Iterative Regularization of a Parameter Identification Problem occurring in Polymer Crystallization

Martin Burger *

Abstract

This paper is devoted to the mathematical analysis and regularization of an identification problem related to non-isothermal crystallization of polymers, which can be modelled by an initial-boundary value problem for a coupled system of parabolic and hyperbolic partial differential equations. The identification problem consists in estimating a material function of temperature, which appears as a nonlinearity in the equations.

Existence and uniqueness of a solution of the direct problem is shown as well as its stability with respect to the parameter. Furthermore, we develop algorithms for the application of various iterative regularization methods to this particular problem. Their use is justified by verifying the Frèchet-differentiability of the parameter-to-output map, which is needed for their realization. The numerical performance of the iterative methods is compared with respect to speed of convergence, stability and efficiency.

Keywords: Crystallization of Polymers, Iterative Regularization, Parameter Identification, Coupled Hyperbolic-Parabolic Systems

AMS Subject Classification: 35K20, 35L50, 35 Q80, 35R25, 35R30, 65M30

Abbreviated Title: Parameter Identification in Polymer Crystallization

1 Introduction

Various techniques of mathematical modelling have been used to derive reasonable models for non-isothermal crystallization of polymers under typical processing conditions. The use of such models is not only the insight into the crystallization process itself they might provide, but also the prediction of the structure development and final morphology of the solidified material, which basically determines its mechanical properties (cf. [13]).

On a macroscopic scale, non-isothermal crystallization can be modelled by the following system of partial differential equations, describing the evolution of the temperature T, the

*Industrial Mathematics Institute, Johannes Kepler University Linz, Altenbergerstr. 69, A-4040 Linz, Austria

degree of crystallinity ξ and the mean free surface densities u and v in a domain $\Omega \subset \mathbf{R}^d$ ($d = 1, 2, 3$) and some time interval $(0, t_*)$ (cf. [6, 7, 9])

$$\frac{\partial \xi}{\partial t} = \widetilde{G}(T)(1-\xi)u \tag{1.1}$$

$$\frac{\partial u}{\partial t} = \nabla.(\widetilde{G}(T)v) + \mathcal{F}_d[\widetilde{G}, \widetilde{N}, T] \tag{1.2}$$

$$\frac{\partial v}{\partial t} = \nabla(\widetilde{G}(T)u) \tag{1.3}$$

$$c\rho\frac{\partial T}{\partial t} = \nabla.(k\nabla T) + \frac{\partial}{\partial t}(h\xi), \tag{1.4}$$

in $\Omega \times (0, t_*)$, supplemented by the boundary conditions

$$u + \langle v, n\rangle = 0, \tag{1.5}$$

$$\frac{\partial T}{\partial n} = \alpha(T - T_{out}), \tag{1.6}$$

on $\partial\Omega \times (0, t_*)$ and initial values given by

$$\xi = 0 \tag{1.7}$$

$$u = 0 \tag{1.8}$$

$$v = 0 \tag{1.9}$$

$$T = T^0 \tag{1.10}$$

in $\Omega \times \{0\}$, usually with $T^0(x) \geq T_m$ (melting temperature) for all $x \in \Omega$. The source term $\mathcal{F}_d$ is a nonlinear operator dependent upon the spatial dimension:

$$\mathcal{F}_1[\widetilde{G}, \widetilde{N}, T](x,t) := 2\widetilde{N}(T(x,t))_t \tag{1.11}$$

$$\mathcal{F}_2[\widetilde{G}, \widetilde{N}, T](x,t) := 2\pi\widetilde{G}(T(x,t))\left(\widetilde{N}(T(x,t)) - \widetilde{N}(T(x,0))\right) \tag{1.12}$$

$$\mathcal{F}_3[\widetilde{G}, \widetilde{N}, T](x,t) := 4\pi\widetilde{G}(T(x,t))\int_0^t \widetilde{G}(T(x,s))\left(\widetilde{N}(T(x,s)) - \widetilde{N}(T(x,0))\right) ds \tag{1.13}$$

While many experiments show that the parameters $\widetilde{G}$ (the growth rate) and $\widetilde{N}$ (the nucleation rate) depend upon temperature only, all parameters appearing in the heat transfer model such as the density ρ, the heat capacity c, the heat conductivity κ, the latent heat h and the heat transfer coefficient α may also depend upon the degree of crystallinity (cf. [9]). However, since the variance of most parameters with respect to T and ξ is rather low, we will analyze a simplified model, which avoids some technical complications, but still includes the essential nonlinearities and thus serves to obtain some insight into the problem's nature.

In the one-dimensional setup we may assume that Ω is the open interval (x_L, x_R) The growth and nucleation rate are now denoted by a and b and include scaling of the problem.

For the sake of simplicity we assume most parameters in the heat equation to be constant and neglect the temperature-depence of a, which is certainly a simplification, but still provides the basic properties of the system with respect to the parameter b. Eliminating (1.1) and (1.7) we obtain (after simple transformations) the system

$$
\begin{aligned}
T_t &= (DT_x)_x + Le^{-\int_0^t (au)ds} au && \text{in } \Omega \times I && (1.14)\\
u_t &= (av)_x + b(T)_t && \text{in } \Omega \times I && (1.15)\\
v_t &= (au)_x && \text{in } \Omega \times I && (1.16)\\
T_n &= \alpha(T - T^1) && \text{on } \partial\Omega \times I && (1.17)\\
u + \langle v, n \rangle &= 0 && \text{on } \partial\Omega \times I && (1.18)\\
T &= T^0 && \text{in } \Omega \times \{0\} && (1.19)\\
u &= 0 && \text{in } \Omega \times \{0\} && (1.20)\\
v &= 0 && \text{in } \Omega \times \{0\}, && (1.21)
\end{aligned}
$$

using the abbreviations $I := (0, t_*)$ and $Q := \Omega \times I$. We will assume that L and α are positive constants and that there exists a positive real number D_0 such that $D(x) \geq D_0$ for all $x \in \Omega$, which is a reasonable assumptions upon a model for heat conduction.

Besides solving the model equations, an important problem in this context is the identification of the nucleation rate $\widetilde{N}$ (respectively b in the scaled system) from indirect measurements, using data about the temperature at the boundary and the final degree of crystallinity. A first algorithm for the stable solution of this inverse problem has been developed in the case of one spatial dimension in [8]; it was based on several assumptions about the well-posedness of the 'direct problem', i.e., about the existence and uniqueness of a solution of the model equations and its stable dependence upon the parameter in appropriate function spaces. In this paper we will also study Newton-type methods for the solution of the inverse problem and rigourously prove well-posedness of the direct problem and other important properties, such as Frèchet-differentiability of the parameter-to-output map. This is not only an important theoretical justification, but also yields further information about the correct choice of norms in the identification problem, which is crucial for the design of all algorithms. In general, Newton-type methods are expected to be faster than the explicit Landweber iteration, but due to the extremly high computional cost of computing the derivative and consequently the Newton-matrix together with the slow down caused by the ill-posedness of the problem, it will turn out that this is not the case for this application.

The estimation of the material function b is a *parameter identification problem* in a system of partial differential equations (cf. [2, 10, 19] for a general reference on this topic); the majority of such problems are *ill-posed*, i.e., the solution does not depend on the data in a stable way. Due to this inherent instability, *regularization methods* have to be used in order to obtain stable approximations (cf. e.g. [14]). Our aim in this work is to apply iterative regularization methods to the identification problem and to investigate their convergence behaviour with respect to stability and efficiency. The numerical results will show that faster methods do not necessarily perform better if the problem is ill-posed.

We will proceed as follows: in Section 2 the direct initial-boundary value problem (1.14)-(1.21) will be analyzed. The various norms and function spaces which are used there are defined and explained in detail in [23, 24].

In Section 3 the inverse problem of identifying the nucleation rate b will be formulated in a mathematical way in the context of *ill-posed problems*. It can be interpreted as the solution of an ill-posed nonlinear operator equation involving the so-called *parameter-to-output map*. Properties of this map, which are needed for the design of stable solution algorithms for the identification problem are analyzed in Section 4.

Section 5 is devoted to the realization of Newton-type regularization methods for this complicated inverse problem and the development of efficient numerical algorithms. We finally present numerical results and conclusions in Section 6.

2 The direct initial-boundary value problem

In this section we will prove existence and uniqueness of a solution of the initial-boundary value problem (1.14)-(1.21) using fixed point arguments. This analysis will be carried out in different steps: first we investigate separately the parabolic and hyperbolic part of the system in Sections 2.1 and 2.2, in Section 2.3 we finally put the solution operators of these two problems together and transform (1.14)-(1.21) to an equivalent fixed point problem. It will turn out that the resulting nonlinear operator is contractive for small terminal time t_*, which allows to obtain the desired result by an application of Banach's fixed point theorem.

2.1 The parabolic equation

We first consider the problem of solving the parabolic equation for given u and v (respectively w), which allows us to focus on the linear part of this equation. Thus, we start with a standard result about the solution of linear parabolic initial-boundary value problems of the form

$$T_t = (DT_x)_x + f \qquad \text{in } \Omega \times I \tag{2.1}$$
$$T_n = \alpha(T - T^1) \qquad \text{on } \partial\Omega \times I \tag{2.2}$$
$$T = T^0 \qquad \text{in } \Omega \times \{0\}. \tag{2.3}$$

for which the following standard result holds:

Lemma 2.1. [24, p.37] *Let $f \in L^2(Q)$, $T^0 \in H^1(\Omega)$, $T^1 \in H^{\frac{1}{2},\frac{1}{4}}(\partial\Omega \times I)$. Then there exists a solution $T \in H^{2,1}(Q)$ of (2.1)-(2.3), which is unique in $H^{1,0}(Q)$, and a constant c_0 (dependent only on Ω) such that*

$$\|T\|_{H^{2,1}(Q)} \leq c_0 \left(\left\|T^0\right\|_{H^1(\Omega)} + \left\|T^1\right\|_{H^{\frac{1}{2},\frac{1}{4}}(\partial\Omega\times I)} + \|f\|_{L^2(Q)} \right). \tag{2.4}$$

This result enables the definition of the affinely-linear solution operator

$$\begin{array}{rccc} \mathcal{S}_1 : & L^2(Q) & \to & H^{2,1}(Q) \\ & f & \mapsto & T. \end{array} \tag{2.5}$$

The solution of the nonlinear boundary value problem (1.14), (1.17), (1.19) for given u and v can be written as the concatenation of $\mathcal{S}_1$ with the nonlinear operator

$$\begin{array}{rccc} \mathcal{N}_1: & \mathcal{D}(\mathcal{N}_1) & \to & L^2(Q) \\ & (u,v) & \mapsto & e^{-\int_0^t (au)ds} au, \end{array} \tag{2.6}$$

where

$$\mathcal{D}(\mathcal{N}_1) := \left\{ \, (u,v) \in L^\infty(I, L^2(\Omega)) \times L^\infty(I, L^2(\Omega)) \mid v_t = (au)_x \, \right\}. \tag{2.7}$$

Although v does not appear explicitely in the definition of $\mathcal{N}_1$, we write $\mathcal{N}_1$(u,v), since we will see below that the appropriate norm of $\mathcal{N}_1$ will depend upon u as well as on v.

Lemma 2.2. *Let $(u,v) \in \mathcal{D}(\mathcal{N}_1)$, then $w = \int_0^t (au)\ ds \in L^\infty(Q)$ with*

$$\|w\|_{L^\infty(Q)} \;\leq\; \mu(t_* \|a\|_{L^\infty(Q)} \|u\|_{L^\infty(I,L^2(\Omega))} + \|v\|_{L^\infty(I,L^2(\Omega))}), \tag{2.8}$$

where μ is the norm of the embedding operator from $H^1(\Omega)$ into $L^\infty(\Omega)$.

Proof. We immediately obtain

$$w_x = \int_0^t (au)_x ds = \int_0^t v_t ds = v \in L^\infty(I, L^2(\Omega)).$$

Thus, $w \in L^\infty(I, H^1(\Omega))$ and because of the continuous embedding $H^1(\Omega) \hookrightarrow L^\infty(\Omega)$ we obtain $w \in L^\infty(Q)$ with

$$\begin{aligned} \|w\|_{L^\infty(Q)} &\leq \mu \|w\|_{L^\infty(I,H^1(\Omega))} \\ &= \sup_{t\in I} \left(\int_\Omega |w(x,t)|^2 + |v(x,t)|^2 \right)^{\frac{1}{2}} \\ &\leq \mu(t_* \|a\|_{L^\infty(Q)} \|u\|_{L^\infty(I,L^2(\Omega))} + \|v\|_{L^\infty(I,L^2(\Omega))}) \end{aligned}$$

□

This result about the smoothness of w is crucial for estimating the L^2-norm of the nonlinear operator $\mathcal{N}_1$:

Lemma 2.3. *Let $(u,v) \in \mathcal{D}(\mathcal{N}_1)$ and $(\overline{u},\overline{v}) \in \mathcal{D}(\mathcal{N}_1)$ with norms bounded by C. Then the estimate*

$$\|\mathcal{N}_1(u,v) - \mathcal{N}_1(\overline{u},\overline{v})\|_{L^2(Q)} \;\leq\; c_1(t_*,C) \left(\|u-\overline{u}\|_{L^\infty(I,L^2(\Omega))} + \|v-\overline{v}\|_{L^\infty(I,L^2(\Omega))} \right) \tag{2.9}$$

with

$$c_1(t_*,C) := \|a\|_{L^\infty(Q)} \sqrt{t_*}(C+1) e^{\mu \max(1, \|a\|_{L^\infty(Q)} t_*)C} \tag{2.10}$$

holds. Furthermore, $\mathcal{N}_1(0,0) = 0$ holds.

Proof. Employing the triangle inequality we have

$$\|\mathcal{N}_1(u,v)-\mathcal{N}_1(\overline{u},\overline{v})\|_{L^2(Q)} \leq \left\|(e^{-w}-e^{-\overline{w}})au\right\|_{L^2(Q)}+\left\|e^{-\overline{w}}a(u-\overline{u})\right\|_{L^2(Q)}, \tag{2.11}$$

and a straight-forward estimate using the monotonicity of the exponential function yields

$$\begin{aligned}\left\|e^{-\overline{w}}a(u-\overline{u})\right\|_{L^2(Q)} &\leq \left\|e^{-w}\right\|_{L^\infty(Q)}\|a\|_{L^\infty(Q)}\|u-\overline{u}\|_{L^2(Q)}\\ &\leq e^{\|w\|_{L^\infty(Q)}}\|a\|_{L^\infty(Q)}\|u-\overline{u}\|_{L^2(Q)}\\ &\leq \|a\|_{L^\infty(Q)}\sqrt{t_*}e^{\mu\max(1,\|a\|_{L^\infty(Q)}t_*)C}\|u-\overline{u}\|_{L^\infty(I,L^2(\Omega))} \qquad (2.12)\end{aligned}$$

for the second term. The first term may be estimated by

$$\left\|(e^{-w}-e^{-\overline{w}})au\right\|_{L^2(Q)} \leq \left\|e^{-w}-e^{-\overline{w}}\right\|_{L^\infty(Q)}\|a\|_{L^\infty(Q)}\|u\|_{L^2(Q)}.$$

Since the exponential function is continuously differentiable, the mean value theorem implies

$$\begin{aligned}\left|e^{-w(x,t)}-e^{-\overline{w}(x,t)}\right| &\leq \sup_{s\in[w(x,t),\overline{w}(x,t)]} e^s\,|w(x,t)-\overline{w}(x,t)|\\ &\leq e^{\max(\|w\|_{L^\infty(Q)},\|\overline{w}\|_{L^\infty(Q)})}\,|w(x,t)-\overline{w}(x,t)|\\ &\leq e^{\mu\max(1,\|a\|_{L^\infty(Q)}t_*)C}\,|w(x,t)-\overline{w}(x,t)|\end{aligned}$$

for almost all $(x,t)\in Q$, and thus,

$$\begin{aligned}\left\|(e^{-w}-e^{-\overline{w}})au\right\|_{L^2(Q)} &\leq \|a\|_{L^\infty(Q)}Ce^{\mu\max(1,\|a\|_{L^\infty(Q)}t_*)C}\\ &\qquad\left(\|u-\overline{u}\|_{L^2(Q)}+\|v-\overline{v}\|_{L^2(Q)}\right) \qquad (2.13)\\ &\leq \|a\|_{L^\infty(Q)}\sqrt{t_*}Ce^{\mu\max(1,\|a\|_{L^\infty(Q)}t_*)C}\\ &\qquad\left(\|u-\overline{u}\|_{L^\infty(I,L^2(\Omega))}+\|v-\overline{v}\|_{L^\infty(I,L^2(\Omega))}\right) \qquad (2.14)\end{aligned}$$

Inserting (2.12) and (2.14) into (2.11) yields (2.9).

The additional result $\mathcal{N}_1(0,0)=0$ is obvious. □

Now we are able to apply the various results to the solution of (1.14), (1.17), (1.19) with given u and v:

Proposition 2.4. *Let $(u,v)\in\mathcal{D}(\mathcal{N}_1)$ and $(\overline{u},\overline{v})\in\mathcal{D}(\mathcal{N}_1)$ with norms bounded by C. Then $T=\mathcal{S}_1(\mathcal{N}_1(u,v))$ and $\overline{T}=\mathcal{S}_1(\mathcal{N}_1(\overline{u},\overline{v}))$ satisfy*

$$\begin{aligned}\|T\|_{H^{2,1}(Q)} &\leq c_0\left(\left\|T^0\right\|_{H^1(\Omega)}+\left\|T^1\right\|_{H^{\frac{1}{2},\frac{1}{4}}(\partial\Omega\times I)}\right)+\\ &\qquad c_0c_1(t_*,C)\left(\|u\|_{L^\infty(I,L^2(\Omega))}+\|v\|_{L^\infty(I,L^2(\Omega))}\right) \qquad (2.15)\\ \left\|T-\overline{T}\right\|_{H^{2,1}(Q)} &\leq c_0c_1(t_*,C)\left(\|u-\overline{u}\|_{L^\infty(I,L^2(\Omega))}+\|v-\overline{v}\|_{L^\infty(I,L^2(\Omega))}\right) \qquad (2.16)\end{aligned}$$

Proof. The assertions follow directly from Lemma 2.1 and Lemma 2.3. □

2.2 The hyperbolic equations

Now we turn to the problem of solving the hyperbolic initial-boundary value problem (1.15), (1.16), (1.18), (1.20), (1.21) for given T. Since the remaining system is linear, we first show well-posedness of the corresponding linear problem, namely

$$u_t = (av)_x + g \qquad \text{in } \Omega \times I \tag{2.17}$$

$$v_t = (au)_x + h \qquad \text{in } \Omega \times I \tag{2.18}$$

$$u + \langle v, n \rangle = 0 \qquad \text{on } \partial\Omega \times I \tag{2.19}$$

$$u = 0 \qquad \text{in } \Omega \times \{0\} \tag{2.20}$$

$$v = 0 \qquad \text{in } \Omega \times \{0\}, \tag{2.21}$$

where a is supposed to fulfill the following

Assumption 2.5. Let, in the remainder of this section, $a \geq 0$, $a \in W^{1,\infty}(Q)$, $f, g \in L^2(Q)$ and

$$(\|a_x\|_\infty + 1)t_* < 1. \tag{2.22}$$

The system (2.17)-(2.21) is an almost standard hyperbolic problem as e.g. in [30], but there are two unusual effects that require a special treatment. First of all, the boundary condition does not fit into the usual form, which is rather a technical than a conceptual problem, and secondly we do not want to assume that a is bounded away from zero uniformly, which is not realistic for practical applications. A consequence of the latter is that we cannot expect (u, v) to be a classical solution, but only weak solution in $L^2(Q)$.

The weak formulation of (2.17)-(2.21) is given by

$$\langle u', \phi \rangle + \langle av, \phi_x \rangle + [au, \phi] = \langle f, \phi \rangle \tag{2.23}$$

$$\langle v', \psi \rangle + \langle au, \psi_x \rangle - [au, \psi] = \langle g, \psi \rangle \tag{2.24}$$

for all $\phi, \psi \in H^1$, where

$$\langle u, v \rangle := \int_\Omega u(x)\ v(x)\ dx \tag{2.25}$$

$$[u, v] := u(x_L)\ v(x_L) + u(x_R)\ v(x_R) \tag{2.26}$$

and $\Omega = (x_L, x_R)$. In addition we will use the notation $|u| := \sqrt{\langle u, u \rangle}$ in the following.

The crucial point in the proof of existence of a solution is the following a-priori estimate for sufficiently smooth solutions

Lemma 2.6. *Let $u, v \in H^{1,1}(Q)$, be a solution of* (2.20)-(2.23), *then*

$$\|u\|^2_{L^\infty(I,L^2(\Omega))} + \|\sqrt{a}u\|^2_{L^2(I,L^2(\partial\Omega))} + \|v\|^2_{L^\infty(I,L^2(\Omega))} \leq c(\|f\|^2_{L^2(Q)} + \|g\|^2_{L^2(Q)}) \tag{2.27}$$

Proof. Since u and v are sufficiently regular we may choose $\phi = u$ and $\psi = v$. Integration by parts in (2.24) and addition of (2.23) yields

$$\langle u', u\rangle + \langle v', v\rangle + \langle a_x u, v\rangle + [au, un] \;=\; \langle f, u\rangle + \langle g, v\rangle$$

Integrating on t we deduce

$$\begin{aligned}
\frac{1}{2}\left(|u|^2 + |v|^2\right) + \int_0^t \left(\langle a_x u, v\rangle + [au, un]\right)\, d\tau \;&=\; \int_0^t \left(\langle f, u\rangle + \langle g, v\rangle\right)\, d\tau \\
&\le\; \frac{1}{2}\int_0^{t_*} \left(|g|^2 + |u|^2 + |h|^2 + |v|^2\right)\, d\tau \\
&\le\; \frac{1}{2}(\|g\|^2_{L^2(Q)} + \|h\|^2_{L^2(Q)}) + \\
&\quad\; \frac{t_*}{2}(\|u\|^2_{L^\infty(I,L^2(\Omega))} + \|v\|^2_{L^\infty(I,L^2(\Omega))})
\end{aligned}$$

On the left hand side we may estimate

$$\begin{aligned}
\int_0^t \langle a_x u, v\rangle\, d\tau \;&\ge\; -\frac{1}{2}\int_0^{t_*} \|a_x\|_\infty \left(|u|^2 + |v|^2\right)\, dt \\
&\ge\; -\frac{t_*}{2}\|a_x\|_\infty \left(\|u\|^2_{L^\infty(I,L^2(\Omega))} + \|v\|^2_{L^\infty(I,L^2(\Omega))}\right)
\end{aligned}$$

Hence, we conclude

$$\begin{aligned}
(1 - (\|a_x\|_\infty + 1)t_*)(\|u\|^2_{L^\infty(I,L^2(\Omega))} + \|v\|^2_{L^\infty(I,L^2(\Omega))}) + 2\|\sqrt{a}u\|^2_{L^2(I,L^2(\partial\Omega))} \\
\le \|g\|^2_{L^2(Q)} + \|h\|^2_{L^2(Q)}\,,
\end{aligned}$$

which immediately yields the desired estimate with $c = \frac{1}{1-(\|a_x\|_\infty+1)t_*}$. $\square$

The existence of a weak solution can now be shown by applying a standard Galerkin-method similary to [23, Chapter 6.8], which also implies that the stability estimate (2.27) holds for weak solutions, too. This immediately yields uniqueness of the solution due to the linearity of the problem. Summing up, we obtain:

Proposition 2.7. *Let $g, h \in L^2(Q)$. Then problem* (2.17) - (2.21) *admits a unique weak solution $(u, v) \in L^\infty(I, L^2(\Omega)) \times L^\infty(I, L^2(\Omega))$, which satisfies*

$$\|u\|^2_{L^\infty(I,L^2(\Omega))} + \|\sqrt{a}u\|^2_{L^2(I,L^2(\partial\Omega))} + \|v\|^2_{L^\infty(I,L^2(\Omega))} \le c(\|g\|^2_{L^2(Q)} + \|h\|^2_{L^2(Q)}) \qquad (2.28)$$

By $\mathcal{S}_2$ we will denote the solution operator of the linear hyperbolic problem (2.17) - (2.21), given by

$$\begin{array}{rccc}
\mathcal{S}_2: & L^2(Q) \times L^2(Q) & \to & L^\infty(I, L^2(\Omega)) \times L^\infty(I, L^2(\Omega)) \\
& (g, h) & \mapsto & (u, v).
\end{array} \qquad (2.29)$$

The nonlinearity arising in the hyperbolic system is the operator

$$\begin{array}{rccc} \mathcal{N}_2: & H^{2,1}(Q) & \rightarrow & L^2(Q) \\ & T & \mapsto & b(T)_t = b'(T)T_t. \end{array} \tag{2.30}$$

For its analysis we can use realistic a-priori information about the nucleation rate, which is a function of special shape (cf. [6])

Assumption 2.8. In the following we assume

$$b \in C_b^{1,1}(\mathbf{R}) := \left\{ f \in C^1(\mathbf{R}) \mid f' \text{ is Lipschitz continuous and bounded} \right\}, \tag{2.31}$$

$b = 0$ for $T \geq T_{max}$ and b constant for $T \leq 0$.

For the evaluation of $b'(T)$ we need the following embedding result:

Lemma 2.9. *Let $\Omega \subset \mathbf{R}$ be bounded and open, then $H^{2,1}(Q) \hookrightarrow C(Q)$. The norm of the embedding operator will be denoted by κ.*

Proof. From [23, p.19 and p.43] we conclude $H^{2,1}(Q) \hookrightarrow C(I, H^1(\Omega))$. Together with $H^1(\Omega) \hookrightarrow C(\Omega)$, this yields the assertion. □

The operator $\mathcal{N}_2$ is not globally Lipschitz-continuous, but for the sake of constructing a fixed point equation it suffices to consider only a bounded subset, defined by

$$\mathcal{M} := \left\{ T \in H^{2,1}(Q) \mid \|T\|_{H^{2,1}(Q)} \leq M \right\}. \tag{2.32}$$

Lemma 2.10. *$\mathcal{N}_2$ is a Lipschitz-continuous operator from $\mathcal{M} \subset H^{2,1}$ to $L^2(Q)$, satisfying $\mathcal{N}_2(0) = 0$.*

Proof. The L^2-norm of $\mathcal{N}_2$ may be estimated by $\|b'\|_\infty \|T\|_{H^{2,1}(Q)}$, the Lipschitz-continuity follows from

$$\begin{aligned} \left\|\mathcal{N}_2(T) - \mathcal{N}_2(\overline{T})\right\|^2_{L^2(Q)} &\leq 2\left\|(b'(T) - b'(\overline{T}))T_t\right\|^2_{L^2(Q)} + 2\left\|b'(\overline{T})(T_t - \overline{T}_t)\right\|^2_{L^2(Q)} \\ &\leq 2\left\|b'(T) - b'(\overline{T})\right\|^2_{L^\infty(Q)} \|T_t\|^2_{L^2(Q)} + 2\|b'\|^2_{C^0(\mathbf{R})} \left\|T_t - \overline{T}_t\right\|^2_{L^2(Q)} \\ &\leq 2\|b\|^2_{C^{1,1}} \left\|T - \overline{T}\right\|^2_{L^\infty(Q)} \|T_t\|^2_{L^2(Q)} + 2\|b\|^2_{C^{1,1}(\mathbf{R})} \left\|T_t - \overline{T}_t\right\|^2_{L^2(Q)} \\ &\leq 2\|b\|^2_{C^{1,1}(\mathbf{R})} (M\kappa + 1) \left\|T - \overline{T}\right\|^2_{H^{2,1}(Q)}. \end{aligned}$$

Finally, $\mathcal{N}_2(0) = 0$ follows obviously from $b'(0) = 0$. □

Combining the results about the linear part and the nonlinearity $\mathcal{N}_2$ we obtain:

Proposition 2.11. *Let $(u, v) = \mathcal{S}_2(\mathcal{N}_2(T), 0)$ and $(\overline{u}, \overline{v}) = \mathcal{S}_2(\mathcal{N}_2(\overline{T}), 0)$ with $T, \overline{T} \in \mathcal{M}$, then*

$$\|u - \overline{u}\|^2_{L^\infty(I,L^2(\Omega))} + \|v - \overline{v}\|^2_{L^\infty(I,L^2(\Omega))} \leq 2c\|b\|^2_{C^{1,1}} (M\kappa + 1) \left\|T - \overline{T}\right\|^2_{H^{2,1}(Q)} \tag{2.33}$$

$$\|u\|^2_{L^\infty(I,L^2(\Omega))} + \|v\|^2_{L^\infty(I,L^2(\Omega))} \leq 2c\|b\|^2_{C^{1,1}} (M\kappa + 1)M^2 \tag{2.34}$$

Proof. The first inequality (2.33) follows immediately by combination of Proposition 2.7 and Lemma 2.10. The second estimate (2.34) is a direct consequence of (2.33) with $\overline{T} = 0$, since $\mathcal{S}_2(\mathcal{N}_2(0)) = \mathcal{S}_2(0) = 0$. □

2.3 Well-posedness of the direct problem

Using the preliminary results about its parabolic and hyperbolic part we are now able to investigate the complete nonlinear problem (1.14)-(1.21) as a fixed point equation. With the above notations we may write

$$T = \mathcal{S}_1(\mathcal{N}_1(u,v)) \tag{2.35}$$

$$(u,v) = \mathcal{S}_2(\mathcal{N}_2(T),0), \tag{2.36}$$

which we may also consider as a fixed point equation for T, since

$$T = \mathcal{P}(T) := \mathcal{S}_1(\mathcal{N}_1(\mathcal{S}_2(\mathcal{N}_2(T),0))). \tag{2.37}$$

All the assumptions on the parameter and the maximal length of the time interval needed are summarized in

Assumption 2.12. In the following we always choose

$$M := 2c_0\left(\left\|T^0\right\|_{H^1(\Omega)} + \left\|T^1\right\|_{H^{\frac{1}{2},\frac{1}{4}}(\partial\Omega\times I)}\right) \tag{2.38}$$

with t_* such that

$$2c_0c_1(t_*,C)C \le M \text{ with } C = \sqrt{2(M\kappa+1)}\,\|b\|_{C^{1,1}}\,M. \tag{2.39}$$

Furthermore, let the Assumptions 2.5 and 2.8 be satisfied.

We note that besides the smoothness required for the coefficients a and b, Assumption 2.12 holds for small final time t_*. It seems that in a practical setup the bound upon t_* is still large enough for a typical process. If a result for a longer time interval is desired one may apply a similar technique successively to the intervals $(t_*,2t_*)$, $(2t_*,3t_*)$, which is avoided here for the sake of simplicity.

Lemma 2.13. *The operator $\mathcal{P}$ is contractive and maps $\mathcal{M}$ to itself.*

Proof. Combining the Propositions 2.4 and 2.11 we obtain

$$\|\mathcal{P}(T)\|_{H^{2,1}(Q)} \le \frac{M}{2} + c_0c_1(t_*,C)C \le \frac{M}{2} + \frac{M}{2} = M$$

and

$$\left\|\mathcal{P}(T) - \mathcal{P}(\overline{T})\right\|_{H^{2,1}(Q)} \le c_0c_1(t_*,C)\frac{C}{M}\left\|T-\overline{T}\right\|_{H^{2,1}(Q)} \le \frac{1}{2}\left\|T-\overline{T}\right\|_{H^{2,1}(Q)}.$$

□

Now we are in position to prove our main result:

Theorem 2.14. *Problem* (1.14)-(1.21) *admits a unique solution*

$$(u,v,T) \in L^\infty(I,L^2(\Omega)) \times L^\infty(I,L^2(\Omega)) \times \mathcal{M}.$$

Proof. We have seen above that solving (1.14)-(1.21) for T is equivalent to solving the fixed-point equation (2.37). In Lemma (2.13) we have shown that $\mathcal{P}$ maps $\mathcal{M}$ into itself and is contractive, hence Banach's fixed point theorem (cf. [12, Thm.7.1]) implies the existence and uniqueness of a solution $T \in \mathcal{M}$. For known T, the pair (u,v) is given uniquely by (2.36). □

3 The inverse problem

The identification of the nucleation rate b in (1.14)-(1.21) can be interpreted as the solution of the nonlinear operator equation

$$F(b) = y^\delta := (T_\Gamma^\delta, \xi_*^\delta), \tag{3.1}$$

where the noisy data are perturbations of

$$\begin{aligned} T_\Gamma &= T|_{\Gamma \times I}, \qquad \Gamma \subset \partial\Omega & (3.2)\\ \xi_* &= \xi(t_*) = 1 - e^{-\int_0^{t_*}(au)ds}. & (3.3) \end{aligned}$$

F denotes the implicitely defined operator

$$\begin{array}{rcl} F: \ \mathcal{D}(F) \subset H^\eta([z_1, z_2]) & \to & L^2(I \times \Gamma) \times L^2(\Omega) \\ b & \mapsto & \left(T_b|_{\Gamma \times I}, 1 - e^{-\int_0^{t_*}(au_b)ds}\right), \end{array} \tag{3.4}$$

where $H^\eta([z_1, z_2])$ is a Sobolev space of order η on the sufficiently large temperature interval $[z_1, z_2]$ and (u_b, v_b, T_b) denotes the solution of (1.14)-(1.21) for the particular parameter b. Theorem 2.14 guarantees existence and uniqueness of (u_b, v_b, T_b), in Section 4 we will also show that the trace-type map

$$(u_b, v_b, T_b) \mapsto (T_b|_{\Gamma \times I}, 1 - e^{-\int_0^{t_*}(au_b)ds})$$

is well-defined and thus F is well-defined, too. In practice, it is not possible to measure T_Γ and ξ_* exactly, but only perturbed data T_Γ^δ and ξ_*^δ with noise bound

$$\left\|y - y^\delta\right\| = \left\|T_\Gamma - T_\Gamma^\delta\right\|_{L^2(\Gamma \times I)} + \left\|\xi_* - \xi_*^\delta\right\|_{L^2(\Omega)} \leq \delta. \tag{3.5}$$

Since the problem (3.1) is most likely ill-posed (and numerical results confirm instabilities), we employ regularization methods to obtain stable approximations of the solution. The most famous direct method is *Tikhonov regularization*, which would consist in minimizing the functional

$$J_\alpha^\delta(b) = \left\|F(x) - y^\delta\right\|^2 + \alpha \left\|b - b^*\right\|^2, \tag{3.6}$$

where b^* represents an a-priori guess of the solution. Stability, convergence and convergence rates for this method as $\delta \to 0$ applied to nonlinear problems have been shown by Seidman et al. [28] and Engl et al. [15]. One obtains convergence to the solution $b^\dagger$ of minimal distance to b^* as $\delta \to 0$ for appropriate choice of $\alpha = \alpha(\delta, y^\delta)$. In practice, the minimization of J_α^δ is not a simple task, since it can have many local minima, which stop any minimization procedure.

A natural alternative are *iterative regularization methods* (since iterative methods are usually employed for the minimization problem arising in Tikhonov regularization anyway), where the main regularizing effect comes from an appropriate early termination of the iteration procedure. A popular method of choosing the stopping index $k_* = k(\delta, y^\delta)$, which is easy to implement, is the so-called *generalized discrepancy principle*:

$$\left\| F(b^\delta_{k_*}) - y^\delta \right\| \le \tau\delta < \left\| F(b^\delta_k) - y^\delta \right\|, \qquad 0 \le k < k_*, \tag{3.7}$$

with appropriately chosen $\tau > 1$. The basic idea behind this principle is that due to noise in the data one should not distinguish between solutions which yield a residual less than δ. Thus, the iteration procedure can be stopped at the first time, where the residual is of the order of the noise level.

Since all common iterative methods use derivatives of F, continuity and Frèchet-differentiability of F is a fundamental requirement for their well-definedness and convergence. In the following we sum up some of the most important methods (see [16] for a detailed overview):

- **The Landweber iteration** is an explicit method defined by

  $$b^\delta_{k+1} = b^\delta_k - \omega F'(b^\delta_k)^*(F(b^\delta_k) - y^\delta), \tag{3.8}$$

  where ω is an appropriate damping factor that has to satisfy $\omega \left\| F'(b) \right\| < 1$ in a sufficiently large neigbourhood of the starting value. Convergence for nonlinear problems has been shown by Hanke et al. [18] (see also [3, 26, 27]) under the conditions

  $$b^\dagger \in B_\rho(b_0) \tag{3.9}$$
  $$\left\| F(\bar{b}) - F(b) - F'(b)(\bar{b} - b) \right\| \le \eta \left\| F(\bar{b}) - F(b) \right\|, \qquad b, \bar{b} \in B_\rho(b_0) \tag{3.10}$$
  $$\tau > 2\tfrac{1+\eta}{1-2\eta} > 2. \tag{3.11}$$

  As for any other regularization method, convergence rates can only be shown under additional smoothness assumptions, so-called *source conditions.* In this case the source condition

  $$b^\dagger - b_0 = F'(b^\dagger)w, \tag{3.12}$$

  where $b^\dagger$ is the solution of minimal distance to b_0, and a slightly stronger condition than (3.10) imply

  $$\left\| b^\delta_{k_*} - b^\dagger \right\| = \mathcal{O}(\sqrt{\delta}) \tag{3.13}$$
  $$k_* = k_*(\delta, y^\delta) = \mathcal{O}(\delta^{-1}). \tag{3.14}$$

- **The Levenberg-Marquardt method** (LM) is the implicit iteration

  $$b^\delta_{k+1} = b^\delta_k - (F'(b^\delta_k)^* F'(b^\delta_k) + \alpha_k I)^{-1} F'(b^\delta_k)^*(F(b^\delta_k) - y^\delta), \tag{3.15}$$

  which is equivalent to computing b^δ_{k+1} as the minimizer of

  $$\left\| F(b^\delta_k) - y^\delta + F'(b^\delta_k)(b - b^\delta_k) \right\|^2 + \alpha_k \left\| (b - b^\delta_k) \right\|^2 = \min_b . \tag{3.16}$$

Convergence of this method has been shown by Hanke [17] under the additional conditions

$$b^\dagger \in B_\rho(b_0) \tag{3.17}$$
$$\left\|F(\bar{b}) - F(b) - F'(b)(\bar{b} - b)\right\| \le C \left\|\bar{b} - b\right\| \left\|F(\bar{b}) - F(b)\right\|, b, \bar{b} \in B_\rho(b_0) \tag{3.18}$$
$$0 < \rho < \tfrac{1}{C}, \qquad \tau > \tfrac{C}{\rho}. \tag{3.19}$$

So far, no result about convergence rates is available.

- **The iteratively regularized Gauss-Newton Method** (IRGN) is defined by

$$b_{k+1}^\delta = b_k^\delta - (F'(b_k^\delta)^* F'(b_k^\delta) + \alpha_k I)^{-1}(F'(b_k^\delta)^*(F(b_k^\delta) - y^\delta) + \alpha_k(b_k - b_0)), \tag{3.20}$$

with the variational characterization of b_{k+1}^δ as the minimizer of

$$\left\|F(b_k^\delta) - y^\delta + F'(b_k^\delta)(b - b_k^\delta)\right\|^2 + \alpha_k \left\|(b - b_0)\right\|^2 = \min_b. \tag{3.21}$$

The only difference to the Levenberg-Marquardt iteration is the choice of the stabilizer, now the prior b_0 is kept fixed during the iteration, which improves the stability properties during the iteration. This method was first proposed by Bakushinskii [1], convergence and convergence rates for several choices of the stopping index has been shown by Kaltenbacher et al. [5, 20, 21] under assumptions on the nonlinearity of F similar to (3.10) and (3.18). If a source condition of the form

$$b^\dagger - b_0 = (F'^*(b^\dagger)F'(b^\dagger))^\nu w, \tag{3.22}$$

with $\frac{1}{2} \le \nu < 1$ is satisfied, local Lipschitz continuity of F' suffices to prove the convergence rate

$$\left\|b^\dagger - b_{k_*}^\delta\right\| = \mathcal{O}(\delta^{\frac{2\nu}{2\nu+1}}), \tag{3.23}$$

respectively

$$\left\|b^\dagger - b_k\right\| = o(\alpha_n^\nu)$$
$$\left\|y - F(b_k)\right\| = o(\alpha_n^{\nu+\frac{1}{2}})$$

for the noise-free case.

- **Broyden's Method** for the noise-free case reads,

$$b_{k+1} = b_k - B_k^\dagger(F(b_k) - y) \tag{3.24}$$
$$B_{k+1} = B_k + \frac{\langle s_k, . \rangle}{\left\|s_k\right\|^2}(y - F(b_k)), \tag{3.25}$$

where $s_k = b_{k+1} - b_k$. In the noisy case, it seems advisable to replace the pseudo-inverse $B_k^\dagger$ by a stable approximation, e.g. analogously to the iteratively regularized Gauss-Newton Method. Convergence and convergence rates for Broyden's Method applied to inverse problems have been shown by Kaltenbacher [22].

- **Frozen Newton-type Methods** are variants of the Levenberg-Marquardt method, the iteratively regularized Gauss-Newton Method or any other Newton-type method, based on keeping the Newton-matrix fixed during several iteration steps. For the iteratively regularized Gauss-Newton Method, the frozen version reads

$$b_{k+1}^\delta = b_k^\delta - (F'(b_0)^* F'(b_0) + \alpha_k I)^{-1} (F'(b_0)^* (F(b_k^\delta) - y^\delta) + \alpha_k (b_k^\delta - b_0)) \tag{3.26}$$

  This method has been proposed by Kaltenbacher [4] and has similar properties as the non-frozen version of the iteration procedure.

We finally note that most of the above methods are affinely invariant and can also be analyzed under corresponding conditions (cf. [11]).

4 The parameter-to-output map

In the following we examine the continuity properties of the parameter-to-output map F. We first note that by using the linear trace operator

$$\begin{array}{rccc} \Lambda_1: & \mathcal{D}(\mathcal{N}_2) \times H^{2,1}(Q) & \rightarrow & L^2(I \times \Gamma) \\ & (u, v, T) & \mapsto & T|_\Gamma \end{array} \tag{4.1}$$

and the nonlinear operator

$$\begin{array}{rccc} \Lambda_2: & \mathcal{D}(\mathcal{N}_2) \times H^{2,1}(Q) & \rightarrow & L^2(\Omega) \\ & (u, v, T) & \mapsto & \xi|_{t=t_*} = 1 - e^{-\int_0^{t_*} (au) dt}, \end{array} \tag{4.2}$$

we may write

$$F = (\Lambda_1, \Lambda_2) \circ R, \tag{4.3}$$

where R is the *parameter-to-solution map*

$$\begin{array}{rccc} R: & \mathcal{D}(F) \subset H^\eta([z_1, z_2]) & \rightarrow & \mathcal{D}(\mathcal{N}_2) \times H^{2,1}(Q) \subset L^\infty(I, L^2(\Omega))^2 \times \times H^{2,1}(Q) \\ & b & \mapsto & (u, v, T). \end{array} \tag{4.4}$$

Since Λ_1 is a continuous linear operator (c.f. e.g. [24]), it is also Frèchet-differentiable with Lipschitz-continuous derivative. Thus, it remains to investigate Λ_2 and R, we start with Λ_2.

Proposition 4.1. *The operator Λ_2 is continuous and Frèchet-differentiable with Lipschitz-continuous derivative*

$$\Lambda_2'(u, v, T)(\varphi, \psi, \theta) = e^{-\int_0^{t_*} (au) dt} \int_0^{t_*} (a\varphi) dt \tag{4.5}$$

Proof. Similar to Lemma 2.2 one can show that the function $w := \int_0^t (au)\, ds$ satisfies $w \in C(I, L^\infty(\Omega))$ with norm depending continuously upon u and v. Hence,

$$\Lambda_2(u, v, T) = 1 - e^{-w|_{t=t_*}}$$

depends continuously on (u, v, T) in $L^\infty(\Omega)$ and consequently in the weaker norm of $L^2(\Omega)$, too. The Frèchet-differentiability of Λ_2 and Lipschitz-continuity of Λ_2' follows from a standard argument using the smoothness of the exponential function, which is the only nonlinear term contained in Λ_2. □

Now we turn to the analysis of the implicitely defined operator R. For this reason we write $\mathcal{G}(T, b)$ instead of $\mathcal{N}_2(T)$, more precisely we define

$$\begin{array}{rccc} \mathcal{G}: & H^{2,1}(Q) \times C^{1,1}(Q) & \to & L^2(Q) \\ & (T, b) & \mapsto & b(T)_t = b'(T)T_t. \end{array} \tag{4.6}$$

Similary we write $\mathcal{Q}(T, b)$ for $\mathcal{P}(T)$ with particular parameter b.

As a consequence of the fixed point technique we have used for proving existence and uniqueness we may conclude almost without further effort the following statement about the stable dependence upon b:

Theorem 4.2. *Let t_* be such that Assumption 2.12 is satisfied for $b \in \mathcal{M}$, then the solutions $T = \mathcal{Q}(T, b)$ and $\overline{T} = \mathcal{Q}(\overline{T}, \overline{b})$ satisfy an estimate of the form*

$$\left\|T - \overline{T}\right\|_{H^{2,1}(Q)} \leq \gamma M \|b - \overline{b}\|_{C^1(Q)}, \tag{4.7}$$

for all $b, \overline{b} \in \mathcal{M}$, where γ is a positive constant.

Proof. We may estimate

$$\left\|\mathcal{Q}(T, b) - \mathcal{Q}(\overline{T}, \overline{b})\right\|_{H^{2,1}(Q)} \leq \left\|\mathcal{Q}(T, b) - \mathcal{Q}(\overline{T}, b)\right\|_{H^{2,1}(Q)} + \left\|\mathcal{Q}(\overline{T}, b) - \mathcal{Q}(\overline{T}, \overline{b})\right\|_{H^{2,1}(Q)}$$

Under the above assumptions Lemma 2.13 implies that $\mathcal{Q}(., b)$ is Lipschitz-continuous with module less or equal $\frac{1}{2}$, hence

$$\begin{aligned} \left\|T - \overline{T}\right\|_{H^{2,1}(Q)} &= \left\|\mathcal{Q}(T, b) - \mathcal{Q}(\overline{T}, \overline{b})\right\|_{H^{2,1}(Q)} \\ &\leq 2\left\|\mathcal{Q}(\overline{T}, b) - \mathcal{Q}(\overline{T}, \overline{b})\right\|_{H^{2,1}(Q)} \end{aligned}$$

Since $\mathcal{Q} = \mathcal{S}_1 \circ \mathcal{N}_1 \circ \mathcal{S}_2 \circ \mathcal{G}$ and because of the Lipschitz-continuity of the operators $\mathcal{S}_1$, $\mathcal{N}_1$ and $\mathcal{S}_2$, there exists a constant γ such that

$$\begin{aligned} \left\|T - \overline{T}\right\|_{H^{2,1}(Q)} &\leq \gamma \|\mathcal{G}(\overline{T}, b) - \mathcal{G}(\overline{T}, \overline{b})\|_{H^{2,1}(Q)} \\ &= \gamma \|(b'(\overline{T}) - \overline{b}'(\overline{T}))\overline{T}_t\|_{H^{2,1}(Q)} \end{aligned}$$

and since $\overline{T} \in \mathcal{M}$ we may further estimate

$$\left\|T - \overline{T}\right\|_{H^{2,1}(Q)} \leq \gamma M \|b - \overline{b}\|_{C^1(Q)}$$

□

Finally, we verify the Fréchet-differentiability of the operator $\mathcal{Q}$ with respect to the parameter b; therefore we have to assume $b \in C_b^2(\mathbf{R})$ in the following.

Lemma 4.3. *The operator $\mathcal{G}$ is Frèchet-differentiable with partial derivatives*

$$\begin{aligned} \mathcal{G}_T(T,b)\theta &= b'(T)\theta_t + b''(T)\theta T_t && (4.8)\\ \mathcal{G}_b(T,b)h &= h'(T)T_t. && (4.9) \end{aligned}$$

If furthermore $b \in C^{2,1}(\mathbf{R})$, $\mathcal{G}'$ is Lipschitz-continuous.

Proof. A straight-forward estimate yields:

$$\begin{aligned} &\|\mathcal{G}(T+\theta,b+h) - \mathcal{G}(T,b) - \mathcal{G}_T(T,b)\theta - \mathcal{G}_b(T,b)h\| \\ &\quad \le \|\mathcal{G}(T+\theta,b+h) - \mathcal{G}(T+\theta,b) - \mathcal{G}_b(T,b)h\| + \|\mathcal{G}(T+\theta,b) - \mathcal{G}(T,b) - \mathcal{G}_T(T,b)\theta\| \\ &\quad \le \|\mathcal{G}(T+\theta,b+h) - \mathcal{G}(T+\theta,b) - \mathcal{G}_b(T+\theta,b)h\| + \\ &\qquad \|\mathcal{G}_b(T+\theta,b)h - \mathcal{G}_b(T,b)h\| + \|\mathcal{G}(T+\theta,b) - \mathcal{G}(T,b) - \mathcal{G}_T(T,b)\theta\|. \end{aligned}$$

Since for fixed T, $\mathcal{G}$ is linear and continuous in b, the partial derivative is obviously given by (4.9) and we may estimate

$$\|\mathcal{G}(T+\theta,b+h) - \mathcal{G}(T+\theta,b) - \mathcal{G}_b(T+\theta,b)h\|_{L^2(Q)} = o(\|h\|_{C^2(\mathbf{R})}).$$

The second term is bounded by

$$\begin{aligned} &\|\mathcal{G}_b(T+\theta,b)h - \mathcal{G}_b(T,b)h\|_{L^2(Q)} \\ &\quad \le \|(h'(T+\theta) - h'(T))T_t\|_{L^2(Q)} + \|h'(T+\theta)\theta_t\|_{L^2(Q)} \\ &\quad \le \kappa \|h\|_{C^2(\mathbf{R})} \|\theta\|_{H^{2,1}(Q)} \|T\|_{H^{2,1}(Q)} + \|h\|_{C^2(\mathbf{R})} \|\theta\|_{H^{2,1}(Q)} \\ &\quad = o(\|\theta\|_{H^{2,1}(Q)} + \|h\|_{C^2(\mathbf{R})}). \end{aligned}$$

Because of $b \in C_b^2(\mathbf{R})$, for any $\epsilon > 0$ there exists a $\delta > 0$ such that for $\|\theta\|_{H^{2,1}(Q)} \le \delta$

$$\begin{aligned} &\|\mathcal{G}(T+\theta,b) - \mathcal{G}(T,b) - \mathcal{G}_T(T,b)\theta\|_{L^2(Q)} = \\ &\quad \|(b'(T+\theta) - b'(T))(T_t+\theta_t) - b''(T)T_t\|_{L^2(Q)} \\ &\quad \le \|(b'(T+\theta) - b'(T) - b''(T)\theta)\|_{L^\infty(Q)} \|T_t\|_{L^2(Q)} \\ &\quad + \|(b'(T+\theta) - b'(T))\theta_t\|_{L^2(Q)} \\ &\quad \le \epsilon \|\theta\|_{H^{2,1}(Q)} \|T\|_{H^{2,1}(Q)} + \kappa \|b\|_{C^2(\mathbf{R})} \|\theta\|^2_{H^{2,1}(Q)} \\ &\quad = o(\|\theta\|_{H^{2,1}(Q)}). \end{aligned}$$

Thus, since obviously $\mathcal{G}_T$ and $\mathcal{G}_b$ are continuous linear operators, $\mathcal{G}$ is F-differentiable.

The Lipschitz-continuity of $\mathcal{G}'$ under the additional assumption $b \in C^{2,1}(\mathbf{R})$ follows from a straight-forward estimate. □

In a similar way we analyze the second nonlinear operator involved:

Lemma 4.4. *The operator $\mathcal{N}_1$ is F-differentiable with derivative*

$$\mathcal{N}_1'(u,v)(\varphi,\psi) = e^{-\int_0^t auds} a\varphi - e^{-\int_0^t auds} au \int_0^t a\varphi ds, \tag{4.10}$$

which is also locally Lipschitz-continuous.

Proof. The triangle inequality yields

$$\begin{aligned}
&\|\mathcal{N}_1(u+\varphi, v+\psi) - \mathcal{N}_1(u,v) - \mathcal{N}_1'(u,v)(\varphi,\psi)\|_{L^2(Q)} \\
&\leq \left\| e^{-\int_0^t auds} \left(e^{-\int_0^t a\varphi ds} - 1 - \int_0^t a\varphi ds \right) a(u+\varphi) \right\|_{L^2(Q)} + \\
&\quad \left\| e^{-\int_0^t auds} a\varphi \int_0^t a\varphi ds \right\|_{L^2(Q)} \\
&\leq \left\| ae^{-\int_0^t auds} \right\|_{L^\infty(Q)} \left\| e^{-\int_0^t a\varphi ds} - 1 - \int_0^t a\varphi ds \right\|_{L^\infty(Q)} \|u+\varphi\|_{L^2(Q)} + \\
&\quad \left\| ae^{-\int_0^t auds} \right\|_{L^\infty(Q)} \|\varphi\|_{L^2(Q)} \left\| \int_0^t a\varphi ds \right\| - L^\infty(Q).
\end{aligned}$$

Because of $\int_0^t a\varphi ds \in L^\infty(Q)$ and the continuous differentiability of the exponential function, we may estimate

$$\begin{aligned}
\left\| e^{-\int_0^t a\varphi ds} - 1 - \int_0^t a\varphi ds \right\|_{L^\infty(Q)} &= \mathcal{O}\left(\left\| \int_0^t a\varphi ds \right\|^2_{L^\infty(Q)} \right) \\
&= \mathcal{O}\left((\|\varphi\|_{L^\infty(I,L^2(\Omega))} + \|\psi\|_{L^\infty(I,L^2(\Omega))})^2 \right),
\end{aligned}$$

which finally implies the differentiability of $\mathcal{N}_1$. The local Lipschitz-continuity of $\mathcal{N}_1'$ can be shown by a straight-forward estimate. $\square$

Corollary 4.5. *The operator $\mathcal{Q}$ is Frèchet-differentiable with partial derivatives*

$$\begin{aligned}
\mathcal{Q}_T(T,b)\theta &= \mathcal{S}_1(\mathcal{N}_1'(\mathcal{S}_2(\mathcal{G}(T,h)))\mathcal{S}_2(\mathcal{G}_T(T,h)\theta)) \qquad (4.11)\\
\mathcal{Q}_b(T,b)h &= \mathcal{S}_1(\mathcal{N}_1'(\mathcal{S}_2(\mathcal{G}(T,h)))\mathcal{S}_2(\mathcal{G}_b(T,h)h)). \qquad (4.12)
\end{aligned}$$

Furthermore, $\mathcal{Q}_T$ is contractive and, if $b \in C^{2,1}(\mathbf{R})$, $\mathcal{Q}_b$ is local Lipschitz-continuous.

Proof. We have shown above that the operators $\mathcal{N}_1$ and $\mathcal{G}$ are differentiable, and since $\mathcal{S}_1$ and $\mathcal{S}_2$ are continuous linear operators, they are differentiable, too. Thus, we obtain (4.11) and (4.12) by straight-forward application of the chain rule.

The contractivity of $\mathcal{Q}_T$ follows from the fact that the norm of the derivative is bounded by the Lipschitz-constant of the operator, which is less than one due to Lemma 2.13. $\square$

Theorem 4.6. *Let $T(b)$ denote the solution of the fixed point equation for given h, then the map $b \mapsto T(b)$ is Frèchet-differentiable with derivative*

$$T'(b)h = (I - \mathcal{Q}_T)^{-1}\mathcal{Q}_b(T,b)h. \tag{4.13}$$

Furthermore, the map $h \mapsto T'(b)h$ is locally Lipschitz-continuous if $b \in C^{2,1}(\mathbf{R})$.

Proof. The regularity of $I - \mathcal{Q}_T$ is a direct consequence of Corollary 4.5 and again Banach's fixed point theorem, and the implicit functions theorem [12, Thm.15.1] implies (4.13). Local Lipschitz-continuity of $h \mapsto T'(b)h$ follows from the contractivity of $I - \mathcal{Q}_T$ and the local Lipschitz-continuity of $\mathcal{Q}_b$. □

Corollary 4.7. *The operator R is Frèchet-differentiable and R' is locally Lipschitz-continuous if $b \in C^{2,1}(\mathbf{R})$.*

Proof. From Theorem 4.6 we know that the assertions hold for the map from $b \mapsto T(b)$. Because of (2.36) and the differentiability of $\mathcal{S}_2$ and $\mathcal{G}$ the same holds for (u,v), and thus for the whole operator R. □

Combining all results about R, Λ_1 and Λ_2 we obtain the following result about the parameter-to-output map:

Theorem 4.8. *Let $\eta > \frac{5}{2}$, then the parameter-to-output map F is Frèchet-differentiable. The derivative F' is locally Lipschitz-continous if $\eta > \frac{7}{2}$.*

Proof. A standard embedding result (cf. [23]) yields

$$\begin{aligned} H^\eta([z_1,z_2]) &\hookrightarrow C^2([z_1,z_2]) && \text{if } \eta > \frac{5}{2} \\ H^\eta([z_1,z_2]) &\hookrightarrow C^{2,1}([z_1,z_2]) && \text{if } \eta > \frac{7}{2}. \end{aligned}$$

Hence, the assertion follows from Proposition 4.1 and Corollary 4.7 together with

$$F'(b)h = (\Lambda_1(R'(b)h), \Lambda_2'(R(b))R'(b)h). \tag{4.14}$$

□

5 Iterative Regularization of the Inverse Problem

In this section we sketch the development of numerical algorithms for the iterative regularization methods, which have been stated for abstract operator equations in Section 3, when applied to the operator F defined by (3.4).

In any case we have to evaluate the adjoint of F', when using Landweber-iteration, and F' itself, when using a Newton-type method. Any of these methods uses the residual $y^\delta - F(b_k^\delta)$ and thus, also the nonlinear operator F has to be evaluated. From (4.3) we observe that we can evalute the explicitely defined operators Λ_1 and Λ_2 after having computed $R(b_k^\delta)$. The latter is realized by solving the direct problem (1.14)-(1.21) for $b = b_k^\delta$, i.e., we have to solve a nonlinear initial-boundary value problem at each iteration step.

For the evaluation of F' we use the operator splitting introduced in (4.14). Again we can first compute $R'(b_k^\delta)h$ and then evaluate Λ_1 and Λ_2'. In Section (4) we have shown that the derivative R' exists and we have computed it in an abstract way. For the sake of its numerical computation we will need a more concrete formulation, which can be derived by linearizing the direct problem around a solution (u, v, T) of (1.14)-(1.21). A straightforward calculation shows that the directional derivative $(\phi, \psi, \theta) = R'(b)h$ is determined as the solution of the linear initial-boundary value problem

$$\theta_t = (D\theta_x)_x + Le^{-\int_0^t (au)ds} a\left(\phi - u\int_0^t (a\phi)ds\right) \qquad \text{in } \Omega \times I \tag{5.1}$$

$$\phi_t = (a\psi)_x + (b'(T)\theta)_t + h(T)_t \qquad \text{in } \Omega \times I \tag{5.2}$$

$$\psi_t = (a\phi)_x \qquad \text{in } \Omega \times I \tag{5.3}$$

$$\theta_n = \alpha\theta \qquad \text{on } \partial\Omega \times I \tag{5.4}$$

$$\phi + \langle \psi, n\rangle = 0 \qquad \text{on } \partial\Omega \times I \tag{5.5}$$

$$\theta = 0 \qquad \text{in } \Omega \times \{0\} \tag{5.6}$$

$$\phi = 0 \qquad \text{in } \Omega \times \{0\} \tag{5.7}$$

$$\psi = 0 \qquad \text{in } \Omega \times \{0\}, \tag{5.8}$$

which has to be solved for given b and $(u, v, T) = R(b)$.

We note that the evaluation of the adjoint $F'(b)^*$ involves the solution of an initial-boundary value problem similar to (5.1)-(5.8); for further details we refer to [8], where $F'(b)^*$ has been computed and used for the realization of the Landweber iteration.

Finally, we turn to the numerical realization of iterative regularization methods restricting our attention to Newton-type methods (cf. [8] for a realization of the Landweber iteration). A numerical algorithm does not only involve a discretization of the input b and the output (T_Γ, ξ_*), but also a discretization of the initial-boundary value problems one has to solve for the evaluation of F and F'. Thus we start our considerations with an algorithm for the solutions of the systems (1.14)-(1.21) and (5.1)-(5.8). An obvious choice for the heat equation (1.14) is the classical Crank-Nicholson scheme, which is unconditionally stable, i.e., the semi-discretization (in time) is given by

$$T^{j+1} = T^j + (t_{j+1} - t_j)\left(D\left(\frac{T^{j+1} + T^j}{2}\right)_x\right)_x + \xi^{j+1} - \xi^j, \tag{5.9}$$

where T^j denotes the temperature at time $t = t_j$. Since ξ^{j+1} occurs on the right-hand side of (5.9) we aim at using an explicit time step for (1.15) and (1.16) to avoid the solution of a coupled nonlinear system. In the simplest case, the semi-discretization in time reads

$$u^{j+1} = u^j + (t_{j+1} - t_j)(av^j)_x + 2(b(T^j) - b(T^{j-1}) \tag{5.10}$$

$$v^{j+1} = v^j + (t_{j+1} - t_j)(av^j)_x, \tag{5.11}$$

for the spatial discretization Lax's method can be used, which imposes an acceptable stability constraint of the form

$$t_{j+1} - t_j < \frac{a\Delta x}{2}, \tag{5.12}$$

where Δx denotes the spatial step size. In order to be consistent with the second-order method in the heat equation an extension of the Lax-Wendroff scheme for hyperbolic problems with sources (cf. e.g. [25, 29]) can be used, which is a second-order method in time under the stability constraint (5.12). This provides an efficient method for the numerical solution of (1.14)-(1.21) and similary for (5.1)-(5.8). The trace-type operators Γ_1 and Γ_2 (as well as its derivative) can be realized in a straight-forward way by interpolation on a time- and spatial grid

$$x_L = x_1 < x_1 < \ldots < x_m = x_R \tag{5.13}$$

$$0 = t_1 < t_1 < \ldots < t_p = t_*, \tag{5.14}$$

using splines of order zero or one.

The remaining part in the discretization of the inverse problem is the discretization of the parameter b in the space $H^\eta([z_1, z_2])$, which can be performed using splines s_i of order η on an appropriate grid $z_1 = \zeta_1 < \zeta_2 < \ldots < \zeta_m = z_2$, i.e., we may write

$$b(z) = \sum_{i=1}^{m} \beta_i s_i(z), \tag{5.15}$$

and use the s_i as the basis in the discrete subspace. For the computation of the Newton-matrix $A_k = F'_{m,n,p}(b_k^\delta)$, which is the discretization of $F'(b_k^\delta)$ this means that we have to solve the system for each s_i $(i = 1, \ldots, n)$ and then compute the coefficients of the spline representation of $\Theta|_\Gamma$ and ξ_*. More precisely, under

$$T_\Gamma(t) = \sum_{j=1}^{p} c_i s_j^I(t) \tag{5.16}$$

$$\xi_*(x) = \sum_{j=1}^{m} d_i s_j^\Omega(t), \tag{5.17}$$

where s_j^I and s_j^Ω are the splines centered at t_j and x_j, respectively, A_k is the $n \times (p+m)$-matrix with entries

$$(A_k)_{ji} = \begin{cases} \langle \Gamma_1 R'_{n,p}(b_k)s_j, s_j^I \rangle & \text{if } j \leq p \\ \langle \Gamma_2'(R_{n,p}(b_k))R'_{n,p}(b_k)s_i, s_j^\Omega \rangle & \text{if } j > p \end{cases},$$

where $R_{n,p}$ and $R'_{n,p}$ denote the evaluation of R and R' by numerical solution of the corresponding initial-boundary value problems. This means that the allocation of A_k needs

n solutions of the initial-value problem (5.1)-(5.8), which is an enormous computational effort.

Once A_k is known, one can easily compute the discrete adjoint of $F'_{m,n,p}$ using A_k^T. Let Q denote the matrix

$$(Q)_{ij} = \langle s_i, s_j \rangle_\eta. \tag{5.18}$$

Then we have

$$\begin{aligned}
\langle F'_{m,n,p}(b_k^\delta)h, (T_\Gamma, \xi_*)\rangle &= (c^T, d^T)A_k\beta \\
&= ((c^T, d^T)A_kQ^{-1})Q\beta \\
&= \langle (c^T, d^T)A_kQ^{-1}(s_i(.))_{i=1,\dots,n}, h\rangle_\eta \\
&= \langle F'_{m,n,p}(b_k^\delta)^*(T_\Gamma, \xi_*), h\rangle_\eta,
\end{aligned}$$

and thus we may write

$$A_k^* := F'_{m,n,p}(b_k^\delta) = Q^{-1}A_k^T,$$

where A_k^* denotes the adjoint taken with respect to the discretized H^η-norm. Knowing A_k and A_k^* we can now easily perform a step of the Levenberg-Marquardt or of the iteratively regularized Newton-method by solving a linear system of dimension $n \times n$.

We finally note that the results of Section 4 imply local convergence of the iteratively regularized Gauss-Newton method for initial values b^0, which satisfy the source condition

$$b^\dagger - b^0 = F'(b^\dagger)^* w. \tag{5.19}$$

We recall that in this case all methods yield the convergence rate

$$\left\| b_{k_*}^\delta - b^\dagger \right\| = \mathcal{O}\left(\sqrt{\delta}\right), \tag{5.20}$$

if the stopping index is chosen according to (3.7) with appropriate τ.

The adjoint of the operator $F'(b)$ can be calculated as follows. First, we observe that the solution $(\tilde\theta, \tilde\phi, \tilde\psi)$ of the backward problem

$$\tilde\theta_t = -(D\tilde\theta_x)_x + b'(T)\tilde\phi_t \qquad \text{in } \Omega \times I \tag{5.21}$$

$$\begin{aligned}\tilde\phi_t = {} & (a\tilde\psi)_x - Le^{-\int_0^t (au)ds} a\tilde\theta \\ & + L\int_t^{t_*} au\tilde\theta e^{-\int_0^\tau (au)ds} d\tau + Lae^{-\int_0^t (au)ds} w_2 \end{aligned} \qquad \text{in } \Omega \times I \tag{5.22}$$

$$\tilde\psi_t = (a\tilde\phi)_x \qquad \text{in } \Omega \times I \tag{5.23}$$

$$\tilde\theta_n = \alpha\tilde\theta + \frac{1}{D}w_1 \qquad \text{on } \partial\Omega \times I \tag{5.24}$$

$$\tilde\phi - \langle \tilde\psi, n\rangle = 0 \qquad \text{on } \partial\Omega \times I \tag{5.25}$$

$$\tilde\theta = 0 \qquad \text{in } \Omega \times \{t_*\} \tag{5.26}$$

$$\tilde\phi = 0 \qquad \text{in } \Omega \times \{t_*\} \tag{5.27}$$

$$\tilde\psi = 0 \qquad \text{in } \Omega \times \{t_*\}, \tag{5.28}$$

for $w = (w_1, w_2) \in L^2(\partial\Omega \times I) \times L^2(\Omega)$, satisfies

$$
\begin{aligned}
\langle h(T)_t, -\tilde{\phi} \rangle_{L^2(Q)} &= -\int_0^{t_*} \int_\Omega \tilde{\phi}(x,t) \frac{\partial}{\partial t} h(T(x,t)) \; dx \; dt \\
&= \int_0^{t_*} \int_{\partial\Omega} \theta(x,t) w_1(x,t) \; d\sigma(x) \; dt \\
&\quad + \int_\Omega e^{-\int_0^{t_*}(au)(x,s)ds} \left(-\int_0^{t_*} (a\phi)(x,s) ds \right) w_2(x) \; dx \\
&= \langle F'(b)h, w \rangle \qquad (5.29) \\
&= \langle h, F'(b)^* w \rangle, \qquad (5.30)
\end{aligned}
$$

where (θ, ϕ, ψ) is the solution of the linearized problem (5.1)-(5.8). Thus, to compute $F'(b)^* w$ it remains to find a function $f \in H^\eta[z_1, z_2]$, such that

$$
\langle h(T)_t, -\tilde{\phi} \rangle_{L^2(Q)} = \langle h, f \rangle_{H^\eta([z_1,z_2])} \qquad \forall \; h \in H^\eta([z_1, z_2]). \tag{5.31}
$$

If we assume for the moment that the cooling is strong enough such that $T_t \leq T_0 < 0$ in $\Omega \times I$, we can transform the variables to x and T and obtain

$$
\begin{aligned}
\langle h(T)_t, -\tilde{\phi} \rangle_{L^2(Q)} &= -\int_0^{t_*} \int_\Omega \tilde{\phi}(x,t) h'(T(x,t)) T_t(x,t) \; dx \; dt \\
&= \int_\Omega \int_{T(x,t_*)}^{T(x,0)} \tilde{\phi}(x,z) h'(z) \; dz \; dx \\
&= \int_{z_1}^{z_2} \int_{\Omega(z)} \tilde{\phi}(x,z) \; dx h'(z) \; dz,
\end{aligned}
$$

where

$$
\Omega(z) = \{ \; x \in \Omega \mid \exists \; t \in I : T(x,t) = z \; \}.
$$

By defining

$$
\begin{aligned}
g(z) &:= -\frac{\partial}{\partial z} \int_{\Omega(z)} \tilde{\phi}(x,z) \; dx, \qquad (5.32) \\
f(z) &:= (L_\eta^* L_\eta)^{-1} g, \qquad (5.33)
\end{aligned}
$$

where L_η is the linear operator that generates the norm in $H^\eta([z_1, z_2)]$, we obtain the identity

$$
\begin{aligned}
\langle h(T)_t, -\tilde{\phi} \rangle_{L^2(Q)} &= \langle g, h \rangle_{L^2([z_1,z_2])} \\
&= \langle f, h \rangle_{H^\eta([z_1,z_2])}.
\end{aligned}
$$

Thus, we have found $F'(b)^* w = f$ by this construction.

Since $\tilde{\phi}$ is the solution of a hyperbolic equation we need not expect strong regularity of $\tilde{\phi}$ and g defined by (5.32). Nevertheless, if only $g \in L^2([z_1, z_2])$ holds, we obtain $f \in$

$H^{2\eta}([z_1, z_2])$ and, in addition, we have to know boundary values for f up to order $\eta - 1$. Hence, the source condition (5.19) is a condition upon the smoothness and upon boundary values of $b^\dagger - b^0$, which means that we have to anticipate nonsmooth parts of the solution in order to obtain faster convergence. Furthermore, if the interval $[z_1, z_2]$ is chosen such that for some $z \in [z_1, z_2]$

$$T(x,t) \neq z, \qquad \forall\ (x,t) \in \Omega \times I,$$

we always have $g(z) = 0$ at this point. For the source condition (5.19), this is an additional restriction upon $b^\dagger - b^0$, i.e., either each value in $[z_1, z_2]$ occurs in the experiment, or we must know the value of $b^\dagger$ a-priori at temperature values that do not occur. In practice, this problem can often be avoided, since the initial temperature is uniform and the cooling is so strong that the temperature decreases in a monotone way. Hence, we may choose z_2 as the value of the initial temperature and z_1 as the minimum of the final temperature at the boundary, which is measured anyway. By monotonicity principles for parabolic equations we can claim that all temperatures that occur will be in this range. For further details about the computation of the adjoint, which is needed for the Landweber iteration, and about the source condition we refer to [8].

The conditions upon the nonlinearity, needed for a convergence statement about the algorithm in absence of (5.19) could not be verified, nevertheless the numerical results indicate the convergence of all methods.

6 Numerical Results and Conclusions

In order to test and compare the behaviour of the different iterative methods, we have performed numerical simulations using artificial data for T_Γ^δ and ξ_*^δ. The data have been generated by solving the direct problem on a very fine grid and artificially perturbed with high-frequency noise. The numerical solution of the inverse problem has been performed on a different grid, the values of the data there have been obtained by interpolation. Since we know the exact solution $b^\dagger$ for these data, we can compute error $\left\| b^\dagger - b_k^\delta \right\|$, e.g. in $H^2([z_1, z_2])$. For Broyden's method, we start with the Newton-matrix, for the frozen Gauss-Newton method we perform a restart with the Newton-matrix after every 5 iterations.

The first Figure 1 shows the development of the error for fixed noise level ($\delta = 2\%$ and $\delta = 5\%$) during the Newton-type iterations, i.e., a plot of $\left\| b^\dagger - b_k^\delta \right\|$ vs. the iteration number k. These plots illustrate the behaviour of the methods with respect to convergence speed and stability. The results confirm the general idea about the iterative regularization of ill-posed problems: during the first iterations the error decreases, but then starts to increase again (cf. e.g. [14]). Since the main regularizing effect comes from the choice of a stopping index, it seems to be of advantage if several iterates are of a 'good quality', i.e., if several choices of the stopping index yield an approximate solution with small error. This means that the curve arising from the plot error vs. iteration number should not be to steep around the index with minimal error.

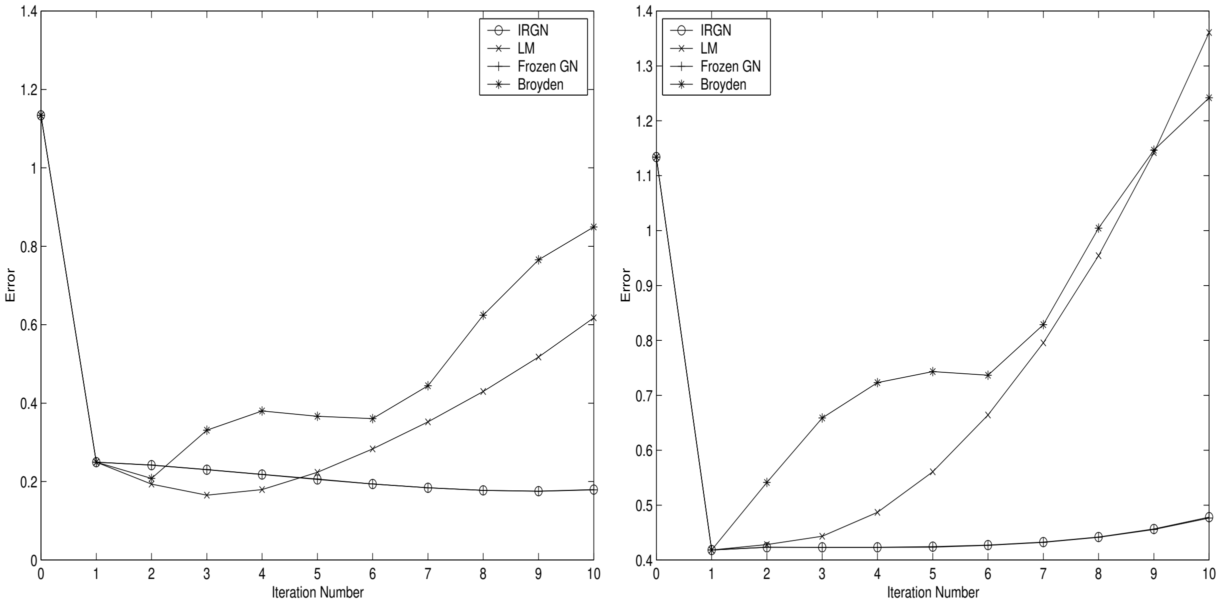

Figure 1: Development of the error $\left\|b_k^\delta - b^\dagger\right\|$ during the iteration (k) for fixed noise level $\delta = 2\%$ (left) and $\delta = 5\%$ (right).

Since for all methods, the first step is the same, differences between the methods arise from the second step on. It turns out that in general, the Levenberg-Marquardt method (LM) is faster than the iteratively regularized Gauss-Newton method (IRGN) and its frozen version (Frozen GN), i.e., it reaches the index with minimal error in less iterations, but as one can observe, the iteratively regularized Gauss-Newton method has better stability properties, since it produces several iterates with 'acceptable' error. This means that the choice of the termination index is more critical for the Levenberg-Marquardt method. A similar statement holds for Broyden's method, which is in general slower, but still critical with respect to the choice of the stopping index.

The behaviour as the noise level tends to zero is illustrated in Figures 2 and 3. The first shows a plot of the minimal error obtained during the iteration procedure (for fixed noise level) $\left\|b_{k_*(\delta)}^\delta - b^\dagger\right\|$ vs. the noise level δ. Obviously, the stopping index cannot be chosen by this criteria in practice, since the exact solution is not known in general. Nevertheless, it turns out, that the discrepancy principle (3.7) with $\tau = 1.3$ yields the same stopping index in almost all cases. This can be observed from Figure 3, which shows a plot of the residual $\left\|F(b_{k_*(\delta)}^\delta) - y^\delta\right\|$; the resulting curves fit the line $r(\delta) = 1.3\delta$ very well. It turns out that all Newton-type methods perform similary with respect to the minimal error that can be achieved, while the Landweber iteration is more successfull for large noise level, but $\left\|b^\dagger - b_{k_*}^\delta\right\|$ seems to converge slower as δ tends to zero.

The efficiency of the iteration methods can be compared from Figures 4 and 5, which show the stopping index $k_*(\delta)$ (according to (3.7) with $\tau = 1.3$) and the number of floating

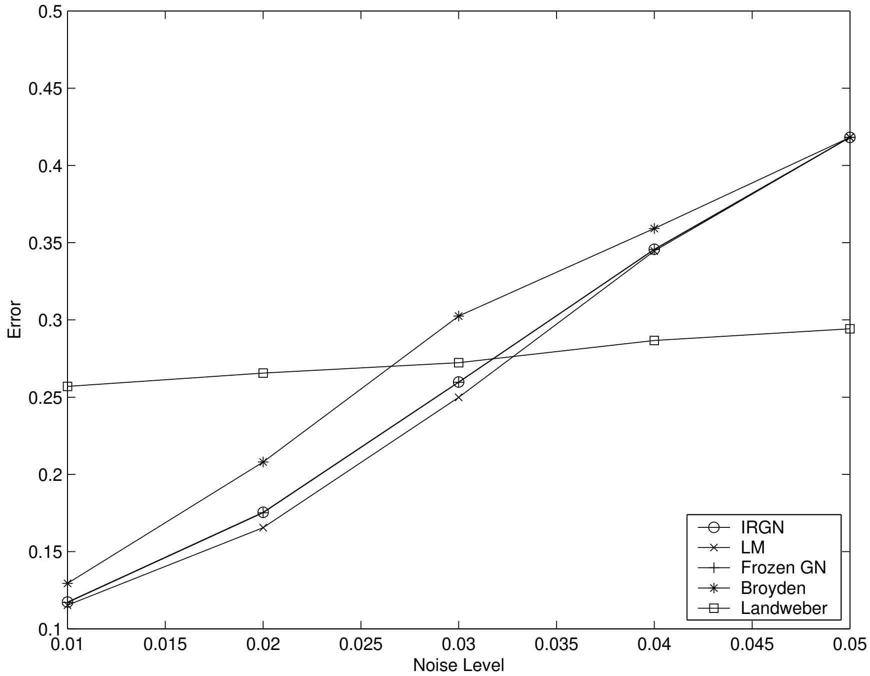

Figure 2: Minimal error $\left\| b_{k_*(\delta)}^{\delta} - b^{\dagger} \right\|$ vs. noise level δ.

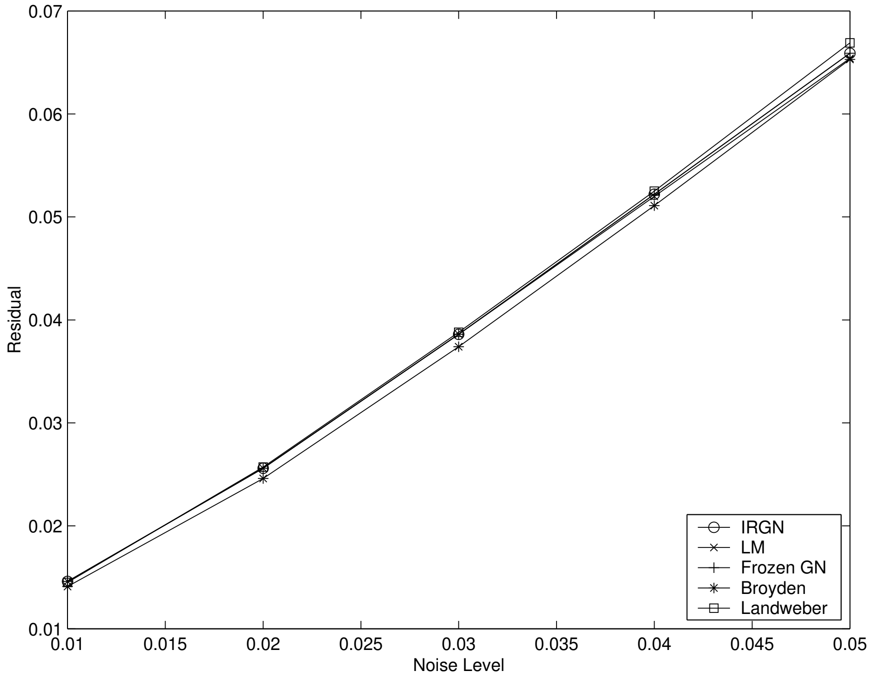

Figure 3: Residual at iteration of minimal error $\left\| F(b_{k_*(\delta)}^{\delta}) - y^{\delta} \right\|$ vs. noise level δ.

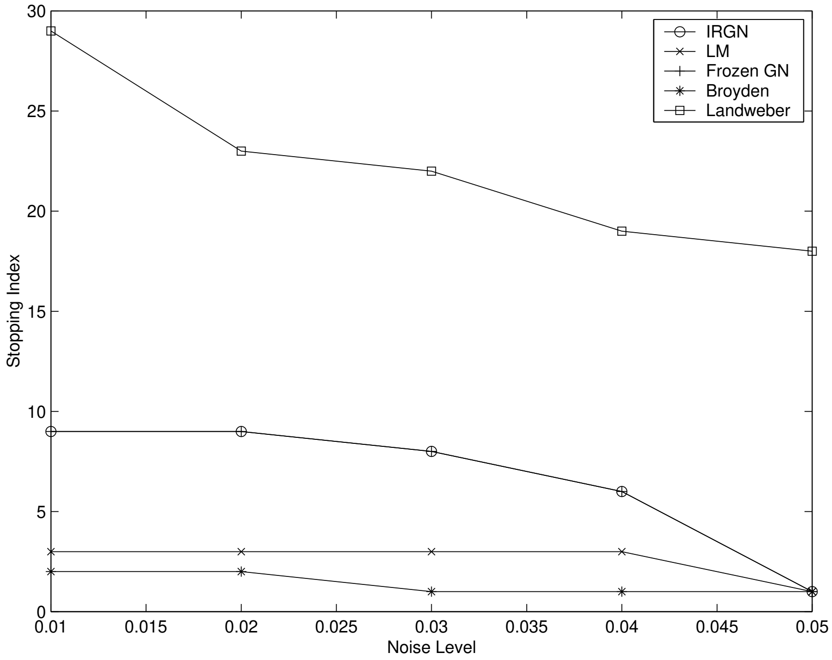

Figure 4: Stopping index $k_*(\delta)$ vs. noise level δ.

point operations $FlOps(k_*(\delta))$ needed until the iteration procedure is stopped, plotted vs. the noise level δ. Obviously all Newton-type methods need less iterations than the Landweber iteration, but with much higher effort in each step. For the discretization we used, the Landweber iteration is almost as efficient as the Levenberg-Marquardt method and the iteratively regularized Gauss-Newton method. With growing number of nodes (m) in the discretization this effect becomes stronger, since in one step of the Landweber iteration only one adjoint initial-boundary value problem has to be solved, while the allocation of the Newton-matrix requires m^2 solutions of the linearized initial-boundary value problem.

Finally, the quality of the approximations obtained using different methods is shown in Figures 6 and 7, which confirm the above statements about the behaviour of the iterative methods investigated.

A possible conclusion from our results, is that Newton-type methods do not perform better than Landweber iteration in many cases. Although the number of iterations needed is much less, the high effort in the allocation of the Newton matrix can lead to a higher number of floating point operations. Especially for high noise level, the Landweber iteration proves to be an efficient method, which yields better results than any other method.

Another outcome of our numerical experiments is that frozen Newton-type methods do not perform much worse than their non-frozen equivalents for this problem; it seems that $F'(b)$ does not change strongly if b is close to the solution $b^\dagger$. Since the numerical effort can be reduced significantly, frozen Newton-type method seem to be a good choice.

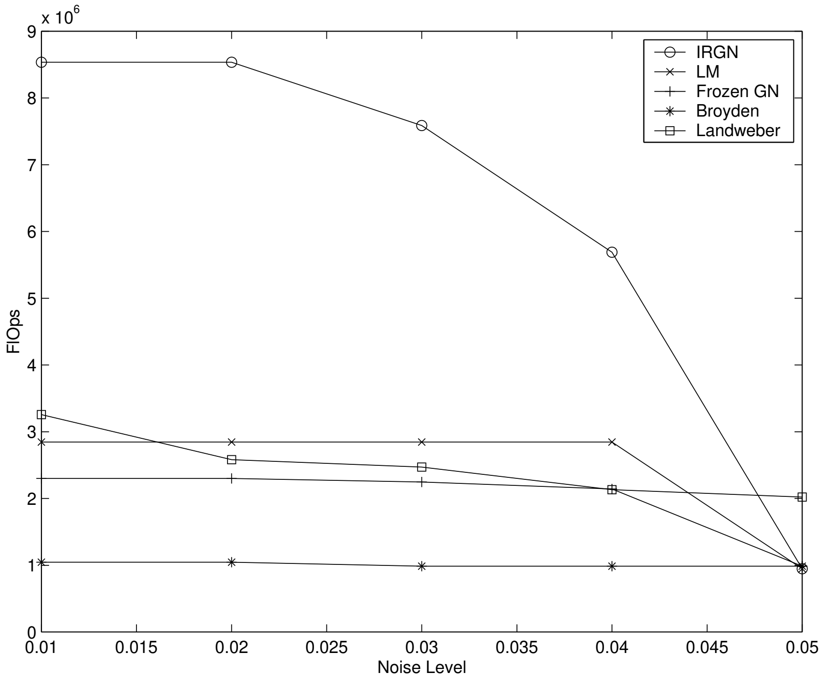

Figure 5: Number of floating point operations $FlOps(k_*(\delta))$ needed until termination of the iteration vs. noise level δ.

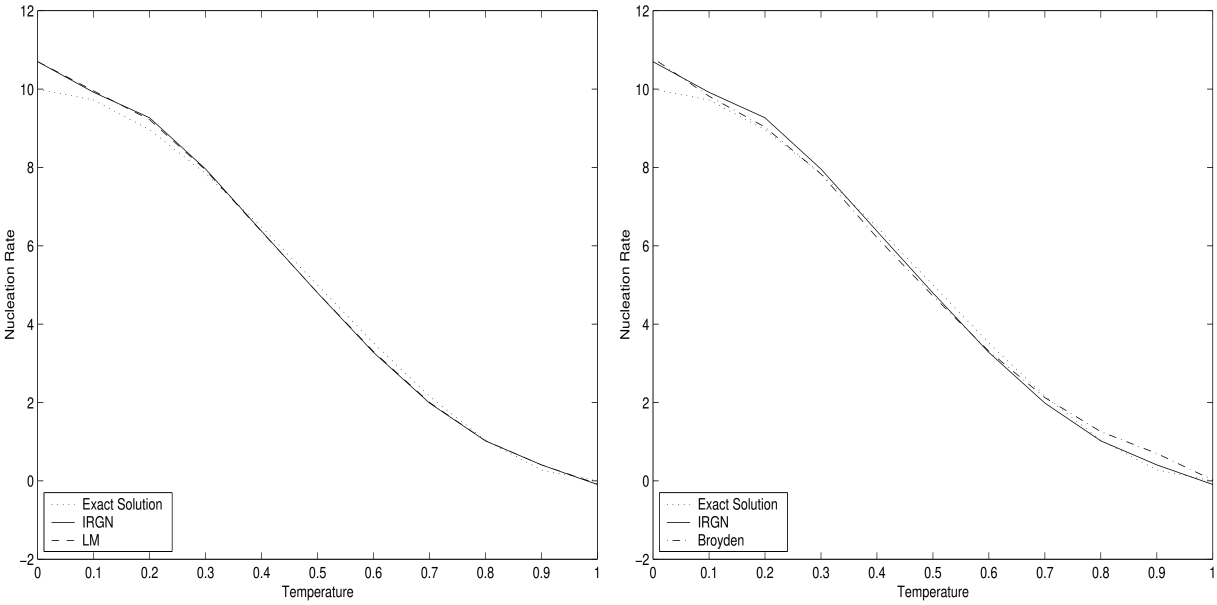

Figure 6: Exact solution (dotted) and closest iterates obtained with the iteratively regularized Gauss-Newton method (solid), the Levenberg-Marquardt method (dashed, left) and Broyden's method (dash-dotted, right), for noise level $\delta = 2\%$.

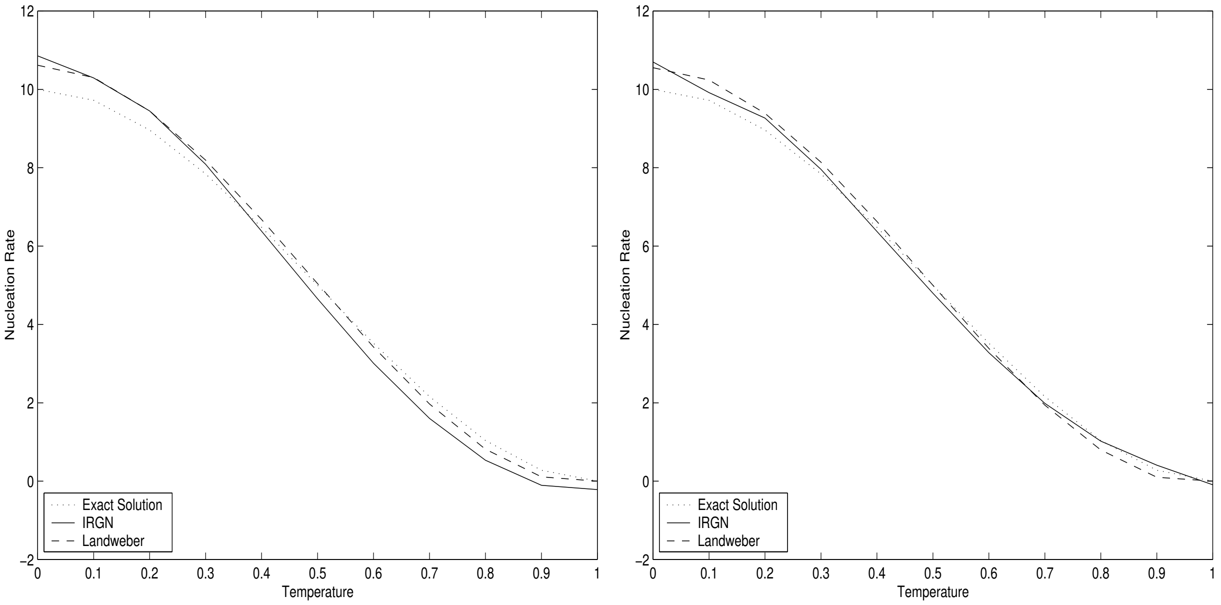

Figure 7: Exact solution (dotted) and closest iterates obtained with the iteratively regularized Gauss-Newton method (solid) and the Landweber iteration (dashed), for noise level $\delta = 5\%$ (left) and $\delta = 2\%$ (right) .

Acknowledgements

This work has been supported by the *Austrian Fonds zur Förderung der Wissenschaftlichen Forschung*, projects SFB F 1308 and P 13478-INF. Fruitful and stimulating discussions are acknowledged to Prof. Heinz W. Engl, University of Linz.

References

[1] A.G.Bakushinskii, *The problem of the convergence of the iteratively regularized Gauss-Newton method*, Comput. Math. Math. Phys. **32** (1992), 1353-1359.

[2] H.T.Banks, K.Kunisch, *Estimation Techniques for Distributed Parameter Systems*, (Birkhäuser, Basel, Boston, Berlin,1989).

[3] A.Binder, M.Hanke, O.Scherzer, *On the Landweber iteration for nonlinear ill-posed problems*, J.Inv.Ill-Posed Problems **4**, 381-389 (1996).

[4] B.Blaschke(-Kaltenbacher), *Some Newton Type Methods for the Regularization of Nonlinear Ill-posed Problems*, PhD Thesis (Linz University, 1996).

[5] B.Blaschke(-Kaltenbacher), A.Neubauer, O.Scherzer, *On convergence rates for the iteratively regularized Gauss-Newton method*, IMA J.Numer.Anal. **17** (1997), 421-436.

[6] M.Burger, V.Capasso, G.Eder, *Modelling crystallization of polymers in temperature fields* (1998), submitted.

[7] M.Burger, V.Capasso, G.Eder, H.W.Engl, *Modelling and parameter identification in non-isothermal crystallization of polymers*, in L.Arkeryd, J.Bergh, P.Brenner, R. Pettersson, eds., *Progress in Industrial Mathematics at ECMI 98* (Teubner, Stuttgart, Leipzig, 1999), 114-121.

[8] M.Burger, V.Capasso, H.W.Engl, *Inverse problems related to crystallization of polymers*, Inverse Problems **15** (1999), 155-173.

[9] M.Burger, V.Capasso, C.Salani, *Modelling multi-dimensional crystallization of polymers in interaction with heat transfer* (1998), submitted.

[10] D.Colton, R.Ewing, W.Rundell (eds.), *Inverse Problems in Partial Differential Equations* (SIAM, Philadelphia, 1990).

[11] P.Deuflhard, H.W.Engl, O.Scherzer, *A convergence analysis of iterative methods for the solution of nonlinear ill-posed problems under affinely invariant conditions*, Inverse Problems **14** (1998) 1081-1106.

[12] K.Deimling, Nonlinear Functional Analysis (Springer, Berlin, Heidelberg, 1985).

[13] G.Eder, *Fundamentals of structure formation in crystallizing polymers*, in K.Hatada, T.Kitayama, O.Vogl, eds., *Macromolecular Design of polymeric Materials* (M.Dekker, New York, 1997), 761-782.

[14] H.W.Engl, M.Hanke and A.Neubauer, *Regularization of Inverse Problems* (Kluwer, Dordrecht, 1996).

[15] H.W.Engl, K.Kunisch, A.Neubauer, *Convergence rates for Tikhonov regularization of nonlinear ill-posed problems*, Inverse Problems **5** (1989), 523-540.

[16] H.W.Engl, O.Scherzer, *Convergence rate results for iterative methods for solving nonlinear ill-posed problems*, to appear in: D.Colton, H.W.Engl, J.McLaughlin, A.Louis, W.Rundell, eds., *Solution Methods for Inverse Problems* (Springer, Vienna, New York, 2000).

[17] M.Hanke, *A regularizing Levenberg-Marquardt scheme, with applications to inverse groundwater filtration problems*, Inverse Problems **13** (1997), 79-95.

[18] M.Hanke, A.Neubauer, O.Scherzer, *A convergence analysis of the Landweber iteration for nonlinear ill-posed problems*, Numer. Math. **72** (1995), 21-37.

[19] V.Isakov, *Inverse Problems for Partial Differential Equations* (Springer, New York, 1998).

[20] B.Kaltenbacher, *Some Newton-type methods for the regularization of nonlinear ill-posed problems*, Inverse Problems **13** (1997), 729-753.

[21] B.Kaltenbacher, *A-posteriori parameter choice strategies for some Newton-type methods for the regularization of nonlinear ill-posed problems*, Numer. Math. **79** (1998), No.4, 501-528.

[22] B.Kaltenbacher, *On Broyden's method for the regularization of nonlinear ill-posed problems*, Numer. Funct. Anal. and Optimiz. **19** (1998), 807-833.

[23] J.L.Lions, E.Magenes, *Non-Homogenous Boundary Value Problems and Applications, Vol. I* (Springer, Berlin, Heidelberg, New York, 1972).

[24] J.L.Lions, E.Magenes, *Non-Homogenous Boundary Value Problems and Applications, Vol. II* (Springer, Berlin, Heidelberg, New York, 1972).

[25] H.Niessner, *Stability of Lax-Wendroff methods extended to deal with source terms*, ZAMM **77** (1997), Suppl. 2, S637-S638.

[26] R. Ramlau, *Discretized Landweber iteration methods for nonlinear inverse problems*, ZAMM **78** (1998), Suppl. 1, S85-S88.

[27] O.Scherzer, *Convergence criteria of iterative methods based on Landweber iteration for solving nonlinear problems*, J.Math.Anal.Appl. **194** (1995), 911-933.

[28] T.I.Seidman, C.R.Vogel, *Well-posedness and convergence of some regularization methods for nonlinear ill-posed problems*, Inverse Problems **5** (1989), 227-238.

[29] Y.Zhang; B.Tabarrok, *Modifications to the Lax-Wendroff scheme for hyperbolic systems with source terms*, Int. J. Numer. Methods Eng. **44** (1999), 27-40.

[30] S.Zheng, *Nonlinear Parabolic Equations and Hyperbolic-Parabolic Coupled Systems* (Longman, New York, 1995).