IMA Journal of Applied Mathematics (2023) **00**, 1–27
https://doi.org/10.1093/imamat/hxad013

Identifying a response parameter in a model of brain tumour evolution under therapy

G. Baravdish, B. T. Johansson and O. Svensson*
ITN, Campus Norrköping, Linköping University, Sweden
*Corresponding author. Email: olof.svensson@liu.se

and

W. Ssebunjo
Department of Mathematics, Makarere University, Uganda

[Received on 15 November 2021; revised on 21 October 2022; revised on 19 February 2023]

A nonlinear conjugate gradient method is derived for the inverse problem of identifying a treatment parameter in a nonlinear model of reaction–diffusion type corresponding to the evolution of brain tumours under therapy. The treatment parameter is reconstructed from additional information about the tumour taken at a fixed instance of time. Well-posedness of the direct problems used in the iterative method is outlined as well as uniqueness of a solution to the inverse problem. Moreover, the parameter identification is recasted as the minimization of a Tikhonov type functional and the existence of a minimizer to this functional is shown. Finite-difference discretization of the space and time derivatives are employed for the numerical implementation. Numerical simulations on full 3D brain data are included showing that information about a spacewise-dependent treatment parameter can be recovered in a stable way.

Keywords: parameter identification; inverse problem; nonlinear conjugate gradient method; reaction-diffusion models; brain tumours.

1. Introduction

In Jaroudi *et al.* (2016, 2019, 2020), the inverse ill-posed problem of finding the initial cell distribution for brain tumours was studied for a nonlinear tumour growth model, and a Landweber method was derived for the stable reconstruction of the initial data. We continue this direction of research by adding a treatment parameter into the model and studying the problem of reconstructing this parameter by a conjugate gradient method. This work is inspired by Cao *et al.* (2020), where parameters in parabolic heat transfer models are simultaneously recovered using a conjugate gradient method. We focus on recovering one parameter only but consider additionally a nonlinear model and presenting 3D numerical simulations in the space variables.

We frame the parameter identification problem in the language of tumour detection. However, details on how to realize the proposed method in a concrete medical situation (such as collecting the necessary data) is not presented here, and the results can be adopted to other situations governed by the similar equations.

The model to be studied is the following reaction–diffusion equation

$$\begin{cases} \partial_t u(x,t) - \operatorname{div}(D(x)\nabla u(x,t)) - f(u(x,t)) = -\alpha(x,t)u(x,t), & \text{in } \Omega \times (0,T) \\ D(x)\nabla u(x,t)\cdot n(x) = 0, & \text{on } \partial\Omega \times (0,T) \\ u(x,0) = \varphi(x), & \text{in } \Omega. \end{cases} \tag{1.1}$$

This model is well-known for studying the growth of brain tumours under treatment (see, for example Swanson *et al.* (2003), Powathil *et al.* (2007), Stamatakos & Giatili (2017), Rockne *et al.* (2009), Rockne *et al.* (2010), Martín-Landrove (2017) and Murray (2002)) with $u(x,t)$ the tumour cell density at location x of the brain region Ω (assumed smooth). Furthermore, $\alpha(x,t)u(x,t)$ is the treatment term describing the death of cells due to chemotherapy or radiation therapy, $\operatorname{div}(D(x)\nabla u)$ corresponds to the diffusion term, D is the diffusion coefficient of cells in the brain tissue modelling a random tumour cell movement as a diffusive flux proportional to the cell density gradient. The function $f(u)$ is the reaction term considered here to be the logistic growth model $f(u) = \rho u^a(1-u^b)^c$, where ρ is the proliferation rate and a, b and c are positive real numbers, describing the rapid reproduction of cells in the brain as a combination of birth and death processes of the cells. The boundary condition in (1.1) guarantees that tumour cells do not diffuse outside the brain region (the skull), with n being the outward unit normal to the boundary. The element φ is the known initial tumour cell density obtained at time $t = 0$.

In the model (1.1), the diffusion coefficient D is assumed to comprise of two regions; the white and grey matter as used in Swanson *et al.* (2000), Swanson *et al.* (2002) and Dolgushin *et al.* (2017). This accounts for the spatial heterogeneity of the brain tissue, and is specified as

$$D(x) = \begin{cases} d_w : x \in \text{white matter} \\ d_g : x \in \text{grey matter} \end{cases}$$

where $d_w >> d_g > 0$ stand for the respective diffusion coefficients for the white and grey matter.

Given values of the coefficients and the initial cell distribution in (1.1), it is a well-posed problem to generate the solution u to (1.1) forward in time. We shall instead consider the following: **Inverse problem:** Given the additional data

$$u(x,T) = \psi(x), \quad x \in \Omega \tag{1.2}$$

where $T > 0$, determine the treatment profile $\alpha(x,t)$. This nonlinear inverse problem is ill-posed and we assume that data are compatible such that a solution exists, however, stability cannot be guaranteed. Furthermore, uniqueness is a subtle issue.

It is pointed out in the introduction to Yamamoto & Zou (2001) that uniqueness in finding α is not to be expected when α depends both on space and time. Since it is not either expected to find a time-dependent profile α from final time data, we focus on the spacewise-dependent case. Then, in the linear case when the reaction term f is zero or linear, uniqueness is shown in Rundell (1987); Isakov (2017) provided that $\alpha > 0$. The case of a non-zero reaction term is covered in Choulli (1994) (in the case of Dirichlet boundary conditions) and with general boundary conditions in Prilepko & Kostin (1993); Kaltenbacher & Rundell (2019). Thus, there are cases in which the inverse problem (1.1)–(1.2) has a unique solution when finding a spacewise-dependent treatment profile. However, since a both space and time-dependent coefficient α does not pose any problems in the derivation of a method for the inverse problem (1.1)–(1.2) it is only in the numerical section that we shall specify α to be space-dependent.

Inverse problems of finding coefficients, the initial distribution or sources in parabolic equations is a too vast area to survey here. To guide the reader to some works related to the present work, we refer to Prilepko *et al.* (2000); Isakov (2017) and additionally Hasanov (2007); Johansson & Lesnic (2007a,b); Hao *et al.* (2013); Klibanov (2013); Kerimov & Ismailov (2015); Van Bockstal & Marin (2017); Huntul *et al.* (2018); Cao *et al.* (2020); Slodička (2020) and references therein. Note also the recent development of inverse problems for degenerate parabolic equations Hussein *et al.* (2020). In terms of parameter reconstruction for a growth model of the above type, see Amir *et al.* (2016); Mang *et al.* (2018); Nguyen *et al.* (2019); Sabir & Raissi (2019); Mang *et al.* (2020).

Parameter identification for parabolic equations is thus classical. Novelties of the presented work are the derivation of a nonlinear conjugate gradient method for the realization of the identification together with mathematics to justify the method, along with the actual numerical implementation on full brain data. Comparisons are done with the classical Landweber method.

For the outline of this work, in Section 2 we outline that the inverse problem (1.1)–(1.2) has a unique solution when the logistic growth model is considered. Precise statements of uniqueness is given in Theorem 2.1. In Section 3, we briefly go over the well-posedness of (1.1) rendering Theorem 3.1. In Section 4, the inverse problem (1.1)–(1.2) is formulated as the minimization of a certain Tikhonov type functional. Properties are discussed and it is outlined that there exists a minimizer, see Proposition 4.1. A proof of the existence of a minimizer is given in an appendix. A conjugate gradient method is derived in Section 5 for the minimization. Section 6 is devoted to numerical experiments for a spacewise-dependent treatment profile on full 3D synthetic brain data. Finally, we give some conclusions and remarks about our findings in Section 7.

2. A note on uniqueness of a solution to the inverse problem (1.1)–(1.2)

We will in particular be concerned with a reaction term in the form $f(u) = \rho u(1-u)$ modelling logistic growth. We outline a uniqueness result for the inverse problem (1.1)–(1.2) with this reaction term. Following the case of a linear governing equation studied in Prilepko & Kostin (1993) assume that (1.1)–(1.2) has two solution pairs u and α respectively u_1 and α_1. Define $w = u - u_1$ and $h = \alpha - \alpha_1$. Then, as the reader can check,

$$\begin{cases} \partial_t w(x,t) - \operatorname{div}(D(x)\nabla w(x,t)) - \beta(x,t)w = -h(x)u(x,t) & \text{in } \Omega \times (0,T), \\ D(x)\nabla w(x,t)\cdot n(x) = 0 & \text{on } \partial\Omega \times (0,T), \\ w(x,0) = w(x,T) = 0 & \text{in } \Omega, \end{cases} \tag{2.1}$$

with $\beta = -\alpha_1 + \rho(1 - u - u_1)$. The solution pair w and h satisfying (2.1) can be considered as a solution to the inverse problem of reconstructing w and a spacewise-dependent source term h from final time data in a linear parabolic governing equation with a Neumann boundary condition. According to (Prilepko & Kostin, 1993, Thm. 2) (see also Rundell (1980)) that problem has a unique solution, hence $w = 0$ and $\varphi = 0$. This in turn, by definition, gives $u = u_1$ and $\alpha = \alpha_1$, and uniqueness of a solution to the inverse problem (1.1)–(1.2) has been shown.

In fact, following Prilepko & Kostin (1993), we can make precise conditions for uniqueness. Let $L^\infty(\Omega)$ be the space of measurable essentially bounded functions in Ω, and let the space $H^k(\Omega)$, $k > 0$, be the standard Sobolev space of functions with weak and square integrable derivatives up to order k.

Moreover, $W^{2,1}(\Omega \times (0,T))$ is the standard anisotropic Sobolev space with two weak derivatives in space and one in time.

THEOREM 2.1. Assume that $\psi \in H^2(\Omega)$ with $\psi > 0$. Then the solution to the inverse problem (1.1)–(1.2) is unique in the class of pairs $u \in W^{2,1}(\Omega \times (0,T))$ and $\alpha \in L^\infty(\Omega)$ with $\alpha \geq 0$ with the standard compatibility conditions assumed to be satisfied.

Alternative classes of functions and restrictions can be derived for uniqueness, see, for example Choulli (1994) (Dirichlet boundary condition) and Kaltenbacher & Rundell (2019).

We remark that it is possible to generalize further and consider a time-dependent treatment parameter in the form $\alpha(x,t) = \alpha(x)g(t)$, with $g(t)$ known. Uniqueness of $\alpha(x)$ can still be shown for some suitable g as above and the methods we develop can be directly applied to that case.

3. Well-posedness of the direct treatment model (1.1)

We shall investigate existence and uniqueness of a solution to the treatment model (1.1), and begin by introducing some function spaces. The space $L^2(0,T;X)$, where X is a Hilbert space, consists of those measurable functions $u(\cdot,t) : (0,T) \to X$, with

$$\int_0^T \|u(\cdot,t)\|_X^2 dt < \infty.$$

By $C^k([0,T];X)$, we denote the functions u for which the mapping $u(\cdot,t) : [0,T] \to X$ has continuous and bounded (in the usual norm) derivatives of order up to $k \geq 0$. The trace space of $H^k(\Omega)$, $k > 0$, is $H^{k-1/2}(\Gamma)$. For simplicity, we denote by $\Omega_t = \Omega \times (0,t)$. In the logistic function $f(u) = \rho u^a(1-u^b)^c$, we assume that $a \geq 1$, $b \geq 1$ and $c \geq 1$.

Concerning the existence and uniqueness of a solution to (1.1), we follow the approach given in a previous work Jaroudi *et al.* (2020) when $\alpha = 0$. The treatment problem in (1.1) can be recast in an abstract form as

$$\begin{cases} u_t' + Bu &= f(u(t)) - \alpha(t)u \\ u(0) &= \varphi \end{cases}$$

making it essentially an ordinary differential equation in the time-variable but having values in a function space. Here, B corresponds to the divergence term in (1.1), and generates a semi-group S, see (Pazy, 1983, Theorem 7.2.5). Existence and uniqueness of what is known as a mild solution,

$$u(t) = S(t)\varphi + \int_0^t S(t-s)(f(u(s)) - \alpha(s)u(s))\, ds \tag{3.1}$$

in the space $C(0,T;H^1(\Omega))$ is given by (Pazy, 1983, Theorem 6.1.2). An additional advantage with the abstract formulation is that recovering the treatment parameter α from additional data $u(x,T) = \psi(x)$ can, by using (3.1), be considered as a nonlinear operator equation

$$A(\alpha) = \psi. \tag{3.2}$$

However, since we shall recast the inverse problem (1.1)–(1.2) as a minimization problem we need a more standard weak formulation in order to show properties of the minimization problem.

Multiplying (1.1) by an element $v \in H^1(\Omega)$ and using integration by parts in the space variables (a Green's formula), incorporating the zero flux condition on the boundary, give

$$\int_\Omega u_t'(x,\cdot)v(x)\,dx + \int_\Omega D\nabla u(x,\cdot)\cdot\nabla v(x)\,dx = \int_\Omega (f(u(x,\cdot)) - \alpha(x,\cdot)u(x,\cdot))v(x)\,dx. \tag{3.3}$$

An element u is termed a weak solution to (1.1) if (3.3) holds for every t in $[0,T]$ and $u(x,0) = \varphi(x)$, for every $v \in H^1(\Omega)$.

To approximate the time-derivative in (1.1), we apply the backward difference

$$u_t'(x,t_k) \approx \frac{u_k(x) - u_{k-1}(x)}{\tau},$$

where $t_k = (T/N)k$, $k = 0,1,\ldots,N$, is a uniform mesh on the time interval $[0,T]$ with step size $\tau = T/N$, and $u_k(x) = u(x,t_k)$.

Employing this time-discretization in (3.3) together with evaluating the nonlinear term at the previous mesh point, we derive the identity

$$\begin{aligned}\int_\Omega \frac{u_k(x)-u_{k-1}(x)}{\tau}v(x)\,dx \;+\; &\int_\Omega D\nabla u_k(x)\cdot\nabla v(x)\,dx\\ &= \int_\Omega (f(u_{k-1}(x)) - \alpha(x,t_{k-1})u_{k-1}(x))v(x)\,dx,\end{aligned}$$

or by rewriting this,

$$\begin{aligned}\frac{1}{\tau}\int_\Omega u_k(x)v(x)\,dx \;+\; &\int_\Omega D\nabla u_k(x)\cdot\nabla v(x)\,dx\\ &= \int_\Omega \left(f(u_{k-1}(x)) - \alpha(x,t_{k-1})u_{k-1}(x) + \frac{u_{k-1}(x)}{\tau}\right)v(x)\,dx.\end{aligned}$$

This is then a standard linear elliptic problem for u_k (for $k = 0$ the condition $u(x,0) = \varphi$ is used). Existence and uniqueness of $u_k \in H^1(\Omega)$ is settled via the Lax-Milgram lemma, see (Ciarlet, 1978, Chapter 1). There is scope for further generalizations here in terms of smoothness of the coefficients, in particular the diffusion tensor D can be more general than what is specified in the current work.

The sequence of functions $\{u_k\}$ can be interpolated into a time-dependent approximation simplest by defining it to be piecewise constant in each interval $[t_{k-1},t_k)$, alternatively to be linear in each such interval. This is known as the Rothe approximation. As the step size τ decreases, the interpolated function tends to a solution of (3.3) in the appropriate norms. In this way, existence of a weak solution can be shown. For the linear case, an introduction to the method of Rothe is given in (Kačur, 1985, Chapter 1).

For the uniqueness, assume that there are two weak solutions u and $\tilde{u}$ to (1.1). Put $w = u - \tilde{u}$, then w satisfies a relation of the form (3.3) with $f(u)$ in the right-hand side replaced by $f(u) - f(\tilde{u})$ and $\varphi = 0$. Note that (3.3) can be extended to hold for functions v, which are piecewise constant in time, and by a limiting argument this relation holds for all $v(x,t) \in L^2(0,T;H^1(\Omega))$. Hence, choosing $v = w$ in the weak formulation for w, we have

$$\begin{aligned}\int_\Omega w_t'(x,\cdot)w(x,\cdot)\,dx \;+\; &\int_\Omega D\nabla w(x,\cdot)\cdot\nabla w(x,\cdot)\,dx\\ &= \int_\Omega (f(u(x,\cdot)) - f(\tilde{u}(x,\cdot)))w(x,\cdot)\,dx - \int_\Omega \alpha(x,\cdot)w^2(x,\cdot)\,dx\\ &\le \int_\Omega (f(u(x,\cdot)) - f(\tilde{u}(x,\cdot)))w(x,\cdot)\,dx,\end{aligned}$$

where in the last step it is used that α is non-negative (it is for such α we have uniqueness in the inverse problem, see Theorem 2.1). Integrating in time over $[0,t]$ noting that $\partial_t w^2 = 2w'w$ and estimating the second term in the left-hand side from below, we obtain

$$\begin{aligned}\|w\|^2_{L^2(\Omega_t)} \;+\; &\|\nabla w\|^2_{L^2(\Omega_t)}\\ &\le C\int_{\Omega_t} (f(u(x,t)) - f(\tilde{u}(x,t)))w(x,t)\,dx\,dt.\end{aligned}$$

The function f is Lipschitz continuous, hence we can further estimate using this and Cauchy's inequality,

$$\|w\|^2_{L^2(\Omega_t)} + \|\nabla w\|^2_{L^2(\Omega_t)} \le C\|w\|^2_{L^2(\Omega_t)}. \tag{3.4}$$

A version of Grönwall's lemma, see (Roubíček, 2005, p. 25), implies that the left-hand side is zero, which in turn, since t with $0 < t < T$ was arbitrary, forces $w = 0$ in Ω_T and we have uniqueness.

The above steps albeit on a more general level is performed in (Roubíček, 2005, Chapter 8), where the reader can find full details on the above arguments, which renders the following result (combining (Roubíček, 2005, Theorem 8.33 and Proposition 8.37)).

THEOREM 3.1. Let f be Lipschitz continuous in $L^2(0,T;H^1(\Omega))$, $\varphi \in L^2(\Omega)$ and $\alpha > 0$ be a bounded measurable function in Ω. Then there exists a unique weak solution $u \in L^2(0,T;H^1(\Omega))$ with $u_t' \in L^2(0,T;L^2(\Omega))$ to the treatment reaction–diffusion problem (1.1), and this element u depends continuously on the data.

It is also possible to show existence of a classical solution being Hölder continuous, see, for example (Kaltenbacher & Rundell, 2020, Theorem 2.1).

4. Reformulation of the inverse problem (1.1)–(1.2) as a minimization problem

As pointed out in the introduction, here and in the sequel, we only consider the case of $\alpha(x,t) = \alpha(x)$ in order to have uniqueness in the inverse problem according to Theorem 2.1. Let

$$\mathscr{A} = \{\alpha(x) : \ \alpha \in L^\infty(\Omega),\ 0 < C_1 \le \alpha(x) \le C_2 \text{ a.e.}\}. \tag{4.1}$$

To recover the unknown function α in the inverse problem (1.1)–(1.2) we shall attempt to minimize the following Tikhonov type functional with regularization parameter $\mu \geq 0$,

$$E(\alpha) = \frac{1}{2} \|u(\cdot, T; \alpha) - \psi(\cdot)\|^2_{L^2(\Omega)} + \frac{\mu}{2} \|\alpha(\cdot)\|^2_{L^2(\Omega)}, \tag{4.2}$$

where $\alpha \in \mathscr{A}$ and $u(x, t; \alpha)$ is the weak solution of (1.1) in $L^2(0, T; H^1(\Omega))$ guaranteed by Theorem 3.1. In particular, using integration by parts it follows that the weak solution satisfies,

$$\begin{aligned} &-\int_{\Omega_T} u(x,t)\partial_t\eta(x,t)\,dxdt + \int_{\Omega_T} (D(x)\nabla u(x,t)) \cdot \nabla\eta(x,t)\,dxdt \\ &\qquad = \int_{\Omega_T} (f(u(x,t)) - \alpha(x)u(x,t))\,\eta(x,t)\,dxdt + \int_{\Omega} \varphi(x)\eta(x,0)\,dx \end{aligned} \tag{4.3}$$

for a class of test functions η with $\eta(x, T) = 0$.

We briefly outline that the minimization of (4.2) admits a solution. We point out that (4.2) is well-defined since $L^\infty(\Omega)$ is contained in $L^2(\Omega)$ due to Ω being bounded. It is clear that $\inf_{\alpha\in\mathscr{A}} E(\alpha)$ is finite and non-zero. Thus, there exists a minimizing sequence $\{\alpha_n\}$ such that $\lim_{n\to\infty} E(\alpha_n)$ attains the infimum. Verbatim the arguments given for example in Hao *et al.* (2013); Cao *et al.* (2020) it is possible to extract a subsequence with the limit point being a minimizer of (4.2). We point out that we additionally have a nonlinear term compared with Hao *et al.* (2013); Cao *et al.* (2020) but according to the previous section, (1.1) is a well-posed problem with a standard parabolic estimate of the solution in terms of the data. Thus, the requested estimates needed in (Hao *et al.*, 2013, Thm. 4.5), (Cao *et al.*, 2020, Thm. 3) are still valid in our case.

PROPOSITION 4.1. The problem of minimizating (4.2) over the set $\mathscr{A}$ given by(4.1) and with u being a weak solution of (1.1) has a solution.

Following (Sect. 4 Hao *et al.* (2013)), we show existence of a minimizer together with Lipschitz continuity and Fréchet differentiability of the mapping from α to the corresponding solution $u(\alpha)$, see the Appendix.

The minimization problem (4.2) can be written, to highlight that it is subjected to the constrains of the partial differential equation for u, as

$$\left\{ \begin{array}{l} \min\limits_{\alpha\in\mathscr{A}} \frac{1}{2} \|u(\cdot, T; \alpha) - \psi(\cdot)\|^2_{L^2(\Omega)} + \frac{\mu}{2} \|\alpha(\cdot)\|^2_{L^2(\Omega)} \\ s.t.\ \partial_t u(x,t) - \operatorname{div}(D(x)\nabla u(x,t)) - f(u(x,t)) + \alpha(x)u(x,t) = 0 \end{array} \right. \tag{4.4}$$

subjected to the initial and boundary conditions in (1.1).

We solve (4.4) by an iterative scheme that necessitates the use of the sensitivity function $v(x, t)$ and the Lagrangian multiplier $\lambda(x, t)$ that are obtained from the sensitivity problem and the adjoint problem respectively as shown below.

Specifically, we apply a conjugate gradient method for the minimization, where the descent direction involves the gradient of the functional and will be calculated by an adjoint problem. The step size in the descent direction will be determined by a so-called sensitivity problem

4.1 *The sensitivity problem*

Suppose that the cell density $u(x,t)$ is perturbed by $\varepsilon v(x,t)$ when the therapy parameter $\alpha(x)$ is perturbed by $\varepsilon a(x)$, where $\varepsilon > 0$ is a small number. The perturbed cell density $u(x,t) + \varepsilon v(x,t)$ and perturbed therapy parameter $\alpha(x) + \varepsilon a(x)$ are substituted into the original problem in (1.1) to obtain a perturbed problem. The perturbed problem having the original problem subtracted from it and taking limits yields the so-called sensitivity problem.

LEMMA 4.1. The sensitivity problem corresponding to (1.1) is given by

$$\begin{cases} \partial_t v - \operatorname{div}(D\nabla v) - f'_u v = -\alpha v - au, \text{ in } \Omega_T \\ D\nabla v \cdot n = 0, & \text{on } \partial\Omega_T \\ v(x,0) = 0, & \text{in } \Omega \end{cases} \tag{4.5}$$

where $\Omega_T = \Omega \times (0,T)$, $\partial\Omega_T = \partial\Omega \times (0,T)$ and f'_u is the Fréchet derivative of f at u.

Note that this is a linear problem and well-posedness is classical. The sensitivity problem will be used to determine the step size in the descent direction (corresponding to the parameter a) in the conjugate gradient method.

4.2 *The adjoint problem*

We rewrite the constrained minimization problem (4.4) as an unconstrained minimization using the Lagrange multiplier method:

$$\min_{\alpha \in \mathscr{A}} \mathscr{E}(\alpha)$$

where $\mathscr{E}(\alpha)$ is the Lagrange functional given by

$$\begin{aligned} \mathscr{E}(\alpha) = {} & \frac{1}{2}\,\|u(\cdot,T;\alpha) - \psi(\cdot)\|^2_{L^2(\Omega)} + \frac{\mu}{2}\,\|\alpha(\cdot)\|^2_{L^2(\Omega)} \\ & + \int_{\Omega_T} \lambda(x,t)\Big[f(u(x,t)) + \operatorname{div}(D(x)\nabla u(x,t)) - \partial_t u(x,t) - \alpha(x)u(x,t)\Big]dxdt, \end{aligned} \tag{4.6}$$

subjected to the initial and boundary conditions in (1.1), the element λ will be specified below.

The directional derivative of $\mathscr{E}$ in the direction a is given by

$$\mathscr{E}'_\alpha(a) = \lim_{\varepsilon \searrow 0} \frac{\mathscr{E}(\alpha + \varepsilon a) - \mathscr{E}(\alpha)}{\varepsilon}. \tag{4.7}$$

Direct calculations yield

$$\mathscr{E}(\alpha + \varepsilon a) - \mathscr{E}(\alpha) = \frac{1}{2}\int_\Omega [2\varepsilon((u(x,T) - \psi(x))v(x,T) + \varepsilon^2 v^2(x,T)]\,dx$$

$$+ \frac{\mu}{2}\int_\Omega (2\varepsilon\alpha a + \varepsilon^2 a^2)\,dx$$

$$+\int_{\Omega_T}\lambda(x,t)\Big[f(u(x,t)+\varepsilon v)-f(u(x,t)+\varepsilon\operatorname{div}(D(x)\nabla v)-\varepsilon\partial_t v-\varepsilon(\alpha v+au)-\varepsilon^2 av\Big]\,dxdt.$$

Hence, applying a Green's formula the directional derivative in (4.7) is formally given by

$$\begin{aligned}\mathscr{E}'_\alpha(a) \;=\; &\int_\Omega (u(x,T)-\psi(x))v(x,T)\,dx\\ &+\int_\Omega \lambda(x,0)v(x,0)\,dx-\int_\Omega \lambda(x,T)v(x,T)\,dx\\ &+\int_{\Omega_T}[\lambda(x,t)f'_u(u(\alpha))+\operatorname{div}(D(x)\nabla\lambda(x,t))+\partial_t\lambda(x,t)-\alpha(x)\lambda(x,t)]v\,dxdt\\ &+\int_\Omega\left(\mu\alpha(x)a(x)-\int_0^T a(x)u(x,t)\lambda(x,t)\,dt\right)dx.\end{aligned}$$

Instead of continuing and writing out all the details in finding the directional derivative in (4.7) and the adjoint equation, we make a list of the required steps:

- Use the boundary condition $(D(x)\nabla v)\cdot n(x)=0$ on $\partial\Omega$.
- Restrict the derivations to functions with $v(x,0)=0$.

Carrying out these steps, we finally arrive at

$$\begin{aligned}\mathscr{E}'_\alpha(a) \;=\; &\int_\Omega\left(\mu\alpha(x)-\int_0^T u(x,t)\lambda(x,t)\,dt\right)a(x)\,dx-\int_\Omega\lambda(x,T)v(x,T)\,dx\\ &+\int_\Omega (u(x,t)-\psi(x))v(x,T)+\int_{\Omega_T}\Big[\lambda(x,t)f'_u(u(\alpha))\\ &\qquad+\operatorname{div}(D(x)\nabla\lambda(x,t))+\partial_t\lambda(x,t)-\alpha(x)\lambda(x,t)\Big]v(x,t)\,dxdt.\end{aligned}$$

Hence, it is possible to establish to following result.

THEOREM 4.1. The Fréchet derivative of the Lagrange functional (6) is given by

$$\mathscr{E}'_\alpha=\mu\alpha(x)-\int_0^T u(x,t)\lambda(x,t)\,dt,\quad x\in\Omega, \tag{4.8}$$

where the Lagrange multiplier satisfies the following adjoint problem

$$\begin{cases}\partial_t\lambda(x,t)+\operatorname{div}(D(x)\nabla\lambda(x,t))+(f'_u-\alpha(x))\lambda(x,t)=0, & \text{in } \Omega_T\\ (D(x)\nabla\lambda(x,t))\cdot n(x)=0, & \text{on } \partial\Omega_T\\ \lambda(x,T)=\psi(x)-u(x,T), & \text{in } \Omega.\end{cases} \tag{4.9}$$

As for the sensitivity problem, the adjoint problem is also linear. By change of variables $\tau=T-t$ the adjoint problem can be reformulated as an initial value problem and thus classical well-posedness results can be applied.

Having derived the sensitivity problem and the adjoint, we can then specify a conjugate gradient method for minimizing (4.2).

5. A conjugate gradient method for the inverse problem (1.1)–(1.2)

We propose an iterative method based on the conjugate gradient method for the estimation of α by minimizing $\mathscr{E}(\alpha)$ in (6). A sequence $\{\alpha_k\}$ is generated via

$$\alpha^{(k+1)}(x) = \alpha^{(k)}(x) + \beta^{(k)} d^{(k)}(x), \tag{5.1}$$

where k denotes the number of iterations and $\alpha^{(0)}(x)$ is an initial guess for $\alpha(x)$. The descent direction $d^{(k)}$ and step size $\beta^{(k)}$ have specific values as given below. As mentioned above, the decent direction $d^{(k)}$ will involve the adjoint problem whilst the step size will be determined by the sensitivity problem.

5.1 *Computing the search direction*

The spacewise-dependent descent direction is given by

$$d^{(k)} = \begin{cases} -\mathscr{E}'^{(0)} & \\ -\mathscr{E}'^{(k)} + \gamma^{(k)} d^{(k-1)}, & k = 1, 2, \ldots, \end{cases} \tag{5.2}$$

where for simplicity we introduced the notation $\mathscr{E}'^{(k)} = \mathscr{E}'_{\alpha^{(k)}}$. The Fletcher-Reeves type conjugate gradient coefficients $\gamma^{(k)}$ is specified by

$$\gamma^{(k)} = \frac{\left\| \mathscr{E}'^{(k)} \right\|^2_{L^2(\Omega)}}{\left\| \mathscr{E}'^{(k-1)} \right\|^2_{L^2(\Omega)}}. \tag{5.3}$$

5.2 *Computing the step size*

The search step sizes $\beta^{(k)}$ in (5.1) is obtained by minimizing

$$\begin{aligned} E(\alpha^{(k+1)}) &= \frac{1}{2} \int_\Omega [u(x, T; \alpha^{(k)}(x) + \beta^{(k)} d^{(k)}(x)) - \psi(x)]^2 \, dx \\ &+ \frac{\mu}{2} \int_\Omega (\alpha^{(k)}(x) + \beta^{(k)} d^{(k)}(x))^2 \, dx. \end{aligned} \tag{5.4}$$

Linearizing the functional $E(\alpha^{(k+1)})$ in (5.4) to first order using the Taylor series expansion,

$$\begin{aligned} E(\alpha^{(k+1)}) &\approx \frac{1}{2} \int_\Omega [u(x, T; \alpha^{(k)}) + \beta^{(k)} \delta u(x, T; \alpha^{(k)}) - \psi(x)]^2 \, dx \\ &+ \frac{\mu}{2} \int_\Omega [(\alpha^{(k)}(x))^2 + 2\beta^{(k)} \alpha^{(k)}(x) d^{(k)}(x) + (\beta^{(k)} d^{(k)}(x))^2] \, dx. \end{aligned}$$

Differentiating the right-hand side with respect to β and solving, yields

$$\beta^{(k)} = -\frac{\int_\Omega \left(u(x,T;\alpha^{(k)}) - \psi\right)\delta u(x,T;\alpha^{(k)})\,dx + \mu\int_\Omega \alpha^{(k)}(x)d^{(k)}(x)\,dx}{\int_\Omega \left(\delta u\left(x,T;\alpha^{(k)}\right)\right)^2 dx + \mu\int_\Omega (d^{(k)}(x))^2\,dx}, \tag{5.5}$$

where $v = \delta u$ as in Lemma 4.1.

5.3 *Stopping criterion*

As a stopping rule the following discrepancy principle may be used:

$$E(\alpha^{(k)}) \le \kappa, \tag{5.6}$$

where κ is a small positive number.

We summarize the steps of the minimization procedure.

5.3.1 *Steps of the algorithm.*

(S1). Set $k = 0$ and choose $\mu > 0$ and an initial guess $\alpha^{(0)}(x)$ for the unknown function $\alpha(x)$. Set $E(\alpha^{(0)}) = +\infty$.

(S2). Solve the direct problem in (1.1) to compute $u(x,t;\alpha^{(k)})$.

(S3). If the stopping condition in (5.6) is satisfied, then go to (S7). Else go to (S4).

(S4). Solve the adjoint problem in (4.9) to compute the function $\lambda(x,t;\alpha^{(k)})$ and the gradient $\mathscr{E}_\alpha^{'(k)}$ in (4.8). Compute the conjugate gradient coefficients $\gamma^{(k)}$ in (5.3) and generate $d^{(k)}$.

(S5). Solve the sensitivity problem in (4.5) to compute $\delta u(x,t;\alpha^{(k)}) = v(x,t;\alpha^{(k)})$ by taking $a^{(k)} = d^{(k)}$ and compute the search step size $\beta^{(k)}$ in (5.5).

(S6). Compute $\alpha^{(k+1)}$ via (5.1) and $E(\alpha^{(k+1)})$. Set $k = k+1$ and return to (S2).

(S7). End

5.4 *A Landweber method for the inverse problem (1.1)–(1.2)*

For numerical comparison, we briefly outline iterations based on the more classical Landweber method. The inverse problem under consideration can be recast to find a solution α to

$$A(\alpha) = \psi, \tag{5.7}$$

where $A : L^2(\Omega) \to L^2(\Omega)$ is nonlinear operator.

The solution operator of (1.1) is completely continuous (a weakly converging sequence is turned into a strongly convergent sequence under that operator) viewed as a mapping from $L^2(\Omega)$ to $L^2(0,T;L^2(\Omega))$, see (Precup, 2013, Theorem 10.6 and p. 246) (for a more general result involving also dependence with respect to the coefficients in the equation, see Coclite & Holden (2005)). Since the operator A in (5.7) is a restriction of the solution to (1.1) to a fixed instance in time, we conclude that also A is completely continuous. This implies, according to (Colton & Kress, 2013, Theorem 4.21), that the Fréchet derivative $A'(\alpha)$ is a compact operator. An equation involving a compact linear operator is the prototype of an ill-posed problem. Thus, simply linearizing in (5.7) will not remove the instability.

The Landweber method for solving the operator equation in (5.7) is given by the iterative scheme

$$\alpha_k = \alpha_{k-1} - \gamma (A'(\alpha))^* (A(\alpha_{k-1}) - \psi) \tag{5.8}$$

where $\alpha_0 \in L^2(\Omega)$ is an initial guess and $0 < \gamma < 1/\|A\|^2$. As a stopping rule the discrepancy principle is used. The general definition of the Fréchet derivative is given in, for example, (Zeidler, 1986, Chapter 4.2). It is straightforward to show that the analogue of the Landweber method in (5.8) applied to the inverse problem (1.1) is given by the following iterative scheme.

Let $\alpha_0 > 0$ be arbitrarily. Solve

$$\begin{cases} \partial_t u_1 - \operatorname{div}(D(x)\nabla u_1) - f(u_1) = -\alpha_0 u_1, & \text{in } \Omega \times (0,T) \\ \partial_n u_1 = 0, & \text{on } \partial\Omega \times (0,T) \\ u_1(0) = \varphi, & \text{in } \Omega. \end{cases} \tag{5.9}$$

Assume now that we have constructed u_k, $k = 1, 2 \ldots$. Solve the linear adjoint problem

$$\begin{cases} \partial_t v_k + \operatorname{div}(D(x)\nabla v_k) + f'_u(u_k)v_k = \alpha_1 v_k, & \text{in } \Omega \times (0,T) \\ D\nabla v_k \cdot n = 0, & \text{on } \partial\Omega \times (0,T) \\ v_k(x,T) = u_k(x,T) - \psi(x), & \text{in } \Omega. \end{cases} \tag{5.10}$$

Let

$$\alpha_k = \alpha_{k-1} - \gamma v_k(0).$$

We construct the next approximate solution u_{k+1} by solving the following problem

$$\begin{cases} \partial_t u_{k+1} - \operatorname{div}(D(x)\nabla u_{k+1}) - f(u_{k+1}) = -\alpha_k u_{k+1}, & \text{in } \Omega \times (0,T) \\ D\nabla u_{k+1} \cdot n = 0, & \text{on } \partial\Omega \times (0,T) \\ u_{k+1}(x,0) = \varphi(x), & \text{in } \Omega. \end{cases} \tag{5.11}$$

The iterations continues by repeating the last two steps until a stopping criteria is met.

6. Numerical results

We do a rather direct implementation of the proposed nonlinear conjugate gradient (NCG) method under logistic growth using finite differences. The treatment term will be identified for two different types, one piecewise continuous and one continuous. Ideal reconstructions are not to be expected for this nonlinear inverse ill-posed problem. However, as will be seen, sufficient accuracy can be obtained that can then form the basis of further investigations with more involved regularization techniques. For comparison, we include results with the Landweber method outlined in the previous section.

Model MRI data with chosen parameter values are used to generate synthetic data by applying the forward tumour growth model. This synthetically generated data together with the conjugate gradient scheme above will be used for the inverse problem of reconstructing the treatment parameter. Details and results are given below.

6.1 *Discretization*

Following Jaroudi *et al.* (2019, 2020) we rewrite the reaction–diffusion model (1.1) in the form

$$\partial_t u = \text{div}(D(x)\nabla u) + f^*(u). \tag{6.1}$$

The element $f^*(u) = f(u) - \alpha u = \rho u^a(1-u^b)^c - \alpha u$, where $a, b, c \geq 1$, comprises of the cell proliferation term and the treatment term $\alpha > 0$. The governing equation in (6.1) is discretized and solved iteratively using the scheme

$$u_{i+1} = u_i + hAu_i + hf^*(u_i),$$

where $h > 0$ is the step size, i is the *i-th* iteration. The matrix A is the discretization of the divergence term $\text{div}(D(x)\nabla u)$ as

$$\text{div}(D(x)\nabla u) = \partial_x\left(d_{11}(x,y,z)\partial_x u\right) + \partial_y\left(d_{22}(x,y,z)\partial_y u\right) + \partial_z\left(d_{33}(x,y,z)\partial_z u\right). \tag{6.2}$$

Each of the three terms in (6.2) is approximated using the average of the forward and backward finite-difference operators as in Jaroudi *et al.* (2019, 2020), i.e.

$$\partial_x\left(d_{11}(x,y,z)\partial_x u\right) = \frac{1}{2}\left[\partial_x^+\left(d_{11}(x,y,z)\partial_x^- u\right) + \partial_x^-\left(d_{11}(x,y,z)\partial_x^+ u\right)\right]$$

which is expanded as

$$\begin{aligned}\partial_x\left(d_{11}(x,y,z)\partial_x u\right) \S = &\left(\frac{1}{2}d_{11}(x+1,y,z) + \frac{1}{2}d_{11}(x,y,z)\right)u(x+1,y,z)\\ &-\left(\frac{1}{2}d_{11}(x+1,y,z) + d_{11}(x,y,z) + \frac{1}{2}d_{11}(x-1,y,z)\right)u(x,y,z)\\ &+\left(\frac{1}{2}d_{11}(x,y,z) + \frac{1}{2}d_{11}(x-1,y,z)\right)u(x-1,y,z).\end{aligned}$$

In this derivation, we have used a spatial grid width equal to one, this can be easily adjusted to any width. This adjustment is done in the implementation. The experiments presented below are done with parameters as given in Table 1.

6.2 *Parameters and setup*

In the inverse problem (1.1)–(1.2), we are given two tumours, one before treatment, $\varphi = \psi_{BT}$, see Fig. 1, and a tumour after treatment, $\psi = \psi_{AT}$. The aim is to recover the treatment parameter α in the governing equation in the model (1.1).

The two tumours are generated using the model (1.1). The initial tumour before treatment, ψ_{BT}, is grown by specifying an initial cell density at $t = 0$ and running the model (1.1) forward to a time $T_{ot} > 0$ with the logistic reaction function $f(u) = \rho u(1-u)$ and treatment parameter $\alpha = 0$. The initial cell

TABLE 1 *This table shows the parameters used in the numerical experiments. The values of the exponents a, b and c are given in Tables 2 and 3*

Reaction function			$f = \rho u^a (1 - u^b)^c$
Parameters	Construction	ρ	0.009
		h	0.05
	Reconstruction	ρ	0.007
		h	0.025

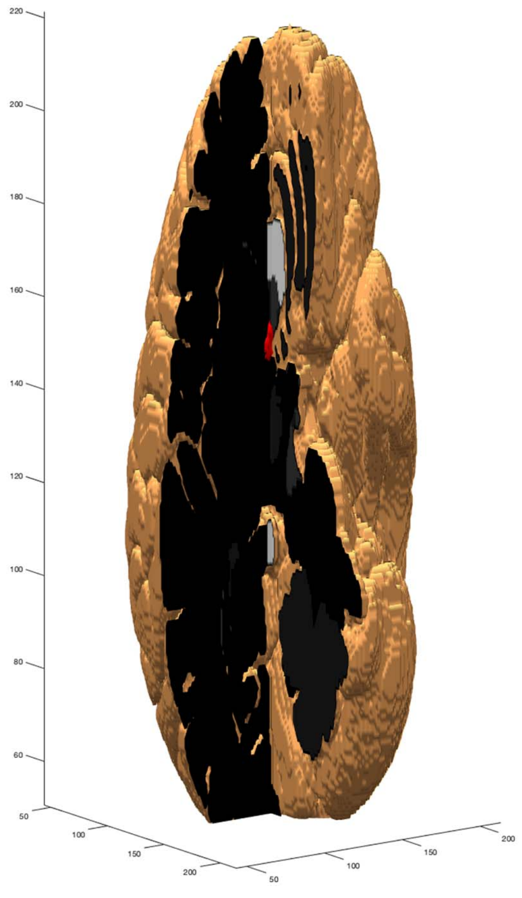


FIG. 1. The brain and the tumour (in red) before treatment.

density is normally distributed with mean value zero and a diagonal covariance matrix. The diffusion parameters in $D(x)$ in the model (1.1) is chosen to be $d_w = 0.5$ and $d_g = 0.25$.

The tumour after treatment, ψ_{AT}, is generated using ψ_{BT} as data at $t = 0$ and running the model (1.1) with the same logistic reaction function but with the treatment parameter varying according to the different experiments as given below.

The treatment term $\alpha(x,t)$ is chosen to be spacewise-dependent only and positive to comply with the uniqueness result of Theorem 2.1 and the well-posedness result Theorem 3.1. The chosen term is one of

$$\alpha_1(x) = \begin{cases} \frac{1}{2} & \text{if } x \in D \\ 0 & \text{otherwise} \end{cases} \tag{6.3}$$

or

$$\alpha_2(x) = \max(0, 1 - \|x-p\|^2/d^2), \tag{6.4}$$

where D is the region where the cell density $\psi_{BT} > 0$. The parameter p is the centre of mass of the tumour ψ_{BT} and $d = 6$.

After a final time $T > 0$ the treatment is terminated and the tumour (cell density) $u(x,T;\alpha)$ is denoted by $\psi_{AT}(x)$ as explained above. The NCG method is then applied to recover the treatment term α given the synthetic data ψ_{BT} and ψ_{AT}. Note that these data are constructed on a coarser mesh according to Table 1, hence the 'inverse crime' is avoided.

We use as reaction term in these numerical experiments one of the functions

$$f_1(u) = \rho u(1-u) \tag{6.5}$$

or

$$f_2(u) = \rho u^{1.2}(1-u^{1.3})^{1.1}. \tag{6.6}$$

We remark that f_1 is used to generate that data, thus using f_2 corresponds to having an error in the model itself.

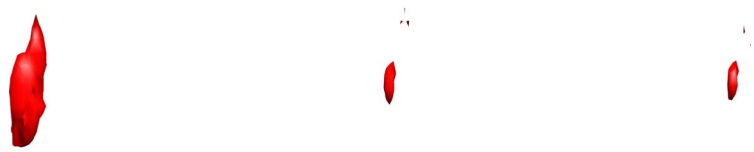
6.3 *Results and error analysis*

In Figure 2 are the 3D images of the two synthetically generated brain tumours $\varphi = \psi_{BT}$ (before treatment) and $\psi = \psi_{AT}$ (after treatment) constructed as described in the previous section, with $T = 3$. Furthermore, in the same figure is the tumour obtained after treatment at time $T_1 = 3.2$ using the reconstructed treatment function $\alpha^{(25)}$ obtained after 25 iterations of the proposed NCG method. Here, the logistic reaction function f_1 in (6.5) is used and $\alpha(x) = \alpha_2$ is as in (6.4).

Note that the time we end the treatment is chosen so we can visibly see, for comparison, some remaining part of the tumour. That is, running the treatment for a sufficiently long time the tumour will eventually vanish.

We point out that, as can be seen from Figure 2, if the treatment parameter α is recovered with a small error then the tumour obtained with this parameter will be close to the tumour with the correct treatment parameter. This is somewhat to be expected since it is possible, building on the steps in the proof of Theorem 2.1, to estimate the difference of two tumours in terms of the parameters of the model.

We then turn to the reconstruction of the treatment parameter itself. Figure 3 shows the error $\|\alpha - \alpha^{(k)}\|_{L^2(\Omega)}$ for $k = 1, 2, \ldots, 25$, with $\alpha^{(k)}$ the approximation obtained in the kth step of the NCG method, and for the two different treatment parameters α given in (6.3) respectively (6.4).

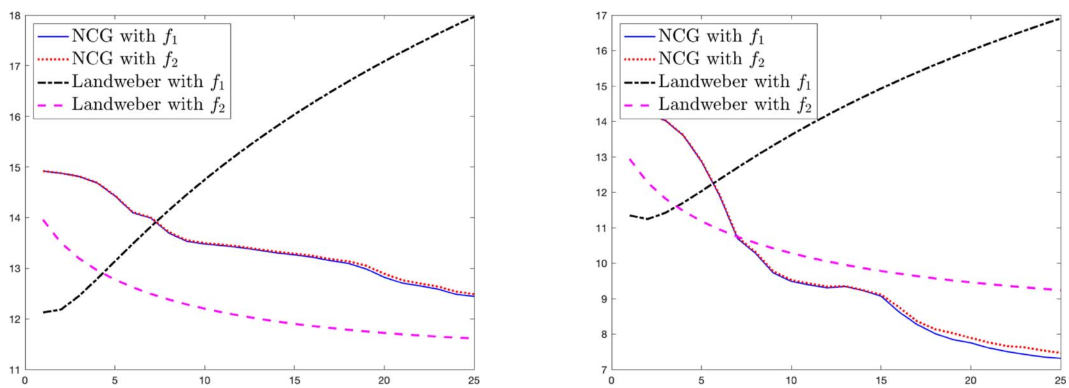


(a) Brain Tumor ψ_{BT}

(b) Brain Tumor ψ_{AT}, true α

(c) Brain Tumor $\psi_{AT}^{(25)}$ after treatment, with recovered α

FIG. 2. The two brain tumours before (a) and after (b) treatment used as data when $\alpha = \alpha_2$ and with the reaction term f_1 together with the tumour (c) obtained from treatment with recovered α. The same zoom factor is used in all figures.



(a) $\alpha = \alpha_1$ piecewise constant

(b) $\alpha = \alpha_2$ being continuous

FIG. 3. The error $\|\alpha - \alpha^{(25)}\|_{L^2(\Omega)}$.

The error keeps decreasing during the iterations for both parameters. It takes longer time to decrease the error with the choice of α being piecewise constant as expected since jumps are more complicated to reconstruct. Also expected is that the error will be less with the correct reaction term f_1, however, it does not make much of a difference using f_2, that is having an error in the model itself.

In Figures 4–11 are the exact parameter α together with the reconstructed parameter $\alpha^{(25)}$ for the two different reaction terms f_1 and f_2 given in (6.5) and (6.6), respectively. The 3D function α is rearranged as a vector, and we select points that are in the vicinity of the tumour. The ill-posedness is exhibited due to oscillations. No choice of the regularization parameter μ or stopping index for the iterations appear to improve the reconstructions much further. Note that we have errors both in the coarseness of the mesh as well as in the model. Thus, continuing to iterate the reconstructions will eventually start to rapidly deteriorate.

An interesting feature is that the reconstructions are of about the same accuracy even when the reaction term f_2 is used corresponding to an error in the model since f_1 is used in generating the synthetic data. We can also see that the reconstructions are more accurate for the continuous treatment parameter α_2.

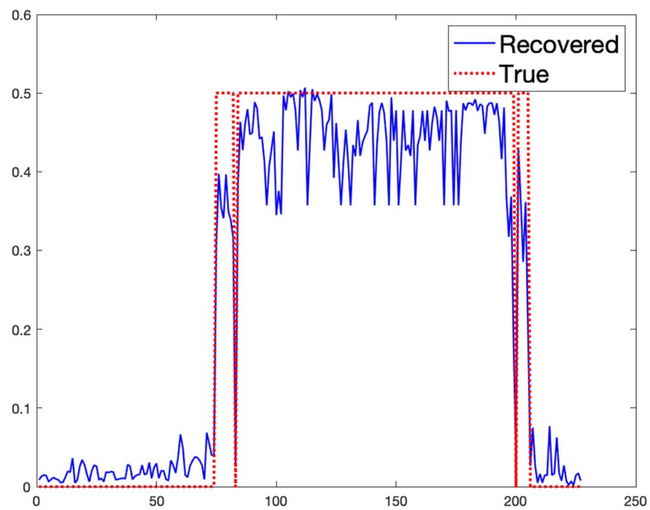


FIG. 4. True α_1 and estimated $\alpha_1^{(25)}$ with f_1 (NCG).

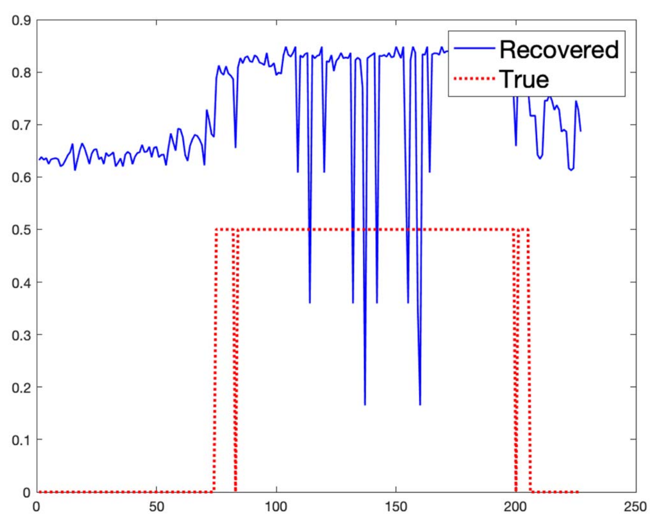


FIG. 5. True α_1 and estimated $\alpha_1^{(25)}$ with f_1 (Landweber).

Since it can be visually difficult to estimate the errors, in Table 2 we give the relative errors for the treatment parameter α as well as errors for the actual tumour itself. Note that the error $\|\psi_{BT} - \psi_{AT}^{(25)}\|_{L^2(\Omega)}$ is included to get an idea of the difference in size of the tumour before and after treatment.

It is clear from the figures and tables that the reconstruction of the treatment term is approaching the correct one. However, the ill-posedness of the nonlinear model makes it difficult to improve the reconstructions much further unless additional special regularization methods and post-processing are

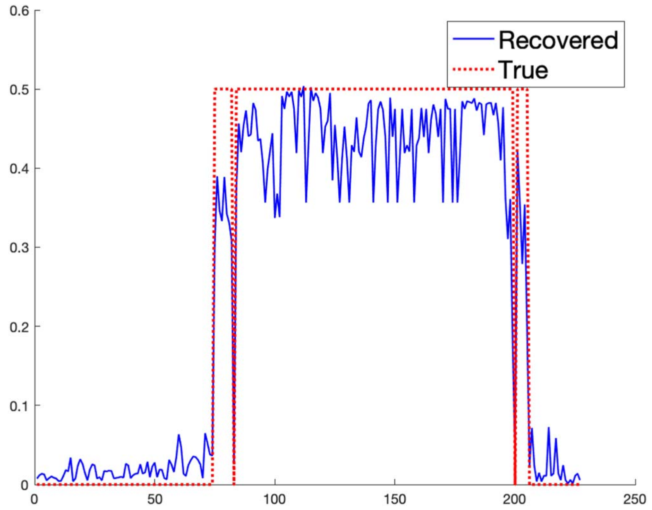


FIG. 6. True α_1 and estimated $\alpha_1^{(25)}$ with f_2 (NCG).

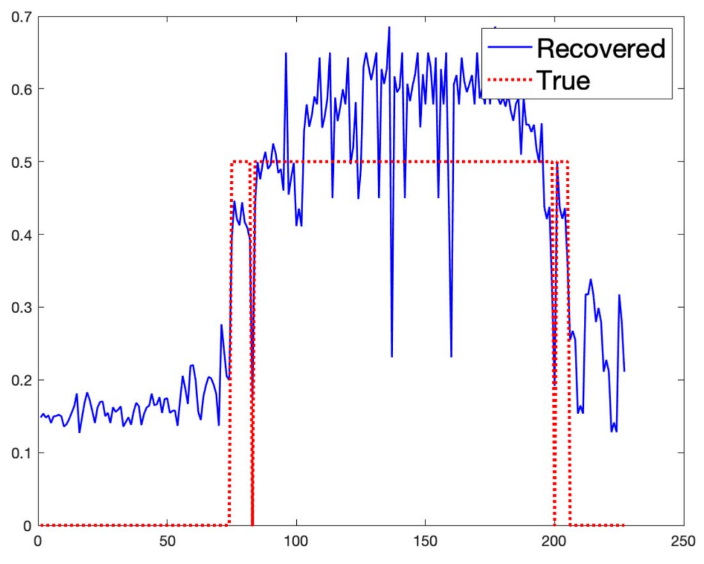


FIG. 7. True α_1 and estimated $\alpha_1^{(25)}$ with f_2 (Landweber).

applied. Since we have errors both due to the mesh size and model, we do not perform additional tests with random errors added into the data. We remark here the recent result Harrach *et al.* (2020) showing how additional errors can be filtered out by repeated measurements.

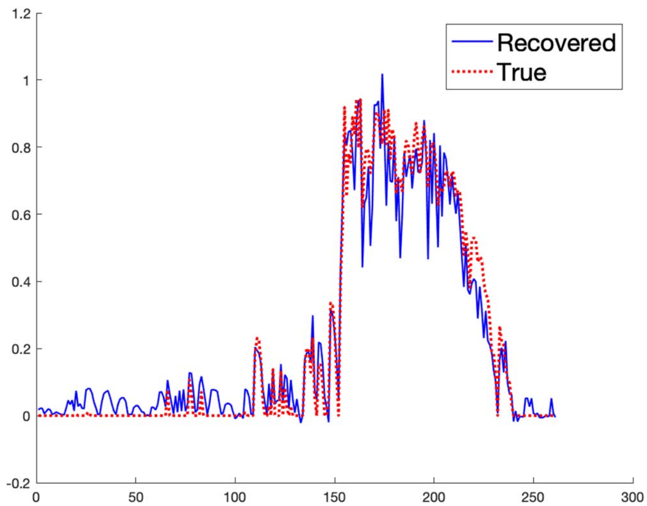


FIG. 8. True α_2 and estimated $\alpha_2^{(25)}$ with f_1 (NCG).

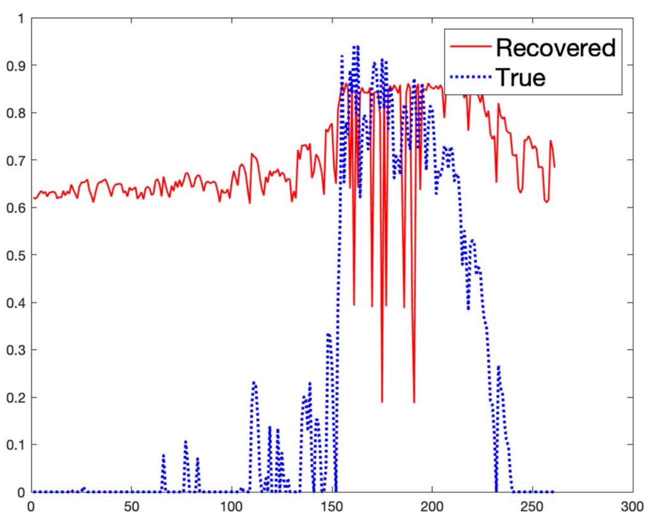


FIG. 9. True α_2 and estimated $\alpha_2^{(25)}$ with f_1 (Landweber).

7. Conclusion

A nonlinear conjugate gradient method has been proposed and investigated for the nonlinear inverse problem of identifying a treatment parameter in a tumour model. Data are the tumour before and after treatment. Uniqueness of a solution to the inverse problem is shown together with well-posedness of the forward models. Finding the treatment parameter is recast as a minimization problem of a Tikhonov type functional. It is shown that this functional has a minimum. Numerical experiments were carried out

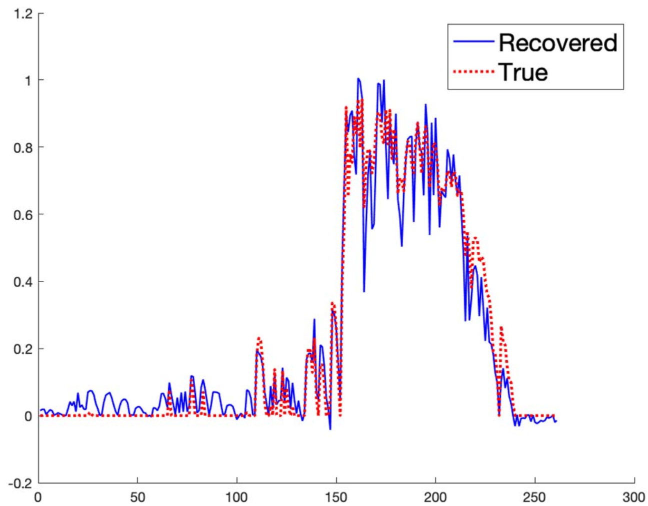


FIG. 10. True α_2 and estimated $\alpha_2^{(25)}$ with f_2 (NCG).

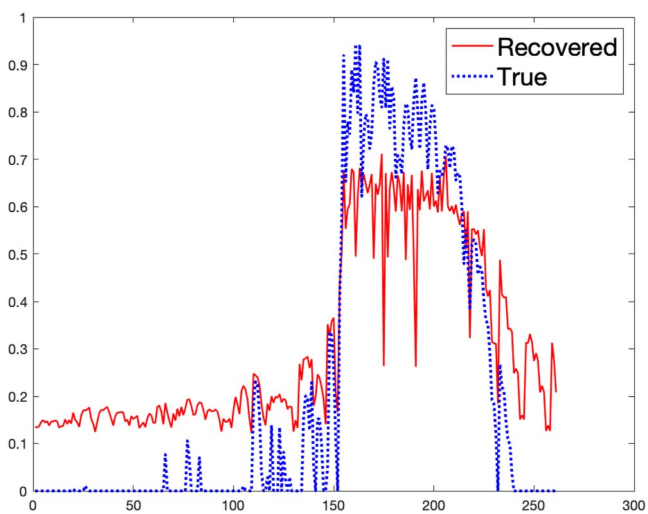


FIG. 11. True α_2 and estimated $\alpha_2^{(25)}$ with f_2 (Landweber).

using a finite-difference scheme together with synthetic MRI data for full 3D tumours. The proposed method can recover information of both piecewise continuous respectively continuous treatment terms also in the case of noise into the model. From the figures and tables, it is clear that the NCG method produces (for most cases) more accurate results than Landweber with fewer iterations. We have not optimized the parameters and in this sense, there is possibility for further improvements.

TABLE 2 *Algorithm performance analysis after 25 iterations with piecewise constant α_1*

Reaction function used in reconstruction	**NCG** $\frac{\|\alpha-\alpha_L^{(25)}\|_{L^2(\Omega)}}{\|\alpha\|_{L^2(\Omega)}}$	$\|\psi_{AT}-\psi_{NCG}^{(25)}\|_{L^2(\Omega)}$	$\|\psi_{BT}-\psi_{NCG}^{(25)}\|_{L^2(\Omega)}$
$f_1(u)=\rho u(1-u)$	0.8324	0.1010	1.2564
$f_2(u)=\rho u^{1.2}(1-u^{1.1})^{1.2}$	0.8354	0.1032	1.2546
	Landweber $\frac{\|\alpha-\alpha_L^{(25)}\|_{L^2(\Omega)}}{\|\alpha\|_{L^2(\Omega)}}$	$\|\psi_{AT}-\psi_L^{(25)}\|_{L^2(\Omega)}$	$\|\psi_{BT}-\psi_L^{(25)}\|_{L^2(\Omega)}$
$f_1(u)=\rho u(1-u)$	1.2020	1.8764	0.6846
$f_2(u)=\rho u^{1.2}(1-u^{1.1})^{1.2}$	0.7766	1.3002	0.3195

TABLE 3 *Algorithm performance analysis after 25 iterations for polynomial α_2*

Reaction function used in reconstruction	**NCG** $\frac{\|\alpha-\alpha_L^{(25)}\|_{L^2(\Omega)}}{\|\alpha\|_{L^2(\Omega)}}$	$\|\psi_{AT}-\psi_{NCG}^{(25)}\|_{L^2(\Omega)}$	$\|\psi_{BT}-\psi_{NCG}^{(25)}\|_{L^2(\Omega)}$
$f_1(u)=\rho u(1-u)$	0.5085	0.0391	1.5339
$f_2(u)=\rho u^{1.2}(1-u^{1.1})^{1.3}$	0.5193	0.0454	1.5391
	Landweber $\frac{\|\alpha-\alpha_L^{(25)}\|_{L^2(\Omega)}}{\|\alpha\|_{L^2(\Omega)}}$	$\|\psi_{AT}-\psi_L^{(25)}\|_{L^2(\Omega)}$	$\|\psi_{BT}-\psi_L^{(25)}\|_{L^2(\Omega)}$
$f_1(u)=\rho u(1-u)$	1.1759	2.0219	0.6301
$f_2(u)=\rho u^{1.2}(1-u^{1.1})^{1.2}$	0.6425	1.4303	0.3582

Funding

Sida-bilateral program with Makerere University in Uganda phase IV 2015–2020 project 316 capacity building in mathematics and its applications.

REFERENCES

AMIR, G., ANDREAS, M. & GEORGE, B. (2016) An inverse problem formulation for parameter estimation of a reaction-diffusion model of low grade gliomas. *J. Math. Biol.*, **72**, 409–433.

CAO, K., LESNIC, D. & LIU, J. (2020) Simultaneous reconstruction of space-dependent heat transfer coefficients and initial temperature. *J. Comput. Appl. Math.*, **375**, 112800.

CHOULLI, M. (1994) An inverse problem for a semilinear parabolic equation. *Inverse Problems*, **10**, 1123–1132.

CIARLET, P. G. (1978) *The Finite Element Method for Elliptic Problems*. Amsterdam: North Holland.

COCLITE, G. M. & HOLDEN, H. (2005) Stability of solutions of quasilinear parabolic equations. *J. Math. Anal. Appl.*, **308**, 221–239.

COLTON, D. & KRESS, R. (2013) *Inverse Acoustic and Electromagnetic Scattering*, 3rd edn. New York, Cham: Springer-Verlag.

DOLGUSHIN, M., KORNIENKO, V. & PRONIN, I. (2017) *Brain Metastases: Advanced Neuroimaging*. Springer.

HAO, D. N., THANH, P. X. & LESNIC, D. (2013) Determination of the heat transfer coefficients in transient heat conduction. *Inverse Probl.*, **29**, 095020.

HARRACH, B., JAHN, T. & POTTHAST, R. (2020) Beyond the Bakushinkii veto: regularising linear inverse problems without knowing the noise distribution. *Numer. Math.*, **145**, 581–603.

HASANOV, A. (2007) Simultaneous determination of source terms in a linear parabolic problem from the final overdetermination: weak solution approach. *J. Math. Anal. Appl.*, **330**, 766–779.

HUNTUL, M., LESNIC, D. & JOHANSSON, B. T. (2018) Determination of an additive time-and space-dependent coefficient in the heat equation. *Internat. J. Numer. Methods Heat Fluid Flow*, **28**, 1352–1373.

HUSSEIN, M., LESNIC, D., KAMYNIN, V. L. & KOSTIN, A. B. (2020) Direct and inverse source problems for degenerate parabolic equations. *J. Inverse Ill-Posed Probl.*, **28**, 425–448.

ISAKOV, V. (2017) *Inverse Problems for Partial Differential Equations* (B. Rosenhahn & B. Andres eds). 3rd edn. Cham: Springer-Verlag.

JAROUDI, R., BARAVDISH, G., ÅSTRÖM, F. & JOHANSSON, B. T. (2016) Source localization of reaction-diffusion models for brain tumors. *Lecture Notes in Computer Science*. Springer International Publishing, pp. 414–425.

JAROUDI, R., BARAVDISH, G., JOHANSSON, B. T. & ÅSTRÖM, F. (2019) Numerical reconstruction of brain tumours. *Inverse Probl. Sci. Eng.*, **27**, 278–298.

JAROUDI, R., ÅSTRÖM, F., JOHANSSON, B. T. & BARAVDISH, G. (2020) Numerical simulations in 3-dimensions of reaction-diffusion models for brain tumour growth. *Int. J. Comput. Math.*, **97**, 1151–1169.

JOHANSSON, B. T. & LESNIC, D. (2007a) A variational method for identifying a spacewise-dependent heat source. *IMA J. Appl. Math.*, **72**, 748–760.

JOHANSSON, T. & LESNIC, D. (2007b) Determination of a spacewise dependent heat source. *J. Comput. Appl. Math.*, **209**, 66–80.

KAČUR, J. (1985) *Method of Rothe in Evolution Equations*. Leipzig: B. G. Teubner.

KALTENBACHER, B. & RUNDELL, W. (2019) On the identification of a nonlinear term in a reaction–diffusion equation. *Inverse Probl.*, **35**, 115007.

KALTENBACHER, B. & RUNDELL, W. (2020) On the simultaneous recovery of the conductivity and the nonlinear reaction term in a parabolic equation. *Inverse Probl. Imaging*, **14**, 939–966.

KERIMOV, N. B. & ISMAILOV, M. I. (2015) Direct and inverse problems for the heat equation with a dynamic-type boundary condition. *IMA J. Appl. Math.*, **80**, 1519–1533.

KLIBANOV, M. V. (2013) Carleman estimates for global uniqueness, stability and numerical methods for coefficient inverse problems. *J. Inverse Ill-Posed Probl.*, **21**, 477–560.

LADYŽENSKAJA, O. A., SOLONNIKOV, V. A. & URAL'CEVA, N. N. (1968) *Linear and Quasilinear Equations of Parabolic Type*, vol. **23**. Providence, R.I.: American Mathematical Society, p. xi+648 Translated from the Russian by S. Smith.

MANG, A., GHOLAMI, A., DAVATZIKOS, C. & BIROS, G. (2018) PDE-constrained optimization in medical image analysis. *Optim. Eng.*, **19**, 765–812.

MANG, A., BAKAS, S., SUBRAMANIAN, S., DAVATZIKOS, C. & BIROS, G. (2020) Integrated biophysical modeling and image analysis: application to neuro-oncology. *Annu. Rev. Biomed. Eng.*, **22**, 309–341.

MARTÍN-LANDROVE, M. (2017) Reaction-diffusion models for glioma tumor growth. arXiv preprint arXiv:1707.09409. pp. 1–41.

MURRAY, J. D. (2002) *Mathematical Biology II. Spatial models and biomedical applications*. Springer-Verlag New York, Inc.

NGUYEN, H. T., KHOA, V. A. & VO, V. A. (2019) Analysis of a quasi-reversibility method for a terminal value quasi-linear parabolic problem with measurements. *SIAM J. Math. Anal.*, **51**, 60–85.

PAZY, A. (1983) *Semigroups of Linear Operators and Applications to Partial Differential Equations*. New York: Springer-Verlag.

POWATHIL, G., KOHANDEL, M., SIVALOGANATHAN, S., OZA, A. & MILOSEVIC, M. (2007) Mathematical modeling of brain tumors: effects of radiotherapy and chemotherapy. *Phys. Med. Biol.*, **52**, 3291.

PRECUP, R. (2013) *Linear and Semilinear Partial Differential Equations*. Berlin: Walter de Gruyter.

PRILEPKO, A. I. & KOSTIN, A. B. (1993) On certain inverse problems for parabolic equations with final and integral observation. *Russian Acad. Sci. Sbornik Math.*, **75**, 473–490.

PRILEPKO, A. I., ORLOVSKY, D. G. & VASIN, I. A. (2000) *Methods for Solving Inverse Problems in Mathematical Physics*. New York: CRC Press.

ROCKNE, R., ALVORD, E., ROCKHILL, J. & SWANSON, K. (2009) A mathematical model for brain tumor response to radiation therapy. *J. Math. Biol.*, **58**, 561.

ROCKNE, R., ROCKHILL, J., MRUGALA, M., SPENCE, A., KALET, I., HENDRICKSON, K., LAI, A., CLOUGHESY, T., ALVORD Jr., E. & SWANSON, K. (2010) Predicting the efficacy of radiotherapy in individual glioblastoma patients in vivo: a mathematical modeling approach. *Phys. Med. Biol.*, **55**, 3271.

ROUBÍČEK, T. (2005) *Nonlinear Partial Differential Equations with Applications*, vol. **153**. Basel: Birkhäuser Verlag, p. xviii+405.

RUNDELL, W. (1980) Determination of an unknown non-homogeneous term in a linear partial differential equation from overspecified boundary data. *Applicable Anal.*, **10**, 231–242.

RUNDELL, W. (1987) The determination of a parabolic equation from initial and final data. *Proc. Amer. Math. Soc.*, **99**, 637–642.

SABIR, S. & RAISSI, N. (2019) Analysis of tumor/effector cell dynamics and decision support in therapy. *Trends in Biomathematics: Mathematical Modeling for Health, Harvesting, and Population Dynamics* (R.P. Mondaini ed). Cham: Springer, pp. 153–164.

SLODIČKA, M. (2020) Uniqueness for an inverse source problem of determining a space dependent source in a non-autonomous parabolic equation. *Appl. Math. Lett.*, **107**, 106395.

STAMATAKOS, G. S. & GIATILI, S. G. (2017) A numerical handling of the boundary coditions imposed by the skull on an inhomogeneous diffusion-reaction model of glioblastoma invasion into the brain: clinical validation aspects. *Cancer Informatics*, **16**, 117693511668482.

SWANSON, K. R., ALVORDJr., E. C. & MURRAY, J. (2000) A quantitative model for differential motility of gliomas in grey and white matter. *Cell Prolif.*, **33**, 317–329.

SWANSON, K. R., ALVORD, E. C. & MURRAY, J. (2002) Virtual brain tumours (gliomas) enhance the reality of medical imaging and highlight inadequacies of current therapy. *Br. J. Cancer*, **86**, 14–18.

SWANSON, K. R., BRIDGE, C., MURRAY, J. & ALVORDJr., E. C. (2003) Virtual and real brain tumors: using mathematical modeling to quantify glioma growth and invasion. *J. Neurol. Sci.*, **216**, 1–10.

VAN BOCKSTAL, K. & MARIN, L. (2017) Recovery of a space-dependent vector source in anisotropic thermoelastic systems. *Comput. Methods Appl. Mech. Engrg.*, **321**, 269–293.

WEIDEMAIER, P. (2002) Maximal regularity for parabolic equations with inhomogeneous boundary conditions in Sobolev spaces with mixed L_p-norm. *Electron. Res. Announc. Amer. Math. Soc.*, **8**, 47–51.

YAMAMOTO, M. & ZOU, J. (2001) Simultaneous reconstruction of the initial temperature and heat radiative coefficient. *Inverse Probl.*, **17**, 1181.

ZEIDLER, E. (1986) *Nonlinear Functional Analysis and its Applications*, vol. **I**. New York: Springer-Verlag.

A. Appendix

We shall give details for a proof of Proposition 4.1 of Section 4 that the minimization of (4.2) has a solution. We follow closely the proofs of the corresponding results of (Hao *et al.*, 2013, Sect. 4) given for a linear governing equation and adjust the arguments here to our nonlinear equation (1.1). There will be three results proved before we at the end of this appendix give a proof of the existence of a minimizer to (4.2).

Define the anisotropic Sobolev space

$$W(0,T) = \{u: \ u \in L^2(0,T;H^1(\Omega)), \ u_t \in L^2(0,T;L^2(\Omega))$$

with the standard inner product and norm.

We prove two results on properties of the mapping of the parameter α to the corresponding weak solution $u(\alpha)$.

LEMMA A.1. The mapping $\alpha \mapsto u(\alpha)$ is Lipschitz continuous from $\mathscr{A}$ to $W(0,T)$, i.e., for any two elements α and $\alpha + \delta\alpha$, both belonging to $\mathscr{A}$, there holds

$$\|u(\alpha+\delta\alpha) - u(\alpha)\|_{W(0,T)} \leq C\|\delta\alpha\|_{L^\infty(\Omega)}.$$

Proof. Let $u(\alpha+\delta\alpha)$ and $u(\alpha)$ be the solutions to

$$\begin{cases} \partial_t u(\alpha+\delta\alpha) - \operatorname{div}\,(D(x)\nabla u(\alpha+\delta\alpha)) - f(u(\alpha+\delta\alpha)) = -(\alpha+\delta\alpha)u(\alpha+\delta\alpha), \\ D(x)\nabla u(\alpha+\delta\alpha)\cdot n = 0, \\ u(\alpha+\delta\alpha)(0) = \varphi, \end{cases} \tag{A.1}$$

and

$$\begin{cases} \partial_t u(\alpha) - \operatorname{div}\,(D(x)\nabla u(\alpha)) - f(u(\alpha)) = -\alpha u(\alpha), \\ D(x)\nabla u(\alpha)\cdot n = 0, \\ u(\alpha)(0) = \varphi, \end{cases} \tag{A.2}$$

respectively. We rewrite the right-hand side in the governing equation for $u(\alpha)$ into

$$\begin{cases} \partial_t u(\alpha) - \operatorname{div}\,(D(x)\nabla u(\alpha)) - f(u(\alpha)) = -(\alpha+\delta\alpha)u(\alpha) + \delta\alpha u(\alpha), \\ D(x)\nabla u(\alpha)\cdot n = 0, \\ u(\alpha)(0) = \varphi. \end{cases} \tag{A.3}$$

Consider then the difference $\nu = u(\alpha+\delta\alpha) - u(\alpha)$. Then ν satisfies the problem

$$\begin{cases} \partial_t \nu - \operatorname{div}\,(D(x)\nabla \nu) - [f(u(\alpha+\delta\alpha)) - f(u(\alpha))] = -(\alpha+\delta\alpha)\nu - \delta\alpha u(\alpha) \\ D(x)\nabla \nu\cdot n = 0, \\ \nu(0) = 0. \end{cases} \tag{A.4}$$

Note that

$$f(u(\alpha+\delta\alpha)) - f(u(\alpha)) = \rho[(u(\alpha+\delta\alpha)(1-u(\alpha+\delta\alpha)) - u(\alpha)(1-u(\alpha))]$$

$$= \rho[\nu - \nu(u(\alpha+\delta\alpha) + u(\alpha))] = \rho\nu(1-u(\alpha+\delta\alpha) - u(\alpha)),$$

thus the solution ν satisfies

$$\begin{cases} \partial_t \nu - \operatorname{div}(D(x)\nabla\nu) - \rho(1-u(\alpha+\delta\alpha)-u(\alpha))\nu = -(\alpha+\delta\alpha)\nu - \delta\alpha u(\alpha), \\ D(x)\nabla\nu\cdot n = 0, \\ \nu(0) = 0, \end{cases} \tag{A.5}$$

that is

$$\begin{cases} \partial_t \nu - \operatorname{div}(D(x)\nabla\nu) - [-(\alpha+\delta\alpha)+\rho(1-u(\alpha+\delta\alpha)-u(\alpha))]\nu = -\delta\alpha u(\alpha), \\ D(x)\nabla\nu\cdot n = 0, \\ \nu(0) = 0. \end{cases} \tag{A.6}$$

This is a linear parabolic problem. Due to the L^∞-smoothness of the parameter α and initial data, and the corresponding smoothness of the solutions $u(\alpha+\delta\alpha)$ and $u(\alpha)$ (see (Roubíček, 2005, Table 3, p. 253) for expected smoothness of a weak solution to (1.1)) standard estimates can be applied for parabolic equations with coefficients in Sobolev spaces see (Ladyženskaja *et al.*, 1968, Chpt III, Thm. 5.1) and for a more recent account (Weidemaier, 2002, Thm. 3.2). Hence,

$$\|\nu\|_{W(0,T)} \le C_\alpha \|\delta\alpha\|_{L^\infty(\Omega)} \|u(\alpha)\|_{W(0,T)}.$$

Since $\|u(\alpha)\|_{W(0,T)}$ can be estimated by the initial data term $\|\varphi\|_{L^2(\Omega)}$ and α is bounded whenever $\alpha \in \mathscr{A}$ (rendering $C_\alpha \le C$) the proof is complete. □

LEMMA A.2. The mapping $\alpha \mapsto u(\alpha)$ from $\mathscr{A}$ to $W(0,T)$ is Fréchet differentiable, that is for any $\delta\alpha \in L^\infty(\Omega)$ such that $\alpha+\delta\alpha \in \mathscr{A}$ there exists a bounded linear operator $\mathscr{U}$ from $\mathscr{A}$ to $W(0,T)$ such that

$$\lim_{\|\delta\alpha\|_{L^\infty(\Omega)}\to 0} \frac{\|u(\alpha+\delta\alpha)-u(\alpha)-\mathscr{U}\delta\alpha\|_{W(0,T)}}{\|\delta\alpha\|_{L^\infty(\Omega)}} = 0. \tag{A.7}$$

Proof. We explicitly construct U as the solution to the linear parabolic problem

$$\begin{cases} \partial_t U - \operatorname{div}(D(x)\nabla U) - [-\alpha+\rho(1-2u(\alpha))]U = -\delta\alpha u(\alpha), & \text{in } \Omega\times(0,T) \\ D(x)\nabla U\cdot n = 0, & \text{on } \partial\Omega\times(0,T) \\ U(x,0) = 0, & \text{in } \Omega \end{cases} \tag{A.8}$$

where $\delta\alpha \in L^\infty(\Omega)$ and $\alpha+\delta\alpha \in \mathscr{A}$. As remarked in the previous proof, due to the Sobolev smoothness of the coefficients, there exists a unique solution $U \in W(0,T)$. Moreover, the map from $\delta\alpha \in L^\infty(\Omega)$ to $U \in W(0,T)$ defines a bounded linear operator $\mathscr{U}$. Let us show that $\mathscr{U}$ indeed qualify to be used in the definition of Fréchet differentiability of the mapping $\alpha \mapsto u(\alpha)$.

Put $w = v - U$ with $v = u(\alpha + \delta\alpha) - u(\alpha)$. Since v satisfies (A.6) and U satisfies (A.8), the element w is a solution to

$$\begin{cases} \partial_t w - \operatorname{div}(D(x)\nabla w) + \alpha w - \rho(1 - 2u(\alpha))w = -\delta\alpha v - \rho v^2, & \text{in } \Omega \times (0,T) \\ D(x)\nabla w \cdot n = 0, & \text{on } \partial\Omega \times (0,T) \\ w(x,0) = 0 & \text{in } \Omega. \end{cases} \tag{A.9}$$

Applying a standard parabolic estimate of the solution in terms of the right-hand side, the element w can be estimated as

$$\|w\|_{W(0,T)} \le C\|\delta\alpha\|^2_{L^\infty(\Omega)},$$

where we in estimate applied Lemma A.1. Dividing by $\|\delta\alpha\|_{L^\infty(\Omega)}$ and taking the limit as $\|\delta\alpha\|_{L^\infty(\Omega)} \to 0$, we obtain (A.7). Thus, the result is proved. □

We then turn to the existence of a minimizer of (4.2) and need the following (recall that weak convergence in L^∞ is denoted weakly-*):

LEMMA A.3. Let $\{\alpha^n\} \subset \mathscr{A}$ be a sequence converging to α^* weakly-* in $L^\infty(\Omega)$. Then the sequence $\{u(\alpha^n)\}$ converges weakly to $u(\alpha^*)$ in the space $W(0,T)$.

Proof. Since $\{\alpha^n\}$ converges weakly-*, the sequence $\{\alpha^n\}$ is in particular bounded. By Theorem 3.1, the solution $u(\alpha^n)$ can in turn be bounded by α^n. Hence, the sequence $\{u(\alpha^n)\}$ is uniformly bounded in $W(0,T)$. Therefore, there is a subsequence, denoted again by $\{u(\alpha^n)\}$, which converges weakly to an element $u^* \in W(0,T)$. We shall show that u^* is a weak solution to (1.1) with parameter α^*.

A weak solution to (1.1) in $W(0,T)$ with $\alpha = \alpha^n$ satisfies $u(x,0;\alpha^n) = \varphi(x)$ and similar to (4.3) for any $v \in L^2(0,T;H^1(\Omega))$:

$$\int_{\Omega_T} u_t(\alpha^n)v\,dxdt + \int_{\Omega_T} D\nabla u(\alpha^n)\cdot\nabla v\,dxdt = \int_{\Omega_T} (f(u(\alpha^n)) - \alpha^n u(\alpha^n))v\,dxdt.$$

Since $\{u(\alpha^n)\}$ converges weakly to $u^* \in W(0,T)$, we have

$$\lim_{n\to\infty}\int_{\Omega_T} u_t(\alpha^n)v\,dxdt = \int_{\Omega_T} u_t^* v\,dxdt$$

and

$$\lim_{n\to\infty}\int_{\Omega_T} D\nabla u(\alpha^n)\cdot\nabla v\,dxdt = \int_{\Omega_T} D\nabla u^*\cdot\nabla v\,dxdt.$$

The space $W(0,T)$ is compactly imbedded into $L^2(\Omega \times (0,T))$, therefore the sequence $u(\alpha^n)$ converges to u^* strongly in $L^2(\Omega \times (0,T))$. This together with the explicit expression of the function f imply

$$\lim_{n\to\infty}\int_{\Omega_T} f(u(\alpha^n))v\,dxdt = \int_{\Omega_T} f(u^*)v\,dxdt$$

(we remark that a general nonlinear function of a weakly converging sequence is not necessary weakly converging). The weak-* convergence of $\{\alpha^n\}$ to $\alpha^* \in L^\infty(\Omega)$ implies weak convergence in $L^2(\Omega)$ since Ω is bounded. This together with the convergence of $\{u(\alpha^n)\}$ to $u^* \in L^2(\Omega \times (0,T))$ render

$$\lim_{n\to\infty} \int_{\Omega_T} \alpha^n u(\alpha^n) v \, dxdt = \lim_{n\to\infty} \int_{\Omega_T} [\alpha^n u^* v - \alpha^n (u^* - u(\alpha^n))] v \, dxdt = \int_{\Omega_T} \alpha^* u^* v \, dxdt,$$

where we used that the product of a weakly converging sequence and a strongly converging sequence in L^2 converges in L^1.

In total, we can conclude that for any $v \in L^2(0,T; H^1(\Omega))$

$$\int_{\Omega_T} u_t^* v + \int_{\Omega_T} D\nabla u^* \cdot \nabla v \, dxdt = \int_{\Omega_T} (f(u^*) - \alpha^* u^*) v \, dxdt.$$

Since u^n converges weakly to u^* in $W(0,T)$ and $W(0,T)$ is compactly imbedded in the space $C([0,T]; (H^1(\Omega))')$, we have that $u^n(\cdot,0)$ converges strongly to $u^*(\cdot,0)$ in $(H^1(\Omega))'$. Hence, $u^*(\cdot,0) = \varphi(x)$. Thus, we have obtained that the limit u^* is a weak solution to (1.1) with parameter α^*.

Taking any other weakly converging subsequence of $\{u(\alpha^n)\}$ the above arguments render that the limit element is a solution to (1.1) with parameter α^*. Due to the uniqueness of a solution of the direct problem (1.1) the limit functions are one and the same, denoted $u^* = u(\alpha^*)$. Thus, since every subsequence of $\{u(\alpha^n)\}$ has a further subsequence (due to the boundedness) converging to the same limit, the whole sequence $\{u(\alpha^n)\}$ then converges weakly. □

We can then prove the existence of a minimizer to (4.2).

THEOREM A.1. The minimization problem (4.2) subject to (1.1) admits a solution.

Proof. Since $E_\mu(\alpha)$ in (4.2) is finite over $\mathscr{A}$, there exists a minimizing sequence $\{\alpha^n\} \subset \mathscr{A}$ such that

$$\lim_{n\to\infty} E_\mu(\alpha^n) = \inf_{\alpha\in\mathscr{A}} E_\mu(\alpha).$$

Furthermore, since $\alpha \in \mathscr{A}$, and $\mathscr{A}$ is weak-* closed, there is a subsequence of $\{\alpha^n\}$, denoted by the same symbol, and an element $\alpha^* \in \mathscr{A}$ such that $\alpha^n \to \alpha^*$ weak-* in $L^\infty(\Omega)$. The element α^* is a natural candidate for a minimizer. Let us show that this element is indeed a minimizer.

Lemma A.3 guarantees the weak convergence of $u(\alpha^n)$ to $u(\alpha^*)$ in the space $W(0,T)$, where $u(\alpha^*)$ is a weak solution to (1.1). From the compactness of the imbedding $W(0,T) \hookrightarrow L^2(\Omega \times (0,T))$, we find that $u(\alpha^n)$ converges strongly to $u(\alpha^*)$ in $L^2(\Omega \times (0,T))$. Hence,

$$\lim_{n\to\infty} \|u(\cdot,T;\alpha^n) - \psi\|_{L^2(\Omega)} = \|u(\cdot,T;\alpha^*) - \psi\|_{L^2(\Omega)}.$$

The weak-* convergence in $L^\infty(\Omega)$ of $\{\alpha^n\}$ implies in particular weak convergence of that sequence in $L^2(\Omega)$, and it therefore follows directly from the weak lower semi-continuity of the norms, together with the previous equality, that

$$\lim_{n\to\infty} E_\mu(\alpha^n) = \frac{1}{2} \left\| u\left(\cdot,T;\alpha^*\right) - \psi(\cdot) \right\|^2_{L^2(\Omega)} + \frac{\mu}{2} \left\| \alpha^*(\cdot) \right\|^2_{L^2(\Omega)}.$$

Thus, α^* is a minimizer of the functional E_μ over the admissible set $\mathscr{A}$. □