On the identification of a nonlinear term in a reaction–diffusion equation

Barbara Kaltenbacher and William Rundell

1 Alpen-Adria-Universität Klagenfurt, Universitätsstraße 65-67, 9020 Klagenfurt, Austria
2 Department of Mathematics, Texas A&M University, TX 77843, United States of America

E-mail: barbara.kaltenbacher@aau.at and rundell@math.tamu.edu

Received 8 March 2019, revised 17 May 2019
Accepted for publication 18 June 2019
Published 3 October 2019

Abstract

Reaction–diffusion equations are one of the most common partial differential equations used to model physical phenomena. They arise as the combination of two physical processes: a driving force \( f(u) \) that depends on the state variable \( u \) and a diffusive mechanism that spreads this effect over a spatial domain. The canonical form is \( u_t - \Delta u = f(u) \). Application areas include chemical processes, heat flow models and population dynamics. As part of model building, assumptions are made about the form of \( f(u) \), and these inevitably contain various physical constants. The direct or forward problem for such equations is now very well developed and understood, especially when the diffusive mechanism is governed by Brownian motion resulting in an equation of parabolic type.

However, our interest lies in the inverse problem of recovering the reaction term \( f(u) \) not just at the level of determining a few parameters in a known functional form, but in recovering the complete functional form itself. To achieve this we set up the standard paradigm for the parabolic equation where \( u \) is subject to both given initial and boundary data; then we prescribe overposed data consisting of the solution at a later time \( T \). For example, in the case of a population model this amounts to census data at a fixed time. Our approach will be two-fold. First we will transform the inverse problem into an equivalent nonlinear mapping from which we seek a fixed point. We will be able to prove important features of this map such as a self-mapping property and give conditions under which it is contractive. Second, we consider the direct map from \( f \) through the partial differential operator to the overposed data. We will investigate Newton schemes for this case.
Classical, Brownian motion diffusion is not the only version, and in recent decades various anomalous processes have been used to generalize this case. Amongst the most popular is one that replaces the standard time derivative by a subdiffusion process based on a fractional derivative of order $\alpha \leq 1$. We will also include this model in our analysis. The final section of the paper will show numerical reconstructions that demonstrate the viability of the suggested approaches. This will also include the dependence of the inverse problem on both $T$ and $\alpha$.

Keywords: reaction–diffusion equation, fractional diffusion, parameter identification, regularization

(Some figures may appear in colour only in the online journal)

1. Introduction

Before the thorough mathematical theory of existence, uniqueness and regularity results for reaction–diffusion equations were developed during the 1960s, there was extensive literature on applications dating from the 1930s. This includes the work of Fisher in considering the Verhulst logistic equation together with a spatial diffusion, $u_t - ku_{xx} = au(1-u)$ and that of Kolmogorov, Petrovskii and Piskunov in similar models now collectively referred to as the Fisher–KPP theory; more complex polynomial terms occurring in combustion theory due to Zeldovich and Frank-Kamenetskii. The class of reaction–diffusion equations describe the behaviour of a wide range of chemical systems where diffusion of material competes with production by a chemical reaction. They also describe systems where heat is produced depending on the local temperature and diffuses away by some mechanism. These systems can also be based on particles, as in the nuclear diffusion equation. For general references here we note the books [2, 9, 22] as well as chapter [19]. These equations and their generalization can be described as follows.

Let $\Omega$ be a bounded, simply connected region in $\mathbb{R}^d$ with a smooth boundary $\partial \Omega$, and let $L$ be a uniformly elliptic operator of second order with $L^\infty$ coefficients.

$$u_t(x, t) - Lu(x, t) = f(u(x, t)) + r(x, t), \quad (x, t) \in \Omega \times (0, T),$$

and subject to the initial and boundary conditions

$$\partial_\nu u(x, t) + \gamma(x)u(x, t) = h(x, t) \quad (x, t) \in \partial \Omega \times (0, T)$$

$$u(x, 0) = u_0 \quad x \in \Omega.$$

In (1) the standard, direct problem is to know both the diffusion operator $L$, the nonhomogeneous linear term $r$ and the reaction term $f(u)$. However, even in the earliest applications it was appreciated that the diffusion term $-k\Delta u$ might have the diffusion coefficient $k$ as an unknown, and while the specific form of $f(u)$ was assumed known from the underlying model, specific parameters may not be. In our case we assume that $f$ is unknown (save that it be a function of $u$ alone). Thus, we are interested in the inverse problem of recovering $f$ from additional data that would be overposed if in fact $f$ were known. There has been previous work along such lines. For example, in a series of papers, [23–25], Pilant and Rundell investigated the recovery of $f$ from overposed data consisting of a time trace of the solution $u$ at a given point $x_0 \in \partial \Omega$, namely $u(x_0, t)$ for all $t > 0$. See also [6]. We will take a very different situation here by assuming that one is able to measure in the spatial direction by taking a snapshot at a fixed later time $T$ and so our overposed data will be
We also note that in the case of such data, if the reaction term is of the form $q(x)f(u)$ where $q$ is unknown but the actual form of the nonlinearity $f(u)$ is known, then it is possible to recover the spatial component in a unique way [17].

Exchanging the time derivative in (1) for one of fractional order adds further complexity but is also of considerable physical interest as we are modifying in a fundamental way the diffusion process (Fick’s law) to one of a so-called anomalous diffusion mechanism that would lead to a non-local, time-fractional equation.

$$D_\alpha^t u(x, t) - Lu(x, t) = f(u(x, t)) + r(x, t), \quad (x, t) \in \Omega \times (0, T)$$

and is subject to the same initial and boundary conditions (2). Here we have also included the possibility of an additional (known) inhomogeneity, which can be used to drive the evolution process.

By $D_\alpha^t$ we mean the Djrbashian–Caputo derivative of order $\alpha$ which is defined to be

$$\alpha D_\alpha^t u = \frac{1}{\Gamma(1-\alpha)} \int_a^t u'(s) (t-s)^{\alpha-1} \, ds$$

and since this is a non-local operator, the designation of the starting point $t = a$ is essential; in our case we shall take $a = 0$. Of course, there are other possibilities for anomalous diffusion models, but the one chosen is both standard in most applications and has an existing rich mathematical theory.

Examples of the above include applications in many of the areas in which traditional reaction diffusion equations have featured over the last 100 years and form a natural generalization. The central feature of a reaction–diffusion model is the interplay between the two central processes that are in some sense in competition. One of the main themes of this paper is to investigate the role the different diffusion models play in our ability to recover the reaction term $f(u)$.

Briefly, the outline of the paper is as follows. In the next section we collect crucial main existence and regularity results for the direct problem. These are well known for the parabolic case but we require an addition to the existing literature for the case of (4). Sections 3 and 4 are devoted to two iterative approaches for numerically solving the inverse problem, a fixed point scheme and a Newton-type method. While these methods are formulated in the general, possibly higher-dimensional setting, the analysis will be restricted to the spatially one-dimensional case, due to the need for certain Sobolev embeddings. In the last section we show some numerical examples illustrating the above and use this to bring out certain features of the problem. In particular we investigate the difference between the classical and fractional diffusion operators and the role that the value of the final time $T$ plays.

2. The direct problem

In this section we provide the results and tools for the direct problem, i.e. solution of the initial boundary value problem (2), (4), which will be needed later on in the reconstruction methods. These are, first of all, a representation of the solution to the inhomogeneous linear problem by means of Mittag-Leffler functions as well as eigenvalues and eigenfunctions of the operator $-L$. Moreover, versions of the familiar Gronwall’s inequality for ordinary differential equations in function space are required for the fractional case; we develop these below, first for the
linear case, then later for the semilinear. These will then be used to provide estimates needed for the semilinear fractional operator.

2.1. The solution representation and the Mittag-Leffler function

Throughout this paper we will assume that \( L = \sum a_{ij} (a_{ij} \partial_j) + c \) is a symmetric uniformly elliptic operator with coefficients \( a_{ij} \in W^{1,\infty}(\Omega), b, c \in L^{\infty}(\Omega) \), and that \( \Omega \) is a bounded and convex or \( C^{1,1} \) domain.

We define an operator \( A \) in \( L^2(\Omega) \) by

\[
(Au)(x) = (-\Delta u)(x) \quad x \in \Omega,
\]

where \( L \) is equipped with impedance boundary conditions, with its domain \( D(A) \subseteq H^2(\Omega) \). Then the fractional power \( A^\sigma \) is defined for any \( \sigma \in \mathbb{R} \). Since \( L \) is a symmetric uniformly elliptic operator, the spectrum of \( A \) is entirely composed of eigenvalues and counting according to the multiplicities, we can set \( 0 < \lambda_1 \leq \lambda_2 \ldots \). By \( \varphi_j \in H^2(\Omega) \), we denote the \( L^2(\Omega) \) orthonormal eigenfunctions corresponding to \( \lambda_j \). Next, we introduce a space \( \tilde{H}^s(\Omega) \) by

\[
\tilde{H}^s(\Omega) = \left\{ v \in L^2(\Omega) : \sum_{j=1}^{\infty} \lambda_j^s |(v, \varphi_j)|^2 < \infty \right\},
\]

which is a Hilbert space with the norm \( \|v\|_{\tilde{H}^s(\Omega)}^2 = \sum_{j=1}^{\infty} \lambda_j^s |(v, \varphi_j)|^2 \). By definition, we have the following equivalent form:

\[
\|v\|_{\tilde{H}^s(\Omega)}^2 = \sum_{j=1}^{\infty} |(v, \lambda_j^s \varphi_j)|^2 = \sum_{j=1}^{\infty} (v, A^s \varphi_j)^2 = \sum_{j=1}^{\infty} (A^s v, \varphi_j)^2 = \|A^s v\|_{L^2(\Omega)}^2.
\]

We have \( \tilde{H}^s(\Omega) \subseteq H^s(\Omega) \) for \( s \in [0, 2] \). Since \( \tilde{H}^s(\Omega) \subseteq L^2(\Omega) \), identifying the dual \( (L^2(\Omega))^\prime \) with itself, we have \( \tilde{H}^s(\Omega) \subseteq L^2(\Omega) \subseteq (H^s(\Omega))^\prime \). Henceforth, we set \( \tilde{H}^{-s}(\Omega) = (\tilde{H}^s(\Omega))^\prime \), which consists of bounded linear functionals on \( \tilde{H}^s(\Omega) \).

The solution \( u \) to the linear problem

\[
\begin{align*}
D_0^\alpha u(x, t) - Lu(x, t) &= r(x, t), \quad (x, t) \in \Omega \times (0, T) \\
\partial_{\nu} u(x, t) + \gamma(x)u(x, t) &= h(x, t), \quad (x, t) \in \partial\Omega \times (0, T) \\
u(x, 0) &= u_0 \quad x \in \Omega,
\end{align*}
\]

can be represented by [13]

\[
u(x, t) = \sum_{j=1}^{\infty} E_{\alpha, 1}(-\lambda_j t^\alpha)(u_0, \varphi_j) \varphi_j(x) + \sum_{j=1}^{\infty} \int_0^t (t-s)^{\alpha-1} E_{\alpha, \alpha}(-\lambda_j (t-s)^\alpha)(r(x, s), \varphi_j) \, ds \varphi_j(x).
\]

Here \( E_{\alpha, \beta}(z) \) is the Mittag-Leffler function [21] defined by

\[
E_{\alpha, \beta}(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\alpha k + \beta)} \quad \alpha > 0, \quad \beta \in \mathbb{R}, \quad z \in \mathbb{C}.
\]
immediately accessible. See also the tutorial on inverse problems for fractional operators [13]. In the case $\alpha = 1$, $E_{\alpha,1}(z) = E_{\alpha,\alpha}(z) = e^z$ and the model (7) as well its solution formula (8) recover the standard parabolic equation.

For our purposes, and in fact for a large range of inverse problems involving a subdiffusion process, one of the key features of the Mittag-Leffler function is its decay for large, real and negative arguments. The asymptotic behaviour as $z \to \infty$ in various sectors of the complex plane $\mathbb{C}$ was first derived by Djrbashian [3], and subsequently refined by many over the years since.

**Lemma 2.1.** Let $\alpha \in (0, 2)$, $\beta \in \mathbb{R}$, and $\mu \in (\alpha \pi/2, \min(\pi, \alpha \pi))$, and $N \in \mathbb{N}$. Then for $|\arg(z)| \leq \mu$ with $|z| \to \infty$, we obtain

$$ E_{\alpha,\beta}(z) = \frac{1}{\alpha} z^{\frac{\beta-\alpha}{\alpha}} e^{\frac{z}{\alpha}} - \sum_{k=1}^{N} \frac{1}{\Gamma(\beta - \alpha k)} \frac{1}{z^k} + O\left(\frac{1}{z^{N+1}}\right), \quad (10) $$

and for $\mu \leq |\arg(z)| \leq \pi$ with $|z| \to \infty$

$$ E_{\alpha,\beta}(z) = -\sum_{k=1}^{N} \frac{1}{\Gamma(\beta - \alpha k)} \frac{1}{z^k} + O\left(\frac{1}{z^{N+1}}\right). \quad (11) $$

In particular, for $x$ real and positive, $E_{\alpha,\beta}(-x)$ decays to zero linearly as $x \to \infty$.

The following useful estimate is a direct corollary of lemma 2.1.

**Corollary 2.1.** Let $0 < \alpha < 2$, $\beta \in \mathbb{R}$, and $\pi \alpha/2 < \mu < \min(\pi, \pi \alpha)$. Then the following estimates hold

$$ |E_{\alpha,\beta}(z)| \leq c_1 (1 + |z|)^{\frac{\beta-\alpha}{\alpha}} e^{\Re\frac{z}{\alpha}} + \frac{c_2}{1 + |z|} |\arg(z)| \leq \mu, $$

$$ |E_{\alpha,\beta}(z)| \leq \frac{c}{1 + |z|} \mu \leq |\arg(z)| \leq \pi. $$

The solution representation (8) can be succinctly rewritten as

$$ u(t) = \mathbb{E}(t)u_0 + \int_0^t \mathbb{E}(t-s)r(\cdot, s) \, ds, \quad (12) $$

where the solution operators $\mathbb{E}$ and $\mathbb{E}$ are defined by, respectively

$$ \mathbb{E}(t)v = \sum_{j=1}^{\infty} E_{\alpha,\alpha}(-\lambda_j t^\alpha)(v, \varphi_j)\varphi_j, \quad (13) $$

and

$$ \mathbb{E}(t)v = \sum_{j=1}^{\infty} t^{\alpha-1} E_{\alpha,\alpha}(-\lambda_j t^\alpha)(v, \varphi_j)\varphi_j. \quad (14) $$

Since $\frac{d}{dt}E_{\alpha,1}(-\lambda_j t^\alpha) = -\lambda_j t^{\alpha-1}E_{\alpha,\alpha}(-\lambda_j t^\alpha)$ we have

$$ \frac{d}{dt} \mathbb{E}(t)v = \mathbb{E}(t)\mathbb{L}v. \quad (15) $$
2.2. Existence, uniqueness and regularity

The differences between the parabolic ($\alpha = 1$) and fractional ($\alpha < 1$) cases is substantial in terms of the regularity of the direct solution even in the homogeneous case $r = 0$. For the parabolic problem, if $u_0 \in L^2(\Omega)$ then $u(t)$ is infinitely differentiable in $t$ and so we have an infinite pick up in regularity of the solution. For the fractional problem, this is not the case; one can only recover at most two weak derivatives; see [26]. This difference is due to the exponential decay of $e^{-\lambda t}$, in the parabolic case but the only linear decay of $E_{\alpha,1}(-\lambda t^\alpha)$ for $0 < \alpha < 1$.

There is also a distinction for the case of a linear nonhomogeneous term as shown in [26] and lemma 2.3 below, which we will use to obtain a corresponding result for the nonlinear situation (4). As in the parabolic case, the main tool for this is Gronwall’s inequality; we will need a version of this for the fractional operator case.

Now we introduce the concept of weak solution for problem (7).

**Definition 2.1.** We call $u$ a solution to problem (7) if (7) holds in $L^2(\Omega)$ and $u(\cdot, t) \in H^1_0(\Omega)$ for almost all $t \in (0,T)$ and $u \in C([0,T]; \dot{H}^{-\sigma}(\Omega))$, with

$$\lim_{t \to 0} \|u(\cdot, t) - u_0\|_{\dot{H}^{-\sigma}(\Omega)} = 0,$$

with some $\sigma \geq 0$, which may depend on $\alpha$.

In the homogeneous case, existence and uniqueness of a solution to (7) can be concluded from (12), and the following stability estimate on the solution operator $E(t)$ defined in (13).

**Lemma 2.2.** For the operator $E(t)$ defined in (13), we have, for any $t > 0$,

$$\|E(t)v\|_{\dot{H}^r(\Omega)} \leq c t^{\frac{\alpha}{\alpha + 2}} \|v\|_{\dot{H}^r(\Omega)},$$

where $0 \leq q \leq 2$ and $q \leq p \leq q + 2$, and the constant $c$ depends only on $\alpha$ and $p - q$.

**Proof.** By the definition of the operator $E(t)$, there holds

$$\|E(t)v\|_{\dot{H}^r(\Omega)}^2 = \sum_{j=1}^{\infty} \lambda_j^p |E_{\alpha}(-\lambda_j t^\alpha)|^2 (v, \phi)\^2 = \sum_{j=1}^{\infty} \lambda_j^p \alpha j t^{\alpha(\alpha - p)} |E_{\alpha}(-\lambda_j t^\alpha)|^2 (v, \phi)\^2.$$

By corollary 2.1, we have

$$|E_{\alpha}(-\lambda_j t^\alpha)| \leq \frac{c}{1 + \lambda_j^p t^{\alpha}$$

where the constant $c$ depends only on $\alpha$. Consequently

$$\|E(t)v\|_{\dot{H}^r(\Omega)}^2 \leq c t^{(q-p)\alpha} \sup_j \frac{\lambda_j^p \alpha j t^{\alpha(\alpha - p)}}{(1 + \lambda_j^p t^{\alpha})^2} \sum_{j=1}^{\infty} \lambda_j^p (v, \phi)\^2 \leq c t^{(q-p)\alpha} \|v\|_{\dot{H}^r(\Omega)}^2,$$

where the last inequality follows from the choice of $p$ and $q$ such that $0 \leq p - q \leq 2$ and Young’s inequality $ab \leq \frac{1}{2} a^p + \frac{p}{p-1} b^{\frac{p}{p-1}}$, which yields
\[
\lambda_j^{p-q}(p-q)\alpha \leq \frac{p-q}{2} (\lambda_{j}^{\rho \alpha})^2 + 1 - \frac{p-q}{2} \leq (1 + \lambda_j^{\rho \alpha})^2. 
\]

\[\]

**Remark 2.1.** The choice \( p \leq q + 2 \) in lemma 2.2 indicates that the operator \( \tilde{E}(t) \) has at best a smoothing property of order 2 in space, which contrasts sharply with the classical parabolic case, for which the following estimate holds [29]

\[
\|\tilde{E}(t)v\|_{H^p(\Omega)} \leq ct^{(q-p)/2} \|v\|_{H^p(\Omega)},
\]

for \( q \geq 0 \) and any \( p \geq q \). This is reflected in the exponential decay of \( e^{-\lambda \tau} \), instead of the linear decay \( E_{\alpha,1}(-\lambda_{j}^{\rho \alpha}) \) for \( 0 < \alpha < 1 \).

Using Young’s inequality for convolution

\[
\|g * h\|_{L^p(\mathbb{R})} \leq \|g\|_{L^p(\mathbb{R})} \|h\|_{L^q(\mathbb{R})},
\]

(16)

for \( g \in L^q(\mathbb{R}), h \in L^r(\mathbb{R}), p^{-1} + 1 = q^{-1} + r^{-1}, p, q, r \geq 1 \), one can similarly prove the following result on the solution operator \( \tilde{E}(t) \).

**Lemma 2.3.** For the operator \( \tilde{E}(t) \) defined in (14), there holds, for any \( t > 0 \),

\[
\|\tilde{E}(t)v\|_{H^p(\Omega)} \leq ct^{-1+\frac{2p}{q}} \|v\|_{H^p(\Omega)},
\]

where \( 0 \leq q \leq 2 \) and \( q \leq p \leq q + 2 \), and the constant \( c \) depends only on \( \alpha \) and \( p - q \).

Gronwall’s inequalities are a basic tool for differential equations; that shown here is the classical integral formulation that does not require interior differentiability of the function \( u(t) \).

**Theorem 2.1.** Let \( \beta \) and \( u \) be real-valued continuous functions defined on the interval \([a, b]\) while \( \alpha \) is integrable on every subinterval of \([a, b]\). If \( \beta(t) \geq 0 \) and if

\[
u(t) \leq \alpha(t) + \int_a^t \beta(s)u(s) \, ds \quad \text{for all } t \in [a, b],
\]

then

\[
u(t) \leq \alpha(t) + \int_a^t \alpha(s)\beta(s)e^{\int_s^t \beta(\tau) \, d\tau} \, ds, \quad \text{for all } t \in [a, b].
\]

(17)

If, in addition, the function \( \alpha \) is non-decreasing, then

\[
u(t) \leq \alpha(t) e^{\int_a^t \beta(\tau) \, d\tau}, \quad \text{for all } t \in [a, b].
\]

(18)

We will need the corresponding results for fractional-order equations.

**Theorem 2.2.** Let \( \beta > 0 \). Suppose that \( a(t) \) is a nonnegative, locally integrable function on \((0, T)\) for some \( T > 0 \), while \( g(t) \) is a nonnegative, nondecreasing continuous function on \((0, T)\). If \( u(t) \) is nonnegative and locally integrable on \((0, T)\) such that

\[
u(t) \leq a(t) + g(t) \int_0^t (t-\tau)^{\beta-1} u(\tau) \, d\tau, \quad 0 < t < T.
\]

(19)
Then
\[ u(t) \leq a(t) + \int_0^t \sum_{n=1}^{\infty} \frac{(g(t)\Gamma(\beta))^n}{\Gamma(n\beta)} (t-\tau)^{n\beta-1} a(\tau) \, d\tau. \]  
(20)

In addition, if \( a(t) \) is nondecreasing on \([0, T]\) then
\[ u(t) \leq a(t) E_{\beta,1}(\Gamma(\beta)t^\beta g(t)), \quad 0 < t < T. \]  
(21)

**Proof.** For locally integrable functions \( \phi(t) \) on \([0, T]\) define the operator \( B \) by
\[ B\phi(t) = g(t) \int_0^t (t-\tau)^{\beta-1} \phi(\tau) \, d\tau. \]  
Then for each integer \( n \) by successively applying (19) we obtain
\[ u(t) \leq \sum_{k=0}^{n} B^k a(t) + B^n u(t). \]  
(22)

We claim that
\[ B^n u(t) \leq [g(t)]^n \int_0^t \frac{\Gamma(\beta)^n}{\Gamma(n\beta)} (t-\tau)^{n\beta-1} u(\tau) \, d\tau. \]  
(23)

The proof of this is found by induction, and the inequality (23) is clearly true for \( n = 1 \). Assume now it is true for \( n = j \), then we have
\[ B^{j+1} u(t) = B(B^j u(t)) \]
\[ \leq g(t) \int_0^t (t-\tau)^{\beta-1} \left[ \int_0^\tau \frac{(t-\tau)^{\beta-1} a(s) \, ds}{\Gamma(j\beta)} \right] \, d\tau \]
\[ \leq (g(t))^{j+1} \int_0^t (t-\tau)^{\beta-1} \left[ \int_0^\tau \frac{\Gamma(\beta)^j}{\Gamma(j\beta)} (t-\tau)^{j\beta-1} u(\tau) \, d\tau \right] \, d\tau \]
\[ \leq (g(t))^{j+1} \int_0^t \left[ \int_0^\tau (t-\tau)^{j-1} (t-\tau)^{\beta-1} u(\tau) \, d\tau \right] \frac{\Gamma(\beta)^j}{\Gamma(j\beta)} \, d\tau \]
\[ = (g(t))^{j+1} \int_0^t \frac{\Gamma(\beta)^{j+1}}{\Gamma((j+1)\beta)} (t-\tau)^{(j+1)\beta-1} u(\tau) \, d\tau, \]  
(24)

where the second line follows from the first from the fact that \( g(t) \) is nondecreasing, and the third line follows by an obvious change of variables. Writing the integrand as a beta function and then invoking the beta–gamma function duplication formula yields the identity
\[ \int_0^\tau (t-\tau)^{j\beta-1} (t-\tau)^{\beta-1} \, d\tau = \frac{\Gamma(j\beta)\Gamma(\beta)}{\Gamma((j+1)\beta)} (t-s)^{(j+1)\beta-1}. \]  
This now shows (23).

With \( M = \sup_{[0,\tau]} g(t) \) we obtain
\[ B^n u(t) \leq \int_0^t \frac{(M\Gamma(\beta))^n}{\Gamma(n\beta)} (t-\tau)^{n\beta-1} u(\tau) \, d\tau, \]  
(25)

and it is easily seen that
\[ B^n u(t) \to 0 \quad \text{as} \quad n \to \infty. \]  
(26)
From the above we now obtain (20).

If \(a(t)\) is nondecreasing on \([0, T]\) then we can write (20) as

\[
u(t) \leq a(t) \left[ 1 + \int_0^t \sum_{n=1}^{\infty} \frac{(g(t)\Gamma(\beta))^n}{\Gamma(n\beta)} (t - \tau)^{n\beta - 1} a(\tau) \, d\tau \right]
\]

\[
\leq a(t) \sum_{n=0}^{\infty} \frac{(g(t)\Gamma(\beta))^n}{\Gamma(n\beta + 1)} = a(t)E_{\beta,1}\left(\Gamma(\beta)t^{\beta}g(t)\right),
\]

where we have used the fact that \(\int_0^t (t - \tau)^{n\beta - 1} \, d\tau = \frac{\Gamma(n\beta)}{\Gamma(n\beta + 1)}\).

(27)

An important particular case occurs when \(g\) is constant and the expression on the right-hand side of the estimate can be written as a Mittag-Leffler function.

**Lemma 2.4.** Suppose \(b \geq 0\), \(\beta > 0\) and \(a(t)\) is a nonnegative function locally integrable on \(0 \leq t < T\) (for some \(T < +\infty\)), and suppose \(u(t)\) is nonnegative and locally integrable on \(0 \leq t < T\) with

\[
u(t) \leq a(t) + b \int_0^t (t - s)^{\beta-1} u(s) \, ds \quad 0 \leq t < T.
\]

Then

\[
u(t) \leq a(t) + b \int_0^t (t - s)^{\beta-1} E_{\beta,\beta}(b(t - s)^{\beta})a(s) \, ds, \quad 0 \leq t < T.
\]

Moreover, the following convolution version of Gronwall’s inequality will be employed.

**Lemma 2.5.** Suppose \(c \geq 0\) and \(b(t)\) is a nonnegative function locally integrable on \(0 \leq t < T\) (for some \(T < +\infty\)), and suppose \(u(t)\) is nonnegative and locally integrable on \(0 \leq t < T\) with

\[
u(t) \leq b(t) + c \int_0^t e^{-\lambda(t-s)} u(s) \, ds.
\]

Then

\[
u(t) \leq b(t) + c \int_0^t e^{-(\lambda-c)(t-s)} b(s) \, ds.
\]

**Proof.** The proof proceeds as in that of theorem 2.2, but in this case, it is easier since there is no singularity involved. Define \((Ba)(t) = c \int_0^t e^{-\lambda(t-s)} a(s) \, ds\), then the inequality for \(u\) reads

\[
u(t) \leq b(t) + (Bu)(t)\]

and implies \(u(t) \leq \sum_{k=0}^{\infty} (B^k b)(t) + (B^k u)(t)\), where (by induction)

\[
(B^k a)(t) = c^k \int_0^t \frac{(t-s)^{k-1}}{(k-1)!} e^{-\lambda(t-s)} a(s) \, ds \to 0 \text{ as } k \to \infty,
\]



hence

\[
u(t) \leq \sum_{k=0}^{\infty} (B^k b)(t) = b(t) + c \int_0^t e^{-(\lambda-c)(t-s)} b(s) \, ds.
\]

\[\square\]
2.3. Semilinear problems

Now we briefly discuss the slightly more complex case, i.e. semilinear problems, using a fixed point argument.

Consider the following initial boundary value problem for a semilinear subdiffusion equation

\[
\begin{aligned}
D^\alpha_t u &= \mathbb{L}u + f(u) + r \quad \text{in } \Omega \times (0, T), \\
\partial_{\nu}u + \gamma u &= 0 \quad \text{on } \partial\Omega \times (0, T), \\
u(t, 0) &= u_0 \quad \text{in } \Omega.
\end{aligned}
\]  

(28)

In the model, we assume

\[
\sup_{\xi \in \mathbb{R}} \left| \frac{d^pf}{d\xi^k} \right| < \infty, \quad k = 0, 1 \quad \text{and} \quad r \in C^1([0, T]; \mathbb{L}^2(\Omega)).
\]  

(29)

The argument below is based on the operator theoretic approach in \( \mathbb{L}^2(\Omega) \); see also [12, theorem 3.1].

**Theorem 2.3.** Let \( u_0 \in \mathbb{H}^2(\Omega) \), \( f(u_0) \in \mathbb{L}^2(\Omega) \), and (29) hold. Then the solution \( u \) to problem (28) belongs to \( C([0, T]; \mathbb{H}^2(\Omega)) \cap C^1((0, T]; \mathbb{L}^2(\Omega)) \) such that \( D^\alpha_t u \in C((0, T]; \mathbb{L}^2(\Omega)) \).

**Proof.** For the unique existence of \( u \), it suffices to discuss the integral equation; see (12)

\[
u(t) = \mathbb{E}(t)u_0 + \int_0^t \mathbb{E}(t-s)f(u(s)) \, ds + \int_0^t \mathbb{E}(t-s)r(s) \, ds, \quad 0 < t < T.
\]

First we estimate \( u_i(t) \), which is given by

\[
u_i(t) = \partial_t \mathbb{E}(t)u_0 + \partial_t \left( \int_0^t \mathbb{E}(s)f(u(t-s)) \, ds \right) + \partial_t \left( \int_0^t \mathbb{E}(s)r(t-s) \, ds \right)
\]

\[
= \mathbb{E}(t)(f(u_0) + r(0) - Au_0) + \int_0^t \mathbb{E}(s) f'(u(t-s))u_i(t-s) \, ds + \int_0^t \mathbb{E}(s)r_i(t-s) \, ds,
\]

where we have used (15). Consequently, for any \( 0 < t < T \), by lemma 2.3 with \( p = q = 0 \), there holds

\[
\|u_i(t)\|_{\mathbb{L}^2(\Omega)} \leq \int_0^t \|\mathbb{E}(s)\|_{\mathbb{L}^2(\Omega)} \|f'(u(t-s))\|_{L^\infty(\Omega)} \|u_i(t-s)\|_{\mathbb{L}^2(\Omega)} \, ds
\]

\[
+ \|\mathbb{E}(t)(r(0) + f(u_0) - Au_0)\|_{\mathbb{L}^2(\Omega)} + \int_0^t \|\mathbb{E}(s)\|_{\mathbb{L}^2(\Omega)} \|r_i(t-s)\|_{\mathbb{L}^2(\Omega)} \, ds
\]

\[
\leq c \int_0^t s^{\alpha-1} \|f'(u(t-s))\|_{L^\infty(\Omega)} \|u_i(t-s)\|_{\mathbb{L}^2(\Omega)} \, ds
\]

\[
+ c \int_0^t s^{\alpha-1} \|r_i(t-s)\|_{\mathbb{L}^2(\Omega)} \, ds + \|\mathbb{E}(t)(r(0) + f(u_0) - Au_0)\|_{\mathbb{L}^2(\Omega)}.
\]

Using assumption (29) on \( f \), we have for \( 0 < t < T \)

\[
\|u_i(t)\|_{\mathbb{L}^2(\Omega)} \leq c \int_0^t s^{\alpha-1} \|u_i(t-s)\|_{\mathbb{L}^2(\Omega)} \, ds + c \int_0^t s^{\alpha-1} \, ds \|r(0)\|_{\mathbb{L}^2(\Omega)} + \|f(u_0)\|_{\mathbb{L}^2(\Omega)} + \|u_0\|_{\mathbb{H}^2(\Omega)}.
\]
By lemma 2.4, we obtain
\[ \|u(t)\|_{L^2(\Omega)} \leq c t^{-\alpha} \left( \|f(u_0)\|_{L^2(\Omega)} + \|r\|_{C^1([0,T];L^2(\Omega))} + \|u_0\|_{H^1(\Omega)} \right) \]
and \( u \in C^1((0, T]; L^2(\Omega)) \) and thus
\[ \|D^\alpha u(t)\| \leq \int_0^T (t-s)^{-\alpha} \|u(s)\| \, ds \leq c \left( \|f(u_0)\|_{L^2(\Omega)} + \|r\|_{C^1([0,T];L^2(\Omega))} + \|u_0\|_{H^1(\Omega)} \right). \]

i.e. \( D^\alpha u \in C([0, T]; L^2(\Omega)) \). Further, using again (15), for any \( \sigma \in [0, 2] \)
\[ A^\sigma u(t) = A^\sigma E(t)u_0 + \int_0^t A^\sigma E(t-s)f(u(s)) \, ds + \int_0^t A^\sigma E(t-s)r(s) \, ds \]
\[ = A^\sigma E(t)u_0 + A^\sigma E(t-s)(f(u(t)) + A^\sigma E(t-s)(r(s) - r(t))) \]
\[ := \sum_{i=1}^N I_i. \]

Clearly, \( \|I_1\|_{L^2(\Omega)} \leq c \|u_0\|_{H^\sigma(\Omega)} \). By lemma 2.2, for \( I_2 \) and \( I_3 \), we have
\[ \|I_2\|_{L^2(\Omega)} \leq c \|f(u(t))\|_{L^2(\Omega)} \leq c \|f\|_{L^\infty(\Omega)}, \]
\[ \|I_3\|_{L^2(\Omega)} \leq c \|r(t)\|_{L^2(\Omega)} \leq c \|r\|_{C^1([0,T];L^2(\Omega))}. \]

By the mean value theorem, we have
\[ f(u(s)) - f(u(t)) = \int_0^1 f'(u(t) + \theta(u(s) - u(t))) \, d\theta(u(s) - u(t)). \]

Consequently,
\[ \|f(u(s)) - f(u(t))\|_{L^2(\Omega)} \leq \|f'\|_{L^\infty(\Omega)} \|u(s) - u(t)\|_{L^2(\Omega)} \]
\[ \leq c \int_t^s \|u_\sigma(\tau)\|_{L^2(\Omega)} \, d\tau \leq c \int_0^t \|u_\sigma(\tau)\|_{L^2(\Omega)} \, d\tau \]
\[ \leq c \int_0^t \tau^{\alpha-1} \, d\tau (\|f(u_0)\|_{L^2(\Omega)} + \|r\|_{C^1([0,T];L^2(\Omega))} + \|u_0\|_{H^1(\Omega)}) \]
\[ \leq c (t^{\alpha} - \sigma^\alpha)(\|f(u_0)\|_{L^2(\Omega)} + \|r\|_{C^1([0,T];L^2(\Omega))} + \|u_0\|_{H^1(\Omega)}). \]

Hence, by lemma 2.3 with \( p = \sigma, q = 0 \), we deduce
\[ \|I_4\|_{L^2(\Omega)} \leq \int_0^s |A^\sigma E(t-s)| ((t^{\alpha} - \sigma^\alpha) \, ds (\|f(u_0)\|_{L^2(\Omega)} + \|r\|_{C^1([0,T];L^2(\Omega))} + \|u_0\|_{H^1(\Omega)}) \]
\[ \leq c \int_0^T (s^{\alpha-1} - \sigma^\alpha) \, ds (\|f(u_0)\|_{L^2(\Omega)} + \|r\|_{C^1([0,T];L^2(\Omega))} + \|u_0\|_{H^1(\Omega)}) \]
\[ \leq c (\|f(u_0)\|_{L^2(\Omega)} + \|r\|_{C^1([0,T];L^2(\Omega))} + \|u_0\|_{H^1(\Omega)}). \]
Likewise, since \( r(s) - r(t) = \int_t^s r(\tau) d\tau \),
\[
\| I_s \| \leq c \| r \|_{C([0,T];L^2(\Omega))}.
\]
Therefore, we have \( A^T u \in C([0,T];L^2(\Omega)) \), i.e. \( u \in C([0,T];H^s(\Omega)) \). This completes the proof of the theorem. \( \square \)

3. An iteration scheme to recover \( f \)

A natural scheme to recover \( f \) is to evaluate equation (4) on the overposed boundary \( t = T \). That is, we define a map \( T : f \mapsto u(x,T;f) \) by
\[
Tf(g) = D_t^0 u(x;f) - Lg - r(x,T), \quad x \in \Omega,
\]
where \( g \) is the given data and define a sequence of approximations by the fixed point iteration
\[
x_{k+1} = T(x_k),
\]
for \( t \leq T \):
\[
\text{Range } u(x,t;f) \subset \text{Range } u(x,T;f) \quad 0 \leq t < T.
\]

The operator \( T \) is a concatenation \( T = P \circ S \) of the operators \( P : Z \to X \) and \( S : X \to Z \) defined by
\[
\begin{align*}
\text{Py such that } & (Py)(g(x)) = y(x), \quad x \in \Omega, \text{ more generally} \\
\text{Py} \in \text{argmin} \{ \| f^t(g) - y \|_Z : f^t \in X \text{ and } \| f^0(u_0) + r(\cdot,0) - Lu_0 \|_{H^s(\Omega)} \leq \rho_0 \},
\end{align*}
\]
where we choose \( \rho_0 \geq \| f_{\text{act}}(u_0) + r(\cdot, T) - Lu_0 \|_{H^s(\Omega)} \) to make sure that a fixed point of \( T \) solves the original inverse problem (35), (3).

We use a bounded interval \( I = [g_{\text{min}}, g_{\text{max}}] \) with \( g_{\text{min}} = \min \{ g(x) : x \in \Omega \} \), \( g_{\text{max}} = \max \{ g(x) : x \in \Omega \} \), in order to be able to make use of embedding theorems. Then we use the function space setting
\[
X = \{ f \in W^{1,\infty}(I) : f(u_0) \in \dot{H}^s(\Omega) \}, \quad Z = \dot{H}^s(\Omega),
\]
with \( \sigma \) such that \( \dot{H}^s(\Omega) \) continuously embeds into \( W^{1,\infty}(\Omega) \) and with the norm
\[
\| f \|_X = \| f \|_{W^{1,\infty}(I)} + \| f(u_0) \|_{\dot{H}^s(\Omega)}.
\]
Moreover, we employ the projection \( P(z) = \max \{ \min \{ z, g_{\text{max}} \}, g_{\text{min}} \} \) on \( I \) to define \( u(x,t;f) \) as a solution to
\[
\begin{align*}
D_t^0 u(x,t) - Lu(x,t) &= f(Pu(x,t)) + r(x,t), \quad (x,t) \in \Omega \times (0,T) \\
\partial_t u(x,t) + \gamma(x)u(x,t) &= 0 \quad (x,t) \in \partial\Omega \times (0,T) \\
u(x,0) &= u_0 \quad x \in \Omega.
\end{align*}
\]
Hence, based on the range condition (31), which we need to assume holding for the actual nonlinearity \( f_{\text{act}} \) only, and the fact that \( u(x, T; f_{\text{act}}) \) solves (4), we can replace the original model (4) by the equation containing the projection (35).

In order to establish well-definedness of the operator \( \mathbb{P} \) we will assume throughout this section that

\[
g \in \dot{H}^\sigma(\Omega) \subseteq W^{1,\infty}(\Omega), \quad g \geq |\nabla g(x)| \geq \bar{g} > 0, \quad x \in \Omega, \quad \| g \|_{\mathbb{Z}} \in \mathbb{Z},
\]

(36)

for some \( 0 < \bar{g} < \bar{g} \), so that there exist \( c(g), C(g) > 0 \) such that

\[
c(g)\| f(g) \|_{W^{1,\infty}(\Omega)} \leq \| f \|_{X} \leq \bar{C}(g)\| f(g) \|_{\dot{H}^\sigma(\Omega)} \text{ for all } f \in X.
\]

(37)

Indeed, we have in case \( \sigma = 2 \)

\[
\| f \|_{\dot{H}^\sigma(\Omega)} = \left( \int_\Omega (|f|^2 + |f'|^2 + |f''|^2) \, dx \right)^{1/2}
\]

\[
\leq \bar{C}(g) \left( \int_\Omega \left( |f(g)|^2 + |f''(g)|\nabla g|^2 + |f'(g(x))\Delta g|^2 \right) \, dx \right)^{1/2}
\]

\[
= \bar{C}(g) \left( \int_\Omega \left( |f \circ g(x)|^2 + |\Delta(f \circ g)(x)|^2 \right) \, dx \right)^{1/2}
\]

\[
\leq \bar{C}(g)\| f \|_{\dot{H}^\sigma(\Omega)},
\]

as well as in case \( \sigma = 1 \)

\[
\| f \|_{\dot{H}^\sigma(\Omega)} = \left( \int_\Omega (|f|^2 + |f'|^2) \, dx \right)^{1/2}
\]

\[
\leq \bar{C}(g) \left( \int_\Omega \left( |f(g)|^2 + |f'(g)|\nabla g|^2 \right) \, dx \right)^{1/2}
\]

\[
= \bar{C}(g) \left( \int_\Omega \left( |f \circ g(x)|^2 + |\nabla(f \circ g)(x)|^2 \right) \, dx \right)^{1/2}
\]

\[
\leq \bar{C}(g)\| f \|_{\dot{H}^\sigma(\Omega)},
\]

from which the general case \( q \in [1, 2] \) follows by interpolation. Moreover,

\[
\| f(g) \|_{W^{1,\infty}(\Omega)} = \sup_{x \in \Omega} |f(g(x))| + \sup_{x \in \Omega} |f'(g(x))| |\nabla g(x)| \leq \min \left\{ \bar{g}^{-1}, 1 \right\} \| f(g) \|_{\dot{H}^\sigma(\Omega)} \leq \bar{C}(g)\| f(g) \|_{\dot{H}^\sigma(\Omega)}.
\]

(38)

**Lemma 3.1.** Under condition (36), the operator \( \mathbb{P} : Z \to X \) is well defined by (32) between the spaces defined by (33), and satisfies

\[
\| \mathbb{P} y \|_{X} \leq 2\bar{C}(g)\| y \|_{Z} \text{ for all } y \in Z.
\]

**Proof.** Existence of a minimizer follows from the fact that \( \{ \| f^\sigma(g) - y \|_{Z} : f^\sigma \in X \text{ and } \| f(u_0) \|_{\dot{H}^\sigma(\Omega)} \leq \rho_0 \} \) is a nonempty closed convex (hence weakly∗ closed) subset of the space \( X \) which is the dual of a separable space. Moreover, from (37) and the triangle inequality as well as by minimality of \( \mathbb{P} f \), comparing with \( f^\sigma = 0 \), we get
\[\|Py\|_X \leq C(g)\|Py\|_Z \leq C(g)\left(\|(Py)(g) - y\|_Z + \|y\|_Z\right) \leq C(g)\left(\|0 - y\|_Z + \|y\|_Z\right) = 2C(g)\|y\|_Z.\]  

(39)

\[\square\]

**Remark 3.1.** For \(f \in X\), \(r(\cdot, T) \in Z\), under the attainability condition

\[\|f(u_0) + r(\cdot, 0) - L_0\|_{H^r(\Omega)} < \rho_0 \text{ and}\]

\(y := f(u(\cdot, T; f)) + L(u(\cdot, T; f) - g) \in Z_\xi := \{f^2(g) : f^2 \in X\},\]

any minimizer \(f^+\) of (32) satisfies \(f^+(g) = y\) and is therefore unique, due to the fact that \(g(\Omega) = I\). The second condition in (40) can be shown to be satisfied by the strict monotonicity of \(g\) according to (36), if \(\Omega = (0, L) \subseteq \mathbb{R}^1\) and \(u_0 \in W^{1,\infty}(\Omega) \cap H^r(\Omega)\) is strictly monotone. This can be seen by using the identity \(f(u(\cdot, T; f)) + L(u(\cdot, T; f) - g) = D^g_{\xi}u(\cdot, T; f) - r(\cdot, T) - Lg \in Z\) (this regularity of \(D^g_{\xi}u(\cdot, T; f)\) will be shown at the beginning of the next section) and using the assumed regularity (36) of \(g\), which transfers to its inverse and gives \(f^+ = y \circ g^{-1} \in W^{1,\infty}(I) \cap H^r(I) \subseteq X\), as an easy verification of the chain rule in \(W^{1,\infty}(a, b) \cap H^r(a, b)\) for \((a, b) = \Omega\) and \((a, b) = I\) shows.

In our implementation we use a regularized version of \(P\) by restricting the search for a minimizer of \(\|f^2(g) - y\|_X\) to a finite-dimensional subset of smooth functions in \(X\). This is actually an intrinsic part of the algorithm and to some extent explains its good contraction performance in spite of the fact that in an infinite-dimensional function space setting, contractivity of \(T\) can only be proven in the conditional sense of theorem 3.2 below.

Well-definedness of the operator \(\mathbb{S} : X \rightarrow L^2(\Omega)\) follows from theorem 2.3. The higher regularity \(\mathbb{S}f \in H^r(\Omega)\) will be established in the following section.

### 3.1. Self-mapping property of the fixed point operator

To prove that \(D^\xi_{f}\xi u(x, T; f) \in Z\) we make use of the representation formulae (12)–(14)

\[u(\cdot, t) = \sum_{j=1}^{\infty} \left\{ E_{\alpha,1}(-\lambda_j f^\alpha)(u_0, \varphi_{j}) + \int_0^s \alpha^{-1} E_{\alpha,\alpha}(-\lambda_j s^\alpha)(f(Pu(\cdot, t - s)) + r(\cdot, t - s), \varphi_{j}) \right\} \varphi_{j}.\]

Hence

\[u_t(\cdot, t) = \sum_{j=1}^{\infty} \left\{ -\lambda_j \alpha^{-1} E_{\alpha,\alpha}(-\lambda_j f^\alpha)(u_0, \varphi_{j}) \right\} + \int_0^s \alpha^{-1} E_{\alpha,\alpha}(-\lambda_j s^\alpha)\left(f(Pu(\cdot, t - s))Pu(\cdot, t - s) + r(\cdot, t - s), \varphi_{j}\right) ds} \varphi_{j},\]

which by the triangle inequality and Young’s inequality for convolution (16) (with \(p = \infty, q = Q, r = Q^* = \frac{Q}{Q-1}\)) yields
\[ \|u(t)\|_{H^r(\Omega)} \leq \sum_{j=1}^{\infty} \lambda_j^{\alpha+2} \left( \rho^{\alpha-1} E_{\alpha,\alpha} \left( -\lambda_j s^\alpha \right) \right)^{2} \left( u_0, \varphi_j \right)^{2} \left( 1/2 \right) \]
\[ + \left( \sum_{j=1}^{\infty} \lambda_j^{\alpha} \rho^{\alpha-1} E_{\alpha,\alpha} \left( -\lambda_j s^\alpha \right) \right)^{2} \left( f(Pu_0) + r(\cdot, 0), \varphi_j \right)^{1/2} \]
\[ + \left( \sum_{j=1}^{\infty} \lambda_j^{\alpha} \left( \int_0^t s^{\alpha-1} E_{\alpha,\alpha} \left( -\lambda_j s^\alpha \right) f'(Pu(t-s)) Pu(t-s) + r(t-s), \varphi_j \right) \right)^{1/2} \]
\[ \leq \rho^{\alpha-1} \left( \|f(Pu_0) + r(\cdot, 0) - \mathbb{1} u_0\|_{H^r(\Omega)} + \tilde{C}^2_\alpha Q^r \|f'(Pu) Pu + r\|_{L^r(0,T;L^2(\Omega))} \right) \]

for \( Q \leq \frac{2(2-\theta)}{2-2\theta+\sigma} \), that is, \( Q^* = \frac{Q}{\sigma - 1} \geq \frac{2(2-\theta)}{2-2\sigma} \).

Here we have used the fact that
\[ \int_0^t \left( s^{\alpha-1} E_{\alpha,\alpha} \left( -\lambda_j s^\alpha \right) \right)^{2} ds \leq \tilde{C}_\alpha^{2} \lambda_j^{-\theta} \]
(41)

for \( \theta \in (0, 2 - 1/\alpha) \) and
\[ \left( \int_0^t \left| s^{\alpha-1} E_{\alpha,\alpha} \left( -\lambda_j s^\alpha \right) \right| ds \right)^{\frac{2 - \theta}{\sigma}} \leq \tilde{C}_\alpha^{ \frac{2 - \theta}{\sigma} \lambda_j^{-\theta} \left( 2 - \theta(\theta - 1) \right)} \]

which leads to the requirement \( \sigma \leq \frac{2}{Q} \left( 2 - Q + \theta(Q - 1) \right) = 2 - 2 \frac{Q}{Q - 1} \left( 2 - \theta \right) \) (that is, \( Q \leq \frac{2(2-\theta)}{2-2\theta+\sigma} \)) in order to cancel the powers of \( \lambda_j \) in the term containing \( f'(Pu) \), so that no space derivatives are finally applied to \( f'(Pu) \).

We can therefore further estimate, noting that by the range condition (31) \( u_0 = Pu_0 \)
\[ \|u(t)\|_{H^r(\Omega)} \leq \rho^{\alpha-1} \left( \|f(u_0) + r(\cdot, 0) - \mathbb{1} u_0\|_{H^r(\Omega)} \right) \]
\[ + \tilde{C}^2_\alpha Q^r \left( \|f\|_{L^\infty(\Omega)} \|u\|_{L^r(0,T;L^2(\Omega))} + \|r\|_{L^r(0,T;L^2(\Omega))} \right) \]
\[ \leq \rho^{\alpha-1} \left( \|f(u_0) + r(\cdot, 0) - \mathbb{1} u_0\|_{H^r(\Omega)} \right) \]
\[ + \tilde{C}^2_\alpha Q^r \left( \frac{1}{\lambda_1} \|f\|_{L^\infty(\Omega)} \|u\|_{L^r(0,T;L^2(\Omega))} + \|r\|_{L^r(0,T;L^2(\Omega))} \right) \]
(42)

If \( Q^* < \infty \) then for \( \eta(t) = \|u(t)\|_{H^r(\Omega)}^{Q^*} \) this means
\[ \eta(t) \leq \left( C_0 \rho^{\alpha-1} + C_r(t) + C_1 \left( \int_0^t \eta(s) \right) \right)^{1/2} \]
\[ \leq \tilde{C}_0 \eta(t)^{Q^*(\alpha-1)} + \tilde{C}_r(t) + \tilde{C}_1 \int_0^t \eta(s) \]

for \( C_0 = \|f(u_0) + r(\cdot, 0) - \mathbb{1} u_0\|_{H^r(\Omega)}, C_1 = \frac{\tilde{C}^2_\alpha Q^r}{\lambda_1} \|f\|_{L^\infty(\Omega)}, C_r(t) = \tilde{C}^2_\alpha Q^r \|r\|_{L^r(0,T;L^2(\Omega))} \)
\[ \tilde{C}_1 = 2^{Q^*-1} \tilde{C}^Q_r \quad \text{for } i = 0, 1, r. \]
Gronwall’s inequality (17) yields

\[ \eta(t) \leq \tilde{C}_0 L^2(\alpha-1) + \tilde{C}_1 + \tilde{C}_1 \int_0^t \left( \tilde{C}_0 s^{\alpha-1} + \tilde{C}_1(s) \right) e^{\tilde{C}_1(t-s)} \, ds, \]

provided \( \alpha \) is sufficiently close to one: \( \alpha > 1 - \frac{1}{2^\theta} = \frac{1}{2} \). That is,

\[ \|u(t)\|_{H^r(\Omega)} \leq 2^{1-1/\theta'} \left( \left( C_0 T^{\alpha-1} + C_1(T) + C_1 \left( \int_0^T \left( \tilde{C}_0 s^{\alpha-1} + \tilde{C}_1(s) \right) e^{\tilde{C}_1(T-s)} \, ds \right) \right)^{1/\theta'} \right) \]

\[ =: \Phi \left( \|f(u_0) + r(\cdot, 0) - L u_0\|_{H^r(\Omega)}, T, \|f\|_{L^\infty(I)}, \|r\|_{L^\infty(0,T,L^1(\Omega))} \right) \]

(note that \( C_0, \tilde{C}_0 \) contain \( \|f(u_0) + r(\cdot, 0) - L u_0\|_{H^r(\Omega)} \) and \( C_1, \tilde{C}_1 \) contain \( \|f\|_{L^\infty(I)} \)).

Note that \( Q^* = \infty \) does not work in the case \( \alpha < 1 \), because of the singularity at \( t = 0 \).

Also note that the condition on \( \alpha \) arising here is compatible with the previous one,

\[ Q \leq \frac{1}{2(\alpha - 1)} \quad \text{with} \quad \theta \in (0, 2 - \frac{1}{\alpha}), \]

since both together mean \( \alpha > 1 \) and can therefore be satisfied provided \( \sigma < 2 \) and \( \alpha \) is sufficiently close to one: \( \alpha > 1 - \frac{2-\sigma}{2} \). That is, since we need \( \sigma > \frac{1}{2} \) to guarantee continuity of the embedding \( H^r(\Omega) \to W^{1,\infty}(\Omega) \), this implies \( \alpha > \frac{3}{2} \).

The function \( \Phi \) appearing in the estimate (43) is an increasing function of the initial data \( \|f(u_0) + r(\cdot, 0) - L u_0\|_{H^r(\Omega)} \) of \( \|f\|_{L^\infty(I)} \) and of \( \|r\|_{L^\infty(0,T,L^1(\Omega))} \): moreover for fixed \( T \) and bounded \( \|f\|_{L^\infty(I)} \), the value of \( \Phi \) can be made arbitrarily small by making \( \|f(u_0) + r(\cdot, 0) - L u_0\|_{H^r(\Omega)} \) and \( \|r\|_{L^\infty(0,T,L^1(\Omega))} \) small.

We thus define

\[ B = \{ f \in X : \|f\|_{W^{1,\infty}(I)} \leq \rho, \|f(u_0) + r(\cdot, 0) - L u_0\|_{H^r(\Omega)} \leq \rho_0 \}, \]

and apply the above estimate (43) to get, for any \( f \in B \),

\[ \|\mathbb{T}(f)\|_{W^{1,\infty}(I)} = \|P(S(f))\|_{W^{1,\infty}(I)} \leq C(\rho) C_{H^r(\Omega),W^{1,\infty}} \left( \Phi(\rho_0, T, \rho, \|r\|_{L^\infty(0,T,L^1(\Omega))}) + \|r(\cdot, T) - Lg\|_2 \right) \leq \rho, \]

where the second inequality in (45) can be achieved by choosing \( \rho \) sufficiently large, \( \rho > C(g) C_{H^r(\Omega),W^{1,\infty}} \|Lg\|_2 \), and then choosing \( \rho_0, \|r\|_{L^\infty(0,T,L^1(\Omega))} \) small enough so that \( \Phi(\rho_0, T, \rho, \|r\|_{L^\infty(0,T,L^1(\Omega))}) \leq \frac{1}{c_{H^r(\Omega),W^{1,\infty}}} \rho - \|Lg\|_2 \). Moreover, by the constraint in the definition (32) of \( \mathbb{P} \), we have

\[ \|\mathbb{T}(f)(u_0) + r(\cdot, 0) - L u_0\|_{H^r(\Omega)} \leq \rho_0, \]

i.e. altogether \( \mathbb{T}(f) \in B \), provided \( f \in B \).

### 3.2. Weak' compactness of \( B \) and weak’ continuity of \( \mathbb{T} \)

The set \( B \) is by definition weakly’ compact and convex in the Banach space \( X \) with norm \( \|f\|_X = \|f\|_{W^{1,\infty}(I)} + \|f(u_0)\|_{H^r(\Omega)} \).

For any sequence \( (f_n)_{n \in \mathbb{N}} \in B \) converging weakly’ in \( X \) to \( f \in X \) we have that the sequence of images under \( \mathbb{T} \), i.e. \( \mathbb{T}(f_n) \), is contained in the weakly’ compact set \( B \). Using this and compactness of the embeddings \( X \to L^\infty(I) \) (where boundedness of the interval \( I \) is crucial) we can extract a subsequence with indices \( (n_k)_{k \in \mathbb{N}} \) and an element \( f^+ \in B \) such that
\[ T(f_n) \xrightarrow{\ast} f^+ \text{ in } X, \]
\[ T(f_n) \to f^+ \text{ in } L^\infty(I), \]
\[ f_n \to f \text{ in } L^\infty(I). \]  
\[ (47) \]

It remains to prove that \( f^+ = T(f) \). For this purpose, we use the fact that the difference \( \tilde{u}_n := u(x,t;f) - u(x,t;f) \) of solution to (35) solves
\[
\mathcal{D}_t^\alpha \tilde{u}_n - \Delta \tilde{u}_n - q_n \tilde{u}_n = f_n(Pu), \]
\[ (48) \]
with homogeneous initial and impedance boundary conditions, where \( q_n(x,t) = \int_0^t \int_{\Omega} (P(u(x,t';f) + \sigma \tilde{u}_n(x,t'))) \, dr \, dP \). \( \hat{f}_n = f_n - f \), and \( u = u(x,t;f) \). From the representation (12)–(14)
\[
\tilde{u}_n(x,t) = \sum_{j=1}^{\infty} \int_0^t (t-s)^{\alpha-1} E_{\alpha,\alpha}(-\lambda_j(t-s)^\alpha)((q_n \hat{u}_n + \hat{f}_n(Pu))(.s), (\varphi)) \, ds \varphi(x)
\]
\[ = \sum_{j=1}^{\infty} \tilde{u}_n^j(t) \varphi(x). \]  
\[ (49) \]
Young’s inequality (16) and (41), as well as the fact that \( f_n \in B \), allows us to obtain the crude estimate
\[
\| \tilde{u}_n(\cdot,t) \|_{L^2(\Omega)} \leq C_\alpha \| q_n \hat{u}_n + \hat{f}_n(Pu) \|_{L^2(0,T;H^{-\sigma}(\Omega))}
\]
\[ \leq C_\alpha \lambda_1^{-\rho} \left( \| q_n \|_{L^\infty(I)} \| \hat{u}_n \|_{L^2(0,T;L^2(\Omega))} + \| \hat{f}_n(Pu) \|_{L^2(0,T;L^2(\Omega))} \right)
\]
\[ \leq C_\alpha \lambda_1^{-\rho} \left( \rho \| \hat{u}_n \|_{L^2(0,T;L^2(\Omega))} + \sqrt{T} \sqrt{\| \hat{f}_n \|_{L^\infty(I)}} \right), \]

which by Gronwall’s inequality (17) yields
\[
\| \tilde{u}_n(\cdot,t) \|_{L^\infty(0,T;L^2(\Omega))} \leq C(\rho, T, |\Omega|) \| \hat{f}_n \|_{L^\infty(I)}. \]

Using the fact that \( \tilde{u}_n^j \) as defined in (49) satisfies the fractional ODE \( \mathcal{D}_t^\alpha \tilde{u}_n^j(t) + \lambda_j \tilde{u}_n^j(t) = ((q_n \hat{u}_n + \hat{f}_n(Pu))(.s), (\varphi)) \) we obtain
\[
\mathcal{D}_t^\alpha \tilde{u}_n(x,t) = \sum_{j=1}^{\infty} \left( \left( q_n \hat{u}_n + \hat{f}_n(Pu))(.s), (\varphi) \right) - \lambda_j \int_0^t (t-s)^{\alpha-1} E_{\alpha,\alpha}(-\lambda_j(t-s)^\alpha)((q_n \hat{u}_n + \hat{f}_n(Pu))(.s), (\varphi)) \, ds \right) \varphi(x).
\]

From Young’s inequality (16) and (41) we therefore obtain an estimate of \( \mathcal{D}_t^\alpha \tilde{u}_n \) in a rather weak norm
\[
\| \mathcal{D}_t^\alpha \tilde{u}_n(.t) \|_{H^{-\sigma}(\Omega)} \leq \| (q_n \hat{u}_n + \hat{f}_n(Pu))(.t) \|_{H^{-\sigma}(\Omega)} + \| (q_n \hat{u}_n + \hat{f}_n(Pu)) \|_{L^2(0,T;L^2(\Omega))}
\]
\[ \leq \| C_2H^{-\sigma,L^2} + \sqrt{T} \| (q_n \hat{u}_n + \hat{f}_n(Pu)) \|_{L^\infty(0,T;L^2(\Omega))}
\]
\[ \leq \| C_2H^{-\sigma,L^2} + \sqrt{T} \| (\| q_n \|_{L^\infty(I)} \| \hat{u}_n \|_{L^\infty(0,T;L^2(\Omega))} + \| \hat{f}_n(Pu) \|_{L^\infty(0,T;L^2(\Omega))})
\]
\[ \leq \| C_2H^{-\sigma,L^2} + \sqrt{T} \| (\rho C(\rho, T, |\Omega|) + \sqrt{\| \hat{f}_n \|_{L^\infty(I)}}) \| \hat{f}_n \|_{L^\infty(I)}.
\]  
\[ (50) \]
Thus, under the attainability condition (40)
\[ \| T(f_n) - T(f)(g) \|_{H^{-(1-\theta)}(\Omega)} = \| D^\theta \tilde{u}_{n}(\cdot, T) \|_{H^{-(1-\theta)}(\Omega)} \to 0 \text{ as } k \to \infty, \]
by (47). On the other hand, (47) also implies
\[ \| T(f_n) - f^+ (g) \|_{L^\infty(\Omega)} \leq \| T(f_n) - f^+ \|_{L^\infty(\Omega)} \to 0 \text{ as } k \to \infty. \]
Hence, the two limits need to coincide
\[ T(f)(g) = f^+ (g) \in Z \subseteq C(\Omega) , \]
thus
\[ T(f) = f^+. \]
A subsequence–subsequence argument yields weak* convergence in X of the whole sequence T(f_n) to T(f).

Invoking Schauder’s fixed point theorem in locally convex topological spaces (see [7]), we have proven the following result.

**Theorem 3.1.** Let \( \alpha \in (\frac{4}{5}, 1] \), \( \Omega \subseteq \mathbb{R}^1 \), an open bounded interval \( \sigma \in (\frac{3}{2}, 2] \) and assume that \( g \) is strictly monotone and satisfies (36) with (33), that \( \rho_0 \) as well as \( \| r \|_{L^Q(0,T;L^2(\Omega))} \) are sufficiently small.

Then for large enough \( \rho > 0 \) the operator \( T \) is self-mapping on the bounded, closed and convex set \( B \) as defined in (44) and \( T \) is weakly* continuous in X as defined in (33). Hence \( T \) has a fixed point \( f \in B \). If this fixed point \( f \) is monotonically decreasing and satisfies (31), then \( f \) solves the inverse problem (2)–(4).

**Proof.** It only remains to prove that under conditions (31) and (40), a fixed point \( f \) solves (2)–(4). By remark 3.1 we have \( f(g) = y = f(Pu(\cdot, T; f) + L(u(\cdot, T; f) - g) \); hence the difference \( w = u(\cdot, T; f) - g \) satisfies the elliptic PDE
\[ Lw + \tilde{y}w = 0, \]
with \( \tilde{y} = \int_0^1 f'(g + \theta Pu)P d\theta \leq 0 \) and homogeneous boundary conditions, and therefore has to vanish.

Below we will make a series of short remarks.

The question arises whether a similar result can be achieved for \( \alpha < \frac{4}{5} \). We have not investigated all possibilities to prove this, but also have no evidence that it cannot be done.

The assumption of strict monotonicity of \( g \) is used to guarantee attainability (40); see remark 3.1. Our iteration algorithm can be implemented without it and in practice our scheme worked perfectly well also for non-monotone \( g \).

Estimate (43) shows that the influence of the initial data indeed decreases for larger \( T \). This confirms the computational observation of faster convergence for larger \( T \).

### 3.3. Contractivity for monotone nonlinearities in the parabolic case—uniqueness

To establish a contraction property of \( T \) in the parabolic case, we will make use of the fact that the solution \( u \) to the forward problem (4), (2) and even its time derivative decays exponentially, provided \( f \) is monotonically decreasing. In order to prove this, we restrict ourselves to \( L = \nabla \cdot (A \nabla) \), where \( A \in \mathbb{R}^{n \times n} \) is a symmetric positive definite matrix, i.e. the coefficients of the elliptic differential operator \( L \) are supposed to be constant, and \( r = 0 \).

Assume that \( f' \leq 0 \), then from
\[ u_t - \Delta u = f(Pu), \quad (x, t) \in \Omega \times (0, T) \]
\[ \partial_t u(x, t) + \gamma(x)u(x, t) = 0 \quad (x, t) \in \partial\Omega \times (0, T) \]
\[ u(x, 0) = u_0, \quad x \in \Omega, \]
we obtain that \( v = (u_t)^2 \) satisfies
\[ v_t - \Delta v = 2f'(Pu)(u_t)^2 - 2\nabla u_t^T A \nabla u_t, \quad (x, t) \in \Omega \times (0, T) \]
\[ \partial_t v(x, t) + 2\gamma(x)v(x, t) = 0 \quad (x, t) \in \partial\Omega \times (0, T) \]
\[ v(x, 0) = v_0, \quad x \in \Omega, \]
where \( v_0 = (\|u_0 + f(u_0)\|^2) \). Now if
\[ v_t - \Delta v = 0, \quad (x, t) \in \Omega \times (0, T) \]
\[ \partial_t v(x, t) + 2\gamma(x)v(x, t) = 0 \quad (x, t) \in \partial\Omega \times (0, T) \]
\[ v(x, 0) = v_0, \quad x \in \Omega, \]
then the maximum principle shows that \( \|v \| \geq 0 \) but more crucially, \( \|v \| \leq \|v_0\|. \) For \( v \) we can use the representation formula (12)–(14), which yields
\[
\|u(\cdot, t)\|^{2}_{L^1(\Omega)} \leq \|v(\cdot, t)\|^{2}_{L^1(\Omega)} = \left( \sum_{j=1}^{\infty} e^{-2\lambda_j^2(\|v_0\|, \varphi_j)^2} \right)^{1/2} \leq e^{-\lambda_t\|v_0\|^2_{L^1(\Omega)}},
\]
and
\[
\|u(\cdot, t)\|^2_{L^\infty(\Omega)} \leq \|v(\cdot, t)\|^2_{L^\infty(\Omega)} \leq C^2_{H^\infty, L^\infty} \|v(\cdot, t)\|^2_{H^\infty(\Omega)}
\]
\[
= C^2_{H^\infty, L^\infty} \left( \sum_{j=1}^{\infty} e^{-2\lambda_j^2(\|v_0\|, \varphi_j)^2} \right)^{1/2} \leq C^2_{H^\infty, L^\infty} e^{-\lambda_t\|v_0\|^2_{H^\infty(\Omega)}}
\]
\[
= C^2_{H^\infty, L^\infty} e^{-\lambda_t\|\|u_0 + f(u_0)\|^2\|_{H^\infty(\Omega)}}.
\]
Now we will employ (52) in order to show that provided we have \( f_1, f_2 \in W^{2, \infty}(I) \), then 
\[ \|T(f_1) - T(f_2)\|_{W^{1, \infty}(I)} \leq C e^{-\lambda_I f_1} \|f_1 - f_2\|_{W^{1, \infty}(I)} \] holds.

For \( f_1, f_2 \in X \), when taking the difference \( T(f_1) - T(f_2) = f_1^+ - f_2^+ \), the terms containing \( g \) explicitly and \( r \) cancel, so that with \( u^{(1)}(x, t) = u(x, t; f_1) \) we have \( f_1^+(g) - f_2^+(g) = D_{\alpha}^\rho(z^{(1)} - u^{(2)}) = D_{\alpha}^\rho z \) where \( z \) solves
\[
D_{\alpha}^\rho z - Lz = f_1(Pu^{(1)}) - f_2(Pu^{(2)})
\]
\[ = \int_0^t \int_{\partial\Omega} \left( P(u^{(1)} + \theta z) \right) \, d\Theta z + (f_1 - f_2)(Pu^{(2)}), \quad (x, t) \in \Omega \times (0, T) \]
\[ \partial_t z(x, t) + \gamma(x)z(x, t) = 0 \quad (x, t) \in \partial\Omega \times (0, T) \]
\[ z(x, 0) = 0, \quad x \in \Omega, \]
and we can estimate \( \|\|T(f_1) - T(f_2)\|z\| \) by estimating \( D_{\alpha}^\rho z \).

In the parabolic case \( \alpha = 1 \) this can be done by differentiating (53) with respect to time to obtain that \( w = D_{\alpha}^\rho z \) satisfies
\[ D_{tw} - Lw = f'_1(Pu^{(1)})Pu^{(1)}_t - f'_2(Pu^{(2)})Pu^{(2)}_t \]
\[ = f'_1(Pu^{(1)})Pw + \bar{y} u^{(2)}_t, \quad (x, t) \in \Omega \times (0, T) \]
\[ \partial_x w(x, t) + \gamma(x) z(x, t) = 0 \quad (x, t) \in \partial\Omega \times (0, T) \]
\[ w(x, 0) = f_1(u_0) - f_2(u_0) \quad x \in \Omega, \]

with \( \bar{y} = (f'_1(Pu^{(1)}) - f'_2(Pu^{(2)}))P = (\int_0^t f'_1(Pu^{(1)} + \theta z)) \, d\theta P \, z + (f_1 - f_2)'(Pu^{(2)})P. \)

We assume that there exist potentials \( \tilde{q}, \tilde{q} \in L^\infty(\Omega), \tilde{q}, \tilde{q} \geq 0 \) such that
\[ \| \int_0^t f'_1(Pu^{(1)} + \theta z) \, d\theta + \tilde{q} \|_{L^\infty(\Omega)} \leq \tilde{c}, \]
\[ \| f'_2(Pu^{(1)}) + \tilde{q} \|_{L^\infty(\Omega)} \leq \tilde{c}, \] (54)

for some \( \tilde{c}, \tilde{c}, \tilde{c}, \tilde{c} \), such that \( 4 \tilde{c} < \tilde{\lambda}_1^2, 4 \tilde{c} < \tilde{\lambda}_2^2 \), where \( \tilde{\lambda}_1 \) is the smallest eigenvalue of \( L_{\tilde{q}} = L + \tilde{q} \) (and analogously for \( \tilde{q} \)). We rewrite both initial boundary value problems in terms of \( L_{\tilde{q}}, L_{\tilde{q}} \), respectively, and use corresponding eigensystems \( \tilde{ \lambda}_j, \tilde{ \varphi}_j, \tilde{ \lambda}_j, \tilde{ \varphi}_j \). This yields
\[ w(\cdot, t) = \sum_{j=1}^\infty \left\{ e^{-\tilde{\lambda}_j t}((f_1 - f_2)(u_0), \tilde{\varphi}_j) \right. \\
+ \left. \int_0^t e^{-\tilde{\lambda}_j(t-s)}((f'_1(Pu^{(1)}(s) + \tilde{q})w(s) + \tilde{y}(s) u^{(2)}_t(s), \tilde{\varphi}_j) \, ds \right\} \tilde{\varphi}_j, \] (55)

hence abbreviating \( df_j(s) = |(\tilde{y}(s) u^{(2)}_t(s), \tilde{\varphi}_j)| \)
\[ \|w(\cdot, t)\|^2_{L^2(\Omega)} \leq 2 \sum_{j=1}^\infty \tilde{\lambda}_j \left( \int_0^t e^{-\tilde{\lambda}_j(t-s)}((f'_1(Pu^{(1)}(s) + \tilde{q})w(s), \tilde{\varphi}_j) + df_j(s) \, ds \right)^2 \\
+ 2 \sum_{j=1}^\infty \tilde{\lambda}_j e^{-2\tilde{\lambda}_j t}((f_1 - f_2)(u_0), \tilde{\varphi}_j)^2, \]

where the first term on the right-hand side can be estimated by
\[ \sum_{j=1}^\infty \tilde{\lambda}_j \left( \int_0^t e^{-\tilde{\lambda}_j(t-s)}((f'_1(Pu^{(1)}(s) + \tilde{q})w(s), \tilde{\varphi}_j) + df_j(s) \, ds \right)^2 \\
\leq \sum_{j=1}^\infty \int_0^t e^{-\tilde{\lambda}_j(t-s)} ds \tilde{\lambda}_j \int_0^t e^{-\tilde{\lambda}_j(t-s)} \left( ((f'_1(Pu^{(1)}(s) + \tilde{q})w(s), \tilde{\varphi}_j) + df_j(s) \right)^2 \, ds \\
= \int_0^t \sum_{j=1}^\infty e^{-\tilde{\lambda}_j(t-s)} \tilde{\lambda}_j^{-1} \left( (f'_1(Pu^{(1)}(s) + \tilde{q})w(s), \tilde{\varphi}_j) + df_j(s) \right)^2 \, ds \\
\leq \int_0^t e^{-\tilde{\lambda}_j(t-s)} \sum_{j=1}^\infty \tilde{\lambda}_j^{-1} \left( (f'_1(Pu^{(1)}(s) + \tilde{q})w(s), \tilde{\varphi}_j) + df_j(s) \right)^2 \, ds \\
\leq 2 \int_0^t \int_0^t \left( \int_0^s \right) e^{-\tilde{\lambda}_j(t-s)} \| (f'_1(Pu^{(1)}(s) + \tilde{q})w(s), \tilde{\varphi}_j) + df_j(s) \|^2_{L^2(\Omega)} \, ds \\
\leq 2 \int_0^t e^{-\tilde{\lambda}_j(t-s)} \left( \tilde{c} \|w(s)\|^2_{L^2(\Omega)} + \| \tilde{y}(s) u^{(2)}_t(s) \|^2_{L^2(\Omega)} \right) ds. \]

This shows that we cannot take an \( r \) larger than 1 here, in order to avoid derivatives acting on the term \( \tilde{y}(s) u^{(2)}_t(s) \) (note that the maximum principle argument for exponential decay of
\( u_t \) works for Lebesgue norms but not for higher-order Sobolev norms. As a consequence, an embedding into \( W^{1,\infty}(\Omega) \) will not be possible here and we have to work with a weaker norm on the differences \( f_1 - f_2 \) which, however, should still be strong enough to capture the \( f_1 - f_2 \) difference term appearing in \( \tilde{y} \). We will use the norm \( \|(f_1 - f_2)(u)\|_{H^p(\Omega)} := \|(f_1 - f_2)(g)\|_{H^p(\Omega)} \) and make the assumptions (36) and \( u_0 \in W^{1,\infty}(\Omega) \), which, together with the fact that by theorem 3.1 \( \|\nabla u^{(2)}\|_{L^\infty(0,T,\mathcal{L}^\infty(\Omega))} \leq \mathcal{T}c_0^{(}\|\mathcal{T}\|_{\mathcal{L}^\infty(\Omega)} \), implies the existence of constants \( C_{-1}(g), C_0(g), \tilde{C}_1(g) \) such that

\[
\begin{align*}
\|(f_1 - f_2)(u_0)\|_{H^p(\Omega)} &\leq C_0(g)\|(f_1 - f_2)(g)\|_{H^p(\Omega)} \\
\|(f_1 - f_2)(u^{(2)})\|_{L^\infty(0,T,\mathcal{L}^\infty(\Omega))} &\leq \tilde{C}_{-1}(g)\|(f_1 - f_2)(g)\|_{H^p(\Omega)} \\
\|(f_1 - f_2)(u^{(2)})\|_{L^\infty(0,T,\mathcal{L}^\infty(\Omega))} &\leq \tilde{C}_1(g)\|(f_1 - f_2)(g)\|_{H^p(\Omega)}. \tag{56}
\end{align*}
\]

Altogether we have

\[
\|w(\cdot,t)\|^2_{H^p(\Omega)} \leq 2e^{-2\lambda_1 t}\|(f_1 - f_2)(u_0)\|_{H^p(\Omega)} \\
+ 4 \int_0^t e^{-\lambda_1 (t-s)} \left( \hat{c}_1\|w(s)\|^2_{H^{-1}(\Omega)} + \|\tilde{y}(s) u_t^{(2)}(s)\|^2_{H^{-1}(\Omega)} \right) ds.
\]

For \( \eta(s) = \|w(s)\|^2_{H^p(\Omega)} + \frac{4}{\lambda_1^2}\|\tilde{y}(s) u_t^{(2)}(s)\|^2_{H^{-1}(\Omega)} \) with \( \hat{c}_3 = \frac{4\lambda_1}{\lambda_1^2} \), the above inequality becomes

\[
\eta(t) \leq \frac{4}{\hat{c}_3} \left[ \int_0^t e^{-\lambda_1 (t-s)} \left( \frac{4}{\lambda_1^2}\|\tilde{y}(s) u_t^{(2)}(s)\|^2_{H^{-1}(\Omega)} + \|\tilde{u}_t^{(2)}(s)\|^2_{H^{-1}(\Omega)} \right) ds \right] + \hat{c}_3 \int_0^t e^{-\lambda_1 (t-s)} \eta(s) ds. \tag{57}
\]

The convolution version of Gronwall’s inequality lemma 2.5 yields

\[
\eta(t) \leq \frac{4}{\hat{c}_3} \left[ \int_0^t e^{-\lambda_1 (t-s)} \left( \frac{4}{\lambda_1^2}\|\tilde{y}(s) u_t^{(2)}(s)\|^2_{H^{-1}(\Omega)} + \|\tilde{u}_t^{(2)}(s)\|^2_{H^{-1}(\Omega)} \right) ds \right] + \hat{c}_3 \int_0^t e^{-\lambda_1 (t-s)} \eta(s) ds.
\]

hence

\[
\|w(t)\|^2_{H^p(\Omega)} \leq e^{-2\lambda_1 t}\|(f_1 - f_2)(u_0)\|_{H^p(\Omega)} \\
+ 4 \int_0^t e^{-\lambda_1 (t-s)} \|\tilde{y}(s) u_t^{(2)}(s)\|^2_{H^{-1}(\Omega)} ds \\
+ \hat{c}_3 \int_0^t e^{-\lambda_1 (t-s)} e^{-2\lambda_1 s} ds \|(f_1 - f_2)(u_0)\|_{H^p(\Omega)}. \tag{58}
\]

Analogously to (55), we obtain for (53) the identity

\[
z(\cdot,t) = \sum_{j=1}^{\infty} \int_0^t e^{-\lambda_1 (t-s)} (\int_0^s f_0^j(P(u^{(1)}(s) + \theta z(s))) d\theta z(s) + (f_1 - f_2)(u^{(2)}(s), \varphi_j) ds) \varphi_j.
\]

Hence, using the replacements in (58)

\[
w \to z, \quad (f_1 - f_2)(u_0) \to 0,
\]

\[
f_0^j(Pu^{(1)}) + \tilde{q} \to \int_0^1 f_0^j(P(u^{(1)}(s) + \theta z(s))) d\theta + \tilde{q}
\]

\[
\tilde{y} u^{(2)}_t \to (f_1 - f_2)(u^{(2)}), \quad r \to \tilde{r},
\]

\[21\]
we obtain
\[ \|z(\cdot,t)\|^2_{H^1(\Omega)} \leq 4 \int_0^t e^{-(\lambda/2) (t-s)} \| (f_1 - f_2)(u^{(2)}(s)) \|^2_{H^{-1}(\Omega)} ds, \]
which for sufficiently large \( t \) would even yield
\[ \|z(\cdot,t)\|_{L^\infty(\Omega)} \leq C_{\text{IP},L^\infty} \left( \int_0^t e^{-(\lambda/2) (t-s)} \| (f_1 - f_2)(u^{(2)}(s)) \|^2_{H^{-1}(\Omega)} ds \right)^{1/2}, \]
but will actually only need \( \hat{r} = 0 \) below. This allows us to estimate the term \( \tilde{y}_u^{(2)} \) as follows
\[
\|\tilde{y}(t)u_t^{(2)}(t)\|_{L^2(\Omega)} \leq \|\tilde{y}(t)\|_{L^2(\Omega)}\|u_t^{(2)}(t)\|_{L^\infty(\Omega)} \\
\leq \left( \|u_t^{(2)}\|_{L^\infty(\Omega)} \|\tilde{y}(t)\|_{L^2(\Omega)} + \|f_1 - f_2\|_{L^1(0, T; H^{-1}(\Omega))} \right) \|u_t^{(2)}(t)\|_{L^\infty(\Omega)} \\
\leq \left( \|u_t^{(2)}\|_{L^\infty(\Omega)} + \|f_1 - f_2\|_{L^1(0, T; H^{-1}(\Omega))} \right) \|u_t^{(2)}(t)\|_{L^\infty(\Omega)}.
\]
Inserting this together with (52) into (58), setting \( r = 1 \), and using our assumption (56), we obtain
\[
\|\mathcal{T}(f_1) - \mathcal{T}(f_2)\|^2 = \|\mathcal{T}(f_1) - \mathcal{T}(f_2)(g)\|^2_{H^1(\Omega)} = \|w(T)\|^2_{H^1(\Omega)} \\
\leq e^{-\lambda T} \|f_1 - f_2\|_{L^2(\Omega)}^2 + 4 \left( \|u_t^{(2)}\|_{L^\infty(\Omega)} \|f_1 - f_2\|_{L^1(0, T; H^{-1}(\Omega))} \right)^2 \\
\times \left( C_{\text{IP},L^\infty}^2 \|u_t^{(2)}\|_{L^2(\Omega)} \int_0^T e^{-(\lambda/2) (T-s)} ds \right) \\
+ \hat{c}_3 \int_0^T e^{-(\lambda/2) (T-s)} ds \|f_1 - f_2\|_{L^2(\Omega)}^2 \\
\leq Ce^{-\lambda T} \|f_1 - f_2\|_{H^1(\Omega)}^2 = Ce^{-\lambda T} \|f_1 - f_2\|_{H^1(\Omega)}^2,
\]
for \( C = C_1C_2 + \left( 1 + e^{-(\lambda/2) T} \right) \tilde{C}_0(g)^2 \) where \( C_1 = 4 \left( \|u_t^{(2)}\|_{L^\infty(\Omega)} \tilde{C}_1(g) + \tilde{C}_1(g) \right)^2 \), and
\( C_2 = \left( C_{\text{IP},L^\infty}^2 \right)^2 \|\mathcal{L}u_0 + f(u_0)\|_{L^2(\Omega)}^2 \) \( H^1(\Omega) \), and
\( \hat{C}_1 = \frac{1}{\lambda_1+\tilde{c}_1} \).
Altogether we have proven:

**Theorem 3.2.** Let the assumptions of theorem 3.1 and additionally \( \alpha = 1, \ r = 0, \ f_1, f_2 \in W^{2,\infty}_2(I) := \{ f \in W^{2,\infty}(I) : f' \leq 0 \ and \ (\mathcal{L}u_0 + f(u_0))^2 \in H^0(\Omega) \} \) hold.

Then there exists a constant \( C > 0 \) depending only on \( \|u_t^{(2)}\|_{L^\infty(\Omega)} \), \( \|\mathcal{L}u_0 + f(u_0)\|_{H^1(\Omega)} \), and the constants \( \tilde{C}_0(g), \ C_0(g), \ C_1(g) \) in (56), such that
\[
\|\mathcal{T}(f_1) - \mathcal{T}(f_2)(g)\|_{H^1(\Omega)} \leq Ce^{-\lambda T} \|f_1 - f_2\|_{H^1(\Omega)}^2,
\]
with \( \hat{c}_3 \) as in (54).

As a consequence of theorem 3.2 we have the following conditional uniqueness result

**Corollary 3.1.** Let the assumptions of theorem 3.2 be satisfied, then

there is a \( T^* > 0 \) such that if \( T > T^* \), then there exists at most one fixed point of \( \mathcal{T} \) in \( W^{2,\infty}_2(I) \). Thus a solution of the inverse problem (4), (2), (3) is unique in \( W^{2,\infty}_2(I) \).
3.4. Some thoughts and comments

It is instructive to compare the inverse problem for (1) with the initial boundary conditions (2) studied in this paper with earlier work in [23–25] for similar equations and initial boundary conditions. Both approaches used an iterative method obtained by projecting the equation onto a boundary where overposed data are available. In the current case this represents a spatial measurement of $u(t)$ at a final time $t = T$, $u(x, T)$ for $x \in \Omega$, and in the former a time measurement $u(x_0, t)$ for $x_0 \in \partial \Omega$. In each case the method relies on a mapping of $f$ to the differential operator projected onto the overposed boundary. In the current case this is $f \rightarrow u(t)$ and in the previous it is (in one space dimension) $f \rightarrow u_x(x_0, t)$. The analysis of the iteration relies on quite different tools for the representation of solutions. Here we work with series representations and estimates can to some extent be carried over from the parabolic case to the subdiffusion setting. In [23–25] the proofs (in a parabolic setting) rely on a representation by means of a Green’s function and do not directly extend to subdiffusion, since the Green’s function there—defined by the Mainardi-Wright function—exhibits a singularity at the origin.

3.5. Data smoothing and noise propagation

Typically, we will be given noisy data $g^\delta$ and sometimes also some information on the statistics of the noise or, in the deterministic setting considered here, on the noise level with respect to the $L^2$ norm

$$\|g^\delta - g\|_{L^2(\Omega)} \leq \delta. \quad (60)$$

Application of the iterative scheme requires sufficiently smooth data, so instead of $g^\delta$ we employ a smoothed version $g^\tilde{\delta}$ satisfying the conditions on $g$ of theorems 3.1, 3.2 as well as

$$\|g^\tilde{\delta} - g\|_{\mathcal{H}^m(\Omega)} \leq \tilde{\delta}, \quad (61)$$

for some $m > 0$, where $\delta \to 0$ as $\tilde{\delta} \to 0$; see, e.g. [18, section 6.1].

In the situation of theorem 3.1 with $m = \sigma$, we can make use of uniform boundedness by $\rho$ of the fixed point $f^\delta$ of the operator $\mathcal{T}_{f^\delta}$ corresponding to data $g^\delta$ as $\delta \to 0$, as well as weak* continuity of $\mathcal{T}$ to prove subsequential weak* convergence in $X$ of $f^\tilde{\delta}$ to $f_{\text{act}}$ as $\tilde{\delta} \to 0$, where $f^\tilde{\delta}$ is the fixed point corresponding to data $g^\tilde{\delta}$. Namely, uniform boundedness of $\|f^\delta\|_{X} \leq \tilde{\delta}$ implies existence of weakly* $\tilde{f}^\delta \in W^{1,\infty}(I)$ and strongly in $L^\infty(I)$ convergent subsequences $\tilde{f}_k := f_k^\tilde{\delta}, \mathcal{T}_{f_k^\tilde{\delta}}(f_k^\tilde{\delta})$ with limits $\tilde{f}$ and $\bar{f}$ respectively. It remains to prove that $\tilde{f} = \mathcal{T}_{\tilde{f}}(\tilde{f})$, which means that $\tilde{f}$ is a solution to the inverse problem with exact data $g$, provided it satisfies the range condition (31). To this end, we abbreviate $\hat{f}_k := f_k^\tilde{\delta}$ and $\tilde{g}_k := g_k^\tilde{\delta}$ and consider the following decomposition

\[
\mathcal{T}_{\tilde{f}}(\tilde{f}) - \tilde{f} = \mathcal{T}_{\tilde{f}}(\tilde{f}) - \mathcal{T}_{\tilde{f}}(\tilde{f}_k) + \mathcal{T}_{\tilde{f}}(\tilde{f}_k) - \mathcal{T}_{\tilde{g}}(\tilde{f}_k) + \tilde{f}_k - \tilde{f} = \mathcal{T}_{\tilde{f}}(\tilde{f}) - \mathcal{T}_{\tilde{g}}(\tilde{f}_k) + \mathcal{P}_\tilde{g}(\tilde{g}_k) - \mathcal{P}_\tilde{g}(\tilde{g}_k) + \mathcal{P}_\tilde{g}(\tilde{g}_k) - \mathcal{P}_\tilde{g}(\tilde{g}_k) + \tilde{f}_k - \tilde{f}, \quad (62)
\]

where $\tilde{y}_k = D^r_f u(\cdot, \tilde{f}_k) - Lg - r(\cdot, T) \in D^r_f u(\cdot, \tilde{f}_k) - Lg - r(\cdot, T)$, $\tilde{y}_k = D^r_f u(\cdot, \tilde{f}_k) - Lg - r(\cdot, T) \in \argmin\{\|f\|_{Y} - y\}_{Y} : f \in X$ and $\|f\|_{Y} - r(\cdot, T) \|g\|_{\mathcal{H}^m(\Omega)} \leq \rho_0$, and we have used the fact that $\mathcal{T}_{\tilde{f}}(\tilde{f}_k) = \mathcal{P}_{\tilde{g}}(\tilde{g}_k) = \tilde{f}_k$. The first and last difference $\mathcal{T}_{\tilde{f}}(\tilde{f}) - \mathcal{T}_{\tilde{g}}(\tilde{f}_k)$ and $\tilde{f}_k - \tilde{f}$ tend to zero weakly* in $X$, due to the weak* continuity of $\mathcal{T}_{\tilde{f}}$. For the second and third differences, we abbreviate the projections defined by (32) by $\tilde{f}_k^+ = \mathcal{P}_\tilde{g}(\tilde{g}_k), \tilde{f}_k^+ = \mathcal{P}_\tilde{g}(\tilde{g}_k)$, i.e. $f^+_k(g) = \tilde{y}_k$ and $\tilde{f}_k^+(\tilde{g}_k) = \tilde{y}_k = \tilde{f}_k^+$. Hence, the second difference in (62) satisfies
\[ \| P_x(y_k) - P_x(y_k) \| = \| f_{k+}^x - f_{k+}^x + x \| = C(\| f_{k+}^x(g) - \tilde{f}_{k+}^x(g) \|) + C(\| g - \tilde{g}_k \|) \leq C(\| L(g - \tilde{g}_k) \|) \leq C(\| \tilde{g}_k \|) \]

and the third difference can be estimated by
\[ \| P_x(y_k) - P_x(y_k) \|_{L^\infty(I)} = \| f_{k+}^x - \tilde{f}_{k+}^x \|_{L^\infty(I)} \leq \| f_{k+}^x(g) - \tilde{f}_{k+}^x(g) \|_{L^\infty(I)} \]

So altogether, the right-hand side of (62) tends to zero as \( k \to \infty \). A subsequence–subsequence argument therefore yields stability.

**Proposition 3.1.** Let \( (\tilde{g}^\delta)_{\delta \in (0, \delta)} \subseteq L^2(\Omega) \) be a family of data satisfying (60) with corresponding smoothed versions \( \tilde{g}^\delta \) satisfying (36) and (61) with \( m = \sigma \) such that \( \delta \to 0 \) as \( \delta \to 0 \) and \( \sigma > \frac{1}{2} \).

Then the sequence of fixed points \( \tilde{f}^{\tilde{g}^\delta} \) of the perturbed operators \( T_{\tilde{g}^\delta} \) converges weakly to the set of fixed points of the unperturbed operator \( T = T_0 \) as \( \delta \to 0^+ \).

Under the conditions of theorem 3.2 with \( T \) sufficiently large so that \( C e^{-\lambda T/2} = c < 1 \) and \( m = 3 \), we can even provide a quantitative stability estimate as follows.

\[
\tilde{\delta} \geq \| \tilde{g}^\delta - Lg \|_{H^1(\Omega)} \geq \| Df_{\text{act}}(\cdot, T; \tilde{f}^{\tilde{g}^\delta}) - Df_{\text{act}}(\cdot, T; \tilde{f}^{\tilde{g}^\delta}) \|_{H^1(\Omega)} + C(\| \tilde{g}^\delta - \tilde{g} \|_{H^1(\Omega)})
\]

Hence
\[
\| \tilde{f}^{\tilde{g}^\delta}(\tilde{g}^\delta) - f_{\text{act}}(\tilde{g}^\delta) \|_{H^1(\Omega)} \leq \frac{1}{1 - c} \left( \tilde{\delta} + (\| f_{\text{act}}(\tilde{g}) \|_{L^\infty(I)} + \| f_{\text{act}}''(\tilde{g}) \|_{L^1(I)} C_{H^1(\Omega)}^{\tilde{g}^\delta} \right)\delta^{\tilde{\beta}}.
\]

**4. Newton-type methods**

4.1. Operator formulation of nonlinearity identification in reaction–diffusion equation

Denote by \( u_{\text{act}}(x, t) = u(x, t; f_{\text{act}}) \) the exact solution of the inverse problem (2)–(4).

Again, we assume the range condition (31) to hold for \( f = f_{\text{act}} \) and use the projection \( P \) in the reformulation (35) of the direct problem. Moreover, as in the previous section, we use \( Y = L^2(\Omega) \) as data space and a continuously embedded subspace \( X \) of \( W^{1,\infty}(I) \) as preimage.
space, e.g. just \( X = W^{1,\infty}(I) \), or, in order to work with a Hilbert space setting, \( X = H^\sigma(I) \) with \( \sigma > \frac{1}{2} \) for \( F : X \to Y, f \mapsto u(\cdot, T) \) where \( u \) solves
\[
D_t^\phi u - \Delta u = f(Pu) + r, \quad (x, t) \in \Omega \times (0, T)
\]
\[
\partial_n u(x, t) + \gamma(x)u(x, t) = 0, \quad (x, t) \in \partial\Omega \times (0, T)
\]
\[
u(x, 0) = u_0, \quad x \in \Omega.
\] (63)

The operator \( F \) is well defined due to theorem 2.3 (replacing \( f \) there by \( f \circ P \) here).

This allows one to write the inverse problem as
\[ F(f) = g. \] (64)

We denote the exact solution by \( f_{\text{act}} \) and some initial guess by \( f_0 \).

The derivative of the forward operator is given by \( F'(f)h = v(\cdot, T) \) where \( v \) solves
\[
D_t^\phi v - \Delta v - f'(Pu)v = h(Pu), \quad (x, t) \in \Omega \times (0, T)
\]
\[
\partial_n v(x, t) + \gamma(x)v(x, t) = 0, \quad (x, t) \in \partial\Omega \times (0, T)
\]
\[
v(x, 0) = 0, \quad x \in \Omega.
\] (65)

The well-definedness of \( v \) can be concluded from results in [26], which has to be applied together with some fixed point argument, since the potential \( f'(Pu) \) depends on \( x \) and \( t \). Its adjoint (as appearing in the explicit version of Tikhonov-regularized Newton or in source conditions) can be derived by means of the standard integration by the parts procedure using the co-area formula. However, this leads to quite opaque formulas and hence we avoid it by stating Newton’s method and source conditions in a variational manner below.

Further regularity of \( F \), such as Lipschitz continuity at \( f_{\text{act}} \), can be verified by using the fact that \( F'(f_{\text{act}})^{-1} F'(f) = w(\cdot, T) \) where \( w \) solves
\[
D_t^\phi w - \Delta w - f'(Pu)w = \text{rhs}, \quad (x, t) \in \Omega \times (0, T)
\]
\[
\partial_n w(x, t) + \gamma(x)w(x, t) = 0, \quad (x, t) \in \partial\Omega \times (0, T)
\]
\[
w(x, 0) = 0, \quad x \in \Omega,
\] (66)

where
\[
\text{rhs} = \left(f'(Pu) - f'(Pu)ight)v + f_{\text{act}} - fPu - h(Pu) + h(Pu) - h(Pu).
\] (67)

\( u, v \) solve (63), (65), and \( u_{\text{act}} \) solves (63), (65) with \( f \) replaced by \( f_{\text{act}} \), respectively; see Subsection 4.4 below.

### 4.2. Formulation of Newton-type methods

For simplicity we again just use a deterministic noise bound
\[ \|g^\delta - g\|_Y \leq \delta, \] (68)

while the noise itself might of course be random.

Consider two versions of regularized Newton iterations; see, e.g. [1, 11, 14–16, 28, 30] and the references therein.

**Tikhonov-regularized Newton:**

\[ f_{k+1} \text{ minimizer of } \|F'(f_k)(f - f_k) + F(f_k) - g^\delta\|_Y^2 + \gamma_k\|f - f_0\|_X^2, \] (69)

with *a priori* choice of the regularization parameter \( \gamma_k = \gamma_0 \kappa^k \) for some \( \kappa \in (0, 1) \).
Ivanov-regularized Newton:

\[ f_{k+1} \text{ minimizer of } \|F'(f_k)(f - f_k) + F(f_k) - g^\delta \|_Y \text{ under the constraint } \|f - f_0\|_X \leq \varrho, \]

with \( \varrho \) chosen so that the exact solution is admissible, i.e. \( \|f_{\text{act}} - f_0\| \leq \varrho \).

In order to guarantee the existence of the minimizers of (69) or (70) respectively, it is crucial to use a space \( X \) that guarantees weak(*) compactness of sublevel sets as well as weak(∗) lower semicontinuity of the cost functional (and constraint). Therefore a choice \( X = H^\sigma(I) \) or \( X = W^{1,\infty}(I) \) is admissible, while \( X = C^4(I) \) would not satisfy these requirements.

Both versions of Newton’s method are stopped according to the discrepancy principle:

\[ k_*(\delta) = \min\{k \mid \|F(f_k) - g^\delta\| > \tau \delta\}, \]

for some \( \tau > 1 \) (independent of \( \delta \)).

The advantage of (70) as compared to (69) is that it always guarantees a bounded sequence of iterates, no matter if further conditions guaranteeing convergence are satisfied; see section 4.3 below. Its drawback is that an upper bound of the distance between \( f_0 \) and \( f_{\text{act}} \) of iterates, no matter if further conditions guaranteeing convergence are satisfied; see section 4.3.

4.3. Convergence of Newton-type methods under a Lipschitz condition at \( f_{\text{act}} \) and a variational source condition

Structural constraints on the forward operator, such as the tangential cone condition, are often not satisfied for parameter identification problems in PDEs unless complete observations of the state (i.e. in our case observations of \( u \) on all of \( \Omega \times (0, T) \)) are available. So they are likely not satisfied for the inverse problem under consideration here.

In contrast to this, a Lipschitz condition on \( F \)

\[ \|F'(f) - F'(f_{\text{act}})\|_{L(X,Y)} \leq L\|f - f_{\text{act}}\|_X \text{ for } f \in B_\varrho(f_0), \]

can usually be verified in a much more standard way. Note that one of the two points where \( F' \) is evaluated is fixed to \( f_{\text{act}} \), which enables one to establish (72) without the need to impose a stronger space \( X \) (which would make the problem more ill-posed), but just by assuming additional regularity on \( f_{\text{act}} \); see section 4.4.

A convergence proof of Newton-type methods without a structural condition (so with just the Lipschitz condition on \( F' \)) requires \( f_{\text{act}} \) to satisfy a source condition. The explicit version of the source condition would involve an adjoint, which we avoid by considering the variational version (see, e.g. [10])

\[ -(h, f_{\text{act}} - f_0)_X \leq \beta \|F'(f_{\text{act}})h\|_Y \text{ for all } h \text{ such that } h + f_{\text{act}} \in B_\varrho(f_0), \]

for some \( \beta < \frac{1}{L} \). Note that (73) implies that \( f_{\text{act}} - f_0 \) lies in the orthogonal complement of the kernel of \( F'(f_{\text{act}}) \) so its verification is related to the uniqueness question for the linearized problem. We here restrict ourselves to the case of a Hilbert space \( X \) with inner product \((\cdot, \cdot)_X\) for simplicity. The considerations below can be extended to a Banach space setting by using Bregman distances (see, e.g. [10]).

Convergence of (69) under conditions (72) and (73) so that the stopping index defined by the discrepancy principle (71) is reached after a finite number of steps follows existing results (e.g. [15, lemma 4.1]; see also [1] for the Hilbert space setting and [28, 30] in Banach spaces).
In case of (70), we can argue as follows, assuming that $\varrho$ has been chosen such that
\[
\|f_{\text{act}} - f_0\|^2 \leq \varrho \leq \|f_{\text{act}} - f_0\|^2 + C\delta
\]
and omitting the norm subscripts to simplify the notation.

Minimality of $f_{k+1}$ and admissibility of $f_{\text{act}}$ implies
\[
\|F'\left(f_k(x_{k+1} - f_k) + F(f_k) - g^0\right) \leq \|F'\left(f_k(x_{k+1} - f_k) + F(f_k) - f_{\text{act}}\right) - \delta
\]
where the left- and right-hand sides can further be estimated by means of the Lipschitz condition on $F'$
\[
F_p(k) := \|F'\left(f_k(x_{k+1} - f_k) + F(f_k) - g^0\right) \geq \|F'\left(f_k(x_{k+1} - f_k) + F(f_k) - f_{\text{act}}\right) - \delta
\]
\[
\geq \|F(f_{k+1}) - F(f_{\text{act}})\| - \|\left(F'\left(f_k\right) - F'\left(f_{\text{act}}\right)\right)
\]
\[
+ \int_0^1 \left(F'\left(f_{\text{act}}\right) - F'(f_k + \vartheta(f_{k+1} - f_k))\right) d\vartheta (f_{k+1} - f_k)\| - \delta
\]
\[
\geq \|F(f_{k+1}) - F(f_{\text{act}})\| - L\left(f_k - f_{\text{act}}\right)
\]
\[
+ \int_0^1 \|\left(1 - \vartheta\right)(f_{\text{act}} - f_k) + \vartheta(f_{\text{act}} - f_{k+1})\| \|f_k - f_{\text{act}}\| - \delta
\]
\[
\geq \|F(f_{k+1}) - F(f_{\text{act}})\| - \frac{5L}{2}\|f_k - f_{\text{act}}\|^2 - \frac{3L}{2}\|f_{k+1} - f_{\text{act}}\|^2 - \delta,
\]
and
\[
F_{p\mathcal{C}}(k) \doteq \|F'\left(f_k(x_{k+1} - f_k) + F(f_k) - g^0\right) \leq \|F'\left(f_k(x_{k+1} - f_k) + F(f_k) - f_{\text{act}}\right) + \delta
\]
\[
\leq \|\left(F'(f_k) - F'(f_{\text{act}})\right) + \int_0^1 \left(F'\left(f_{\text{act}}\right) - F'(f_k + \vartheta(f_k - f_{\text{act}}))\right) d\vartheta (f_{k+1} - f_k)\| + \delta
\]
\[
\leq L\left(f_k - f_{\text{act}}\right) + \int_0^1 \|f_k - f_{\text{act}}\| d\vartheta \|f_{\text{act}} - f_k\| + \delta \leq \frac{3L}{2}\|f_k - f_{\text{act}}\|^2 + \delta,
\]
so inserting $F_p(k)$ and $F_{p\mathcal{C}}(k)$ into (74) we get
\[
\|F(f_{k+1}) - F(f_{\text{act}})\| \leq 2\delta + 4L\|f_k - f_{\text{act}}\|^2 + \frac{3L}{2}\|f_{k+1} - f_{\text{act}}\|^2.
\]
Here the quadratic terms can be estimated by means of the variational source condition and the fact that $\|f_k - f_0\|^2 \leq \varrho \leq \|f_0 - f_{\text{act}}\|^2 + C\delta$
\[
\frac{1}{2}\|f_k - f_{\text{act}}\|^2 = \frac{1}{2}\|f_k - f_0\|^2 - \frac{1}{2}\|f_0 - f_{\text{act}}\|^2 - (f_k - f_{\text{act}}) f_{\text{act}} - f_0
\]
\[
\leq \frac{C}{2} \delta + \beta \|f_{\text{act}}(f_k - f_{\text{act}})\|
\]
\[
\leq \frac{C}{2} \delta + \beta \left(\|f_k - f_{\text{act}}\| + \frac{L}{2}\|f_k - f_{\text{act}}\|^2\right),
\]
where we have again used the Lipschitz condition in the last estimate, hence
\[
\frac{1 - \beta L}{2} \|f_k - f_{\text{act}}\|^2 \leq \frac{C}{2} \delta + \beta \|F(f_k) - F(f_{\text{act}})\|.
\]
Using this for $\ell = k$ and $\ell = k + 1$ altogether yields
\[
\|F(f_{k+1}) - F(f_{\text{act}})\| \leq \frac{2 - 2\beta L + 11CL}{1 - \beta \delta} \delta + \frac{8\beta L}{1 - \beta \delta} \|F(f_k) - F(f_{\text{act}})\| + \frac{8\beta L}{1 - \beta \delta} \|F(f_{k+1}) - F(f_{\text{act}})\|,
\]
i.e. an estimate of the form
\[
\|F(f_{k+1}) - F(f_{\text{act}})\| \leq m\|F(f_k) - F(f_{\text{act}})\| + \tilde{C}\delta,
\]
with $m = \frac{8\beta L}{1 - 4\beta \delta}$ smaller than one if $\beta$ is sufficiently small, so
\[
\|F(f_k) - g^\delta\| \leq \|F(f_k) - F(f_{\text{act}})\| + \delta \leq m^k\|F(f_0) - F(f_{\text{act}})\| + \left(\frac{C}{1 - m} + 1\right) \delta \leq \tau \delta,
\]
for
\[
k \geq \frac{1}{\log \left(\frac{1}{m}\right)} \left(\log \left(\frac{1}{\delta}\right) + \log(\|F(f_0) - F(f_{\text{act}})\|) - \log \left(\tau - 1 - \frac{C}{1 - m}\right)\right),
\]
which proves that $k_\ast(\delta)$ is finite provided $\tau$ has been chosen sufficiently large, $\tau > 1 + \frac{C}{1 - m}$.

Moreover, (71) together with (75) for $\ell = k_\ast(\delta)$ yields the convergence rate
\[
\|f_{k_\ast(\delta)} - f_{\text{act}}\|^2 \leq \frac{1}{1 - \beta L} \left(C\delta + 2\beta\|F(f_{k_\ast(\delta)}) - F(f_{\text{act}})\|\right) \leq \frac{C + 2\beta(\tau + 1)}{1 - \beta L} \delta.
\]

**Proposition 4.1.** Let $Y$ be a Banach space, $X$ a Hilbert space, and $(g^\delta)_{\delta \in [0,\bar{\delta}]} \subseteq Y$ a family of data satisfying (68). Let, for each $\delta \in (0,\bar{\delta})$, the regularized approximation $f_{k_\ast(\delta)}$ be defined by an early stopped Newton-type iteration according to either (69), (71) or (70), (71), with $\tau$ sufficiently large and for some differentiable forward operator $F : X \rightarrow Y$ satisfying (72), and some initial guess satisfying (73).

Then for each $\delta \in (0,\bar{\delta})$, the stopping index is finite and the regularized approximations converge to the exact solution $f_{\text{act}}$ of $F(f) = g$ as appearing in (73) at the rate (76), i.e.
\[
\|f_{k_\ast(\delta)} - f_{\text{act}}\| = O(\sqrt{\delta}), \|F(f_{k_\ast(\delta)}) - g\| = O(\delta).
\]

**4.4. Back to the reaction–diffusion equation**

We now verify the assumptions of proposition 4.1 for the forward operator $F$ defined by (63).

**4.4.1. Lipschitz continuity of $F'$**. Since the right-hand side of (66) (see (67)), can be rewritten as
\[
\text{rhs} = \int_0^1 f''(f_{\text{act}})P(u_{\text{act}} + \vartheta(u - u_{\text{act}}))\,d\vartheta P(u_{\text{act}} - u)\,P\nu
\]
\[
+ \int_0^1 h'(f_{\text{act}})P(u_{\text{act}} + \vartheta(u - u_{\text{act}}))\,d\vartheta P(u_{\text{act}} - u) + \langle f_{\text{act}} - f(Pu)\rangle P\nu,
\]
we can remain with the assumption $X \subseteq W^{1,\infty}(I)$ on the preimage space, while assuming twice differentiability only on $f_{\text{act}}$. Assuming that $f'(P_{\text{act}})P$ and $f'(Pu)P$
are not too far away from an only space-dependent function, so that there exists
\(-q(x) \in \mathcal{D}(F) := \left\{ q \in L^2(\Omega) : \exists \epsilon > 0 : \|\tilde{q} - q\|_{L^2(\Omega)} < \epsilon \right\} = \bigcup_{q > 0} B^L_\mu(q)\) (see [17]) with
\[
\|f'_\text{act}(P(u_{\text{act}} + \vartheta(u - u_{\text{act}})))P + q\|_{L^\infty(0,T;\Omega)} \leq c_1, \quad \|f'(P_u)P + q\|_{L^\infty(0,T;\Omega)} \leq c_1,
\]
for all \(\vartheta \in [0,1]\) and some \(c_1\) sufficiently small, we can rewrite (66), (65), and the equation for the difference \(\tilde{u} := u_{\text{act}} - u\) as
\[
D^s_w w - Lw + q(x)w = (f'_\text{act}(P(u_{\text{act}}))P + q)w + \text{rhs}, \quad (x,t) \in \Omega \times (0,T)
\]
\[
D^s_v v - Lv + q(x)v = (f'(P_u)P + q)v + h(P_u), \quad (x,t) \in \Omega \times (0,T),
\]
and
\[
D^s_\tilde{u} \tilde{u} - L\tilde{u} + q(x)\tilde{u} = (f(T) + q(P_{\text{act}} - f)(P_u)), \quad (x,t) \in \Omega \times (0,T),
\]
where these functions are equipped with homogeneous initial and impedance boundary conditions, where \(f(T) = \int_0^1 f'_\text{act}(P(u_{\text{act}} + \vartheta(u - u_{\text{act}})))P d\vartheta\). Note that (78) holds for all \(f \in B_q(f_0)\) provided
\[
\|f'_\text{act}(P_{\text{act}})P + q\|_{L^\infty(0,T;\Omega)} \leq c_0, \quad \|f'(P_{\text{act}})P + q\|_{L^\infty(0,T;\Omega)} \leq c_0,
\]
and \(\vartheta\) is sufficiently small. Using eigenvalues and eigenfunctions of \(-L_q = -L + q(x)\) we apply the representation formula (8). Together with (16) with \(p = \infty, q = r = 2\) and (41) for \(\theta \in (0,2 - 1/\alpha)\), this yields, with \(p^* = \frac{p}{p-1}, p^{**} = \frac{2p}{p-2}\),
\[
\|w\|_{L^2(0,T;L^2(\Omega))} \leq \tilde{C}_p \left( \|f'_\text{act}(P_{\text{act}})P + q\|_{L^2(0,T;L^2(\Omega))} + \|\text{rhs}\|_{L^2(0,T;L^2(\Omega))} \right)
\leq \tilde{C}_p \tilde{C}_p \|P_{\text{act}}\|_{L^\infty(0,T;L^\infty(\Omega))} \left( \|P_{\text{act}} P u\|_{L^2(0,T;L^2(\Omega))} \right)
\leq \tilde{C}_p \tilde{C}_p \|P_{\text{act}}\|_{L^\infty(0,T;L^\infty(\Omega))} \left( \|\tilde{f} + q\|_{L^\infty(0,T;L^2(\Omega))} \right).
\]
For \(\|f'_\text{act}(P_{\text{act}})P + q\|_{L^2(0,T;L^\infty(\Omega))}\) sufficiently small, this gives an estimate on the term \(\|w\|_{L^2(0,T;L^2(\Omega))}\). Similarly, we get estimates on \(\|w\|_{L^2(0,T;L^2(\Omega))}\) appearing in the right-hand side, using (16) with \(p = r = 2, q = 1\) and \(\int_0^1 |r^{1-\alpha} E_{\alpha,\alpha}(-\lambda r^\alpha)| d\lambda \leq \frac{1}{\alpha}\),
\[
\|P_{\text{act}}\|_{L^2(0,T;L^2(\Omega))} \leq \|P_{\text{act}}\|_{L^2(0,T;L^2(\Omega))} \leq \tilde{C}_p \|P_{\text{act}}\|_{L^\infty(0,T;L^\infty(\Omega))} \left( \|\tilde{f} + q\|_{L^\infty(0,T;L^2(\Omega))} \right)
\leq \tilde{C}_p \|P_{\text{act}}\|_{L^\infty(0,T;L^\infty(\Omega))} \left( \|\tilde{f} + q\|_{L^\infty(0,T;L^2(\Omega))} \right).
\]
and likewise
\[
\|P_{\text{act}}\|_{L^2(0,T;L^2(\Omega))} \leq \tilde{C}_p \|P_{\text{act}}\|_{L^\infty(0,T;L^\infty(\Omega))} \left( \|\tilde{f} + q\|_{L^\infty(0,T;L^2(\Omega))} \right).
\]

4.4.2. Variational source condition. Recall that \(F'(f_{\text{act}})h = v(\cdot, T) := j\) where \(v\) solves (65). Thus, using the operators \(A, B, C\) defined by \(Aw = -Lw + f'_\text{act}(g)w\) with impedance boundary conditions and \(g(x) = u_{\text{act}}(x, T), Bh := h(g),\) and \(Ch = D^s_z z(\cdot, T),\) where \(z\) solves
\[
D^s_z z - Lz - f'_\text{act}(P_{\text{act}})Pz = h(P_{\text{act}}) \quad z(0) = 0,
\]
we can write
\[ j = \mathbb{A}^{-1}(\mathbb{B} - \mathbb{C})h. \]

Here we assume \( q := -f'_{act}(g) \) to be nonnegative and contained in \( L^\infty(\Omega) \) so that \( \mathbb{A} \) is an isomorphism between \( H^1_0(\Omega) \) and \( L^2(\Omega) \), where \( H^1_0(\Omega) \) is defined analogously to \( H^1(\Omega) \); see (6) with \( \|Dg\|_1 := L - q \) in place of \( L \).

If \( \mathbb{C} \) can be shown to be a contraction in the sense that
\[
\|\mathbb{C}h\|_{V(\Omega)} \leq \kappa \|\mathbb{B}h\|_{V(\Omega)} \quad \text{and} \quad \|h\|_{W(I)} \leq C_1 \|\mathbb{B}h\|_{V(\Omega)} \quad \text{and} \quad \|z\|_{W(I)} \leq C_2 \|\mathbb{A}^{-1}y\|_{L^2(\Omega)},
\]
then (73) holds with
\[
\beta = \frac{C_1C_2}{1 - \kappa} |f_{act} - f_0|_{W(I)^*}.
\]
The spaces \( V(\Omega) \) and \( W(I) \) need to be appropriately chosen for this purpose. In particular we can choose the norms on \( W(I) \) and \( V(\Omega) \) to be weak in order to enable all these estimates to hold. (The weaker we choose \( W(I) \), the stronger we have to pay for it by the smoothness condition of \( |f_{act} - f_0|_{W(I)^*} \) being finite and small enough.)

Using a system of eigenvalues and eigenfunctions \( \lambda_j, \varphi_j \) of \(-L + q\) we get
\[
z(\cdot, t) = \sum_{j=1}^{\infty} \int_0^t s^{\alpha-1} E_{\alpha, \alpha}(-\lambda_j^2 s^\alpha) \left( (f'_{act}(Pu_{act}) + q)Pu_{act} + h(Pu_{act}) (t-s) \right) \varphi_j \, ds \cdot \varphi_j.
\]
Hence assuming \( P(0) = 0 \), i.e. \( g_{\min} \leq 0 \leq g_{\max} \)
\[
z_t(\cdot, t) = \sum_{j=1}^{\infty} \int_0^t s^{\alpha-1} E_{\alpha, \alpha}(-\lambda_j^2 s^\alpha) \left( f''_{act}(Pu_{act})Pu_{act}, Pz + (f'_{act}(Pu_{act}) + q)Pu_{act} \right) \, ds \cdot \varphi_j.
\]
We therefore get, analogously to (42) and estimating
\[
\begin{align*}
\|f'_{act}(Pu_{act})Pu_{act}, Pz\|_{L^\infty(0,T;L^2(\Omega))} &= \|f'_{act}(Pu_{act})Pu_{act}, P \int_0^T z(s) \, ds\|_{L^\infty(0,T;L^2(\Omega))} \\
&\leq \|f''_{act}(Pu_{act})Pu_{act}, P\|_{L^\infty(0,T;L^\infty(\Omega))} C_{H^1 \rightarrow H^1/0} \|z\|_{L^\infty(0,T;H^1(\Omega))},
\end{align*}
\]
that
\[
\|z_t(T)\|_{H^1(\Omega)} \leq t^{\alpha-1} \|h(u_0)\|_{H^1(\Omega)} + \|h'(Pu_{act})Pu_{act}\|_{L^\infty(0,T;L^2(\Omega))} + \|f''_{act}(Pu_{act})Pu_{act}, P\|_{L^\infty(0,T;L^\infty(\Omega))} + \|f''_{act}(Pu_{act})Pu_{act}, P\|_{L^\infty(0,T;L^\infty(\Omega))} \|z\|_{L^\infty(0,T;H^1(\Omega))},
\]
for \( Q < \frac{2(2-\theta)}{2-2\theta+\theta} \). Gronwall’s inequality (17) then yields
\[
\|z(T)\|_{H^1(\Omega)} \leq 2^{1-1/Q} \left( a(T) + 2^{Q-1}c_2 \left( \int_0^T a(s)Q^e c_2^{Q-1} \left( z(s)^{Q^e} \right) \, ds \right)^{1/Q^e} \right),
\]
where
\[ a(t) = t^{\alpha-1} \| h(u_0) \|_{H^\alpha_q(\Omega)} + \| h'(P_{\text{act}})P_{\text{act}} \|_{L^\infty(0,t;L^2(\Omega))}, \]

\[ \leq t^{\alpha-1} \| h(u_0) \|_{H^\alpha_q(\Omega)} + \| h' \|_{L^\infty(I)} \| P_{\text{act}} \|_{L^\infty(0,t;L^2(\Omega))}, \]

and

\[ c_2 = \tilde{C}_{\alpha}^{2^\alpha} \left( \frac{1}{\tilde{C}_{\alpha}^{2^\alpha}} \| f'_{\text{act}}(u_{\text{act}}) + q \|_{L^\infty(0,T;L^\infty(\Omega))} + \| (f''_{\text{act}}(P_{\text{act}})P_{\text{act}}) \|_{L^\infty(0,T;L^\infty(\Omega))} T^{1/2} \right). \]

Assuming

\[ \| h(u_0) \|_{H^\alpha_q(\Omega)} \leq c_3 \| h(g) \|_{H^\alpha_q(\Omega)}, \]

for all \( h \in X \) and \( c_2 \) or \( c_3 \) as well as \( \| P_{\text{act}} \|_{L^\infty(0,T)} \) sufficiently small so that

\[ 2^{1-1/2^\alpha} \left( \int_0^T \frac{2^\alpha}{\tilde{C}_{\alpha}^{2^\alpha}} (\int_0^T \frac{2^\alpha}{\tilde{C}_{\alpha}^{2^\alpha}} (T-s) ds)^{1/2^\alpha} c_3 \right. \]

\[ + \left. 2^{1-1/2^\alpha} \left( 1 + 2^\alpha - 1 \right) c_2 \left( \frac{2^\alpha}{\tilde{C}_{\alpha}^{2^\alpha}} (\int_0^T (T-s) ds)^{1/2^\alpha} \right) C(g) \| P_{\text{act}} \|_{L^\infty(0,T)} \right) =: \kappa < 1, \]

we obtain contractivity (80) with \( V(\Omega) = H^\alpha_q(\Omega) \).

From proposition 4.1 we can conclude the following convergence result.

**Corollary 4.1.** Assume that \( \Omega \subseteq \mathbb{R}_+^1, \alpha \in (\frac{1}{2}, 1], f'_{\text{act}} \leq 0, f_{\min} \leq 0 \leq f_{\max} \| f_{\text{act}} - f_0 \| \leq q \) sufficiently small and that (79), (81) holds with \( c_2 \) or \( c_3 \) and \( \| P_{\text{act}} \|_{L^\infty(0,T)} \) sufficiently small so that (82) is satisfied. Let \( F : X = H^\alpha_q(\Omega) \rightarrow Y = L^2(\Omega) \) with \( \sigma > \frac{1}{2} \) and \( q \) as in (79) be defined by \( F(f) = u(\cdot, T) \) where \( u \) solves (63), and let \( (\delta^\alpha)_{\delta \in [0,\delta]} \subseteq Y \) be a family of data satisfying (68) and let, for each \( \delta \in (0,\delta] \), the regularized approximation \( f_{\text{act}}(\delta) \) be defined by an early stopped Newton-type iteration according to either (69), (71) or (70), (71), with \( \tau \) sufficiently large.

Then for each \( \delta \in (0,\delta] \), the stopping index is finite and the regularized approximations converge to the exact solution \( f_{\text{act}} \) of \( F(f) = g \) as appearing in (73) at the rate (76), i.e.

\[ \| f_{\text{act}}(\delta) - f_{\text{act}} \| = O(\sqrt{\delta}), \| F(f_{\text{act}}(\delta)) - g \| = O(\delta). \]

**5. Reconstructions**

**5.1. Computational algorithms used**

Here we describe some of the computational logistics behind the two approaches analysed in the previous sections: that of the iteration scheme based on (30) and those Newton-type methods based on computing the map and the associated Jacobian from (64) and (65). In both methods the forward solver was a standard Crank–Nicolson scheme for the parabolic equation while in the fractional case the time stepping and computation of \( D^\alpha_q \) at the final time was based on a second-order method introduced in [8].

The former of these methods is by far the simplest and without question performed both exceedingly well and in a much superior way to those based on the more complex Newton methods. The algorithm can be summarized in the following steps:
Obtain the (noisy) data $g(x) = u(x, T; f)$ from some unknown $f(u)$ but given the values of the initial and boundary conditions under which the solution $u$ of the direct problem was constrained.

Since we require the quantity $Lg$ this must be computed in a stable manner. This was accomplished by using the fact that $u(x, T)$ must share the prescribed boundary conditions and so the exact value can be approximated as $\sum_n g_n \phi_n(x)$ where $\{\phi_n(x)\}$ is the orthogonal set of eigenfunctions of $L$ on $\Omega$. Then $Lg$ can be written as $\sum_n \lambda_n g_n \phi_n(x)$. Now we require that $Lg$ be at least as smooth as the desired value for $f(g)$ to lie in $W^{1,\infty}$ and we do this by a standard least squares fit with an appropriate penalty term to ensure this level of smoothness in $Lg$. Hence we used the $H^s(\Omega)$ norm where $s$ must be chosen in accordance with the Sobolev embedding theorem value (see previous sections).

At each iteration we must recover $f$ from $f(g) = u_t(x, T; f) - Lg$. We do so by representing $f$ at the (now smoothed) function $g$ in a basis set of functions. This is essentially the projections $P$ and $\tilde{P}$ noted in (32) and (35). Since we make no constraints on the form of $f$ other than sufficient regularity we do not choose a basis with in-built restrictions as would be obtained from an eigenfunction expansion. Instead we used a radial basis of shifted Gaussian functions $b_j(x) := e^{-(x-x_j)^2/\sigma}$ centred at nodal points $\{x_j\}$ and with width specified by the parameter $\sigma$. Since we use a finite number of such analytic functions, the computed $f$ at each iteration will be likewise smooth. We compute the basis coefficients from a given $f$ in the standard way using the Gram matrix of $\{e^{-(x-x_j)^2/\sigma}\}$. It should be noted that no smoothing is necessary at this step – the entire regularization of the problem being achieved by the smoothing of the noisy data $g$.

Figure 1 shows the reconstruction of the reaction term $f(u) = 2u(1-u)(u-a)$ with $a = 0.75$ from 1% and 5% noise.

**Figure 1.** Reconstructions of $f(u) = 2u(1-u)(u-a)$ with $a = 0.75$ from 1% and 5% noise.
this resulted in under-smoothing. In all numerical runs based on the iteration scheme (30) the initial approximation for $f$ was taken to be the zero function.

Of course there are more challenging possible reaction terms, and figure 2 shows the reconstruction of the Lipschitz function given by

$$f^{(b)}(u) = \begin{cases} 8u^2 & u \leq \frac{1}{2} \\ (1 + \cos(5(u - \frac{1}{2})))e^{-(u-\frac{1}{2})} & u > \frac{1}{2} \end{cases}.$$  

The data $g(x)$ had 1% added noise and again effective numerical convergence was obtained within five iterations and even iteration 2 was almost as close to the original.

For the case of Newton schemes we of course must compute the derivative of the map $F$ and its associated derivative through equations (64) and (65). The computations for these rely on the same solver families used for the direct equation. As usual, there are several versions of Newton schemes available. For example, we utilized the option of freezing the Jacobian at the initial approximation to obviate what is the single most computationally intensive item in each iteration. In common with many nonlinear problems, the initial approximation is quite important. Choosing one at considerable distance from the actual function frequently led to a failure to converge. Even with a reasonable initial approximation it was often prudent to implement some control over the length of the Newton step. Although the convergence under such conditions was fairly rapid, perhaps surprisingly, the number of iterations required was almost always larger than that taken by the iteration scheme (30).

Figures 3 show reconstructions of $f(u) = f^{(a)}(u)$ and $f(u) = f^{(b)}(u)$ using Newton iteration. In both cases the data $g$ was subject to 1% noise. The initial guess was taken to be of similar form: $f^{(a)}_{\text{init}}(u) = u(1 - u)(u - \frac{1}{2})$ and $f^{(b)}_{\text{init}} = 1 + \sin(4u)$. However, the norms of $f$ and $f_{\text{init}}$ were quite far apart in both cases.

The figures show the 10th iteration although the rate clearly slowed down by the fourth or fifth iteration, which were already close to the one shown.

The use of a frozen Newton gave almost similar results; only a slight lag in the convergence rate was noticed. This of course shows that, provided the initial approximation is sufficiently good, this version has a significantly lower computational cost. Conversely, if the initial approximation is poorer, then a hybrid approach with initial shorter Newton steps followed by a holding of the derivative after a certain point becomes a viable option.
Figure 3. Reconstructions of $f^*(u), f^b(u)$ from 1% noise using Newton iteration.

Figure 4. Variation of $\|f^{(a)} - f^{(b)}\|$ with $T$ and $\alpha$: left showing iteration scheme and right showing the Newton.

We note that the issue of regularization parameter choice is far from trivial in this case. Our regularization of the (singular) Jacobian required a stronger norm than just $L^2$ and in fact we used a penalty term of the form $\epsilon_0 I + \epsilon_2 R$ where $R$ was a smoothing matrix formed from $R_{jk} = \int_{u_{\text{min}}}^{u_{\text{max}}} b_j(u) b_k(u) \lambda(u) \, du$ where the weight function $\lambda(u)$ allowed for differential smoothing over the interval.

5.2. Dependence on the final time $T$ and on $\alpha$

The sign of $f(u)$ and of $f'(u)$ in the reaction–diffusion model will make a difference to the rate of convergence as will the time of measurement $T$. For example, theorem 3.2 shows that the contractive nature of the map $T$ depends on $T$, and the larger the value $T$ the smaller the contraction constant. In the parabolic case this decay is in fact exponential but we cannot expect the same in the fractional situation where the large time decay of the exponential is replaced by the linear decay of the Mittag-Leffler function.

The leftmost figure in figure 4 shows the dependence on the choice of $T$ and of $\alpha$ of the convergence rate as measured by the improvement of the first iteration over the initial approximation $f = 0$, namely $\|f^{(a)} - f^{(b)}\|$. This is a more definitive measure of the effect of the parameters $T$ and $\alpha$ than the number of iterations in achieving effective convergence as the latter involves a degree of subjectivity. However both measures give consistent results.
In the parabolic case the results are as predicted by theorem 3.2; the number of steps required for a given accuracy decreases as \( T \) increases. What is perhaps surprising given the only linear asymptotic behaviour of the Mittag-Leffler function is the fact the convergence for large values of \( T \) is at least as good in the fractional case and, indeed, it is considerably better for small values of the final time \( T \).

The right side of figure 4 shows the corresponding results for Newton’s method. Here we again only look at the difference in the \( L^\infty \) norm of the initial approximation of \( f^{(a)}_{\text{init}} \) of \( f^{(a)}(u) \) obtained after the first iteration.

In the parabolic case there is a definite improvement in the convergence rate as \( T \) increases. However, this feature diminishes with decreasing \( \alpha \) and by \( \alpha = 0.25 \) was almost imperceptible. Note that the improvements obtained by the first approximations shown in the above figures should not be directly compared as the iteration scheme (30) started from an initial approximation of \( f^{(a)}_{\text{init}} = 0 \) whereas the Newton scheme started from \( f^{(a)}_{\text{init}}(u) = u(1-u)(u-\frac{1}{4}) \).

6. Epilogue

Reaction–diffusion equations are building blocks for mathematical modelling in a wide range of applications. In many cases these models make a specific-form ansatz for both the diffusion and reaction mechanisms and contain parameters that have to be determined in order to prime the system for subsequent predictive behaviour of the quantity \( u \) of prime interest. This has traditionally been accomplished by a least squares fit to a relatively small number of parameters. In this paper we have sought to go beyond this by making no prior assumptions about the qualitative behaviour of the unknown reaction term \( f(u) \), but by recovering a full functional form from measured information consisting of the variable \( u \) at a fixed time. Such data are very natural in many applications; for example, in the case of ecology models, it corresponds to a population census.

We have assumed that the diffusive process is known, including any parameters involved in this component. Thus, the diffusion coefficient \( a(x) \) and any transport term \( b(x) \) is assumed known as well as whether diffusion is through a Brownian or a particular version of an anomalous process. Other random walk models are possible and are frequently utilized. This paper has concentrated on the case of a process where the probability density function for the mean waiting time may be unbounded leading to a time derivative of the fractional order \( \alpha < 1 \). In the limiting case when the mean waiting time is bounded, that is, \( \alpha = 1 \), we of course obtain the regular parabolic operator, and this is also covered by our analysis. The recovery of \( f \), in particular the convergence rate of the numerical algorithms, depends on \( \alpha \) and also on the final time of measurement \( T \), and the interplay between these quantities is quite complex.

We also noted the possibility of a fractional power in the space variable. However, this was through the fractional power of the elliptic operator \(-L\) rather than through the similar process of fractional-order spatial derivatives that would arise if our assumption on the random walk process took a spatial probability density function with infinite variance. While the latter model has a definite physical basis, the replacement of \(-L\) by its fractional power \((-L)^\sigma\) has a less clear physical connection. It does have mathematical elegance and convenience for analysis and computation and this has lead to its popularity. Discovering the interplay among the fractional power constants \( \alpha, \sigma \), the diffusion coefficient \( a(x) \) and the inversion of \( f(u) \) would be very interesting.
The fact that the model considered allows for spatially dependent coefficients is important beyond mathematical generality; this could be an essential factor in modelling. For example, in ecology, certain areas could be either rich in food or less hospitable and thus support a different \( f(u) \) term. In this respect the current paper is complementary to [17], where it was assumed that the reaction term \( f(u) \) was known but the coefficient that quantified this environmental factor had to be determined.

From a mathematical perspective, the standard parabolic reaction diffusion equation has considerable complexity. For example, without a growth constraint on \( f(u) \) the solution \( u \) can ‘blow up’ in finite time. Even if we measure our data—the solution \( u \)—at a time \( T \), where \( T \) is less than this blow-up time, we are likely to have extreme values occurring if the leading term in \( f(u) \) is asymptotically near the critical exponent. Our reconstruction methods had to be set in a suitable space to prevent such an event occurring during the iterative process. The regularity of the non-homogeneous parabolic equation is well understood and there are critical gaps in the Schauder space regularity of the solution in the time variable that prevented a straightforward analysis of the convergence of the iteration schemes used in reconstructing \( f(u) \). In Sobolev spaces the embedding schemes between these spaces forced much of the analysis to hold only in one space variable. In the case of time-fractional diffusion the situation is more complicated. The subdiffusion (\( \alpha < 1 \)) equation has even more restricted regularity properties due to the singularity at \( t = 0 \) that causes at most a two-derivative pick up in smoothness of the solution \( u(\cdot,t) \) over that of \( f \) whereas the parabolic equation has an infinite smoothness pick up. For reasons such as these, this paper is quite technical in parts and, as a compromise on length, has a rather dense exposition, which we have tried to ameliorate as much as possible.

In conclusion, this paper has only touched the surface of an exceedingly nontrivial problem. There remain challenging mathematical questions in incorporating more complex diffusion models into the framework as well as the myriad of cases that could occur in terms of finding optimal driving conditions such as initial boundary data to best recover a given type of function \( f \). Clearly, in higher space dimensions and for space- and time-fractional models there are significant obstacles in developing both effective numerical methods and providing the required numerical analysis. In seeking a solution to the inverse problem of reconstructing \( f(u) \) we have concentrated on the featured iteration scheme based on the map involving \( u_\ell(\cdot,T;f) \). An alternative approach is to extend the scope of Newton-type methods based on the map from \( f(u) \) to the data \( g(x) = u(x,T) \) instead. This will require showing the injectivity of the resulting Jacobian map as well as finding regularization strategies for its effective inversion. We view this as a nontrivial step, but believe it allows for multiple options. Beyond the indicated analysis, mathematical opportunity beckons.

**Acknowledgment**

The work of the first author was supported by the Austrian Science Fund FWF under Grants I2271 and P30054.

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

The authors wish to thank the reviewers for their careful reading of the manuscript and their detailed reports with valuable comments and suggestions that have led to an improved version of the paper.
ORCID iDs

Barbara Kaltenbacher https://orcid.org/0000-0002-3295-6977
William Rundell https://orcid.org/0000-0001-9579-2183

References

[1] Bakushinsky A B 1992 The problem of the convergence of the iteratively regularized Gauss–Newton method Comput. Math. Math. Phys. 32 1353–9
[2] Cantrel R S and Cosner C 2003 Spatial Ecology via Reaction–Diffusion Equations (Wiley Series in Mathematical and Computational Biology) (Chichester: Wiley)
[3] Djrbashian M M 1966 Integral Transformations and Representation of Functions in a Complex Domain (Moscow: Nauka) (in Russian)
[4] Djrbashian M M 1970 Fractional derivatives and the cauchy problem for differential equations of fractional order Izv. Akad. Nauk Armajan. SSR 75 71–96
[5] Djrbashian M M 1993 Harmonic Analysis and Boundary Value Problems in the Complex Domain (Basel: Birkhäuser)
[6] DuChateau P and Rundell W 1985 Uncity in an inverse problem for an unknown reaction term in a reaction–diffusion equation J. Differ. Equ. 59 155–64
[7] Fan K 1952 Fixed-point and minimax theorems in locally convex topological linear spaces Proc. Natl Acad. Sci. USA 38 121–6
[8] Gao G H, Sun Z Z and Zhang H W 2014 A new fractional numerical differentiation formula to approximate the Caputo fractional derivative and its applications J. Computat. Phys. 259 33–50
[9] Grindrod P 1996 The Theory and Applications of Reaction–Diffusion Equations (Oxford Applied Mathematics and Computing Science Series) 2nd edn (New York: Clarendon) (Patterns and waves)
[10] Hofmann B, Kaltenbacher B, Pöschl C and Scherzer O 2007 A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators Inverse Problems 23 987–1010
[11] Hohage T 1997 Logarithmic convergence rates of the iteratively regularized Gauss–Newton method for an inverse potential and an inverse scattering problem Inverse Problems 13 1279–99
[12] Jin B, Li B and Zhou Z 2018 Numerical analysis of nonlinear subdiffusion equations SIAM J. Numer. Anal. 56 1–23
[13] Jin B and Rundell W 2015 A tutorial on inverse problems for anomalous diffusion processes Inverse Problems 31 035003
[14] Jin Q 2010 On a class of frozen regularized Gauss–Newton methods for nonlinear inverse problems Math. Comput. 79 2191–211
[15] Kaltenbacher B, Neubauer A and Scherzer O 2008 Iterative Regularization Methods for Nonlinear Ill-Posed Problems (Radon Series on Computational and Applied Mathematics) (Berlin: de Gruyter)
[16] Kaltenbacher B and Previatti de Souza M L 2018 Convergence and adaptive discretization of the IRGNM Tikhonov and the IRGNM Ivanov method under a tangential cone condition in Banach space Numer. Math. 140 449–78
[17] Kaltenbacher B and Rundell W 2019 On an inverse potential problem for a fractional reaction–diffusion equation Inverse Problems 35 065004
[18] Kaltenbacher B and Rundell W 2019 Regularization of a backwards parabolic equation by fractional operators Inverse Problems Imaging 13 401–30
[19] Kuttler C 2017 Reaction–diffusion equations and their application on bacterial communication Disease Modelling and Public Health (Handbook of Statistics) ed A Rao et al (Amsterdam: Elsevier)
[20] Mainardi F and Gorenflo R 2000 On Mittag-Leffler-type functions in fractional evolution processes J. Comput. Appl. Math. 118 283–99 (Higher transcendental functions and their applications)
[21] Mittag-Leffler G M 1903 Sur la nouvelle function E
c
Acad. Sci. Paris 137 554–5
[22] Murray J D 2002 Mathematical Biology. I (Interdisciplinary Applied Mathematics) vol 17) 3rd edn (New York: Springer) (An introduction)
[23] Pilant M and Rundell W 1987 Iteration schemes for unknown coefficient problems in parabolic equations Numer. Methods PDE 3 313–25

37
[24] Pilant M and Rundell W 1988 Fixed point methods for a nonlinear parabolic inverse coefficient problem Commun. PDE 13 469–93
[25] Pilant M S and Rundell W 1986 An inverse problem for a nonlinear parabolic equation Commun. PDE 11 445–57
[26] Sakamoto K and Yamamoto M 2011 Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems J. Math. Anal. Appl. 382 426–47
[27] Samko S G, Kilbas A A and Marichev O I 1993 Fractional Integrals and Derivatives (Yverdon: Gordon and Breach)
[28] Schuster T, Kaltenbacher B, Hofmann B and Kazimierski K S 2012 Regularization Methods in Banach Spaces (Berlin: Walter de Gruyter)
[29] Thomée V 2006 Galerkin Finite Element Methods for Parabolic Problems 2nd edn (Berlin: Springer)
[30] Werner F 2015 On convergence rates for iteratively regularized Newton-type methods under a Lipschitz-type nonlinearity condition J. Inverse Ill-Posed Problems 23 75–84