\widehat{u}\|_X\}}+\mathbb{1}_{\{|\widehat{u}-\mathcal{P}^n\widehat{u}|<\delta\|\widehat{u}\|_X\}})\big)\right\|_{L^1((0,T),H)}\|\widehat{u}\|_X\\
&\le\|V'(u)\|_{\mathcal{L}(L^\infty(D)^\ell,\mathcal{L}(H))}\frac{\|\widehat{u}-\mathcal{P}^n\widehat{u}\|_{L^\infty(D)^\ell}}{\|\widehat{u}\|_X}\left\|(y'+Qy)\mathbb{1}_{\{|\widehat{u}-\mathcal{P}^n\widehat{u}|\ge\delta\|\widehat{u}\|_X\}}\right\|_{L^1((0,T),H)}\|\widehat{u}\|_X\\
&\quad+\|V'(u)\|_{\mathcal{L}(L^\infty(D)^\ell,\mathcal{L}(H))}\delta\left\|(y'+Qy)\mathbb{1}_{\{|\widehat{u}-\mathcal{P}^n\widehat{u}|<\delta\|\widehat{u}\|_X\}}\right\|_{L^1((0,T),H)}\|\widehat{u}\|_X.
\end{aligned}$$

Due to $\|\widehat{u}-\mathcal{P}^n\widehat{u}\|_{L^\infty(D)^\ell}\le(1+C_{\mathcal{P}})\|\widehat{u}\|_{L^\infty(D)^\ell}\le(1+C_{\mathcal{P}})\|\widehat{u}\|_X$ by (S2) and $C_X=1$ in (17), we obtain by dropping the complementary indicator function in the bottom line above

$$\begin{aligned}
&\|V'(u)\big(\widehat{u}-\mathcal{P}^n\widehat{u}\big)(y'+Qy)\|_{L^1((0,T),H)}\\
&\le(1+C_{\mathcal{P}})\|V'(u)\|_{\mathcal{L}(L^\infty(D)^\ell,\mathcal{L}(H))}\left\|(y'+Qy)\mathbb{1}_{\{|\widehat{u}-\mathcal{P}^n\widehat{u}|\ge\delta\|\widehat{u}\|_X\}}\right\|_{L^1((0,T),H)}\|\widehat{u}\|_X\\
&\quad+\|V'(u)\|_{\mathcal{L}(L^\infty(D)^\ell,\mathcal{L}(H))}\delta\,\|y'+Qu\|_{L^1((0,T),H)}\|\widehat{u}\|_X.
\end{aligned}$$

The middle line here can also be directly expressed in terms of δ according to our higher integrability assumption and applying Hölder's inequality with exponent $q/2>1$ to

$$\begin{aligned}
&\left\|(y'+Qy)(\mathbb{1}_{\{|\widehat{u}-\mathcal{P}^n\widehat{u}|\ge\delta\|\widehat{u}\|_X\}}\right\|_{L^1((0,T),H)}\\
&=\int_0^T\left(\int_D\left|y'(t)+Qy(t)\right|^2(x)\mathbb{1}_{\{|\widehat{u}-\mathcal{P}^n\widehat{u}|\ge\delta\|\widehat{u}\|_X\}}(x)\,\mathrm{d}x\right)^{1/2}\mathrm{d}t\\
&\le\int_0^T\left(\left\|y'(t)+Qy(t)\right\|_{H^q}^2\operatorname{vol}\big(\{|\widehat{u}-\mathcal{P}^n\widehat{u}|\ge\delta\|\widehat{u}\|_X\}\big)^{(q-2)/(2q)}\right)^{1/2}\mathrm{d}t\\
&=\|y'+Qy\|_{L^1((0,T),H^q)}\operatorname{vol}\big(\{|\widehat{u}-\mathcal{P}^n\widehat{u}|\ge\delta\|\widehat{u}\|_X\}\big)^{(q-2)/q}\\
&\le\|y'+Qy\|_{L^1((0,T),H^q)}\left(\frac{\|\widehat{u}-\mathcal{P}^n\widehat{u}\|^2_{L^2(D)^\ell}}{\delta^2\|\widehat{u}\|_X^2}\right)^{(q-2)/q}.
\end{aligned}$$

In the last step we employed Tschebyscheff's inequality, see, e.g., [7]. Applying Lemma 3.4 componentwise, we can further estimate

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

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

$$\|y'+Qy\|_{L^1((0,T),H)}\le\|y'+Qy\|_{L^1((0,T),H^q)}\operatorname{vol}_d(D)^{(q-2)/q}$$

that

$$\|V'(u)\big(\widehat{u}-\mathcal{P}^n\widehat{u}\big)(y'+Qy)\|_{L^1((0,T),H)}$$

$$\leq \|\widehat{u}\|_X \|V'(u)\|_{\mathcal{L}(L^\infty(D)^\ell,\mathcal{L}(H))} \|y' + Qy\|_{L^1((0,T),H^q)}$$
$$\times \left((1 + C_{\mathcal{P}}) \left(\frac{C\left(2^{-(s/d)}\right)^n}{\delta^2} \right)^{(q-2)/q} + \delta \operatorname{vol}_d(D)^{(q-2)/q} \right)$$

which holds for all $\delta > 0$. Optimization in δ then yields $\delta \propto \left(2^{-s(q-2)/(3dq-4d)}\right)^n$. Since $\|y' + Qy\|_{L^1((0,T),H^q)} < C'$ in a neighborhood of u^+ and V' is continuous by assumption, we can indeed find $C^+ < \infty$ in (34) which proves the lemma. $\square$

Finally, we show that under reasonable modeling assumptions the necessary higher integrability condition from the previous lemma can indeed be fulfilled in the acoustic regime.

Lemma 3.6. *Let D be C^1-domain with $d \in \{1, 2, 3\}$ and assume that $f = (f_1, f_2) \in W^{2,1}([0,T], H)$ with $f_2 \in C([0,T], L^{\widetilde{q}}(D, \mathbb{R}^d))$ for some $\widetilde{q} > 2$ such that $f(0) = f'(0) = 0$ and $y_0 = 0$. Then for any interior point $u^+ \in \mathcal{D}(F)$ there exists a neighborhood such that corresponding solutions $y = F(u)$ to*

$$\begin{aligned} \frac{1}{\rho\nu^2} p' - \operatorname{div}(v) &= f_1 \quad \text{in } D \times (0,T), \\ \rho v' - \nabla p &= f_2 \quad \text{in } D \times (0,T), \end{aligned} \tag{35}$$

satisfy $\|y'\|_{L^1([0,T],H^q)} < C' < \infty$ and $\widetilde{q} \geq q > 2$ only depends on D and the ratio $\rho_{\max}/\rho_{\min} < \infty$ from the definition of $\mathcal{D}(F)$.

Proof. The proof makes use of converting higher time regularity to higher spatial integrability. By Lemma 3.2 and our compatible source and initial data, we know that for $u^+ \in \mathcal{D}(F)$ we have $y' = (p', v') \in C([0,T], \mathcal{D}(A)) \cap C^1([0,T], H)$, in particular $p' \in L^1([0,T], H^1(D))$. By Sobolev embedding, see [5], we obtain at least $p' \in L^1([0,T], L^6(D))$ for $d \in \{1, 2, 3\}$. Since the constant in (24) depends continuously on B and thus on u, we can actually conclude $\|p'\|_{L^1([0,T],L^6(D))} < \infty$ uniformly in a neighborhood of u^+. Concerning v', we only have integrability information about its divergence. Therefore, we first note that by (35) we have

$$\operatorname{div}\left(\frac{1}{\rho}\nabla p\right) = \operatorname{div}(v') = \frac{p'' - f_1'}{\rho\nu^2} + \operatorname{div}(f_2)$$

in the sense of distributions. As $(\rho, \nu) \in \mathcal{D}(F)$ is bounded away from 0 uniformly, the non-divergence summand on the right hand side is in $C([0,T], L^2(D))$ by assumption. Hence by well-posedness of the Laplace equation in free space for fixed $t \in [0,T]$, there exists $g = g(t) \in H^1_{\mathrm{loc}}(\mathbb{R}^d, \mathbb{R}^d)$ such that $\operatorname{div}(g) = \mathbb{1}_D (p'' - f_1')/(\rho\nu^2)$ and

$$\|g\|_{H^1(D,\mathbb{R}^d)} \leq C \left\| \frac{p'' - f_1'}{\rho\nu^2} \right\|_{L^2(D)}$$

for some d- and D-dependent constant $C > 0$. Again, by Sobolev embedding we obtain $g \in L^6(D, \mathbb{R}^d)$ and the second order equation for p reduces to

$$\operatorname{div}\left(\frac{1}{\rho}\nabla p\right) = \operatorname{div}(f_2 + g)$$

with $(f_2 + g) \in L^{\min\{\widetilde{q},6\}}(D, \mathbb{R}^d)$. Now Meyers' estimate, see [18], implies that

$$\|\nabla p\|_{L^q(D,\mathbb{R}^d)} \leq C \|f_2 + g\|_{L^q(D,\mathbb{R}^d)}$$

for some $\min\{\widetilde{q}, 6\} \geq q > 2$ and a constant C both of which only depend on D and $\rho_{\max}/\rho_{\min} < \infty$. In particular, $\|\nabla p\|_{L^1([0,T],L^q(D,\mathbb{R}^d))} < \infty$ uniformly in a neighborhood of u^+ and using (35) we finally get that also $\|v'\|_{L^1([0,T],L^q(D,\mathbb{R}^d))} < \infty$ locally uniform in u. This completes the proof. □

4. Numerical Results

We present numerical experiments[2] on multi-parameter reconstruction to demonstrate the operation of $\texttt{REGINN}^\infty$ in an easy test scenario where all assumptions required for our analysis in the previous sections are satisfied.

Recall from Corollary 2.4 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 $T = 1$:

$$\begin{aligned} \frac{1}{\rho\nu^2}p' - \partial_x v &= f_1 \quad \text{in } (0,1)\times(0,1),\\ \rho v' - \partial_x p &= f_2 \quad \text{in } (0,1)\times(0,1),\\ v(0,\cdot) = p(0,\cdot) &= 0 \quad\ \ \text{on } (0,1),\\ p(\cdot,0) = p(\cdot,1) &= 0 \quad\ \ \text{on } (0,1). \end{aligned} \tag{36}$$

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

$$\begin{aligned} f_1(t,x) &= 100\Big(x(x-1)\frac{1}{\rho(x)\nu(x)^2} - \frac{\pi}{2}\cos\left(\frac{\pi}{2}x\right)\Big),\\ f_2(t,x) &= 100\Big(-t(2x-1) + \sin\left(\frac{\pi}{2}x\right)\rho(x)\Big), \end{aligned} \tag{37}$$

where

$$\rho(x) = 1 + \frac{1}{5}\,\mathbb{1}_{[7/30,17/30]}(x) \quad \text{and} \quad \nu(x) = 1 - \frac{1}{10}\,\mathbb{1}_{[13/30,23/30]}(x). \tag{38}$$

The corresponding exact data, that is, the solution of (36) and (37), are given by

$$p(t,x) = 100tx(x-1) \quad \text{and} \quad v(t,x) = 100t\sin\left(\frac{\pi}{2}x\right). \tag{39}$$

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. So, conditions (S1)-(S3) are fulfilled. Note

[2]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 Figure 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.

that the dimension of X_1^n is 2^n. In view of Remark 2.3 we set $n_{\max} = 8$ yielding the semi-discrete parameter-to-state map

$$F_{n_{\max}} \colon \mathcal{D}(F_{n_{\max}}) \subset X^{n_{\max}} \to L^2([0,1], H), \quad (\rho, \nu) \mapsto (p, v), \tag{40}$$

where $(p, v) \in L^2([0,1], H)$, $H = L^2(0,1)^2$, solves (36) with (37) and $(\rho, \nu) \in \mathcal{D}(F_{n_{\max}}) = X^{n_{\max}} \cap \mathcal{D}(F)$, see (29) 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 $\texttt{REGINN}^\infty$ applied to the inverse problem

$$\text{find } (\rho, \nu) \in X^{n_{\max}} : \quad F_{n_{\max}}(\rho, \nu) \approx (p^\delta, v^\delta). \tag{41}$$

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 $\texttt{REGINN}^\infty$ (Algorithm 1) to (41) where we choose $q_n = n / \log_2 C_\infty$ for $J_{n,m}$ from (9) in accordance with the lower bound in (47) 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 in a loop over n until (8) is met. We adapt μ_m during iteration according to the rule proposed in [21]: we start with $\mu_1 = \mu_0$ and set

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

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 $x_{m+1} = x_m + \widetilde{s}_m$ with the last update candidate $\widetilde{s}_m \in X^{n_{\max}}$ before abortion. We emphasize that $\widetilde{s}_m$ is a not Newton update in the sense of (8), but the corresponding x_{m+1} might fulfill the discrepancy principle unlike x_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 the error of any reconstruction of ρ and ν will always be at least $(\max\rho - \min\rho)/2 = 1/10$ and $(\max\nu - \min\nu)/2 = 1/20$ 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 (39). 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, ν) given by (38) 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 $n_{\max}$. Figure 1 displays the exact parameter functions ρ and ν (orange) 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

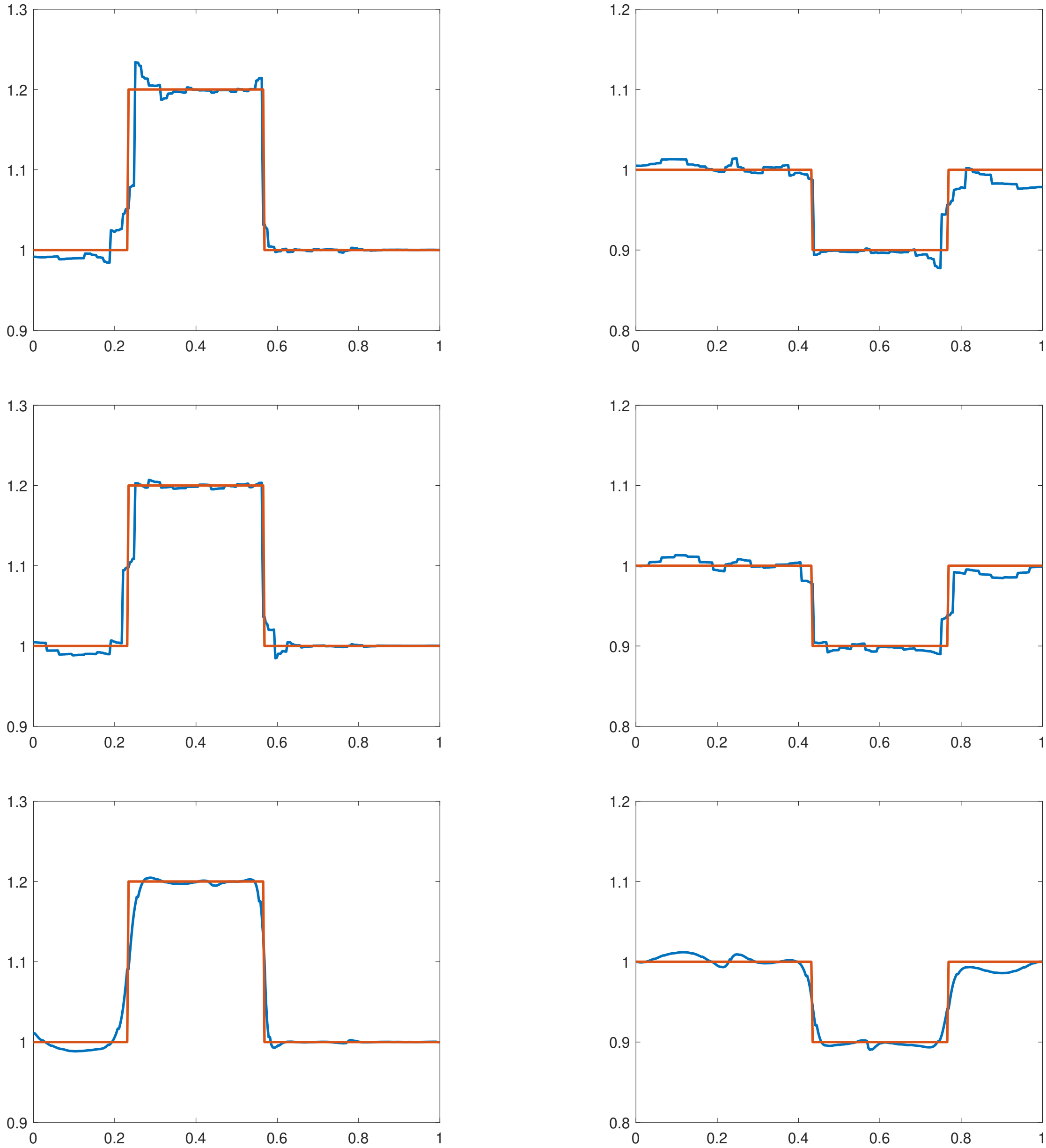


FIGURE 1. Approximate solutions ρ_M (blue, left column) and ν_M (blue, right column) by Algorithm 1 with initial spaces X^2, X^5, and X^8 (top to bottom).

chosen large enough to locate discontinuity points sufficiently precise while at the same time it should not be too large to prevent oversmoothing. 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 $\|\zeta\| = \delta\|(p, v)\|$. 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

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

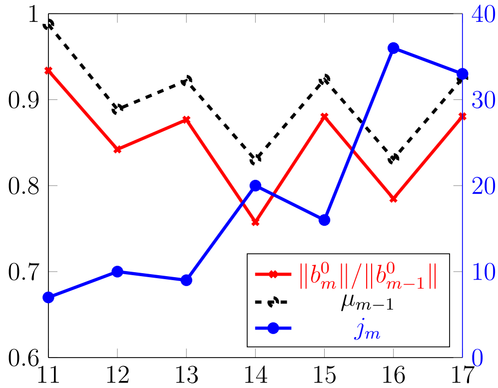


FIGURE 2. Left: Convergence history for exact data case $n_0 = 2$ from Figure 1. Peaks for j_m arise whenever the discretization level n_m is increased as cumulative contribution. Right: Graphical presentation of the values j_m (blue) and μ_{m-1} (black dashed) as functions of $m \in \{11, \dots, 17\}$ 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 (13) holds for a tiny ω.

reconstructions. The corresponding results for ρ_M and ν_M are shown in Figure 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 last update fulfills the discrepancy principle afterwards. Altogether, the plots illustrate the regularization property of $\texttt{REGINN}^\infty$ with respect to weak-$\star$ in $L^\infty(D)$ as ensured by Corollary 2.4.

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 $\texttt{REGINN}^\infty$ 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.

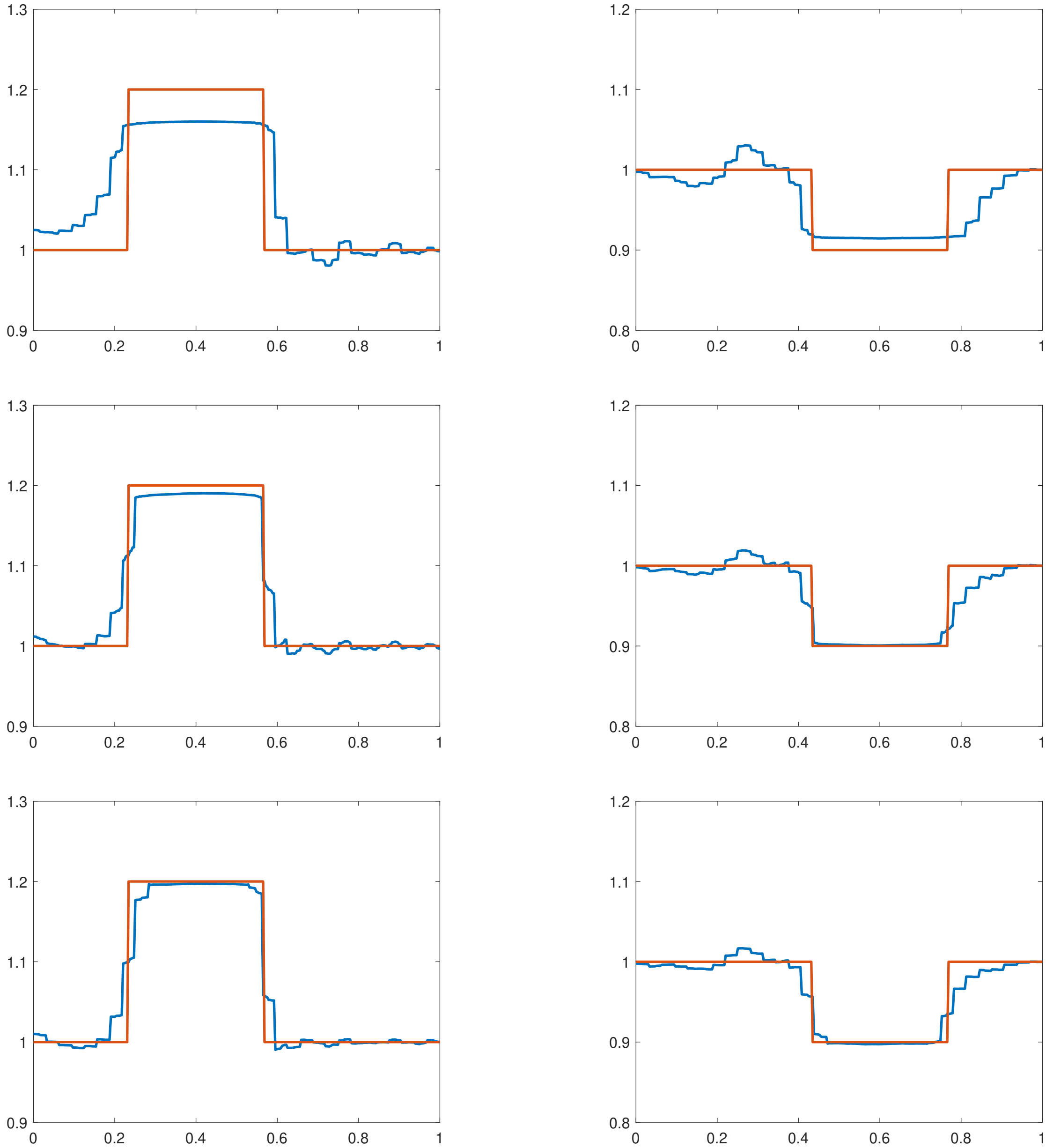


FIGURE 3. Regularized solutions ρ_M (blue, left column) and ρ_M (blue, right column) by Algorithm 1 for relative noise levels of 5%, 2% and 1% (top to bottom).

Future research may include a convergence rate analysis under higher regularity assumptions as in (18) or under more general variational source conditions with respect to a Bregman distance [10]. We could even incorporate a metric to overcome that $L^\infty(D)^\ell$ is non-separable; an approach proposed in [20]. 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., [16, 17]. This fact demands an implementation of our method under the KR-norm on Y. Moreover, our theory allows more general distance functions on Y. Indeed, any distance concept is admissible which is convex in one of its two arguments (e.g. Bregman distances).

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 Section 2 for D an open and bounded subset of $\mathbb{R}^d$. Using Cartesian products in case of $\ell > 1$, cf. (25), we restrict our attention to $\ell = 1$ here.

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

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

It obeys the scaling relation

$$\varphi_N(t) = 2^{1-N} \sum_{k=0}^{N} \binom{N}{k} \varphi_N(2t-k). \tag{42}$$

Further properties are

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

$$1 = \sum_{m\in\mathbb{Z}} \varphi_N(t-m), \tag{43}$$

see, e.g., [23].

Using the tensor product B-Spline $\Phi(x) := \prod_{i=1}^d \varphi_N(x_i)$, $x = (x_1,\dots,x_d)^\top \in \mathbb{R}^d$, 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

$$X^n = X^n_N(D) := \operatorname{span}\big\{\Phi_{n,k}|_D : k \in \mathcal{I}_n(D)\big\} \tag{44}$$

with

$$\mathcal{I}_n = \mathcal{I}_n(D) := \big\{k \in \mathbb{Z}^d : \operatorname{supp}\Phi_{n,k} \subset \overline{D}\big\}. \tag{45}$$

These spaces are nested due to (42), 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

$$\mathcal{P}^n u = \mathcal{P}^n_N u := \sum_{k\in\mathcal{I}_n} \langle u, \widetilde{\Phi}_{n,k}\rangle_{L^2(D)} \Phi_{n,k} \quad \text{ for } u \in L^\infty(D) \tag{46}$$

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 [3]. 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

$$\begin{aligned}\|\mathcal{P}^n_N u\|_{L^\infty(D)} &= \sup_{x\in D}\Big|\sum_{k\in\mathcal{I}_n} \langle u, \widetilde{\Phi}_{n,k}\rangle_{L^2(D)} \Phi_{n,k}(x)\Big| = \Big|\sum_{k\in\mathcal{I}_n} \langle u, \widetilde{\Phi}_{n,k}\rangle_{L^2(D)} \Phi_{n,k}(x^*)\Big| \\ &\le \sum_{k\in\mathcal{I}_n} |\langle u, \widetilde{\Phi}_{n,k}\rangle_{L^2(D)}|\, \Phi_{n,k}(x^*) \le \|u\|_{L^\infty(D)} \sum_{k\in\mathcal{I}_n} \|\widetilde{\Phi}_{n,k}\|_{L^1(D)}\, \Phi_{n,k}(x^*).\end{aligned}$$

Since

$$\|\widetilde{\Phi}_{n,k}\|_{L^1(D)} \le 2^{-nd/2} \int_{\mathbb{R}^d} |\widetilde{\Phi}(x)|\,\mathrm{d}x \quad \text{and} \quad \sum_{k\in\mathcal{I}_n} \Phi_{n,k}(x^*) \le \sum_{k\in\mathbb{Z}^d} \Phi_{n,k}(x^*) \overset{(43)}{=} 2^{nd/2}$$

we have established (S2) with $C_{\mathcal{P}} \le \|\widetilde{\Phi}\|_{L^1(\mathbb{R}^d)}$. Observe that $\|\widetilde{\Phi}\|_{L^1(\mathbb{R}^d)} \ge 1$ 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,

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

for

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

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 $\delta_q \le 1/C_\infty$ for any given $C_\infty > 1$. Hence, $1 \le C_\infty \|u\|_{L^q(D)}$ 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, $\delta_q = \|u_q\|_{L^q(D)} \le \mathrm{vol}_d(D)^{1/q} \to 1$ as $q \to \infty$. Therefore $\{\delta_q\}$ is bounded and admits a convergent subsequence, say, $\lim_{i\to\infty} \delta_{q_i} = \delta^* \le 1$. 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

$$\|u_q\|_{L^q(D)} \ge \left(2^{-nd}\right)^{1/q} \to 1 \tag{47}$$

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 $|u_q|_{B_\delta(x_q^*)\cap D}| \ge 1-\varepsilon$ 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

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

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

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

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

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

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., [28, Prop. 10.13(2)].

A different approach. The functions of $X_N^n(D)$ from the above construction vanish on the set $D \setminus \cup_{k\in\mathcal{I}_n} \operatorname{supp} \Phi_{n,k}$ which is non-empty in general. Here we present briefly an alternative approach to overcome this drawback. Basically, we extend the preimage-space of the map F of Section 2 while keeping all its necessary properties to carry over Theorem 2.1 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

$$\begin{aligned}\|\widetilde{F}(u) - \widetilde{F}(\widetilde{u}) - \widetilde{F}'(\widetilde{u})(u - \widetilde{u})\|_Y &= \|F(Ru) - F(R\widetilde{u}) - F'(R\widetilde{u})(Ru - R\widetilde{u})\|_Y \\ &\leq \omega \,\|F(Ru) - F(R\widetilde{u})\|_Y = \omega\,\|\widetilde{F}(u) - \widetilde{F}(\widetilde{u})\|_Y.\end{aligned}$$

For this $\widetilde{F}$ we can define spaces $X^n = X^n_N(\widetilde{D})$ as above 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.

Condition (5) does not transfer directly to the new construction but we have, for $\widetilde{u} \in \mathcal{D}(\widetilde{F})$ and $\widehat{u} \in L^\infty(\widetilde{D})$, that

$$\|\widetilde{F}'(\widetilde{u})\big(\widehat{u} - \widetilde{\mathcal{P}}^n_N\widehat{u})\|_Y \leq \|F'(R\widetilde{u})(R\widehat{u} - \mathcal{P}^n_N R\widehat{u})\|_Y + \|F'(R\widetilde{u})(\mathcal{P}^n_N R\widehat{u} - R\widetilde{\mathcal{P}}^n_N\widehat{u})\|_Y \tag{48}$$

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 (48) tends to 0 for $n \to \infty$ by (5). The right norm converges to 0, for instance, if $F'(x)\colon L^\infty(D) \to Y$ 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., [23, Chap. 12.3] and [19, 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 (46) 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 [23] 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 (46), however, with respect to $\mathcal{I}_n(\widetilde{\square})$ $X^n(\widetilde{\square})$. Hence, for any $s \in D$ 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

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

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

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

and the assertion follows. $\square$

Appendix C. On the tangential cone condition for the operator in (40)

Here we argue that the semi-discrete non-linear operator $\boldsymbol{F}_{n_{\max}}$ defined in (40) satisfies the TCC (3).

The underlying abstract system is (21) with the concrete settings (27), $u_0 = 0$, and (28) 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 (36) which has a unique classical solution under $(\rho, \nu) \in \mathcal{D}(\boldsymbol{F}_{n_{\max}})$ and for the sources (37). In view of Lemma 3.1, $\boldsymbol{F}_{n_{\max}}$ is Fréchet-differentiable and we have $\boldsymbol{F}'_{n_{\max}}(\rho,\nu)(\widehat{\rho},\widehat{\nu}) = (\overline{p},\overline{v})$ where $(\overline{p},\overline{v})$ weakly solves

$$
\begin{aligned}
\frac{1}{\rho\nu^2}\overline{p}' - \partial_x \overline{v} &= -\widehat{\rho}p' && \text{in } (0,1)\times(0,1),\\
\rho\overline{v}' - \partial_x \overline{p} &= -\widehat{\nu}v' && \text{in } (0,1)\times(0,1),\\
\overline{v}(0,\cdot) = \overline{p}(0,\cdot) &= 0 && \text{on } (0,1)\\
\overline{v}(\cdot,0) = \overline{p}(\cdot,1) &= 0 && \text{on } (0,1).
\end{aligned}
\tag{49}
$$

In a first step we validate injectivity of $\boldsymbol{F}'_{n_{\max}}(\rho,\nu)$ for any $(\rho,\nu) \in \mathcal{D}(\boldsymbol{F}_{n_{\max}})$. To this end, assume $\boldsymbol{F}'_{n_{\max}}(\rho,\nu)(\widehat{\rho},\widehat{\nu}) = (0,0)$. From (49) we get

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

Assume $0 \neq \widehat{\rho} \in \boldsymbol{X}^{n_{\max}}$. 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, $p' = 0$ in $[0,1]\times[a,b]$. By the first equation in (36), $-\partial_x v = f_1$ 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 $x \in [a,b]$. So, $\widehat{\rho} = 0$ in $(0,1)$. One argues analogously to validate $\widehat{\nu} = 0$ in $(0,1)$. Hence, $\boldsymbol{F}'_{n_{\max}}(\rho,\nu)$ is one-to-one which implies the TCC at any interior point of $\mathcal{D}(\boldsymbol{F}_{n_{\max}})$ due to Lemma C.1 of [4].

References

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

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

[3] A. Cohen, I. Daubechies, and J.-C. Feauveau, *Biorthogonal bases of compactly supported wavelets*, Comm. Pure Appl. Math., 45 (1992), pp. 485–560, https://doi.org/10.1002/cpa.3160450502.

[4] M. Eller and A. Rieder, *Tangential cone condition and Lipschitz stability for the full waveform forward operator in the acoustic regime*, Inverse Problems, 37 (2021), p. 085011, https://doi.org/10.1088/1361-6420/ac11c5.

[5] L. C. Evans, *Partial differential equations*, vol. 19 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, second ed., 2010, https://doi.org/10.1090/gsm/019.

[6] A. Fichtner, *Full Seismic Waveform Modelling and Inversion*, Advances in Geophysical and Environmental Mechanics and Mathematics, Springer-Verlag Berlin Heidelberg, 2011, https://doi.org/10.1007/978-3-642-15807-0.

[7] O. C. Ibe, *Markov processes for stochastic modeling*, Elsevier, Inc., Amsterdam, second ed., 2013, https://doi.org/10.1016/B978-0-12-407795-9.00001-3.

[8] A. N. Iusem and E. Resmerita, *A proximal point method in nonreflexive Banach spaces*, Set-Valued Var. Anal., 18 (2010), pp. 109–120, https://doi.org/10.1007/s11228-009-0126-z.
[9] B. Kaltenbacher and M. L. P. de Souza, *Convergence and adaptive discretization of the IRGNM Tikhonov and the IRGNM Ivanov method under a tangential cone condition in Banach space*, Numer. Math., 140 (2018), pp. 449–478, https://doi.org/10.1007/s00211-018-0971-5.
[10] B. Kaltenbacher and A. Klassen, *On convergence and convergence rates for Ivanov and Morozov regularization and application to some parameter identification problems in elliptic PDEs*, Inverse Problems, 34 (2018), pp. 055008, 24, https://doi.org/10.1088/1361-6420/aab739.
[11] B. Kaltenbacher, A. Klassen, and M. L. Previatti de Souza, *The Ivanov regularized Gauss-Newton method in Banach space with an a posteriori choice of the regularization radius*, J. Inverse Ill-Posed Probl., 27 (2019), pp. 539–557, https://doi.org/10.1515/jiip-2018-0093.
[12] B. Kaltenbacher, A. Neubauer, and O. Scherzer, *Iterative regularization methods for nonlinear ill-posed problems*, vol. 6 of Radon Series on Computational and Applied Mathematics, Walter de Gruyter GmbH & Co. KG, Berlin, 2008, https://doi.org/10.1515/9783110208276.
[13] A. Kirsch and A. Rieder, *Inverse problems for abstract evolution equations* II*: higher order differentiability for viscoelasticity*, SIAM J. Appl. Math., 79 (2019), pp. 2639–2662, https://doi.org/10.1137/19M1269403.
[14] A. Lechleiter and A. Rieder, *Towards a general convergence theory for inexact Newton regularizations*, Numer. Math., 114 (2010), pp. 521–548, https://doi.org/10.1007/s00211-009-0256-0.
[15] F. J. Margotti, *On Inexact Newton Methods for Inverse Problems in Banach Spaces*, PhD thesis, Karlsruher Institut für Technologie (KIT), 2015, https://doi.org/10.5445/IR/1000048606.
[16] J. Messud, R. Poncet, and G. Lambaré, *Optimal transport in full-waveform inversion: analysis and practice of the multidimensional Kantorovich–Rubinstein norm*, Inverse Problems, 37 (2021), p. 065012, https://doi.org/10.1088/1361-6420/abfb4c.
[17] L. Métivier, R. Brossier, Q. Mérigot, E. Oudet, and J. Virieux, *An optimal transport approach for seismic tomography: application to 3D full waveform inversion*, Inverse Problems, 32 (2016), pp. 115008, 36, https://doi.org/10.1088/0266-5611/32/11/115008.
[18] N. G. Meyers, *An L^p-estimate for the gradient of solutions of second order elliptic divergence equations*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3), 17 (1963), pp. 189–206.
[19] P. Oswald, *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. Theory and applications.
[20] C. Pöschl, E. Resmerita, and O. Scherzer, *Discretization of variational regularization in Banach spaces*, Inverse Problems, 26 (2010), pp. 105017, 18, https://doi.org/10.1088/0266-5611/26/10/105017.
[21] A. Rieder, *On the regularization of nonlinear ill-posed problems via inexact Newton iterations*, Inverse Problems, 15 (1999), pp. 309–327, https://doi.org/10.1088/0266-5611/15/1/028.
[22] V. S. Rychkov, *On restrictions and extensions of the Besov and Triebel–Lizorkin spaces with respect to Lipschitz domains*, Journal of the London Mathematical Society, 60 (1999), pp. 237–257, https://doi.org/10.1112/S0024610799007723.
[23] L. L. Schumaker, *Spline functions: basic theory*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, third ed., 2007, https://doi.org/10.1017/CBO9780511618994.
[24] T. Schuster, B. Kaltenbacher, B. Hofmann, and K. S. Kazimierski, *Regularization Methods in Banach Spaces*, De Gruyter, 2012, https://doi.org/10.1515/9783110255720.
[25] T. Schuster, A. Rieder, and F. Schöpfer, *The approximate inverse in action: IV. Semi-discrete equations in a Banach space setting*, Inverse Problems, 28 (2012), pp. 104001, 19, https://doi.org/10.1088/0266-5611/28/10/104001.
[26] H. Triebel, *Theory of function spaces.* II, vol. 84 of Monographs in Mathematics, Birkhäuser Verlag, Basel, 1992, https://doi.org/10.1007/978-3-0346-0419-2.
[27] J. Virieux and S. Operto, *An overview of full-waveform inversion in exploration geophysics*, Geophysics, 74 (2009), pp. 1–26, https://doi.org/10.1190/1.3238367.
[28] E. Zeidler, *Nonlinear functional analysis and its applications* I*: Fixed-point theorems*, Springer-Verlag, New York, 1986.

Department of Mathematics, Karlsruhe Institute of Technology (KIT), D-76128 Karlsruhe, Germany
Email address: lukas.pieronek@kit.edu
Email address: andreas.rieder@kit.edu