

Determining the nonlinearity in an acoustic wave equation

Barbara Kaltenbacher*

William Rundell†

Abstract

We consider an undetermined coefficient inverse problem for a nonlinear partial differential equation describing high intensity ultrasound propagation as widely used in medical imaging and therapy. The usual nonlinear term in the standard model using the Westervelt equation in pressure formulation is of the form pp_t . However, this should be considered as a low order approximation to a more complex physical model where higher order terms will be required. Here we assume a more general case where the form taken is $f(p)p_t$ and f is unknown and must be recovered from data measurements. Corresponding to the typical measurement setup, the overposed data consists of time trace observations of the acoustic pressure at a single point or on a one dimensional set Σ representing the receiving transducer array at a fixed time. Additionally to an analysis of well-posedness of the resulting PDE, we show injectivity of the linearized forward map from f to the overposed data and use this as motivation for several iterative schemes to recover f . Numerical simulations will also be shown to illustrate the efficiency of the methods.

Keywords: Inverse problem, nonlinear acoustics, reconstruction algorithms

Ams classification: 35R30, 35L15, 35L71.

1 Introduction

The use of ultrasound is well established in the imaging of human tissue and the propagation of high intensity ultrasound is modeled by nonlinear wave equations. Nonlinearity enters the model via the state equation, which is a constitutive relation between acoustic pressure and mass density. This results in the typical situation of the nonlinear effect appearing as a product of a function of the state variable and its time derivative. A common such model is the Westervelt equation in which the Taylor expansion of this constitutive relation is truncated to its second degree terms and in this case a certain ratio of quantities B/A governs the nonlinearity, cf [7, Chapter 2].

$$u_{tt} - c^2 \Delta u - b \Delta u_t = (\kappa u^2)_{tt} = (2\kappa u u_t)_t \quad (1)$$

where c is the wave speed, b a damping coefficient and $\kappa(x)$ is proportional to B/A and may depend on the spatial variable $x \in \mathbb{R}^d$. Its recovery from overposed data was recently investigated in [11].

In this paper, our focus will be on using a general nonlinear state equation and on identifying this nonlinearity from indirect measurements. In section 2 we will give more details on the above models which result from physical laws governing quantities such as the acoustic particle velocity, the acoustic pressure and the mass density. Combining these laws while including successive higher order terms,

*Department of Mathematics, Alpen-Adria-Universität Klagenfurt. barbara.kaltenbacher@aau.at.

†Department of Mathematics, Texas A&M University, Texas 77843. rundell@math.tamu.edu

(the Blackstock scheme) one arrives at a succession of more complex partial differential equations. In particular, instead of truncating the Taylor coefficients in the model we can leave the nonlinearity as a function thus arriving at the equation

$$u_{tt} - c^2 \Delta u - b \Delta u_t = (f_W(u)u_t)_t. \quad (2)$$

Now the question is whether we can go beyond the finite Taylor series paradigm of recovering a limited number of lower order coefficients and seek to recover the constitutive function f_W in (2).

We will work in a finite domain $\Omega \subset \mathbb{R}^d$ where Ω is a simply connected domain with smooth boundary $\partial\Omega$. It will turn out to be convenient as in [11] to rewrite (2) as a parabolic equation with a non-local memory term

$$u_t - b \Delta u - c^2 \int_0^t \Delta u(\tau) d\tau = f_W(u)u_t.$$

This was the method taken in [11] and relates the problem to be treated to the problem of the identification of an unknown nonlinear specific heat coefficient, see [13]. The traditional assumption that damping is proportional to velocity used in (2) can also be modified to include so-called fractional damping where the time derivative is replaced by a subdiffusion operator of order α as, e.g., in [2, 12] and the references therein. In this situation the integral form of the equation is more convenient as the the standard integral is then replaced by one of Abel fractional type. We assume that the operator $-\Delta$ on Ω is equipped with boundary conditions on $\partial\Omega$ and impose initial conditions.

Typical observations available in such nonlinear acoustic experiments are measurements of the acoustic pressure at an array of transducers or hydrophones. Thus, there are two obvious types of overposed data: one measuring in the spatial dimension and the other in the time domain. In the first case, for some subdomain or curve ω lying in Ω and for some fixed time T we measure

$$g(x) = u(x, T), \quad x \in \omega \subseteq \Omega \quad (3)$$

while in the second for some point $x_0 \in \overline{\Omega}$ we measure

$$h(t) = u(x_0, t), \quad 0 \leq t \leq T. \quad (4)$$

As in all problems where the unknown coefficient depends on the state variable u , we must impose range conditions in the above. Specifically, the values of u occurring in $\Omega \times [0, T]$ must be as subset of those on the measurement domain. This will naturally impose constraints on the type of initial and boundary conditions possible.

The outline of the paper is as follows. We will provide more details on the physical background leading to the equations in order to justify why the recovery of the term f_W represents a significant advancement over models relying on fixed Taylor expansions and (possibly) unknown coefficient terms. The section following will provide an analysis of the forward operator. Section 4 will give the two main reconstruction tools we will use; a (quasi) Newton scheme to recover f_W using the map $u(x, t; f_W)$ where this is projected onto the overposed boundary; fixed point schemes based on an iterative map of Picard type, where the latter will be sped up by Anderson acceleration [1, 4, 16]. We will provide an analysis of the forward problem and of well-definedness of the iteration schemes as well as the results of numerical experiments to demonstrate the effectiveness (and limitations) of these methods for the problem at hand.

2 The model

We first derive a generalized version of equation (1) that results from replacing the second order Taylor expansion underlying the conventional state equation by a general model function. Then following the usual elimination steps to obtain a nonlinear second order wave equation (see, e.g., [7]), we arrive at a generalized version of the Westervelt equation. To this end, we start with the fundamental quantities of acoustics, which are the following space and time dependent functions

- acoustic particle velocity $\vec{v}$;
- acoustic pressure p ;
- mass density ρ ;

which can be decomposed into their constant mean and a fluctuating part

$$\vec{v} = \vec{v}_0 + \vec{v}_\sim, \quad p = p_0 + p_\sim, \quad \rho = \rho_0 + \rho_\sim,$$

where $\vec{v}_0 = 0$ in the absence of a flow.

These quantities are interrelated by the following physical balance and material laws:

- balance of momentum

$$\rho \left(\vec{v}_t + \nabla |\vec{v}|^2 \right) + \nabla p = \tilde{\mu} \Delta \vec{v}; \quad (5)$$

- balance of mass

$$\nabla \cdot (\rho \vec{v}) = -\rho_t; \quad (6)$$

- state equation relating the acoustic pressure and density fluctuations $p_\sim$ and $\rho_\sim$: Only in special cases is this known explicitly, an example being perfect gases, with $\frac{p}{\rho} = \left(\frac{\rho}{\rho_0} \right)^\gamma$ where γ is the adiabatic index (also known as the ratio of specific heats). The commonly used substitute for obtaining the classical Kuznetsov or Westervelt equations of nonlinear acoustics is

$$\rho_\sim = \frac{p_\sim}{c^2} - \frac{1}{\rho_0 c^4} \frac{B}{2A} p_\sim^2 - \frac{\tilde{\kappa}}{c^4} p_{\sim t}. \quad (7)$$

The first two terms in (7) are basically just a polynomial ansatz obtained from a Taylor expansion. We only keep the linear terms, that is, the first term (as this needs to be there to yield a wave type equation) and the last term (as this accounts for attenuation) and generalize the nonlinearity to be an arbitrary function $\tilde{f}$ in

$$\rho_\sim = \frac{1}{c^2} p_\sim - \frac{1}{c^4} \tilde{f}(p_\sim) - \frac{\tilde{\kappa}}{c^4} p_{\sim t}, \quad (8)$$

Subtracting the divergence of (5) from the time derivative of (6) gives

$$\rho_{tt} - \Delta p = -\nabla \cdot \left(\rho_t \vec{v} - \rho \nabla |\vec{v}|^2 + \tilde{\mu} \Delta \vec{v} \right).$$

Inserting (8), and using $\rho_{0,t} = 0$, $\nabla \rho_0 = 0$, $p_{0,t} = 0$, $\nabla p_0 = 0$ as well as the approximations $\nabla \cdot (\rho_{\sim,t} \vec{v}) - \nabla \rho_\sim \cdot \nabla |\vec{v}|^2 \approx 0$, $p_{\sim,ttt} \approx c^2 \Delta p_{\sim,t}$, $\rho_0 c^2 \nabla \cdot \Delta \vec{v} \approx -\Delta p_{\sim,t}$, $\rho_0 c^2 \Delta |\vec{v}|^2 \approx \rho_0 |\vec{v}|_{tt}^2 \approx \frac{1}{\rho_0^2 c^2} (p^2)_{tt}$ that are usually applied in the derivation of the Westervelt equation, setting $b = \kappa + \frac{\mu}{\rho_0}$ we arrive at

$$p_{\sim,tt} - c^2 \Delta p_\sim - b \Delta p_{\sim,t} = \frac{1}{c^2} \left(\frac{1}{\rho_0^2} p_\sim^2 + \tilde{f}(p_\sim) \right)_{tt}$$

Now skipping the subscripts $\sim$, and with the abbreviation $f(p) = \frac{2}{\rho_0^2}p + \tilde{f}'(p)$, we arrive at the following generalisation of the Westervelt equation

$$p_{tt} - c^2 \Delta p - b \Delta p_t = \frac{1}{c^2} (f(p) p_t)_t. \quad (9)$$

It is sometimes convenient (e.g., for computational purposes and extending the model to fractional operators) to transform this to a parabolic equation with memory by applying time integration and using homogeneous initial conditions as well as $\tilde{f}(0) = 0$

$$p_t - b \Delta p - c^2 \int_0^t \Delta p(\tau) d\tau = \frac{1}{c^2} f(p) p_t. \quad (10)$$

Combining the right hand term with the very first on the left hand side we can rewrite this as

$$(1 - \frac{1}{c^2} f(p)) p_t - b \Delta p - c^2 \int_0^t \Delta p(\tau) d\tau = 0 \quad (11)$$

and so the problem is related to the identification of an unknown nonlinear specific heat coefficient, see [13].

Here one might wish to keep track of c^2 , first of all because of its different order of magnitude as compared to b and 1, and secondly in order to possibly consider simultaneous identification of c and f . In this paper, for simplicity of exposition we just merge it into f by replacing $f \leftarrow \frac{1}{c^2} f$. Moreover, similarly to [11], we will add a driving term $r = r(x, t)$ on the right hand side of the PDE that is supposed to model excitation and thus arrive at

$$p_{tt} - c^2 \Delta p - b \Delta p_t = (f(p) p_t)_t + r. \quad (12)$$

and equivalently

$$(1 - f(p)) p_t - b \Delta p - c^2 \int_0^t \Delta p(\tau) d\tau = \tilde{r} = \int_0^t r(\tau) d\tau + (1 - f(p_0)) p_1 - b \Delta p_0. \quad (13)$$

Extension to fractional damping

In case of fractional damping (more precisely, the Caputo-Wisner model see, e.g., [2, 12], we replace (12) by

$$p_{tt} - c^2 \Delta p - b \Delta D_t^\alpha p = \frac{1}{c^2} \left(\frac{1}{\rho_0^2} p^2 + \tilde{f}(p) \right)_t + r \quad (14)$$

or after time integration and using homogeneous initial conditions (so it does not matter whether D_t^α is the Djirbashian-Caputo or the Riemann-Liouville fractional derivative) as well as $\tilde{f}(0) = 0$ and again using $f(p) = \frac{1}{c^2} (\frac{2}{\rho_0^2} p + \tilde{f}'(p))$

$$(1 - f(p)) p_t - b \Delta I_t^{1-\alpha} p - c^2 \int_0^t \Delta p(\tau) d\tau = \tilde{r} \quad (15)$$

which can be rewritten as

$$(1 - f(p)) p_t - \Delta \int_0^t \left(\frac{b}{\Gamma(1-\alpha)} (t - \tau)^{-\alpha} + c^2 \right) p(\tau) d\tau = \tilde{r} \quad (16)$$

The inverse problem

The modeling task of reconstructing f in (12) from overposed data (3) or (4) thus amounts to inverting the forward map $F = \text{tr}_\Sigma \circ G$ which is a composition of the parameter-to-state-map $G : f \mapsto p =$

$p(x, t; f)$ of the initial boundary value problem for (12) with the trace operator on $\Sigma = \{x_0\} \times (0, T)$ in the time trace and $\Sigma = \omega \times \{T\}$ in the final time case. Before showing well-definedness of F in section 3 and devising some reconstruction schemes in section 4, we briefly comment on unique invertibility of its linearisation in the remainder of this section.

Linearised injectivity

Linearisation at $f = 0$ simplifies the equations to

$$\begin{aligned} z_{tt}^0 - c^2 \Delta z^0 - b \Delta z_t^0 &= (\underline{df}(p^0) p_t^0)_t, \\ z^0(0) &= 0, \quad z_t^0(0) = 0 \end{aligned} \quad (17)$$

where

$$\begin{aligned} p_{tt}^0 - c^2 \Delta p^0 - b \Delta p_t^0 &= r, \\ p^0(0) &= p_0, \quad p_t^0(0) = p_1. \end{aligned} \quad (18)$$

We can say more about injectivity as well as ill-posedness of the inverse problem linearised at $f = 0$ for specially chosen driving and initial functions $r(x, t) = \xi(x) \eta''(t) - b \Delta \xi(x) \eta'(t) - c \Delta \xi(x) \eta(t)$, $p_0(x) = \xi(x) \eta(0)$, $p_1(x) = \xi(x) \eta'(0)$ leading to $p^0(x, t) = \xi(x) \eta(t)$. Indeed, in the time trace case, setting $\xi = 1$, and choosing η such that $\eta(0) = 0$ and $\eta'(0) = 0$ we can explicitly write the (unique) solution of (49) as $z^0(x, t) = \underline{df}(\eta(t))$ where $\underline{df}(s) = \int_0^s \underline{df}(\tau) d\tau$, which gives the explicit reconstruction $\underline{df}(s) = \frac{h'(\eta^{-1}(s))}{\eta'(\eta^{-1}(s))}$. Ill-posedness therefore results from differentiation of the data and from $\eta'(0) = 0$.

3 Analysis of the forward problem

The question of well-definedness of the forward operator $F = \text{tr}_\Sigma \circ G$ amounts to proving existence of a unique solution $G(f) = p$ to

$$\begin{aligned} [(1 - f(p)) p_t]_t + c^2 \mathcal{A} p + b \mathcal{A} p_t &= r, \\ p(0) &= p_0, \quad p_t(0) = p_1 \end{aligned} \quad (19)$$

that is regular enough to admit a trace on $\Sigma = \{x_0\} \times (0, T)$ in the time trace and $\Sigma = \omega \times \{T\}$ in the final time case, respectively. Here, we denote by $\mathcal{A}$ the negative Laplacian with homogeneous Dirichlet boundary conditions. One could as well incorporate a potentially varying sound speed $c_0(x)$ and other boundary conditions $\mathcal{A}$ as we did in [11],

Setting $\psi(x, t) = \int_0^t p(x, \tau) d\tau$ so that $p(x, t) = \psi_t(x, t)$ and integrating (19) with respect to time we arrive at

$$\begin{aligned} (1 - f(\psi_t)) \psi_{tt} + c^2 \mathcal{A} \psi + b \mathcal{A} \psi_t &= \tilde{r}, \\ \psi(0) &= 0, \quad \psi_t(0) = p_0 \end{aligned} \quad (20)$$

where

$$\tilde{r}(x, t) = \int_0^t r(x, \tau) d\tau + (1 - f(p_0)) p_1(x) + b \mathcal{A} p_0(x). \quad (21)$$

This is a formulation that avoids differentiating f . We will see that this advantage can be preserved in the well-posedness proof at least in the one space dimensional setting. Physically, ψ corresponds to the acoustic velocity potential with the mass density scaled to unity.

To analyse the nonlinear problem (20), we first of all consider the linear one

$$\begin{aligned} (1 - \sigma)u_{tt} + c^2 \mathcal{A}u + b \mathcal{A}u_t - \eta u_t &= \tilde{r}, \\ u(0) &= u_0, \quad u_t(0) = u_1 \end{aligned} \quad (22)$$

with given $\sigma(x, t)$ bounded away from 1, $\sigma(x, t) \leq \bar{\sigma} < 1$ so that the coefficient of the second time derivative does not degenerate. Later on, we will set $\sigma = f(\psi_t)$, $\eta = 0$, $u_0 = 0$, $u_1 = p_0$ in order to prove well-posedness of (20) by a fixed point argument, as well as $\sigma = f(\psi_t)$, $\eta = f'(\psi_t)\psi_{tt}$ to investigate its linearization. Existence of a unique solution to (22) can be established by the usual Faedo-Galerkin approach of discretisation in space with eigenfunctions of $\mathcal{A}$, deriving energy estimates and taking weak limits. We here only focus on the energy estimates for (22). Some of these can be found scattered over in the cited literature for the special case $f(p) = \kappa p$, however, they will here be tailored to minimize the regularity assumptions of f , which is why we provide their derivation in the appendix.

Lemma 3.1. • If $u_1 \in H_0^1(\Omega)$, $\tilde{r} \in L^2(0, T; L^2(\Omega))$, $\eta \in L^2(0, T; L^q(\Omega))$ with

$$C(\eta) = \begin{cases} (C_{H^1, L^\infty}^\Omega)^2 \|\eta\|_{L^2(0, T; L^2(\Omega))}^2 & \text{if } d = 1 \\ (C_{H^1, L^6}^\Omega)^2 \|\eta\|_{L^2(0, T; L^3(\Omega))}^2 & \text{if } d \in \{2, 3\}. \end{cases} < \frac{b(1 - \bar{\sigma})}{4}, \quad (23)$$

and $\sigma \in L^\infty(\Omega \times (0, T))$ with $\sigma \leq \bar{\sigma}$ a.e., then

$$\begin{aligned} &\max \left\{ \|\nabla u_t\|_{L^\infty(0, T; L^2(\Omega))}, \sqrt{(1 - \bar{\sigma})C_{lo}(\bar{\sigma})} \|u_{tt}\|_{L^2(0, T; L^2(\Omega))} \right\} \\ &\leq \sqrt{bC_{lo}(\bar{\sigma})} e^{C(T)} \left(\|\nabla u_1\|_{L^2(\Omega)}^2 + \frac{2}{b(1 - \bar{\sigma})} \|\tilde{r}\|_{L^2(0, T; L^2(\Omega))}^2 \right)^{1/2}. \end{aligned} \quad (24)$$

with

$$C_{lo}(\bar{\sigma}) = \frac{2(1 - \bar{\sigma})}{b(1 - \bar{\sigma}) - 4C(\eta)}, \quad C(T) = C_{lo}(\bar{\sigma}) \left(c^2 + \frac{c^4 T}{b} \right) T. \quad (25)$$

• If additionally $u_1 \in H^2(\Omega)$, $\nabla \tilde{r} \in L^2(0, T; L^2(\Omega))$, $\eta \in L^2(0, T; H^1(\Omega))$ with

$$C(\eta) := (C_{H^2, L^\infty}^\Omega)^2 \|\nabla \eta\|_{L^2(0, T; L^2(\Omega))}^2 + (C_{H^1, L^6}^\Omega)^2 \|\eta\|_{L^2(0, T; L^3(\Omega))}^2 < \frac{b(1 - \bar{\sigma})}{4}, \quad (26)$$

and $\sigma \in L^\infty(0, T; W^{1,3}(\Omega))$ with

$$C(\sigma) := C_{H^1 \rightarrow L^6}^\Omega \|\nabla \sigma\|_{L^\infty(0, T; L^3(\Omega))} < \frac{1 - \bar{\sigma}}{2} \quad (27)$$

then

$$\begin{aligned} &\max \left\{ \|\mathcal{A}u_t\|_{L^\infty(0, T; L^2(\Omega))}, \sqrt{(1 - \bar{\sigma})C_{hi}(\bar{\sigma})} \|\nabla u_{tt}\|_{L^2(0, T; L^2(\Omega))} \right\} \\ &\leq \sqrt{bC_{hi}(\bar{\sigma})} e^{C(T)} \left(\|\mathcal{A}u_1\|_{L^2(\Omega)}^2 + \frac{2}{b(1 - \bar{\sigma})} \|\nabla \tilde{r}\|_{L^2(0, T; L^2(\Omega))}^2 \right)^{1/2}. \end{aligned} \quad (28)$$

with

$$C_{hi}(\bar{\sigma}) = \frac{2(1 - \bar{\sigma} - 2C(\sigma))}{b(1 - \bar{\sigma}) - 4C(\eta)}, \quad C(T) = C_{hi}(\bar{\sigma}) \left(c^2 + \frac{c^4 T}{b} \right) T.$$

We first of all consider the spatially 1-d case (which would be less relevant for imaging but is in fact of interest for the problem of recovering the univariate function f) since it allows to work with less regular f than the 2- and 3-dimensional setting. Indeed, in one space dimension it is enough to use the low level energy estimate (24) and therefore save on derivatives in f . Using continuity of the embedding $H^1(\Omega) \rightarrow L^\infty(\Omega)$ in one space dimension, the estimate on $\|\nabla u_t\|_{L^\infty(0, T; L^2(\Omega))}$ in (24) allows us to bound $\|u_t\|_{L^\infty(\Omega \times (0, T))}$. By [14, Lemma 3.3] we even get $u_t \in C_w([0, T]; C^\beta(\Omega))$ for any $\beta \in [0, \frac{1}{2})$, where C_w

denotes the space of weakly continuous functions

$$\begin{aligned} \|u_t\|_{L^\infty(\Omega \times (0,T))} &\leq \|u_t\|_{C_w([0,T];C^\beta(\Omega))} \\ &\leq Ke^{C(T)} \left(\|\nabla u_1\|_{L^2(\Omega)}^2 + \frac{1}{b(1-\bar{\sigma})} \|\tilde{r}\|_{L^2(0,T;L^2(\Omega))}^2 \right)^{1/2} \end{aligned} \quad (29)$$

for some constant K depending on the constants $b, c^2, \bar{\sigma}$, $K = K(b, c^2, \bar{\sigma})$ with $K(b, c^2, \bar{\sigma}) \rightarrow \infty$ as $\bar{\sigma} \nearrow 1$; the estimate likewise holds with C^β replaced by $W^{\beta,\infty}$.

Concerning the nonlinear problem, this estimate also tells us that it suffices to assume $f \leq \bar{\sigma}$ and consider f on the interval $[-M, M]$ where $M = M(\bar{\sigma}, \underline{\sigma})$ is an upper bound on the right hand side in (29), that is

$$Ke^{C(T)} \left(\|\nabla p_0\|_{L^2(\Omega)} + \frac{1}{b(1-\bar{\sigma})} \tilde{R} \right) \leq M \quad (30)$$

with

$$\tilde{R} = b\sqrt{T} \|\mathcal{A} p_0\|_{L^2(\Omega)} + (1 + \underline{\sigma})\sqrt{T} \|p_1\|_{L^2(\Omega)} + \|r\|_{(H^1(0,T;L^2(\Omega)))^*} \quad (31)$$

where $-\underline{\sigma}$ is a lower bound for f so that $0 \leq 1 - \bar{\sigma} \leq 1 - f \leq 1 + \underline{\sigma}$.¹ Fixing $\bar{\sigma} < 1$, $\underline{\sigma} \geq 0$, we can thus conclude from (24) that for any

$$f \in C([-M, M]), \quad -\underline{\sigma} \leq f \leq \bar{\sigma}, \quad (32)$$

the operator $\mathcal{T} : v \rightarrow \psi$ where ψ solves (22) with $\sigma = f(v)$, $\eta = 0$, $\tilde{r}$ as in (21), is a self-mapping on the closed convex set

$$\begin{aligned} \mathcal{M} = \{ \phi \in H^2(0, T; L^2(\Omega)) \cap W^{1,\infty}(0, T; H_0^1(\Omega)) : \|\phi_t\|_{L^\infty(0,T;L^\infty(\Omega))} \leq M, \\ \|\nabla \phi_t\|_{L^\infty(0,T;L^2(\Omega))} \leq m_1, \|\phi_{tt}\|_{L^2(0,T;L^2(\Omega))} \leq m_2 \}, \end{aligned}$$

provided

$$e^{C(T)} \left(\|\nabla p_0\|_{L^2(\Omega)} + \frac{1}{b(1-\bar{\sigma})} \tilde{R} \right) \leq \max \left\{ \frac{m_1}{\sqrt{bC_{lo}(\bar{\sigma})}}, m_2 \sqrt{\frac{1-\bar{\sigma}}{b}}, \frac{M}{K} \right\}. \quad (33)$$

The set $\mathcal{M}$ is weak* compact in the Banach space $H^2(0, T; L^2(\Omega)) \cap W^{1,\infty}(0, T; H_0^1(\Omega))$, which is the dual of a separable space, so that Schauder's Fixed Point Theorem in locally convex topological spaces [5] provides existence of a fixed point of $\mathcal{T}$ provided $\mathcal{T}$ is weak* continuous. To prove the latter, we consider an arbitrary sequence $(v_n)_{n \in \mathbb{N}} \subseteq \mathcal{M}$ with weak* limit $v_n \xrightarrow{*} v$. Then, due to the above energy estimates, the sequence defined by $\psi_n = \mathcal{T}(v_n)$ lies in $\mathcal{M}$ and therefore has a weakly* convergent subsequence. The limit ψ^* of any weakly* convergent subsequence $(\psi_{n_k})_{k \in \mathbb{N}}$ needs to coincide with the unique weak solution $\psi = \mathcal{T}(v)$ of (22) with $\sigma = f(v)$, $\eta = 0$, $\tilde{r}$ as in (21). Indeed, for any $\phi \in C^\infty(0, T; C_0^\infty(\Omega))$ we have

$$\begin{aligned} &\int_0^T \int_\Omega \left(\psi_{tt}^* \phi + b \nabla \psi_t^* \cdot \nabla \phi + c^2 \nabla \psi^* \cdot \nabla \phi - f(v_t) \psi_{tt}^* \phi - \tilde{r} \phi \right) dx dt \\ &= \lim_{k \rightarrow \infty} \int_0^T \int_\Omega \left(\psi_{n_k,tt} \phi + b \nabla \psi_{n_k,t} \cdot \nabla \phi + c^2 \nabla \psi_{n_k} \cdot \nabla \phi - f(v_{n_k,t}) \psi_{n_k,tt} \phi - \tilde{r} \phi \right) dx dt = 0, \end{aligned} \quad (34)$$

where the limit in the nonlinear term follows from Lebesgue's Dominated Convergence Theorem, together with the fact that by continuity of f and $v_{n_k,t} \rightarrow v_t$ in $C(\Omega \times (0, T))$ (by the above mentioned continuous embedding) we have pointwise convergence $f(v_{n_k,t}(x, t)) \rightarrow f(v_t(x, t))$ for all $(x, t) \in \Omega \times (0, T)$ and $\int_0^T \int_\Omega |f(v_{n_k,t}) \psi_{n_k,tt} \phi| dx dt \leq \|\psi_{n_k,tt}\|_{L^2(0,T;L^2(\Omega))} \bar{\sigma} \|\phi\|_{L^\infty(\Omega \times (0,T))} < \infty$.

¹For the standard example $f(p) = \kappa p$ we see that on the other hand we need to impose $\kappa M \leq \bar{\sigma} < 1$, that is, M must be small enough in order to generate a nonempty set of admissible nonlinearities f . In accordance with (30), this can always be achieved by making the initial data p_0 , the driving term r , and/or the final time T small enough.

To obtain uniqueness we also prove contractivity of $\mathcal{T}$, which requires more smoothness of f , more precisely, Lipschitz continuity. For any $v, \tilde{v} \in \mathcal{M}$, the difference $\hat{\psi} = \psi - \tilde{\psi} = \mathcal{T}(v) - \mathcal{T}(\tilde{v})$ solves

$$(1 - f(v_t))\hat{\psi}_{tt} + c^2 \mathcal{A} \hat{\psi} + b \mathcal{A} \hat{\psi}_t = (f(v_t) - f(\tilde{v}_t))\tilde{\psi}_{tt}, \\ \hat{\psi}(0) = 0, \quad \hat{\psi}_t(0) = 0$$

Thus, using (24), (29) with $\sigma = f(v_t)$, $\eta = 0$, $\tilde{r} = (f(v_t) - f(\tilde{v}_t))\tilde{\psi}_{tt}$, $u_1 = 0$, and the fact that $\|\tilde{\psi}_{tt}\|_{L^2(0,T;L^2(\Omega))} \leq m_2$ due to the already shown self-mapping property of $\mathcal{T}$, we obtain

$$\max \left\{ \frac{1}{\sqrt{bC_{lo}(\bar{\sigma})}} \|\nabla \hat{\psi}_t\|_{L^\infty(0,T;L^2(\Omega))}, \sqrt{\frac{1-\bar{\sigma}}{b}} \|\hat{\psi}_{tt}\|_{L^2(0,T;L^2(\Omega))}, \frac{1}{K} \|\hat{\psi}_t\|_{L^\infty(0,T;W^{\beta,\infty}(\Omega))} \right\} \\ \leq \sqrt{\frac{2}{b(1-\bar{\sigma})}} e^{C(T)} L m_2 \|v_t - \tilde{v}_t\|_{L^\infty(0,T;L^\infty(\Omega))},$$

where L is a Lipschitz constant for f . Thus, provided

$$\sqrt{\frac{2K^2}{b(1-\bar{\sigma})}} e^{C(T)} L m_2 < 1, \quad (35)$$

the operator $\mathcal{T}$ is a contraction on $\mathcal{M}$ with respect to the norm $\|w\| := \|w_t\|_{L^\infty(0,T;W^{\beta,\infty}(\Omega))} + \|w(0)\|_{L^\infty(\Omega)}$ on $L^\infty(0,T;W^{\beta,\infty}(\Omega))$. The latter is the dual of a separable space and therefore by Alaoglu's Theorem, $\mathcal{M}$ is weak* closed with respect to $L^\infty(0,T;W^{\beta,\infty}(\Omega))$. Banach's Fixed Point Theorem therefore implies existence of a unique fixed point of $\mathcal{T}$ in $\mathcal{M}$.

Theorem 3.1. *For $\Omega \subseteq \mathbb{R}^1$, any fixed $\bar{\sigma} < 1$, $\underline{\sigma} \geq 0$, $M > 0$ and for any $f \in C([-M,M])$ such that $-\underline{\sigma} \leq f \leq \bar{\sigma}$ on $[-M,M]$, $p_0 \in H_0^1(\Omega) \cap H^2(\Omega)$, $p_1 \in L^2(\Omega)$, $r \in (H^1(0,T;L^2(\Omega)))^*$ satisfying (33) with $\tilde{R}$ as in (31), there exists a solution ψ of (20) with (21). This solution satisfies the estimate*

$$\max \left\{ \frac{1}{\sqrt{C_{lo}(\bar{\sigma})}} \|\nabla \psi_t\|_{L^\infty(0,T;L^2(\Omega))}, \sqrt{\frac{1-\bar{\sigma}}{b}} \|\psi_{tt}\|_{L^2(0,T;L^2(\Omega))}, \frac{1}{K} \|\psi_t\|_{C_w([0,T];C^\beta(\Omega))} \right\} \\ \leq e^{C(T)} \left(\|\nabla p_0\|_{L^2(\Omega)}^2 + \frac{2}{b(1-\bar{\sigma})} \tilde{R}^2 \right)^{1/2}.$$

If additionally f is Lipschitz continuous with Lipschitz constant L such that the smallness condition (35) on p_0 and/or T holds, then this solution is unique.

From this we conclude that $p = \psi_t \in L^\infty(0,T;H_0^1(\Omega)) \cap H^1(0,T;L^2(\Omega)) \cap C_w([0,T];C^\beta(\Omega))$ and will therefore infer well-definedness of the forward operator in Corollary 3.1 below.

In higher space dimensions we need higher Sobolev regularity to guarantee $\psi_t(t) \in L^\infty(\Omega)$ and thereby exclude degeneracy. Hence we will rely on (28) to establish a self-mapping property of the operator $\mathcal{T}$ on a set that in view of this estimate will be defined by

$$\mathcal{M} = \{\phi \in H^2(0,T;H_0^1(\Omega)) \cap W^{1,\infty}(0,T;H^2(\Omega)) : \|\phi_t\|_{L^\infty(\Omega \times (0,T))} \leq M, \\ \|\mathcal{A} \phi_t\|_{L^\infty(0,T;L^2(\Omega))} \leq m_1, \|\nabla \phi_{tt}\|_{L^2(0,T;L^2(\Omega))} \leq m_2\}. \quad (36)$$

For this purpose we have to bound $\|\nabla \sigma\|_{L^\infty(0,T;L^3(\Omega))}$ for $\sigma = f(v_t)$ and $\|\nabla \tilde{r}\|_{L^2(0,T;L^2(\Omega))}$ for $\tilde{r}$ as in (21). To this end, we estimate

$$\|\nabla \sigma\|_{L^\infty(0,T;L^3(\Omega))} = \|f'(v_t) \nabla v_t\|_{L^\infty(0,T;L^3(\Omega))} \leq \|f'\|_{L^\infty(-M,M)} C_{H^1 \rightarrow L^3}^\Omega \|\mathcal{A} v_t\|_{L^\infty(0,T;L^2(\Omega))} \\ \|\nabla \tilde{r}\|_{L^2(0,T;L^2(\Omega))} \leq b \|\nabla \mathcal{A} p_0\|_{L^2(\Omega)} + \|\nabla r\|_{(H^1(0,T;L^2(\Omega)))^*} \\ + (1 + \underline{\sigma}) \sqrt{T} \|\nabla p_1\|_{L^2(\Omega)} + \sqrt{T} \|f'\|_{L^\infty(-M,M)} \|p_1 \nabla p_0\|_{L^2(\Omega)} =: \tilde{R} \quad (37)$$

and assume

$$LC_{H^1 \rightarrow L^3}^\Omega C_{H^1 \rightarrow L^6}^\Omega m_1 \leq \frac{1 - \bar{\sigma}}{4} \quad (38)$$

to satisfy (27) (note that with $\eta = 0$, (26) is anyway satisfied)

$$e^{C(T)} \left(\|\mathcal{A} p_0\|_{L^2(\Omega)} + \frac{1}{b(1 - \bar{\sigma})} \tilde{R} \right) \leq \max \left\{ \frac{m_1}{\sqrt{2}}, m_2 \sqrt{\frac{1 - \bar{\sigma}}{2b}}, \frac{M}{K} \right\} \quad (39)$$

with $\tilde{R}$ as in (37).

Since (24) remains valid also in higher space dimensions for $\eta = 0$, contractivity with respect to the $\|\cdot\|$ norm for Lipschitz continuous f with Lipschitz constant satisfying (35) directly carries over to the spatially higher dimensional setting.

Theorem 3.2. *For $\Omega \subseteq \mathbb{R}^d$, $d \in \{1, 2, 3\}$, $\partial\Omega \in C^{2,\alpha}$, any fixed $\bar{\sigma} < 1$, $\underline{\sigma} \geq 0$, $M > 0$ and for any $f \in W^{1,\infty}(-M, M) = C^{0,1}([-M, M])$ such that $-\underline{\sigma} \leq f \leq \bar{\sigma}$ on $[-M, M]$, $p_0 \in H_0^1(\Omega) \cap H^3(\Omega)$, $p_1 \in H^1(\Omega)$, $\nabla r \in (H^1(0, T; L^2(\Omega)))^*$ satisfying (35), (38), (39) with $\tilde{R}$ as in (37), there exists a unique solution ψ of (20) with (21). This solution satisfies the estimate*

$$\begin{aligned} & \max \left\{ \frac{1}{\sqrt{C_{hi}(\bar{\sigma})}} \|\mathcal{A} \psi_t\|_{L^\infty(0, T; L^2(\Omega))}, \sqrt{\frac{1 - \bar{\sigma}}{b}} \|\nabla \psi_{tt}\|_{L^2(0, T; L^2(\Omega))}, \frac{1}{K} \|\psi_t\|_{L^\infty(0, T; W^{\beta, \infty}(\Omega))} \right\} \\ & \leq e^{C(T)} \left(\|\mathcal{A} p_0\|_{L^2(\Omega)}^2 + \frac{2}{b(1 - \bar{\sigma})} \tilde{R}^2 \right)^{1/2}. \end{aligned}$$

With $p = \psi_t$ we thus get

$$p \in U := L^\infty(0, T; H_0^1(\Omega) \cap H^2(\Omega)) \cap H^1(0, T; H_0^1(\Omega)) \cap C_w([0, T]; C^\beta(\Omega)). \quad (40)$$

Corollary 3.1. *Let $\bar{\sigma} < 1$, $\underline{\sigma} \geq 0$, $M > 0$, $L > 0$, $\mathcal{D} = \{f \in W^{1,\infty}(-M, M) : -\underline{\sigma} \leq f \leq \bar{\sigma}, |f'| \leq L \text{ a.e.}\}$, and either*

- (a) $\Omega \subseteq \mathbb{R}^1$, $p_0 \in H_0^1(\Omega) \cap H^2(\Omega)$, $p_1 \in L^2(\Omega)$, $r \in (H^1(0, T; L^2(\Omega)))^*$ satisfying (33), (35) with $\tilde{R}$ as in (31) or
- (b) $\Omega \subseteq \mathbb{R}^d$, $d \in \{2, 3\}$, $p_0 \in H_0^1(\Omega) \cap H^3(\Omega)$, $p_1 \in H_0^1(\Omega)$, $\nabla r \in (H^1(0, T; L^2(\Omega)))^*$ satisfying (38), (39), (35) with $\tilde{R}$ as in (37).

Then the forward operator $F : \mathcal{D} \rightarrow Y$ is well-defined, where

$$Y = L^p(0, T) \text{ for time trace observations and } Y = L^p(\omega) \text{ for final time observations,} \quad (41)$$

respectively, for any $p \in [1, \infty]$.

The following continuity result on F is useful for, e.g., proving that Tikhonov regularisation is well-defined, but we will also use it to establish weak sequential convergence of our iterative reconstruction schemes, cf. Remarks 4.1 and 4.2 below.

Proposition 3.1. *Under the conditions of Corollary 3.1, the operator $F : \mathcal{D} \subseteq W^{1,\infty}(-M, M) \rightarrow Y$ is weakly(*) continuous, that is, for any sequence $(f_n)_{n \in \mathbb{N}} \subseteq \mathcal{D}$ converging weakly* in $W^{1,\infty}(-M, M)$ to f , we have $G(f_n) \xrightarrow{*} G(f)$ in U and $F(f_n) \xrightarrow{*} F(f)$ in Y (see (40), (41) for the definitions of U and Y).*

Proof. Define $\tilde{U} := W^{1,\infty}(0, T; H_0^1(\Omega) \cap H^2(\Omega)) \cap H^2(0, T; H_0^1(\Omega)) \cap \{\psi : \psi_t \in C_w([0, T]; C^\beta(\Omega))\}$. We proceed by proving that $\tilde{G} : W^{1,\infty}(-M, M) \rightarrow \tilde{U}$ is weakly(*) continuous. The result then follows from linearity and boundedness of $\partial_t : \tilde{U} \rightarrow U$ and $\text{tr}_\Sigma : U \rightarrow Y$. Let $(f_n)_{n \in \mathbb{N}} \subseteq \mathcal{D}$ be an arbitrary sequence with $f_n \xrightarrow{*} f$ in $W^{1,\infty}(-M, M)$. Then $f \in \mathcal{D}$ and by Theorem 3.2 $(\psi_n)_{n \in \mathbb{N}} := (\mathcal{G}(f_n))_{n \in \mathbb{N}}$ is bounded in $\tilde{U}$ (more precisely, contained in $\mathcal{M}$). Hence there exists a subsequence $(\psi_{n_k}, f_{n_k})_{k \in \mathcal{N}}$ and

an element $(\psi^*, f^*) \in \tilde{U} \times W^{1,\infty}(-M, M)$ such that $(\psi_{n_k}, f_{n_k}) \xrightarrow{*} \psi^*$ in $\tilde{U} \times W^{1,\infty}(-M, M)$, $(\psi_{n_k, t}, f_{n_k}) \rightarrow (\psi_t^*, f^*)$ in $L^2(0, T; L^2(\Omega)) \times C([-M, M])$ (by the Arzelà-Ascoli Theorem), where by uniqueness of limits, $f^* = f$. Thus, for any $\phi \in C_0^\infty(0, T; C_0^\infty(\Omega))$, similarly to (34) we have

$$\begin{aligned} & \int_0^T \int_\Omega \left(\psi_{tt}^* \phi + b \nabla \psi_t^* \cdot \nabla \phi + c^2 \nabla \psi^* \cdot \nabla \phi - f(\psi_t^*) \psi_{tt}^* \phi - \tilde{r} \phi \right) dx dt \\ &= \lim_{k \rightarrow \infty} \int_0^T \int_\Omega \left(\psi_{n_k, tt} \phi + b \nabla \psi_{n_k, t} \cdot \nabla \phi + c^2 \nabla \psi_{n_k} \cdot \nabla \phi - f_{n_k}(\psi_{n_k, t}) \psi_{n_k, tt} \phi - \tilde{r} \phi \right) dx dt = 0, \end{aligned} \quad (42)$$

where for the convergence of the nonlinear term we argue as follows. We decompose

$$\begin{aligned} & f(\psi_t^*) \psi_{tt}^* - f_{n_k}(\psi_{n_k, t}) \psi_{n_k, tt} \\ &= (f(\psi_t^*) - f(\psi_{n_k, t})) \psi_{tt}^* + (f(\psi_{n_k, t}) - f_{n_k}(\psi_{n_k, t})) \psi_{tt}^* + f_{n_k}(\psi_{n_k, t}) (\psi_{tt}^* - \psi_{n_k, tt}) \end{aligned}$$

and consider the limit in each of the terms (integrated against ϕ over $\Omega \times (0, T)$) separately. For the first term, we can conclude $\int_0^T \int_\Omega (f(\psi_t^*) - f(\psi_{n_k, t})) \psi_{tt}^* \phi dx dt \rightarrow 0$ from Lebesgue's Dominated Convergence Theorem and continuity of f , similarly to the proof of (34) above. The second term can be estimated by $|\int_0^T \int_\Omega (f(\psi_{n_k, t}) - f_{n_k}(\psi_{n_k, t})) \psi_{tt}^* \phi dx dt| \leq \|f - f_{n_k}\|_{C(-M, M)} \|\psi_{tt}^* \phi\|_{L^1(0, T; L^1(\Omega))} \rightarrow 0$. Finally, we have

$$\begin{aligned} & \left| \int_0^T \int_\Omega f_{n_k}(\psi_{n_k, t}) (\psi_{tt}^* - \psi_{n_k, tt}) \phi dx dt \right| \\ &= \left| - \int_0^T \int_\Omega \left(f_{n_k}'(\psi_{n_k, t}) \psi_{n_k, tt} \phi + f_{n_k}(\psi_{n_k, t}) \phi_t \right) (\psi_t^* - \psi_{n_k, t}) dx dt \right| \\ &\leq \left(L m_2 \|\phi\|_{L^\infty(0, T; L^\infty(\Omega))} + \max\{-\underline{\sigma}, \overline{\sigma}\} \|\phi_t\|_{L^2(0, T; L^2(\Omega))} \right) \|\psi_t^* - \psi_{n_k, t}\|_{L^2(0, T; L^2(\Omega))} \rightarrow 0. \end{aligned}$$

From (42), together with a subsequence-subsequence argument and the uniqueness part of Theorem 3.2 we conclude convergence of the whole sequence $(\psi_n)_{n \in \mathbb{N}}$ to $G(f)$. $\diamond$

Remark 3.1. *The fractional attenuation case (14) can be tackled similarly in principle. However, the resulting reformulation via the acoustic velocity potential (that allowed us to minimize smoothness assumptions on f above) does not work any more for $\alpha < 1$. More precisely, the natural approach of multiplying the linearized equation*

$$\begin{aligned} & (1 - \sigma) u_{tt} + c^2 \mathcal{A} u + b \mathcal{A} D_t^\alpha u - \eta u_t = \tilde{r}, \\ & u(0) = u_0, \quad u_t(0) = u_1 \end{aligned}$$

with $D_t^{1+\alpha} u(\tau) = D_t(D_t^\alpha u)(\tau) - \frac{\tau^{-\alpha}}{\Gamma(1-\alpha)} u_1$ and integrating over $(0, t)$ yields the energy identity

$$\begin{aligned} & \int_0^t \langle u_{tt}(\tau), (I^{1-\alpha} u_{tt})(\tau) \rangle_{L^2(\Omega)} d\tau + \frac{b}{2} \|\nabla(D_t^\alpha u)(t)\|_{L^2(\Omega)} \\ &= \frac{b}{2} \|\nabla(D_t^\alpha u)(0)\|_{L^2(\Omega)} + \int_0^t \langle \sigma u_{tt}(\tau), D_t^{1+\alpha} u(\tau) \rangle_{L^2(\Omega)} d\tau + b \int_0^t \frac{\tau^{-\alpha}}{\Gamma(1-\alpha)} \langle u_t(0), D_t^{1+\alpha} u(\tau) \rangle_{L^2(\Omega)} d\tau \\ &+ c^2 \int_0^t \frac{\tau^{-\alpha}}{\Gamma(1-\alpha)} \langle \nabla u_t(0), \nabla u(\tau) \rangle_{L^2(\Omega)} d\tau + c^2 \int_0^t \langle \nabla u_t(\tau), \nabla D_t^\alpha u(\tau) \rangle_{L^2(\Omega)} d\tau \\ &- c^2 \langle \nabla u(t), \nabla D_t^\alpha u(t) \rangle_{L^2(\Omega)} + c^2 \langle \nabla u(0), \nabla D_t^\alpha u(0) \rangle_{L^2(\Omega)} + \int_0^t \langle \tilde{r}(\tau), D_t^{1+\alpha} u(\tau) \rangle_{L^2(\Omega)} d\tau. \end{aligned}$$

Here the first term on the left hand can be estimated from below by means of coercivity of the Abel integral operator [15]

$$\int_0^t \langle u_{tt}(\tau), (I^{1-\alpha} u_{tt})(\tau) \rangle_{L^2(\Omega)} d\tau \geq \cos((1-\alpha)\pi/2) \|u_{tt}\|_{H^{-(1-\alpha)/2}(0, t; L^2(\Omega))}.$$

the second term on the right hand side can be bounded from above by

$$\int_0^t \langle \sigma u_{tt}(\tau), D_t^{1+\alpha} u(\tau) \rangle_{L^2(\Omega)} d\tau \leq \|u_{tt}\|_{H^{-(1-\alpha)/2}(0,t;L^2(\Omega))} \|\sigma D_t^{1+\alpha} u\|_{H^{(1-\alpha)/2}(0,t;L^2(\Omega))}$$

where the second factor can be estimated by means of the Kato-Ponce inequality

$$\|fg\|_{W^{p,r}(0,T)} \lesssim \|f\|_{W^{p,p_1}(0,T)} \|g\|_{L^{q_1}(0,T)} + \|f\|_{L^{p_2}(0,T)} \|g\|_{W^{p,q_2}(0,T)} \quad (43)$$

for $0 \leq \rho \leq \bar{\rho} < 1$, $1 < r < \infty$, $p_1, p_2, q_1, q_2 \in (1, \infty]$, with $\frac{1}{r} = \frac{1}{p_1} + \frac{1}{q_1}$, $i = 1, 2$; see, e.g., [6] as follows

$$\begin{aligned} \|\sigma D_t^{1+\alpha} u\|_{H^{(1-\alpha)/2}(0,t;L^2(\Omega))} &\lesssim \|\sigma\|_{H^{(1-\alpha)/2}(0,t;L^\infty(\Omega))} \|D_t^{1+\alpha} u(\tau)\|_{L^2(0,t;L^2(\Omega))} \\ &\quad + \|\sigma\|_{L^2(0,t;L^2(\Omega))} \|D_t^{1+\alpha} u\|_{H^{(1-\alpha)/2}(0,t;L^2(\Omega))} \end{aligned}$$

where

$$\|D_t^{1+\alpha} u(\tau)\|_{L^2(0,t;L^2(\Omega))} \lesssim \|u_{tt}\|_{H^{-(1-\alpha)}(0,t;L^2(\Omega))}$$

and

$$\|D_t^{1+\alpha} u\|_{H^{(1-\alpha)/2}(0,t;L^2(\Omega))} \lesssim \|u_{tt}\|_{H^{-(1-\alpha)/2}(0,t;L^2(\Omega))}.$$

It can thus be absorbed into the first left hand side term, provided σ is small enough. However, the fifth term on the right hand side yields a nonnegative term

$$c^2 \int_0^t \langle \nabla u_t(\tau), \nabla D_t^\alpha u(\tau) \rangle_{L^2(\Omega)} d\tau = c^2 \int_0^t \langle \nabla u_t(\tau), \nabla(I^{1-\alpha} u_t)(\tau) \rangle_{L^2(\Omega)} d\tau$$

on the right hand side of the energy identity that cannot be dominated by any of the nonnegative left hand side energy contributions. This difficulty has already been observed previously, see [8, Remark 1]. Therefore, one would have to work with the pressure formulation (19) analogously to [10, Section 3.1]. However, this obviously requires higher differentiability of f .

4 Reconstruction Schemes

Now we turn to the inverse problem of identifying f in

$$\begin{aligned} [(1-f(p))p_t]_t + c^2 \mathcal{A}p + b\mathcal{A}p_t &= r, \\ p(0) &= p_0, \quad p_t(0) = p_1 \end{aligned} \quad (44)$$

from time trace

$$h(t) = p(x_0, t), \quad t \in (0, T) \quad (45)$$

or final time

$$g(x) = p(x, T), \quad x \in \omega \subseteq \Omega \quad (46)$$

observations.

4.1 Fixed point iterations

Fixed point formulations of the inverse problem can be obtained by projecting the PDE on the observation manifold and inserting the available measurement data where possible. From this we obtain an iterative reconstruction scheme by applying Picard iteration.

In the case of time trace data (45) using the fact that since $p = p(x, t; f)$ solves (13) we have the identity

$$\begin{aligned} (1 - f(h(t)))h'(t) &= -\mathcal{A} \left(bp(x_0, t; f) + c^2 \int_0^t p(x_0, \tau, f) d\tau \right) \\ &= (1 - f(p(x_0, t; f)))p_t(x_0, t; f) \end{aligned}$$

and we obtain the fixed point scheme

$$f_{k+1}(h(t)) = 1 - \frac{p_t(x_0, t; f_k)}{h'(t)} (1 - f_k(p(x_0, t; f_k))). \quad (47)$$

For final time data (46) using the identity

$$\begin{aligned} (1 - f(g(x)))p_t(x, T; f) &= -\mathcal{A} \left(bg(x) + c^2 \int_0^T p(x, \tau; f) d\tau \right) \\ &= -b\mathcal{A}(g(x) - p(x, T; f)) + (1 - f(p(x, T; f)))p_t(x, T; f) \end{aligned}$$

we get the fixed point scheme

$$f_{k+1}(g(x)) = f_k(p(x, T; f_k)) + \frac{b}{p_t(x, T; f_k)} \mathcal{A}(g(x) - p(x, T; f_k)). \quad (48)$$

With fractional damping (16), the time trace data iteration scheme (47) remains exactly the same as in the strong damping case $\alpha = 1$, just the PDE to be solved in between is modified. For final time data, the possibility of inserting observations becomes very limited, since except for the very first term in (16), p only appears under an integral over time; so we do not pursue the final time data case in the context of (16).

4.2 Newton type schemes

Recall that the forward operator $F = \text{tr}_\Sigma \circ G$ is a composition of the parameter-to-state map $G : f \mapsto p = p(x, t; f)$ where p solves (44) with the trace operator on $\Sigma = \{x_0\} \times (0, T)$ in the time trace and $\Sigma = \omega \times \{T\}$ in the final time case.

The linearisation $z = G'(f)\underline{df}$ solves

$$\begin{aligned} [(1 - f(p))z_t]_t + c^2 \mathcal{A}z + b\mathcal{A}z_t - [f(p)'z p_t]_t &= (\underline{df}(p)p_t)_t, \\ z(0) &= 0, \quad z_t(0) = 0. \end{aligned} \quad (49)$$

Newton's method is therefore defined by the equations

$$\begin{aligned} z(x_0, t; f_k) &= h(t) - p(x_0, t; f_k), \quad t \in (0, T) \text{ in the time trace case} \\ z(x, T; f_k) &= g(x) - p(x, T; f_k), \quad x \in \omega \text{ in the final time case} \end{aligned} \quad (50)$$

where $p(\cdot; f_k)$ solves (44) with $f = f_k$ and $z(\cdot; f_k)$ solves (49) with $f = f_k$, $\underline{df} = f_{k+1} - f_k$.

We will also consider a frozen version of Newton's method, where we linearize at a fixed initial guess, that is, rely on $z = G'(f_0)(f_{k+1} - f_k)$ rather than $z = G'(f_k)(f_{k+1} - f_k)$. Therefore, in the frozen version of (50), $p(\cdot; f_k)$ still solves (44) with $f = f_k$ but $z(\cdot; f_k)$ now solves (49) with $f = f_0$, $\underline{df} = f_{k+1} - f_k$.

These Newton type methods extend to the case of fractional damping (14) in a straightforward manner by replacing $b\mathcal{A}z_t$ by $b\mathcal{A}D_t^\alpha z$ in (49).

4.3 Andersen acceleration

The fact that Picard iteration, namely finding a fixed point of the map $\mathcal{T}$, using $x_{n+1} = \mathcal{T}(x_n)$, can be extremely slow is legendary. The usual approach is to obtain a sufficiently small bound on $\mathcal{T}'$ that the contraction mapping can be used. Even here we are faced with linear convergence. There are a variety of methods that have been used to speed up the convergence of the sequence of iterates and one that is commonly taken is Anderson acceleration. See, for example, [1, 16].

The algorithm with depth m and *damping factors* $\{\beta_k\}$ is as follows.

1. Initial approximation x_0 then compute $\tilde{x}_1 = \mathcal{T}(x_0)$ and set $x_1 = \tilde{x}_1$.
2. For $k = 1, 2, \dots$
 - set $m_k = \min\{k, m\}$.
 - compute $\tilde{x}_{k+1} = \mathcal{T}(x_k)$
 - Solve the minimization problem for $\{a_j\}$ (where a_j depends on k)

$$\min \left\{ \left\| \sum_{j=k-m_k}^k a_j (\tilde{x}_{j+1} - x_j) \right\| : \sum_{j=k-m_k}^k a_j = 1 \right\}$$

- Then set

$$x_{k+1} = \sum_{j=k-m_k}^k a_j ((1 - \beta_k)x_j + \beta_k \tilde{x}_{j+1})$$

Some notes/remarks: The key is the minimization step. Since m is typically small ($3 \leq m \leq 5$) the simplest approach is to use least squares in the sense of $w(k) := \tilde{x}_k - x_{k-1}$ and

$$A = \begin{bmatrix} w(k) \\ \dots \\ w(k-m_k) \end{bmatrix} \quad C = [1 \ 1 \ \dots \ 1] \quad \text{then set} \quad B = \begin{bmatrix} A^T A & C^T \\ C & 0 \end{bmatrix} \quad r = \begin{bmatrix} 0 \\ \vdots \\ 1 \end{bmatrix} \quad (51)$$

to obtain the weights $\{a_k\}$ as the solution of $Ba = r$. For larger values of m , methods based on QR factorisation are advisable but the cost of any aspect of the acceleration step is a very small fraction of the cost of a direct solve of the generalised Westervelt equation. The weights will typically be of mixed sign but an equally simple algorithm allows the constraint $a_j \geq 0$ to be imposed. For the examples involving the fixed point schemes (47), (48) this constraint gave weights that were very close to the original Picard scheme, that is without any effective acceleration. As we will see shortly the unconstrained version gave considerable improvement to the fixed point schemes (47), (48).

It has been shown that Anderson acceleration improves the convergence rate of contractive fixed-point iterations in the vicinity of a fixed-point, [4], but will actually slow the rate of quadratically convergent schemes such as those based on Newton methods.

Based on the analysis in Section 3, we now investigate well-definedness of these iterative schemes.

4.4 Well-definedness of projection based reconstruction schemes

In the time trace data iteration scheme

$$f_{k+1}(h(t)) = 1 - \frac{p_t(x_0, t; f_k)}{h'(t)} (1 - f_k(p(x_0, t, f_k))) \quad (52)$$

we would need $p_t \in L^2(0, T; C(\overline{\Omega}))$ to obtain a well-defined $L^2(0, T)$ trace $p_t(x_0, \cdot)$. The final time data iteration scheme

$$f_{k+1}(g(x)) = f_k(p(x, T; f_k)) + \frac{1}{p_t(x, T; f_k)} b\mathcal{A}(g(x) - p(x, T; f_k)), \quad (53)$$

also requires evaluation of space derivatives. To this end, we point out that by [14, Lemma 3.3] we have $p = p(\cdot, \cdot; f) = G(f) \in C_w([0, T]; H^2(\Omega))$. Therefore, evaluation of $\mathcal{A}p(T) \in L^2(\Omega)$ makes sense. However, since we also need p_t to be at least in $C_w([0, T]; L^\infty(\Omega))$ in order for the term $\frac{1}{p_t(\cdot, T)}$ to make sense and be bounded, we would need a higher order estimate here, which, however, would require higher differentiability of f .

Since we prefer to stay with only Lipschitz continuous f (bearing an mind that the inverse problem is less ill-posed on weaker regularity spaces) we enforce well-definedness of the iteration schemes by projecting

$$f_{k+1}(h(t)) = \mathcal{S} \left[1 - \frac{\mathcal{P}_{[\underline{P}, \overline{P}]}[p_t(x_0, t; f_k)]}{h'(t)} (1 - f_k(p(x_0, t, f_k))) \right] \quad (54)$$

$$f_{k+1}(g(x)) = \mathcal{S} \left[f_k(p(x, T; f_k)) \mathcal{P}_{[\underline{Q}, \overline{Q}]} \left[\frac{1}{p_t(x, T; f_k)} \right] b\mathcal{A}(g(x) - p(x, T; f_k)) \right]. \quad (55)$$

Here $\mathcal{P}_{[a, b]}$ is pointwise defined by $\mathcal{P}_{[a, b]}(z) = \max\{a, \max\{b, z\}\}$ and $\mathcal{S}$ is a smoothing operator mapping into $\mathcal{D}$ as defined in Corollary 3.1 and leaving functions already contained in $\mathcal{D}$ invariant, for example defining, for $y \in L^2(\Sigma)$, its smoothed version $\mathcal{S}[y]$ as a minimiser of

$$\min \|z - y\|_{L^2(\Sigma)}^2 \text{ such that } -\underline{\sigma} \leq f \leq \overline{\sigma}, \quad -L \leq f' \leq L \text{ a.e.}$$

Here $\underline{P}, \overline{P}, \underline{Q}, \overline{Q}, \underline{\sigma}, \overline{\sigma}, L$ should be chosen such that the exact solution $f_{\text{act}} \in \mathcal{D}$ and $\underline{P} \leq p_t(x_0, t; f_{\text{act}}) \leq \overline{P}$ a.e. or $\underline{Q} \leq \frac{1}{p_t(x, T; f_{\text{act}})} \leq \overline{Q}$ a.e., respectively.

Corollary 4.1. *Under the conditions of Corollary 3.1, the iteration schemes*

- (54) with $h' \in L^\infty(\Omega)$, $|h'| \geq \underline{\gamma} > 0$ a.e., $\underline{P} \leq \overline{P}$ and
- (55) with $g \in H^2(\omega)$, $|(g \circ \xi)'| \geq \underline{\gamma} > 0$ a.e., for some curve ξ contained in ω , $\underline{Q} \leq \overline{Q}$

are well-defined.

Remark 4.1. Assume that the inverse problem has a solution $(f_{\text{act}}, p_{\text{act}})$ such that $f_{\text{act}} \in \mathcal{D}$. Since $\mathcal{D}$ is bounded in $W^{1, \infty}(-M, M)$ we can conclude existence of a subsequence f_{k_n} of the projected fixed point iterates according to (54) or (55), respectively that converges weakly* to some $f^* \in \mathcal{D}$, which due to Proposition 3.1 (and its proof) together with $p^* = G(f)$ solves the inverse problem. In case the solution to the inverse problem is unique, a subsequence-subsequence argument yields weak* convergence in $W^{1, \infty}(-M, M)$ (hence norm convergence in $C([-M, M])$ and in $W^{s, p}(-M, M)$ for any $s < 1$, $p \in [1, \infty)$) of the entire sequence of iterates.

4.5 Well-definedness of Newton's method

The linearization of the forward operator $F = \text{tr}_\Sigma \circ G$ at some f is defined by $F'(f) = \text{tr}_\Sigma \circ G'(f)$, where $G'(f)\underline{df} = \underline{dp}$ solves (49) that is, $\underline{dp} = u_t$ where u solves (22) with $\sigma = f(p)$, $\eta = f'(p)p_t$, $\tilde{r} = \underline{df}p_t$, $u_0 = 0$, $u_1 = 0$. In order to obtain a well-defined L^p trace on Σ , we need at least $\underline{dp} = u_t \in \overline{L^p}([0, T]; C(\Omega))$ in the time trace data and $\underline{dp} = u_t \in C_w([0, T]; L^p(\Omega))$ in the final time data case, respectively. We employ Lemma 3.1 and the estimates

$$\begin{aligned} \|\eta\|_{L^2(0, T; L^q(\Omega))} &\leq L\|p_t\|_{L^2(0, T; L^q(\Omega))} \leq \begin{cases} LC_{PF}^\Omega m_2 \text{ with } q = 2 & \text{if } d = 1 \\ LC_{H^1, L^3}^\Omega m_2 \text{ with } q = 3 & \text{if } d \in \{2, 3\} \end{cases} \\ \|\nabla \eta\|_{L^2(0, T; L^2(\Omega))} &\leq \|f''\|_{L^\infty(-M, M)}\|p_t \nabla p\|_{L^2(0, T; L^2(\Omega))} + L\|\nabla p_t\|_{L^2(0, T; L^2(\Omega))} \\ &\leq \|f''\|_{L^\infty(-M, M)}(C_{H^1, L^4}^\Omega)^2 m_1 m_2 + Lm_2 \end{aligned}$$

for $f \in \mathcal{D}$, and m_1, m_2 as in (36), where in the latter case we additionally have to assume $f \in W^{2, \infty}(-M, M)$. From this we deduce that the Newton iteration is well-defined by a minimiser

$$\begin{aligned} f_{k+1} &\in \operatorname{argmin}_{f_+ \in \mathcal{D}} \|\text{tr}_\Sigma(p + \underline{dp}) - y\|_{L^2(\Sigma)}^2 \\ &\text{such that } p \text{ solves (44) with } f = f_k, \text{ and } \underline{dp} \text{ solves (49) with } f = f_k, \underline{df} = f_+ - f_k \end{aligned} \quad (56)$$

where $y = h$ in the time trace and $y = g$ in the final time data case.

Corollary 4.2. *Under the conditions of Corollary 3.1, Newton's method is well-defined by (56) in the final time data case with $d \in \{1, 2, 3\}$ and in the time trace data case with $d = 1$. The latter extends to $d \in \{2, 3\}$ if $\mathcal{D}$ is replaced by $\tilde{\mathcal{D}} := \mathcal{D} \cap \{f \in W^{2, \infty}(-M, M) : \|f''\|_{L^\infty(-M, M)} \leq N\}$ for some $N > 0$.*

A frozen version of (56) with fixed $f^0 \in \mathcal{D} \cap W^{2, \infty}(-M, M)$, $p^0 = G(f^0)$ can be defined by

$$\begin{aligned} f_{k+1} &\in \operatorname{argmin}_{f_+ \in \mathcal{D}} \|\text{tr}_\Sigma(p + \underline{dp}) - y\|_{L^2(\Sigma)}^2 \\ &\text{such that } p \text{ solves (44) with } f = f_k, \text{ and } \underline{dp} \text{ solves (49) with } f = f^0, p = p^0, \underline{df} = f_+ - f_k. \end{aligned} \quad (57)$$

This simplifies both the numerical computations and the analysis.

Corollary 4.3. *For $f^0 \in \mathcal{D} \cap W^{2, \infty}(-M, M)$, $p^0 = G(f^0)$ a frozen Newton method is well-defined by (57) in both data cases and space dimensions $d \in \{1, 2, 3\}$.*

Remark 4.2. *Analogously to Remark 4.1, we can conclude subsequential weak* convergence of the (frozen) Newton sequence to a solution of the inverse problem.*

Noisy data and regularisation

In realistic measurements scenarios, the measured data h or g will typically be contaminated by noise and therefore only approximations $\tilde{h} \approx h$, $\tilde{g} \approx g$ are available. Their distance from the exact data can (if at all) only be estimated in some L^p norm corresponding to our choice of the data space Y , $\|\tilde{h} - h\|_Y \leq \delta$, $\|\tilde{g} - g\|_Y \leq \delta$, and they will typically lack differentiability. Still, the iteration schemes (54), (55) are well-defined upon replacement of $\tilde{h}, \tilde{g}$ by smoothed versions $\hat{h}, \hat{g}$ in such a way that the prerequisites of Corollary 4.1 are preserved, that is, $\hat{h}' \in L^\infty(\Omega)$, $|\hat{h}'| \geq \underline{\hat{\gamma}} > 0$ a.e., in the time trace case and $\hat{g} \in H^2(\omega)$, $|(\hat{g} \circ \xi)'| \geq \underline{\hat{\gamma}} > 0$ a.e., for some curve ξ contained in ω , in the final time case. Indeed, if the exact data satisfies these conditions, the smoothed approximation can be chosen so that its distance from the exact data in the L^p norm is of the order of magnitude of the original noise level. This can for example be achieved by defining $\hat{d} = \hat{h}$ or $\hat{d} = \hat{g} \circ \xi$ from $\tilde{d} = \tilde{h}$ or $\tilde{d} = \tilde{g} \circ \xi$ according to

$$\hat{d} \in \operatorname{argmin}_{b \in B_n := \operatorname{span}\{b_1, \dots, b_n\}} \|b - \tilde{d}\|_{L^p} \text{ such that } sb' \geq \underline{\gamma}/2$$

where $s = \text{sign}(d')$ and $\{b_1, \dots, b_n\}$ is a sufficiently large set of smooth (in the final time case at least H^2) basis functions that allow to approximate the exact data up to precision $\text{dist}^{L^\infty}(d, B_n) \leq \underline{\gamma}/2$. Minimality of $\hat{d}$ and admissibility of d for the above optimization problem immediately yield $\|\hat{d} - \tilde{d}\|_{L^p} \leq \|d - \tilde{d}\|_{L^p}$. The optimization problem itself is well-posed and relatively simple to solve as a minimization of a convex cost function over a finite dimensional set under linear inequality constraints. Clearly, this involves a priori information on the exact data, more precisely on the lower bound $\underline{\gamma} > 0$ and the sign of d' , and additionally on the curve ξ in the final time case.

The regularization strategy that we employ here and in the fixed point schemes (54), (55) themselves as well as in the Newton type iterations (56), (57) is projection or restriction onto a (weakly) compact set. Indeed, considering a family of noisy data with $\delta \rightarrow 0$ and stopping indices $k(\delta) \rightarrow \infty$, analogously to Remarks 4.1, 4.2 we obtain weak subsequential convergence of the iterates $f_{k(\delta)}$ to a solution of the inverse problem with exact data. Thus the proposed schemes are regularization methods in the classical sense of, e.g., [3, Section 3].

5 Reconstructions

In this section we show reconstructions of f from either time trace or final time data. The spatial set will be the interval $[0, 1]$ and we will take the time trace measurement point to be the right-hand endpoint $x = 1$.

Our numerical implementation uses (44) in the integrated version as in (13) and so treat it as a parabolic equation with nonlocal memory term $c^2 \int_0^t \Delta p(\tau) d\tau$. A Crank-Nicolson integrator was used with an inner iteration loop to handle the nonlinear term $f(p) p_t$. A Neumann boundary condition was imposed at the right hand endpoint; the left hand condition could be Dirichlet, Neumann or impedance type. Typically, in the physical model one would have zero initial conditions but this isn't necessary for the mathematical formulation.

Data consisted of the measurements $h(t) = p(1, t)$ or $g(x) = p(x, T)$. As a practical matter we used the above mentioned solver to obtain this data and collected a sample at 50 equally spaced points on the interval $[0, T]$ or $[0, 1]$. Uniformly distributed random noise was then added to these values to obtain $h_{\text{meas}}(t)$ or $g_{\text{meas}}(x)$. This was then pre-filtered by smoothing and up-resolving to the working resolution of the number of points taken (~ 400 for the interval $t \in [0, T]$ and ~ 200 for the interval $x \in [0, 1]$) for the direct solver used in the inversion routine.

For the Newton scheme, the unknown f was represented in terms of given basis functions. Since we wish to make no constraints on the form of f , we do not choose a basis with in-built restrictions as would be obtained from an eigenfunction expansion. Instead we used a sine basis that helped to realize the condition $f(0) = 0$. In all cases the starting approximation for the iterative methods used was the constant function $f \equiv 0$.

We are going to show the results of the methods described in section 4. In particular, we will provide comparisons between frozen Newton and the fixed point iteration schemes from sections 4.2 and 4.1, respectively. In the latter case, we will also show the effect of Anderson acceleration cf. section 4.3. As to the parameters in the Anderson scheme, we simply fixed the depth to be $m = 3$ and the weights $\beta_k = 1$, which corresponds to full acceleration.

It is well-known in undetermined coefficient problems that recovering a function of an independent

variable from data in an orthogonal variable direction leads to severe ill-posedness whereas data given in a parallel direction often yields only mild ill-conditioning. In the case of the unknown depending on the state variable, here $u(x, t)$, one might assume that little difference would be found provided both data sets meet the range condition. However, previous work on such problems has shown that this is not always the case. In particular, for a parabolic type equation time trace data frequently leads to superior reconstructions over measurement data obtained in spatial manner see, e.g., [9]. As we will see below, this is also the situation here. As a result, one is able to reconstruct more complex functions using time trace data for a given level of accuracy. In selecting functions for testing we concentrated on highlighting this difference rather than functions that might correspond to a known physical reality. The chosen functions were

$$f_{\text{act}}(u) = 0.2 \sin(6\pi u)(1 - \exp(u)) \quad (58)$$

for time trace and

$$f_{\text{act}}(u) = 0.1(1 - \exp(-3u)) \cos(0.7u) \quad (59)$$

for final time data.

In Figure 1 we show reconstructions of the functions (58) and (59), while Figure 2 displays the convergence history of the three different methods we use.

The two examples shown above represent smooth functions and a reasonable question is what is the ability to reconstruction a function lying at the boundary of the regularity theorems in section 3, namely a function that is just Lipschitz continuous. In Figure 3 we therefore the results of the fixed point scheme and its Anderson accelerated version using time trace data. For doing this with Newton, the optimal basis would probably have to be piecewise linear, as for example in [10]. To avoid a mismatch between the actual location of the non-differentiability points and the nodes of the basis functions requires a large number of basis functions in general, unless these points are a priori known. Thus without this knowledge, the computational complexity of calculating the Jacobian would make (frozen) Newton much slower than the accelerated fixed point schemes.

We also did reconstructions from time trace data in the fractional case (16), where as mentioned in section 4 the time trace data iteration scheme (47) remains the same as in case $\alpha = 1$, with the dependence on α showing up only in the PDE to be solved. Provided α was not close to zero, the differences in the reconstructions were very little.

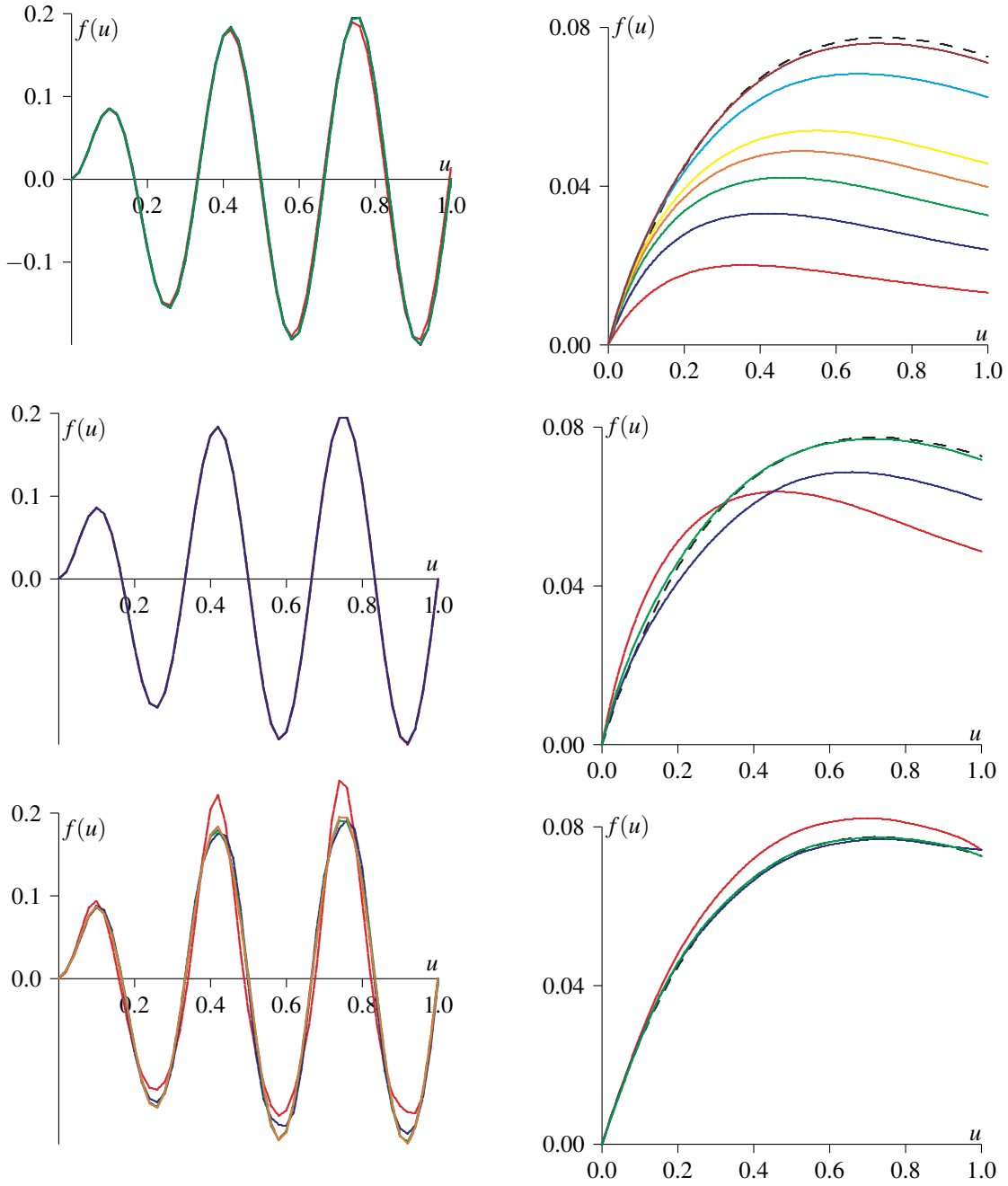


Figure 1: **Reconstructions of $f(u)$ from time trace (left) and final time (right) data with Picard (top row), Anderson accelerated Picard (middle row), frozen Newton (bottom row); first (red), second (blue), third (green), fourth (orange) iterate versus exact f (back dashed)**

Appendix (Proof of Lemma (3.1))

Multiplying (22) with u_{tt} and integrating over $\Omega \times (0, t)$ yields

$$\begin{aligned}
 & \int_0^t \|\sqrt{1 - \sigma(\tau)} u_{tt}(\tau)\|_{L^2(\Omega)}^2 d\tau + \frac{b}{2} \|\nabla u_t(t)\|_{L^2(\Omega)}^2 \\
 &= \frac{b}{2} \|\nabla u_1\|_{L^2(\Omega)}^2 + c^2 \int_0^t \|\nabla u_t(\tau)\|_{L^2(\Omega)}^2 d\tau - c^2 \langle \nabla u_t(t), \nabla u(t) \rangle_{L^2(\Omega)} \\
 & \quad + \int_0^t \langle \tilde{r}(\tau) + \eta(\tau) u_t(\tau), u_{tt}(\tau) \rangle_{L^2(\Omega)} d\tau \\
 &\leq \frac{b}{2} \|\nabla u_1\|_{L^2(\Omega)}^2 + c^2 \int_0^t \|\nabla u_t(\tau)\|_{L^2(\Omega)}^2 d\tau^{18}
 \end{aligned} \tag{60}$$

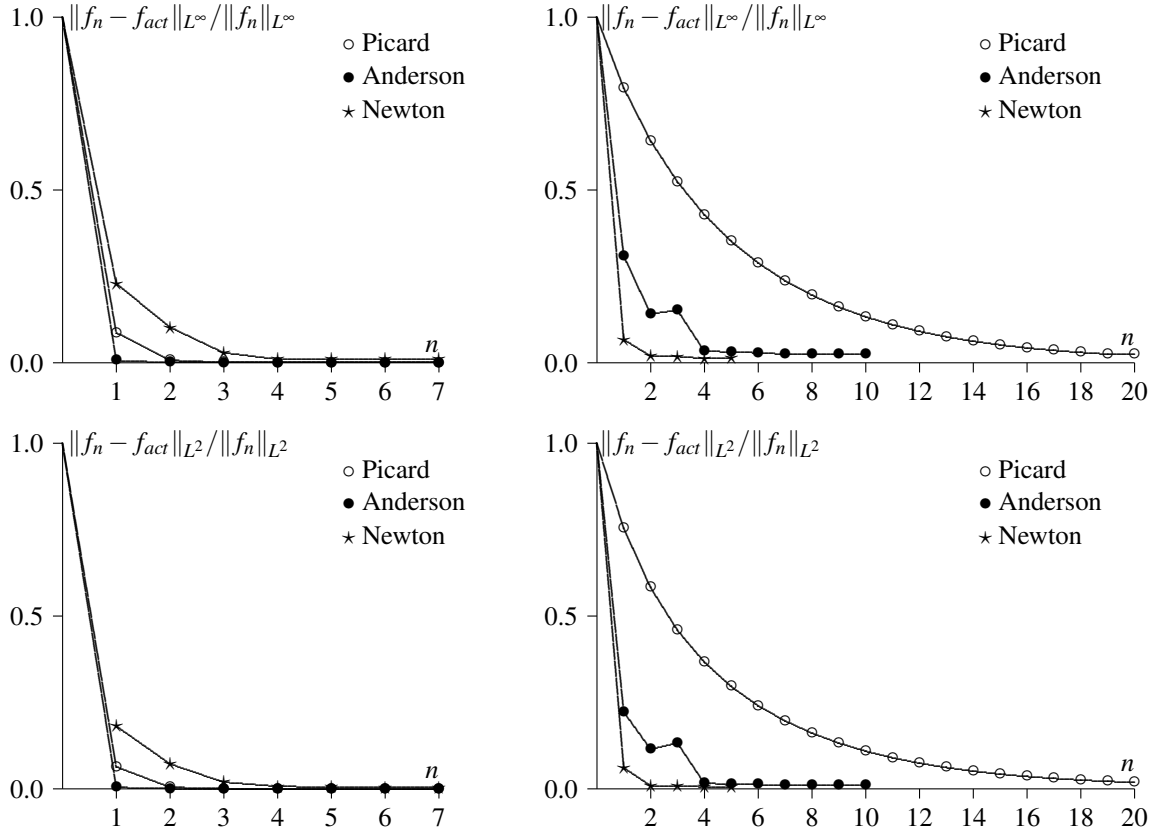


Figure 2: **Relative error norms of the iterates for $f(u)$ from time trace (left) and final time (right) data in L^∞ (top) and L^2 (bottom) norm**

where we have used $\sigma \leq \bar{\sigma} < 1$, $u(0) = 0$, $\|\nabla u(t)\|_{L^2(\Omega)}^2 = \|\int_0^t \nabla u_t(\tau) d\tau\|_{L^2(\Omega)}^2 \leq t \int_0^t \|\nabla u_t(\tau)\|_{L^2(\Omega)}^2 d\tau$ and Young's inequality. From this we extract

$$\begin{aligned} & \frac{1-\bar{\sigma}}{2} \int_0^t \|u_{tt}(\tau)\|_{L^2(\Omega)}^2 d\tau + \frac{b}{4} \|\nabla u_t(t)\|_{L^2(\Omega)}^2 \\ & \leq \frac{b}{2} \|\nabla u_1\|_{L^2(\Omega)}^2 + \frac{1}{2(1-\bar{\sigma})} \int_0^t \|\tilde{r}(\tau) + \eta(\tau)u_t(\tau)\|_{L^2(\Omega)}^2 d\tau \\ & \quad + \left(c^2 + \frac{c^4 T}{b}\right) \int_0^t \|\nabla u_t(\tau)\|_{L^2(\Omega)}^2 d\tau. \end{aligned}$$

To estimate the term containing η we make use of continuity of the embeddings $H^1(\Omega) \rightarrow L^\infty(\Omega)$ in one space dimension or $H^1(\Omega) \rightarrow L^\infty(\Omega)$ in two and three space dimensions, respectively, as well as the Poincaré-Friedrichs inequality

$$\|v\|_{L^p(\Omega)} \leq C_{H^1, L^p}^\Omega \|\nabla v\|_{L^2(\Omega)} \text{ for all } v \in H_0^1(\Omega), \text{ with } \begin{cases} p = \infty & \text{if } d = 1 \\ p \leq 6 & \text{if } d \in \{2, 3\}. \end{cases}$$

to conclude $\int_0^t \|\eta(\tau)u_t(\tau)\|_{L^2(\Omega)}^2 d\tau \leq C(\eta) \|\nabla u_t\|_{L^\infty(0,t;L^2(\Omega))}^2$ where we assume $C(\eta)$ as in (23) to be smaller than $\frac{b(1-\bar{\sigma})}{4}$. Gronwall's inequality for

$$\xi(t) = \max \left\{ \left(\frac{b}{4} - \frac{C(\eta)}{1-\bar{\sigma}} \right) \|\nabla u_t\|_{L^\infty(0,t;L^2(\Omega))}^2, \frac{1-\bar{\sigma}}{2} \|u_{tt}\|_{L^2(0,t;L^2(\Omega))}^2 \right\}$$

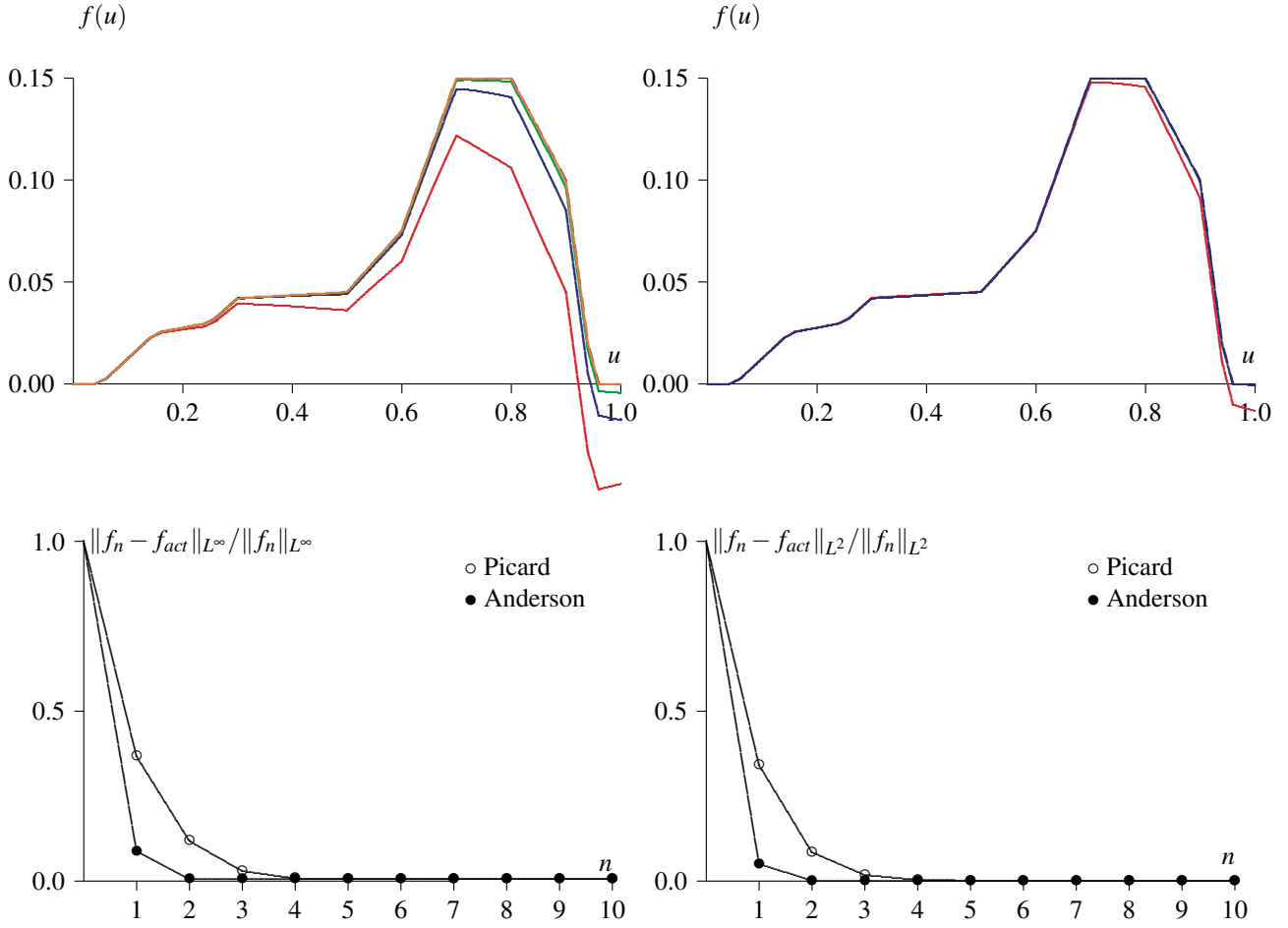


Figure 3: Reconstructions of $f(u)$ for piecewise linear example from time trace data with Picard (top left) and Anderson accelerated Picard (top right); first (red), second (blue), third (green), fourth (orange), fifth (yellow), sixth (cyan), seventh (maroon) iterate versus exact f (back dashed); relative error norms of the iterates for $f(u)$ in L^∞ (bottom left) and L^2 (bottom right) norm

satisfying

$$\xi(t) \leq \frac{b}{2} \|\nabla u_1\|_{L^2(\Omega)}^2 + \frac{1}{1-\bar{\sigma}} \|\tilde{r}\|_{L^2(0,T;L^2(\Omega))}^2 + \left(c^2 + \frac{c^4 T}{b}\right) \frac{4(1-\bar{\sigma})}{b(1-\bar{\sigma}) - 4C(\eta)} \int_0^t \xi(\tau) d\tau$$

yields (24).

To prove (28), we multiply (22) with $\mathcal{A}u_{tt}$ and integrate over $\Omega \times (0, t)$, which similarly to (60), but taking into account

$$\int_{\Omega} (1-\sigma) u_{tt} \mathcal{A} u_{tt} dx = \int_{\Omega} \left((1-\sigma) |\nabla u_{tt}|^2 - u_{tt} \nabla \sigma \cdot \nabla u_{tt} \right) dx$$

yields

$$\begin{aligned}
& \int_0^t \|\sqrt{1-\sigma(\tau)} \nabla u_{tt}(\tau)\|_{L^2(\Omega)}^2 d\tau + \frac{b}{2} \|\mathcal{A}u_t(t)\|_{L^2(\Omega)}^2 \\
&= \frac{b}{2} \|\mathcal{A}u_1\|_{L^2(\Omega)}^2 + c^2 \int_0^t \|\mathcal{A}u_t(\tau)\|_{L^2(\Omega)}^2 d\tau - c^2 \langle \mathcal{A}u_t(t), \nabla u(t) \rangle_{L^2(\Omega)} \\
&\quad + \int_0^t \langle u_{tt} \nabla \sigma + \nabla(\tilde{r} + \eta u_t)(\tau), \nabla u_{tt}(\tau) \rangle_{L^2(\Omega)} d\tau \\
&\leq \frac{b}{2} \|\mathcal{A}u_1\|_{L^2(\Omega)}^2 + c^2 \int_0^t \|\mathcal{A}u_t(\tau)\|_{L^2(\Omega)}^2 d\tau \\
&\quad + \frac{1-\bar{\sigma}}{2} \int_0^t \|\nabla u_{tt}(\tau)\|_{L^2(\Omega)}^2 d\tau + \frac{b}{4} \|\mathcal{A}u_t(t)\|_{L^2(\Omega)}^2 \\
&\quad + \frac{c^4}{b} \int_0^t \|\mathcal{A}u_t(\tau)\|_{L^2(\Omega)}^2 d\tau + \frac{1}{2(1-\bar{\sigma})} \int_0^t \|\nabla(\tilde{r} + \eta u_t)(\tau)\|_{L^2(\Omega)}^2 d\tau \\
&\quad + \|\nabla \sigma\|_{L^\infty(0,T;L^3(\Omega))} \int_0^t \|u_{tt}(\tau)\|_{L^6(\Omega)} \|\nabla u_{tt}(\tau)\|_{L^2(\Omega)} d\tau.
\end{aligned} \tag{61}$$

Using continuity of the embeddings $H_0^1(\Omega) \rightarrow L^6(\Omega)$ and $H^2(\Omega) \rightarrow L^\infty(\Omega)$ in

$$\begin{aligned}
& \int_0^t \|\nabla(\eta u_t)(\tau)\|_{L^2(\Omega)}^2 d\tau \leq C(\eta) \|\mathcal{A}u_t\|_{L^\infty(0,t;L^2(\Omega))}^2 \\
& \|\nabla \sigma\|_{L^\infty(0,T;L^3(\Omega))} \int_0^t \|u_{tt}(\tau)\|_{L^6(\Omega)} \|\nabla u_{tt}(\tau)\|_{L^2(\Omega)} d\tau \leq C(\sigma) \int_0^t \|\nabla u_{tt}(\tau)\|_{L^2(\Omega)}^2 d\tau
\end{aligned}$$

and imposing the smallness conditions (26), (27) on η , $\nabla \eta$, and $\nabla \sigma$ analogously to above we obtain (28).

Acknowledgment

The work of the first author was supported by the Austrian Science Fund FWF under the grant P30054.

The work of the second author was supported in part by the National Science Foundation through award DMS-1620138.

References

- [1] Donald G. Anderson. Iterative procedures for nonlinear integral equations. *J. Assoc. Comput. Mach.*, 12:547–560, 1965.
- [2] Wei Cai, Wen Chen, Jun Fang, and Sverre Holm. A Survey on Fractional Derivative Modeling of Power-Law Frequency-Dependent Viscous Dissipative and Scattering Attenuation in Acoustic Wave Propagation. *Applied Mechanics Reviews*, 70(3), 06 2018.
- [3] Heinz W. Engl, Martin Hanke, and Andreas Neubauer. *Regularization of Inverse Problems*. Kluwer, Dordrecht, 1996.

- [4] Claire Evans, Sara Pollock, Leo G. Rebholz, and Mengying Xiao. A proof that Anderson acceleration improves the convergence rate in linearly converging fixed-point methods (but not in those converging quadratically). *SIAM J. Numer. Anal.*, 58(1):788–810, 2020.
- [5] Ky. Fan. Fixed-point and minimax theorems in locally convex topological linear spaces. *Proc. Nat. Acad. Sci. U.S.A.*, 38:121–126, 1952.
- [6] Loukas Grafakos and Seungly Oh. The Kato–Ponce inequality. *Communications in Partial Differential Equations*, 39(6):1128–1157, 2014.
- [7] Mark F Hamilton and David T Blackstock. *Nonlinear acoustics*, volume 1. Academic press San Diego, 1998.
- [8] Barbara Kaltenbacher and Vanja Nikolić. Time-fractional Moore-Gibson-Thompson equations. 2021. arXiv:2104.13967 [math.AP].
- [9] Barbara Kaltenbacher and William Rundell. The inverse problem of reconstructing reaction-diffusion systems. 2020. submitted and arXiv:2003.00489 [math.NA].
- [10] Barbara Kaltenbacher and William Rundell. On an inverse problem of nonlinear imaging with fractional damping. *arXiv:2103.08965*, 2021.
- [11] Barbara Kaltenbacher and William Rundell. On the identification of the nonlinearity parameter in the westervelt equation from boundary measurements. *Inverse Problems & Imaging*, 0, 2021.
- [12] Barbara Kaltenbacher and William Rundell. Some inverse problems for wave equations with fractional derivative attenuation. *Inverse Problems*, 37(4):045002, 28, 2021.
- [13] Michael Pilant and William Rundell. Recovery of an unknown specific heat by means of overposed data. *Numerical Methods for Partial Differential Equations*, 6(1):1–16, 1990.
- [14] Roger Temam. *Infinite-dimensional dynamical systems in mechanics and physics*, volume 68. Springer Science & Business Media, 2012.
- [15] Urs Vgeli, Khadijeh Nedaiasl, and Stefan Sauter. A fully discrete Galerkin method for Abel-type integral equations. *Advances in Computational Mathematics*, 12 2016.
- [16] Homer F. Walker and Peng Ni. Anderson acceleration for fixed-point iterations. *SIAM J. Numer. Anal.*, 49(4):1715–1735, 2011.