Numer. Math. 59, 473–509 (1991)

Numerische Mathematik

Estimation of a temporally and spatially varying diffusion coefficient in a parabolic system by an augmented Lagrangian technique*

K. Kunisch[1] **and G. Peichl**[2]

[1] Institut für Mathematik, Technische Universität, Kopernikusgasse 24, A-8010 Graz, Austria
[2] Institut für Mathematik, Karl-Franzens-Universität Graz, Heinrichstr. 36, A-8010 Graz, Austria

Received August 7, 1990

Summary. In this paper we apply a hybrid method to estimate a temporally and spatially varying diffusion coefficient in a parabolic system. This technique combines the output-least-squares- and the equation error method. The resulting optimization problem is solved by an augmented Lagrangian approach and convergence as well as rate of convergence proofs are provided. The stability of the estimated coefficient with respect to perturbations in the observation is guaranteed.

Mathematics Subject Classification (1991): 65K10, 35R30, 49B22, 49D29

1 Introduction

This paper is devoted to the development of an estimation procedure for the diffusion coefficient a in

$$(1.1)\qquad \begin{aligned} &u_t = (a(t,x)u_x)_x + f(t,x), && (t,x)\in(0,T]\times(0,1)\,,\\ &u(0,x) = \varphi(x), && x\in(0,1)\,,\\ &u(t,0) = u(t,1) = 0, && t\in(0,T]\,, \end{aligned}$$

from knowledge of the state u in $Q = (0, T]\times(0,1)$. If only point observations of u in Q are available, then we assume that an interpolation of these pointwise data has been carried out to give an observation function z defined on Q. The problem consists in determining a coefficient a^* from some set of admissible parameters $\mathscr{A}$ such that the solution $u(a^*)$ of (1.1) with $a = a^*$ best fits the data z. This problem has received a considerable amount of attention and we only mention selected contributions [BKL, BKM, BL, EL, LE, Y], which also describe applications in

* Supported in part by the Fonds zur Förderung der wissenschaftlichen Forschung, Austria, under project S 3206. K.K. also acknowledges support through AFOSR-F49620-86-C111
Offprint requests to: K. Kunisch

biological modelling and fluid flow in porous media. In mathematical models for biological systems the identification and estimation of the diffusivity parameter and its dependence on spatial as well as temporal (e.g. seasonal) effects is of special importance.

The most commonly studied approach to the inverse problem of determining the coefficient a from knowledge of z corresponding to the state $u(\bar{a})$ at the "true coefficient" $\bar{a}$ is given by

$$(\mathcal{O}) \qquad \min |u(a) - z|_H^2 \text{ over } \mathscr{A} ,$$

where $|\cdot|_H$ denotes some appropriate norm and $u(a)$ is the solution of (1.1) as a function of a. While this formulation is successful in some situations and has the advantage of being easily explained, flexible with respect to different types of observations and output norms H and fairly straightforward to program on a computer, it also has disadvantages which include the fact that $(\mathcal{O})$ may have a small residual at the minimizer a^*, yet a^* is still far from the "true" parameter $\bar{a}$, and furthermore in any iterative solution of $(\mathcal{O})$, the value of $u(a^n)$ is required for the sequence of approximating parameters a^n converging to a^*, so that (1.1) has to be solved frequently at each step of the iterative scheme.

For these reasons some investigators have addressed the question of finding alternative approaches to the output least squares method $(\mathcal{O})$. These include asymptotic embedding and adaptive control techniques developed in [AHS, BS, S] and, for only spatially varying coefficients a, a marching scheme [EL] and a time series analysis approach [LE].

Before we describe the method that we propose, let us briefly recall the equation error technique to estimate a in (1.1). Proceeding formally, let $\tilde{e}(a, v) = v_t - (av_x)_x - f$. The equation error technique then consists in solving $\tilde{e}(a, z) = 0$ for a. This could alternatively be formulated as the optimization problem

$$(\mathscr{E}) \qquad \min |\tilde{e}(a, z)|^2 \quad \text{over } a \in \mathscr{A} ,$$

which, differently from $(\mathcal{O})$ has the convenient property of being quadratic in the unknown variable. However, it has the disadvantages of requiring differentiation of the data.

In this paper we propose a hybrid method, combining the attractive features of the equation error and the output least squares approach. With a and u both treated as independent variables, the Eq. (1.1) is considered as an explicit constraint in a regularized least squares functional:

$$(1.2) \qquad \min |u - z|_H^2 + \beta |a|_V^2 \quad \text{subject to } \mathscr{S}\tilde{e}(a, u) = 0, \quad a \in \mathscr{A} .$$

Here β is a small positive scalar, $\mathscr{S}$ denotes a linear smoothing operator and V is a norm on the parameter space. The choice of $\mathscr{S}$, H, V and the accommodation of the boundary- and initial conditions of (1.1) in (1.2) shall be explained below. Problem (1.2) will be solved by an augmented Lagrangian technique, i.e. by solving

$$(1.3) \qquad \min \mathscr{J}(a, u) \quad \text{over } (a, u) \quad \text{with } a \in \mathscr{A} ,$$

where

$$\mathscr{J}(a, u) = |u - z|_H^2 + \beta |a|_V^2 + \langle \lambda^*, \mathscr{S}\tilde{e}(a, u) \rangle + c |\mathscr{S}\tilde{e}(a, u)|^2 ,$$

with c being an appropriately chosen fixed positive constant and λ^* the Lagrange multiplier associated with the equality constraint $\mathscr{S}\tilde{e}(a, u) = 0$. Observe that $\mathscr{J}$ is quadratic in a for fixed u and similarly quadratic in u for fixed a. Since λ^* is unknown, (1.3) can only be solved iteratively with respect to λ. Given a start-up value λ_0, λ_n is obtained from λ_{n-1} via

$$\lambda_n = \lambda_{n-1} + c\mathscr{S}\tilde{e}(a^n, u^n) ,$$

where (a^n, u^n) is the solution of (1.3) with λ^* replaced by $\lambda_{n-1}, n = 1, 2, \ldots$. We shall show that this iterative process is convergent. In fact we shall also treat a norm constraint which is involved in defining $\mathscr{A}$ by means of an augmented Lagrangian term. The essential technical tool for the convergence proof is a coercivity estimate for $\mathscr{J}(a, u)$. It requires that the quadratic form determined by the second Fréchet derivative of $\mathscr{J}$ at a solution (a^*, u^*) of (1.3) is uniformly positive. This coercivity estimate is quite similar to the one used in the stability analysis described in [CK] and guarantees that the solutions of (1.3) depend Hölder continuously on z.

A similar procedure as the one described above was developed for the estimation of the diffusion coefficient in an elliptic boundary value problem. The analytical results are established in [IK1, IK2]. The resulting numerical algorithm proved to be very successful with many test examples. A discussion of the implementation and of the hybrid nature of the algorithm combining least squares and equation error features can be found in [IKK].

The paper is organized as follows. In Sect. 2 we summarize some results concerning (1.1) and describe the set of admissible parameters. Section 3 is devoted to a description of the algorithm and the convergence results. In Sect. 4 we discuss the stability of the solution (a^*, u^*) of (1.3) with respect to perturbations in z. Sufficient conditions for the coercivity assumption that is required in Sects. 3 and 4 are given in Sect. 5. The proofs of Sect. 3 are given in the Appendix.

The notation that we employ is rather standard and we make only a few comments. We refer to [A, LM] for the theory of Sobolev spaces. For $T > 0$ and B a Banach space $L^2(0, T; B)$ denotes the (equivalence class) of square integrable functions in the sense of Bochner. At times we shall omit the domain $(0, T)$ in the notation of $L^2(0, T; B)$ and we always drop B if it is just $\mathbb{R}$. Generally we use an index with the notation of norms and inner products, but it can be dropped if it is L^2. For $C([0, T]; B)$, the space of continuous functions from $[0, T]$ to B, we simply write $C(0, T; B)$. In the estimates the dependence of constants on arguments that are determined by the problem statement is always a continuous one and it is such that the constant may go to infinity if any of its arguments tends to infinity in its natural norm; e.g. if $z \in L^2(H_0^1)$, then $K(z)$ depends continuously on z and possibly $K(z) \to \infty$ if $|z|_{L^2(H_0^1)} \to \infty$. For Hilbert spaces $X_i, i = 1, \ldots, r$ the product space $\otimes_{i=1}^r X_i$ is endowed with the natural Hilbert space topology.

2. Preliminaries

In this section we establish some preliminaries concerning the equation

$$(2.1) \qquad \begin{aligned} u_t &= (au_x)_x + f && \text{in } Q , \\ u(0, x) &= \varphi(x) && \text{for } x \in \Omega , \\ u(t, 0) &= u(t, 1) = 0 && \text{for } t \in (0, T] , \end{aligned}$$

where $\Omega = (0, 1)$, $Q = (0, T] \times \Omega$, $\varphi \in L^2$ and $f \in L^2(0, T; H^{-1})$. The parameter space is chosen as

$$W_a = \{a \in L^2(Q): a_x, a_t, a_{tx} \in L^2(Q)\}$$

which is endowed with the natural scalar product

$$\langle a, b\rangle_{W_a} = \langle a, b\rangle_{L^2(Q)} + \langle a_x, b_x\rangle_{L^2(Q)} + \langle a_t, b_t\rangle_{L^2(Q)} + \langle a_{tx}, b_{tx}\rangle_{L^2(Q)} ,$$

and the subordinate Hilbert space norm. We have the following

Lemma 2.1. (i) *W_a is a (separable) Hilbert space,*
(ii) *W_a embeds continuously and compactly into $C(0, T; C(\bar{\Omega}))$.*
(iii) $W_a = \{a: a \in L^2(0, T; H^1(\Omega)), (d/dt)a \in L^2(0, T; H^1(\Omega))\}$.

The proofs of this and the following lemmas are given in [KP], where it is also shown that the embedding constant from W_a into $C(0, T; C(\bar{\Omega}))$ can be chosen to be $k_1 = \sqrt{13/9}(T^{-1} + \frac{4}{9}T)^{1/2}$. The set of admissible parameters is defined as the subset of W_a given by:

$$\mathscr{A} = \{a \in W_a: a(t, x) \geqq \nu \text{ for } (t, x) \in Q, |a|_{W_a} \leqq \gamma\}$$

for constants ν and γ with $0 < \nu\sqrt{T} < \gamma$, so that $\mathscr{A}$ is not empty. Finally we denote

$$W = W(0, T) = \{u: u \in L^2(0, T; H_0^1(\Omega)), \frac{du}{dt} \in L^2(0, T; H^{-1}(\Omega))\} ,$$

which is a Hilbert space when endowed with the scalar product

$$\langle u, v\rangle_W = \int_0^T \langle u(t), v(t)\rangle_{H_0^1} dt + \int_0^T \left\langle \frac{d}{dt} u(t), \frac{d}{dt} v(t) \right\rangle_{H^{-1}} dt ,$$

where

$$\langle \varphi, \psi\rangle_{H_0^1(\Omega)} = \langle \varphi_x, \psi_x\rangle_{L^2(\Omega)} ,$$

$$\langle \varphi, \psi\rangle_{H^{-1}(\Omega)} = \langle \varphi, (-\Delta)^{-1}\psi\rangle_{L^2(\Omega)} ,$$

and where Δ denotes the Laplace operator from H_0^1 to $H^{-1}(\Omega)$.

By standard existence theory (see [L, p. 102], [W, p. 384]) there exists for any $a \in \mathscr{A}$ a unique solution $u \in W(0, T)$ which depends continuously on $(f, \varphi) \in L^2(0, T; H^{-1}(\Omega)) \times L^2(\Omega)$. Moreover u satisfies the following properties:

Lemma 2.2. *Let u denote the solution of* (2.1) *for any $a \in \mathscr{A}$.* Then we have

$$\text{(2.2)} \qquad \frac{1}{\nu}|u(T)|^2_{L^2(\Omega)} + |u|^2_{L^2(H_0^1)} \leqq \frac{1}{\nu}|\varphi|^2_{L^2(\Omega)} + \frac{1}{\nu^2}|f|^2_{L^2(H^{-1})} ,$$

and

$$\text{(2.3)} \qquad |u_t|^2_{L^2(H^{-1})} \leqq \frac{2k_1^2\gamma^2}{\nu}|\varphi|^2_{L^2(\Omega)} + \left(\frac{2k_1^2\gamma^2}{\nu^2} + 2\right)|f|^2_{L^2(H^{-1})} .$$

Concerning the dependence of u on a we have the following result:

Lemma 2.3. *Let a_j be a weakly convergent sequence in $\mathscr{A}$ with limit a. Then the corresponding solutions $u(a_j)$ converge weakly in $W(0, T)$ and strongly in $L^2(H_0^1)$ to the solution $u(a)$ of* (2.1).

3 The augmented Lagrangian algorithm and its convergence

In this section we describe and analyze an augmented Lagrangian algorithm for the estimation of a in (2.1) from data z. First let us fix some notation. We define a cost functional J^β: $W_a \times W \to \mathbb{R}$ by

$$J^\beta(a, u) = \frac{1}{2}|u - z|^2_{L^2(H_0^1)} + \frac{\beta}{2}|a|^2_{W_a} ,$$

where $z \in L^2(H_0^1)$ and $\beta \geqq 0$, and an operator $e = (e_1, e_2)$: $W_a \times W \to L^2(H_0^1) \times L^2$ by

$$(3.1) \quad e(a, u) = (e_1(a, u), e_2(a, u)) = ((-\Delta^{-1})(u_t - (au_x)_x - f), u(0, \cdot) - \varphi) .$$

Since W embeds continuously into $C(0, T; L^2(\Omega))$ [LM, p. 19], the second coordinate of (3.1) is a well defined element in $L^2(\Omega)$. With these operators defined, we consider the regularized least squares optimization problem

$$(P^\beta) \qquad \min J^\beta(a, u) \quad \text{subject to } e(a, u) = 0 \quad \text{and} \quad a \in \mathscr{A} .$$

We refer to a pair (a^*, u^*) as a local solution of (P^β) if there exists a neighborhood $U(a^*, u^*)$ of (a^*, u^*) in $W_a \times W$ such that $J^\beta(a^*, u^*) \leqq J^\beta(a, u)$ for every $(a, u) \in U(a^*, u^*)$ satisfying the constraints of (P^β), and we call (a^*, u^*) a (global) solution of (P^β) if $J^\beta(a^*, u^*) \leqq J^\beta(a, u)$ for all $a \in \mathscr{A}$ and u such that $e(a, u) = 0$.

Let us make a few comments concerning the regularization term in (P^β). Except for special cases it will be necessary to require $\beta > 0$ to establish convergence of the algorithm that we shall propose. Clearly, the solution (a^*, u^*) depends on β, but since we shall concentrate on a fixed value of β in our main results, we do not indicate this dependence. It can be seen from the proof of Lemma 3.1 below that (P^β) also has a solution if the norm constraint in $\mathscr{A}$ is omitted, provided only that $\beta > 0$. Without the norm constraint in the definition of $\mathscr{A}$, (P^β) may have no solution if $\beta = 0$ (see for an example [KW]). We shall not pause here to discuss the behaviour of the solutions (a^*, u^*) as $\beta \to 0^+$, but refer instead to [CK, EKN] in this respect. Some results concerning this aspect will be summarized in Proposition 4.1.

In the following lemma we shall establish existence of a solution to (P^β) and of an associated Lagrange multiplier. For $(a, u) \in W_a \times W$ and $(\lambda, \mu, \eta) = (\lambda_1, \lambda_2, \mu, \eta) \in (L^2(H_0^1) \times L^2(\Omega)) \times \mathbb{R} \times W_a$ we define the Lagrangian

$$L^\beta(a, u; \lambda, \mu, \eta) = J^\beta(a, u) + \langle \lambda, e(a, u) \rangle_{L^2(H_0^1) \times L^2} + \mu g(a) + \langle \eta, l(a) \rangle_{W_a} ,$$

where g: $W_a \to \mathbb{R}$ and l: $W_a \to W_a$ are given by

$$g(a) = \tfrac{1}{2}(|a|^2_{W_a} - \gamma^2) \qquad \text{and} \qquad l(a) = v - a .$$

Moreover, let $W_a^+ = \{a \in W_a : a(t, x) \geqq 0 \text{ on } Q\}$ denote the positive cone in W_a and let $D_{(a,u)}L^\beta(a, u; \lambda, \mu, \eta)$ stand for the Fréchet derivative of L^β with respect to (a, u).

Lemma 3.1. (a) *For any $\beta \geqq 0$ there exists a (global) solution (a^*, u^*) of (P^β).*

(b) *If (a^*, u^*) is a local solution of (P^β) then there exists a Lagrange multiplier $(\lambda^*, \mu^*, \eta^*) \in (L^2(H_0^1) \times L^2(\Omega)) \times \mathbb{R} \times W_a$, i.e.*

(i) $\mu^* \geqq 0$, $\langle \eta^*, a \rangle_{W_a} \geqq 0$ *for* $a \in W_a^+$,

(ii) $D_{(a,u)} L^\beta(a^*, u^*; \lambda^*, \mu^*, \eta^*) = 0$,

(iii) $\mu^*(|a^*|_{W_a}^2 - \gamma^2) = 0$, $\langle \eta^*, v - a^* \rangle_{W_a} = 0$.

As described in the Introduction our approach differs from a common output least squares technique in that both a and u are treated as independent variables which are related here through the constraint $e(a, u) = 0$. This constraint will be realized by an augmented Lagrangian formulation. This means that a Lagrangian- as well as a penalty term involving $e(a, u)$ are added to the cost functional. Under appropriate conditions, a local solution (a^*, u^*) of (P^β) will also be a local minimum of this augmented functional, without the necessity of imposing $e(a, u) = 0$ as an explicit constraint. This is achieved for all sufficiently large penalty parameters, without a requirement that this penalty parameter be increased to infinity.

Since Lagrange multipliers associated with local solutions of (P^β) are essential in our approach we next summarize some of their properties.

Lemma 3.2. (a) *The Lagrange multiplier $(\lambda_1^*, \lambda_2^*, \mu^*, \eta^*)$ associated with a local solution (a^*, u^*) of (P^β), $\beta \geqq 0$, is unique.*

(b) *Moreover, λ_1^* is the (unique) solution of the backwards parabolic equation*

$$v_t = -(a^* v_x)_x - \Delta(u^* - z) \qquad \text{on } Q$$

$$v(T, \cdot) = 0,\ v(t, 0) = v(t, 1) = 0 \qquad \text{for } t \in [0, T)\,,$$

and therefore $\lambda_1^* \in W$. In addition $\lambda_2^* = \lambda_1^*(0)$ holds.

Remark 3.1. The proofs in the Appendix reveal two additional interesting facts concerning the Lagrange multiplier.

(i) If $\beta > 0$ and $|a^*|_{W_a} = \gamma$, then u^* cannot be equal to z in $L^2(H_0^1)$.

(ii) For β sufficiently large, $\mu^* = 0$.

Henceforth it is assumed that (a^*, u^*) is a fixed local solution of (P^β), $\beta \geqq 0$, and that $(\lambda^*, \mu^*, \eta^*)$ is the associated Lagrange multiplier.

In the augmented Lagrangian algorithm that we propose not only the equality constraint $e(a, u) = 0$ but also the norm constraint $|a|_{W_a} \leqq \gamma$ are removed from the explicit constraints. Hence only the simple linear constraint $a \geqq v$ remains explicit. This elimination is accomplished by means of an appropriately defined cost functional. For $(a, u) \in W_a \times W$, $(\lambda, \mu) \in (L^2(H_0^1) \times L^2(\Omega)) \times \mathbb{R}$ and $c \geqq 0$ we put

$$\mathscr{H}_c(a, u; \lambda, \mu) = J^\beta(a, u) + \langle \lambda, e(a, u) \rangle_{L^2(H_0^1) \times L^2(\Omega)} + \frac{c}{2} |e(a, u)|^2_{L^2(H_0^1) \times L^2(\Omega)}$$

$$+ \mu \hat{g}(a, \mu, c) + \frac{c}{2} |\hat{g}(a, \mu, c)|^2,$$

where

$$\hat{g}(a, \mu, c) = \max\left(g(a), -\frac{\mu}{c} \right).$$

Let us consider the problem

$$\text{(3.2)} \qquad \min \mathscr{H}_c(a,u;\lambda^*,u^*) \quad \text{over } (a,u)\in W_a\times W \quad \text{subject to } a\geqq v\,.$$

The relationship between (P^β) and (3.2) will be analyzed in detail at the end of this section. At the moment we content ourselves to assert that a local solution (a^*, u^*) of (P^β) is also a local solution of (3.2) provided the coercivity condition (C) that will be specified below holds at (a^*, u^*). The augmented Lagrangian algorithm determines (a^*, u^*) by solving (3.2). Since the Lagrange multiplier pair (λ^*, μ^*) in (3.2) is not known a priori, it has to be determined iteratively within the algorithm. This will require to solve

$$\text{(3.3)} \qquad \min \mathscr{H}_c(a,u;\lambda,\mu) \quad \text{over } (a,u)\in W_a\times W \quad \text{subject to } a\geqq v \quad \text{on } Q\,,$$

(for (λ, μ) in general different from (λ^*, μ^*)). Unless (λ, μ) is sufficiently close to (λ^*, μ^*) or c is sufficiently large, (3.3) may have no solution. Hence we shall consider as an intermediate step the problem

$$\text{(3.4)} \qquad \min \mathscr{H}_c(a,u;\lambda,\mu) \quad \text{subject to } a\geqq v \quad \text{and} \quad (a,u)\in B_\delta\,,$$

where $B_\delta = \{(a,u)\in W_a\times W : |(a,u)-(a^*,u^*)|_{W_a\times W} \leqq \delta\}$. It is assumed throughout that

$$0<\delta\leqq\delta^* = \frac{1}{2}\sqrt{\frac{3\sigma}{k_2c_0}}\,(2+\gamma)^{-1}\,,$$

where k_2 is a constant that depends continuously on $(\gamma, k_1, v^{-1}, \varphi, f)$ and is given explicitly in Appendix. The constraint $(a,u)\in B_\delta$ reflects the local nature of the convergence analysis that will be given. It need not be implemented in actual computations as will be discussed below. Let us also observe that due to the choice of the specific norms involved in the definition of $\mathscr{H}_c$, the differentiation of u with respect to x is uniformly of first order in all terms of $\mathscr{H}_c$ that contain u.

The algorithm

(i) Choose a nondecreasing sequence of positive numbers $\{c_i\}_{i=1}^\infty$ and start-up values $(\lambda^0,\mu^0)\in(L^2(H_0^1)\times L^2(\Omega))\times\mathbb{R}^+$. Set $n=1$.

(ii) Determine (a^n, u^n) as the solution of

$$(P_{c_n}) \qquad \min \mathscr{H}_{c_n}(a,u;\lambda^{n-1},\mu^{n-1}) \quad \text{subject to } a\geqq v \quad \text{and} \quad (a,u)\in B_\delta\,.$$

(iii) Put

$$\lambda^n = \lambda^{n-1} + \frac{c_n}{2}\,e(a^n,u^n)$$

$$\mu^n = \mu^{n-1} + \frac{c_n}{2}\,\hat{g}(a^n,\mu^{n-1},c_n)\,.$$

(iv) If convergence is achieved stop. Otherwise put $n=n+1$ and return to (ii).

Convergence of this algorithm can be established provided that a certain coercivity condition already mentioned before holds at (a^*, u^*) and the associated Lagrange multiplier $(\lambda^*, \mu^*, \eta^*)$. To describe this condition we require a modified Lagrangian functional which involves a slack variable w which allows to treat the inequality constraint $g(a) \leqq 0$ as an equality constraint. For $(a, u, w) \in W_a \times W \times \mathbb{R}$, $(\lambda, \mu, \eta) \in (L^2(H_0^1) \times L^2(\Omega)) \times \mathbb{R} \times W_a$ and $c \geqq 0$ let

$$(3.5) \qquad J_c^\beta(a, u, w) = J^\beta(a, u) + \frac{c}{2} |e(a, u)|^2_{L^2(H_0^1) \times L^2(\Omega)} + \frac{c}{2} |g(a) + w|^2$$

and define

$$(3.6) \qquad L_c^\beta(a, u, w; \lambda, \mu, \eta) = J_c^\beta(a, u, w) + \langle \lambda, e(a, u) \rangle_{L^2(H_0^1) \times L^2} + \mu g(a) + \langle \eta, l(a) \rangle_{W_a} .$$

Condition (C)
There exist constants $\sigma > 0$ and $c_0 \geqq 0$ such that

$$D^2_{(a,u,w)} L^\beta_{c_0}(a^*, u^*, w^*; \lambda^*, \mu^*, \eta^*)((h, v, y), (h, v, y)) \geqq \sigma |(h, v, y)|^2_{W_a \times W \times \mathbb{R}}$$

holds for all $(h, v, y) \in W_a \times W \times \mathbb{R}$, where $w^* = -g(a^*)$.

Here we denote by $D^2_{(a,u,w)} L^\beta_{c_0}(a^*, u^*, w^*; \lambda^*, \mu^*, \eta^*)((h, v, y), (h, v, y))$ the second Fréchet derivative of $L^\beta_{c_0}$ at $(a^*, u^*, w^*; \lambda^*, \mu^*, \eta^*)$ in directions (h, v, y). Condition (C) will be investigated independently in Sect. 4. If it holds at a solution (a^*, u^*) with $\beta > 0$, then this solution is required to be a global solution of (P^β), whereas, if (C) can be established at a solution (a^*, u^*) with $\beta = 0$ then it suffices for this latter solution to be a local one.

In the statement of the basic convergence result for the augmented Lagrangian algorithm we shall also introduce a constant $\hat{\mu}$ which depends on $(\mu^0, \lambda^0, k_1, \gamma, \nu^{-1}, \varphi, f, z)$. The dependence is rather involved but can be given explicitly. We further put

$$\tilde{c} = \max\left(\frac{\hat{\mu}}{\delta^*}, 2c_0\right) .$$

Theorem 3.1. *Assume that* (C) *holds and that* $(\lambda^0, \mu^0) \in (L^2(H_0^1) \times L^2(\Omega)) \times \mathbb{R}^+$ *and* $c_1 \geqq \tilde{c}$. *Then the quadrupel* $(a^n, u^n; \lambda^{n-1}, \mu^{n-1})$ *defined by the algorithm satisfies the estimate*

$$(3.7) \qquad |a^* - a^n|^2_{W_a} + |u^* - u^n|^2_W + \frac{2}{\sigma c_n}(|\lambda^* - \lambda^n|^2_{L^2(H_0^1) \times L^2(\Omega)} + |\mu^* - \mu^n|^2)$$
$$\leqq \frac{2}{\sigma c_n}(|\lambda^* - \lambda^{n-1}|^2_{L^2(H_0^1) \times L^2(\Omega)} + |\mu^* - \mu^{n-1}|^2)$$

and $|\lambda^n|_{L^2(H_0^1) \times L^2(\Omega)}$ *and* μ^n *are in* $[0, \hat{\mu}]$ *for all* $n = 1, 2 \ldots$.

This theorem implies the convergence of (a^n, u^n) to (a^*, u^*) and boundedness of (λ^n, μ^n). The sequence $\{(a^n, u^n)\}$ furthermore satisfies

Corollary 3.1. *Under the assumptions of Theorem* 3.1

$$\sum_{n=1}^{\infty} c_n(|a^* - a^n|^2_{W_a} + |u^* - u^n|^2_W) \leqq \frac{2}{\sigma}(|\lambda^* - \lambda^0|^2_{L^2(H_0^1) \times L^2(\Omega)} + |\mu^* - \mu^0|^2) .$$

Remark 3.2. Corollary 3.1 also implies the estimate

$$|a^* - a^n|_{W_a}^2 + |u^* - u^n|_W^2 \leqq \frac{2}{\sigma c_n}(|\lambda^* - \lambda^0|^2_{L^2(H_0^1)\times L^2(\Omega)} + |\mu^* - \mu^0|^2),$$

for all $n = 1, 2, \ldots$. Hence, if c_1 is sufficiently large or if $|(\lambda^0, \mu^0) - (\lambda^*, \mu^*)|$ is sufficiently small, then $(a^n, u^n) \in \operatorname{int} B_\delta$ for all n. In this case the constraint $(a, u) \in B_\delta$ in (3.4) is not active and (a^n, u^n) is also a local solution of (3.3) with $(\lambda, \mu, c) = (\lambda^{n-1}, \mu^{n-1}, c_n)$. If the above requirement on (λ^0, μ^0) or c_1 are not satisfied, then nevertheless $(a^n, u^n) \in \operatorname{int} B_\delta$ for all n sufficiently large, due to the convergence of (a^n, u^n) to (a^*, u^*).

Theorem 3.1 gives only uniform boundedness of the updated Lagrange multipliers (λ^n, μ^n). Next we shall show their convergence, which in turn will improve the rate of convergence of (a^n, u^n) to (a^*, u^*). Simultaneously we establish convergence of the Lagrange multiplier η^n associated with the constraint $a \geqq \nu$ in (P_{c_n}). The assumption will be made that $(a^n, u^n) \in \operatorname{int} B_\delta$ for all n. In the theorem below K denotes a constant that depends on $(\nu^{-1}, \gamma, k_1, \delta^*, \beta, \lambda^0, \mu^0, c_0, T, \varphi, f, z, \sup c_n)$ and $\hat{K}(|\langle a^* - \nu, a^* \rangle|^{-1})$ stands for a constant, which in addition depends on the angle $\langle a^* - \nu, a^* \rangle_{W_a}$ between $a^* - \nu$ and a^*. The explicit dependence of K and $\hat{K}$ on the parameters of the problem is given in the proofs. We further introduce the constants

$$\hat{\delta} = \min(\delta^*, \tfrac{1}{3}(\gamma - |a^*|_{W_a})) \qquad \text{and} \qquad \hat{c} = \max\left(\tilde{c}, \frac{3\hat{\mu}}{(\hat{\delta} + \gamma)(\gamma - |a^*|_{W_a})}\right).$$

Theorem 3.2. *Assume that* (C) *holds, that* c_n *is a nondecreasing sequence with* $c_n \in [\hat{c}, c_{\max}]$ *for some* $c_{\max} \geqq \hat{c}$ *and that* $(a^n, u^n) \in \operatorname{int} B_\delta$ *for all* n. *Then we have*:

(i) *If* $|a^*|_{W_a} = \gamma$ *and* $\delta \in (0, \delta^*]$, *then for all* $n = 1, 2, \ldots$

$$|\lambda^* - \lambda^n|^2_{L^2(H_0^1)\times L^2(\Omega)} + |\mu^* - \mu^n|^2 + |\eta^* - \eta^n|^2_{W_a}$$
$$\leqq \frac{[K + \hat{K}(|\langle a^* - \nu, a^* \rangle|^{-1})]^n}{\sigma^n \prod_{i=1}^n c_i}(|\lambda^* - \lambda^0|^2_{L^2(H_0^1)\times L^2(\Omega)} + |\mu^* - \mu^0|^2).$$

(ii) *If* $|a^*|_{W_a} < \gamma$, $\delta \in (0, \hat{\delta}]$, $c_1 \geqq \hat{c}$ *and* $\mu^0 = 0$, *then* $\mu^n = 0$ *for all* n *and*

$$|\lambda^* - \lambda^n|^2_{L^2(H_0^1)\times L^2(\Omega)} + |\eta^* - \eta^n|^2_{W_a} \leqq \frac{K^n}{\sigma^n \prod_{i=1}^n c_i}|\lambda^* - \lambda^0|^2_{L^2(H_0^1)\times L^2(\Omega)}.$$

We now turn to a discussion of the relationship between the optimization problems (P^β) and (3.2). This will be accomplished in several steps. Recall the definition of J_c^β in (3.5) and consider

$$(P_c^\beta) \qquad \begin{aligned} &\min J_c^\beta(a, u, w) \\ &\text{subject to} \\ &e(a, u) = 0, \\ &g(a) + w = 0, \\ &l(a) \leqq 0, \\ &w \geqq 0, \end{aligned}$$

where the nonlinear inequality constraint with finite dimensional image space is reformulated as an equality constraint with a slack variable, and the equality constraints are also considered in penalty terms. It is simple to argue that (a^*, u^*) is a local solution of (P^β) if and only if $(a^*, u^*, w^*) = (a^*, u^*, -g(a^*))$ is a local solution of (P_c^β). The Lagrangian functional associated with (P_c^β) is given by

$$(3.8)\qquad \hat{L}_c^\beta(a, u, w; \lambda, \mu, v, \eta) = J_c^\beta(a, u, w) + \langle \lambda, e(a, u)\rangle_{L^2(H_0^1)\times L^2(\Omega)} + \mu(g(a) + w) - vw + \langle \eta, l(a)\rangle_{W_a},$$

for $(a, u, w) \in W_a \times W \times \mathbb{R}$ and $(\lambda, \mu, v, \eta) \in (L^2(H_0^1) \times L^2(\Omega)) \times \mathbb{R} \times \mathbb{R} \times W_a$.

Lemma 3.3. *Let $(a^*, u^*, w^*) = (a^*, u^*, -g(a^*))$ be a local solution of (P_c^β). Then there exists a unique Lagrange multiplier $(\bar{\lambda}, \bar{\mu}, \bar{v}, \bar{\eta})$ of (P_c^β) in the sense that*

(i) $\bar{v} \geqq 0$, $\langle \bar{\eta}, a\rangle_{W_a} \geqq 0$, *for all* $a \in W_a^+$,

(ii) $D_{(a,u,w)}\hat{L}_c^\beta(a^*, u^*, -g(a^*); \bar{\lambda}, \bar{\mu}, \bar{v}, \bar{\eta}) = 0$,

(iii) $\langle \bar{\eta}, l(a^*)\rangle = 0$, $\bar{v}w^* = 0$.

In fact, $(\bar{\lambda}, \bar{\mu}, \bar{v}, \bar{\eta}) = (\lambda^*, \mu^*, \mu^*, \eta^*)$.

Observe that due to Lemma 3.3, the Lagrangian function for (P_c^β) evaluated at $(\lambda^*, \mu^*, \mu^*, \eta^*)$ simplifies to

$$\hat{L}_c^\beta(a, u, w; \lambda^*, \mu^*, \mu^*, \eta^*) = J_c^\beta(a, u, w) + \langle \lambda^*, e(a, u)\rangle_{L^2(H_0^1)\times L^2(\Omega)} + \mu^* g(a) + \langle \eta^*, l(a)\rangle_{W_a},$$

which prompted the introduction of $L_c^\beta(a, u, w; \lambda, \mu, \eta)$ in (3.6), and which is also the functional for which the coercivity condition (C) is required to hold at $(a^*, u^*, -g(a^*); \lambda^*, \mu^*, \eta^*)$. Our next goal is to eliminate the constraints $e(a, u)$ and $g(a) + w = 0$ from the explicit constraints in (P_c^β). For this purpose we consider

$$(\tilde{P}_c^\beta)\qquad \min \mathscr{L}_c^\beta(a, u, w; \lambda^*, \mu^*) \quad \text{subject to } w \geqq 0,\quad l(a) \leqq 0$$

with

$$(3.9)\qquad \mathscr{L}_c^\beta(a, u, w; \lambda, \mu) = J_c^\beta(a, u, w) + \langle \lambda, e(a, u)\rangle_{L^2(H_0^1)\times L^2(\Omega)} + \mu(g(a) + w),$$

for $(a, u, w) \in W_a \times W \times \mathbb{R}$ and $(\lambda, \mu) \in (L^2(H_0^1) \times L^2(\Omega)) \times \mathbb{R}$. The relationship between $(\tilde{P}_c^\beta)$ and (P_c^β), or equivalently (P^β), is clarified next.

Lemma 3.4. *Assume that* (C) *holds at the local solution (a^*, u^*) of (P^β). Then $(a^*, u^*, -g(a^*))$ is a local solution of $(\tilde{P}_c^\beta)$. Conversely, if $(a^*, u^*, -g(a^*))$ is a local solution of $(\tilde{P}_c^\beta)$, that satisfies the constraints of (P_c^β), then it is also a local solution of (P_c^β) and (a^*, u^*) is a local solution of the original problem (P^β).*

In a final reformulation of the optimization problem the slack variable w is eliminated and we arrive at (3.2), which we repeat for convenence:

$$(3.2)\qquad \min \mathscr{H}_c(a, u; \lambda^*, \mu^*) \quad \text{subject to } l(a) \leqq 0,$$

where

$$\mathscr{H}_c(a, u; \lambda, \mu) = J^\beta(a, u) + \langle \lambda, e(a, u)\rangle_{L^2(H_0^1)\times L^2(\Omega)} + \frac{c}{2}|e(a, u)|^2_{L^2(H_0^1)\times L^2(\Omega)} + \mu\hat{g}(a, \mu, c) + \frac{c}{2}|\hat{g}(a, \mu, c)|^2.$$

It is simple to argue that $(a^*, u^*, \max(0, -g(a^*) - \mu^*/c))$ is a local solution of $(\tilde{P}_c^\beta)$ if and only if (a^*, u^*) is a local solution of (3.2). *In particular, if* (a^*, u^*) *is a local solution of* (P^β) *such that* (C) *holds at* $(a^*, u^*, -g(a^*))$, *then* (a^*, u^*) *is also a solution of* (3.2). This ends our discussion on the relationship between (P^β) and problem (3.2) which is solved in the algorithm that was described above.

We close this section with the remark that the results of this paper can be extended to the case when Ω is a multi-dimensional domain, provided that W_a is endowed with a topology that embeds W_a compactly into $C(0, T; C(\bar{\Omega}))$ and the a priori estimates of Sect. 2 are changed accordingly.

4 Stability analysis

In this section we show that the coercivity condition (C) of Sect. 3 also guarantees stability of the solutions of the problem (3.2) as well as of (P^β) with respect to perturbations in z. Recall that (3.2) is the optimization problem that is solved iteratively by the augmented Lagrangian algorithm. To emphasize the dependence of (3.2) on the observation z we henceforth write in this section $\mathscr{H}_c(a, u, \lambda^*, \mu^*, z)$ for $\mathscr{H}_c(a, u, \lambda^*, \mu^*)$ and we denote by z^* the unperturbed observation function. For convenience we repeat the definition of $\mathscr{H}_c(a, u; \lambda^*, \mu^*, z)$ and of (3.2), whose dependence on z is also indicated by an index:

$$\mathscr{H}_c(a, u; \lambda^*, \mu^*, z) = \frac{1}{2}|u - z|^2_{L^2(H_0^1)} + \frac{\beta}{2}|a|^2_{W_a} + \langle \lambda^*, e(a, u)\rangle_{L^2(H_0^1)\times L^2(\Omega)}$$

$$+ \frac{c}{2}|e(a, u)|^2_{L^2(H_0^1)\times L^2(\Omega)} + \mu^* \hat{g}(a, \mu^*, c) + \frac{c}{2}\hat{g}(a, \mu^*, c)^2$$

and

$$(3.2)_z \qquad \min \mathscr{H}_c(a, u; \lambda^*, \mu^*, z) \quad \text{over } (a, u) \in W_a \times W \quad \text{subject to } a \geqq \nu .$$

Henceforth we fix an unperturbed observation function $z^* \in L^2(H_0^1)$ and *when referring to quantities in Sect. 3 it is understood that z is replaced by* z^*. We have the following result on the stability of the solution of $(3.2)_z$ with respect to perturbations in z:

Theorem 4.1. *Assume that* (C) *holds at a (local or global) solution of* $(3.2)_{z^*}$ *and that* $c \geqq \max(\mu^*(\delta^*)^{-1}, 2c_0)$. *Then there exist neighborhoods* V_1 *of* z^* *in* $L^2(H_0^1)$ *and* V_2 *of* (a^*, u^*) *in* $W_a \times W$, *and a constant* $\kappa > 0$ *such that for every* $z \in V_1$ *there exists a local solution* $(a_z, u_z) \in V_2$ *of* $(3.2)_z$ *and all local solutions of* $(3.2)_z$ *in* V_2 *satisfy*

$$|(a_z - a^*, u_z - u^*)|_{W_a \times W} \leqq \kappa |z - z^*|^{1/2}_{L^2(H_0^1)} .$$

An analogous result holds for the stability of the solutions of the problem (P^β). We recall the definition of (P^β) and indicate the dependence on z by an additional index:

$$(P^\beta)_z \qquad \min \left\{\frac{1}{2}|u - z|^2_{L^2(H_0^1)} + \frac{\beta}{2}|a|^2_{W_a}\right\} \text{ subject to } e(a, u) = 0 \quad \text{and} \quad a \in \mathscr{A} .$$

Theorem 4.2. *Assume that* (C) *holds at a (local or global) solution of* $(P^\beta)_{z^*}$. *Then there are neighborhoods* $\hat{V}_1$ *of* z^* *in* $L^2(H_0^1)$ *and* $\hat{V}_2$ *of* (a^*, u^*) *in* $W_a \times W$, *and* *a constant* $\hat{\kappa} > 0$ *such that for every* $z \in \hat{V}_1$ *there exists a local solution* $(a_z, u_z) \in \hat{V}_2$ *of* $(3.2)_z$ *and all local solutions of* $(P^\beta)_z$ *in* $\hat{V}_2$ *satisfy*

$$|(a_z - a^*, u_z - u^*)|_{W_a \times W} \leqq \hat{\kappa}|z - z^*|^{1/2}_{L^2(H^1_0)} .$$

Proof of Theorem 4.1. The proof is based on Theorems 4 and 6 of [Al]. These theorems require that the solution (a^*, u^*) is a regular point with respect to the constraint $a \geqq v$ and that there exists a neighborhood $\bar{U}$ of (a^*, u^*) and a constant $\bar{\alpha}$ such that

$$\mathscr{H}_c(a, u; \lambda^*, \mu^*, z^*) \geqq \mathscr{H}_c(a^*, u^*; \lambda^*, \mu^*, z^*) + \bar{\alpha}|(a - a^*, u - u^*)|^2_{W_a \times W} \tag{4.1}$$

for all $(a, u) \in \bar{U}$ satisfying $a \geqq v$. Clearly, (a^*, u^*) is a regular point with respect to the constraint $a \geqq v$. In view of the technical Lemma A.2 with $\hat{\mu} = \mu^*$ we have

$$\begin{aligned} J^\beta(a, u) &+ \langle \lambda^*, e(a, u)\rangle_{L^2(H^1_0) \times L^2(\Omega)} + \frac{c_0}{2}|e(a, u)|^2_{L^2(H^1_0) \times L^2(\Omega)} \\ &+ \mu^* \hat{g}(a, \mu^*, c) + \frac{c_0}{2}|\hat{g}(a, \mu^*, c)|^2 + \langle \eta^*, v - a\rangle_{W_a} \\ &\geqq J^\beta(a^*, u^*) + \frac{\sigma}{2}|(a - a^*, u - u^*)|^2_{W_a \times W} \end{aligned} \tag{4.2}$$

for all $c \geqq \max\left(\frac{\mu^*}{\delta^*}, 2c_0\right)$ and $(a, u) \in B_{\delta^*}$. Since $c \geqq c_0$ and since $\langle \eta^*, v - a\rangle \leqq 0$ for $a \geqq v$, (4.2) implies

$$\mathscr{H}_c(a, u; \lambda^*, \mu^*, z^*) \geqq \mathscr{H}_c(a^*, u^*; \lambda^*, \mu^*, z^*) + \frac{\sigma}{2}|(a - a^*, u - u^*)|^2_{W_a \times W} ,$$

for all $c \geqq \max\left(\frac{\mu^*}{\delta^*}, 2c_0\right)$ and $(a, u) \in B_{\delta^*}$, with $a \geqq v$. This is the desired estimate (4.1) with $\bar{\alpha} = \frac{\sigma}{2}$ and $\bar{U} = B_{\delta^*}$. □

Proof of Theorem 4.2. We proceed as in the proof of Theorem 4.1. In the proof of Lemma 3.1 that is given in the Appendix, it is shown that (a^*, u^*) satisfies the regular point condition with respect to the constraints $e(a, u) = 0$ and $a \in \mathscr{A}$. Hence it suffices to argue the existence of a neighborhood $\bar{U}$ of (a^*, u^*) and of a constant $\bar{\alpha} > 0$ such that

$$J^\beta(a, u) \geqq J^\beta(a^*, u^*) + \bar{\alpha}|(a - a^*, u - u^*)|^2_{W_a \times W} \tag{4.3}$$

for all $(a, u) \in \bar{U}$ with $e(a, u) = 0$ and $a \in \mathscr{A}$.

As a consequence of the fact that

$$\mu\hat{g}(a, \mu, c) + \frac{c}{2}\hat{g}(a, \mu, c)^2 = \frac{1}{2c}(|\max(0, cg(a) + \mu)|^2 - \mu^2) ,$$

$a \in \mathscr{A}$ implies $\mu^*\hat{g}(a, \mu^*, c) + \frac{c}{2}\hat{g}(a, \mu^*, c)^2 \leqq 0$. This observation together with (4.2) implies (4.3) and the theorem is proved. □

5 The coercivity condition

In this section we turn to a discussion of the coercivity condition (C) and establish three special cases for which (C) holds. In the first case we make essential use of the regularization term in the cost functional $J^\beta(a, u)$ and determine a range for the regularization parameter within which (C) is fulfilled. This range of regularization parameters is determined, most importantly, by the distance between the observation z and the attainable set $\{u(a): a \in \mathscr{A}\}$. It will be convenient, henceforth, to indicate the dependence of the solutions of (P^β) on β and we shall denote in this section by (a_β^*, u_β^*) a global minimum of (P^β), $\beta \geqq 0$. Furthermore we put

$$U^\beta = \{u_\beta^* : (a_\beta^*, u_\beta^*) \text{ is a global solution of } (P^\beta)\} ,$$

and

$$A^\beta = \{a_\beta^* : (a_\beta^*, u_\beta^*) \text{ is a global solution of } (P^\beta)\} .$$

The following result in [CK] summarizes some properties of (a_β^*, u_β^*) as a function of β. It is essential here, that (a_β^*, u_β^*) be a global solution of (P^β).

Proposition 5.1. *For $\beta \geqq 0$ let (a_β^*, u_β^*) denote a global solution of (P^β) and let $0 \leqq \beta_0 < \beta$. Then the following relationships hold:*

(i) $\sup_{a_\beta^* \in A^\beta} |a_\beta^*|_{W_a} \leqq \inf_{a_{\beta_0}^* \in A^{\beta_0}} |a_{\beta_0}^*|_{W_a}$,

(ii) $\sup_{u_{\beta_0}^* \in U^\beta} |u_{\beta_0}^* - z|^2_{L^2(H_0^1)} \leqq \inf_{u_\beta^* \in U^\beta} |u_\beta^* - z|^2_{L^2(H_0^1)}$.

(iii) *If $\beta_n \to 0^+$ and $\tilde{a}$ is any weak cluster point of $a_{\beta_n}^* \in A^{\beta_n}$, then $(\tilde{a}, u(\tilde{a}))$ is a solution of (P^0). Furthermore we have*

$$|\tilde{a}|_{W_a} = \min_{a_0^* \in A^0} |a_0| W_a = \lim_{n\to\infty} \sup_{a_{\beta_n}^* \in A^{\beta_n}} |a_{\beta_n}^*|_{W_a} .$$

(iv) *There exists a monotonically increasing function ρ with $\lim_{\beta \to 0^+} \rho(\beta) = 0$ such that for $\beta > 0$*

$$\sup_{U^\beta} |u_\beta^* - z|^2_{L^2(H_0^1)} \leqq |u_0^* - z|^2_{L^2(H_0^1)} + \beta\rho(\beta) .$$

In the statement of (iv) of Propositon 5.1 observe that $|u_0^* - z|^2$ is independent of the choice of $u_0^* \in (P^0)$.

Proof of Proposition 5.1. Part (i)–(iii) are from [CK] and (iv) is a consequence of (iii) and the fact that

$$|u_\beta^* - z|^2_{L^2(H_0^1)} \leqq |u_0^* - z|^2_{L^2(H_0^1)} + \beta(|a_0^*|^2 - |a_\beta^*|^2) ,$$

which holds for every solution (a_0^*, u_0^*) of (P^0) and (a_β^*, u_β^*) of (P^β), $\beta > 0$ ([CK]). □

Proposition 5.2. *Let (a_β^*, u_β^*) and (a_0^*, u_0^*) denote global solutions of (P^β) and (P^0) respectively and suppose that*

$$|u_0^* - z|^2_{L^2(H_0^1)} < \beta_0\left(\frac{\nu^2}{2k_1^2}(1 - \beta_0) - \rho(\beta_0)\right) \tag{5.1}$$

holds for some $\beta_0 \in (0, 1)$. Then there exist positive constants c_0, $\tilde{\sigma}$ and $\underline{\beta} \in (0, \beta_0)$ such that

$$D^2_{(a,u,w)} L^\beta_c(a^*_\beta, u^*_\beta, w^*_\beta; \lambda^*, \mu^*, \eta^*)((h, v, y),(h, v, y)) \geqq \beta\tilde{\sigma}|(h, v, y)|^2_{W_a \times W \times \mathbb{R}} \tag{5.2}$$

holds for all $\beta \in [\underline{\beta}, \beta_0]$ and $c \geqq c_0$. If z is attainable, i.e. if $u^*_0 = z$, then (5.2) *is valid for all $\beta \in [0, \beta_0]$ for some $\beta_0 > 0$, $\tilde{\sigma} > 0$.*

Proposition 5.2 implies that Condition (C) holds with $\sigma = \beta\tilde{\sigma}$, for $\beta \in [\underline{\beta}, \beta_0]$, provided that $|u^*_0 - z|_{L^2(H^1_0)}$ is sufficiently small.

Proof of Proposition 5.2. Throughout the proof we shall write λ^* for λ^*_1 and put $\Lambda^* = (\lambda^*, \mu^*, \eta^*)$. To denote differentiation with respect to x we use an index or the operator ∇. The second Fréchet derivative of L^β_c is given by

$$\begin{aligned} D^2_{(a,u,w)} L^\beta_c(a^*_\beta, u^*_\beta, w^*_\beta; \Lambda^*)((h, v, y), (h, v, y)) \\ = |v|^2_{L^2(H^1_0)} + \beta|h|^2_{W_a} \\ - 2\langle \lambda^*, (-\Delta)^{-1}(hv_x)_x\rangle_{L^2(H^1_0)} + \mu^*|h|^2 \\ + c|(-\Delta)^{-1}[v_t - (a^*_\beta v_x)_x - (hu^*_{\beta,x})_x]|^2_{L^2(H^1_0)} + c|v(0,\cdot)|^2_{L^2} + c(\langle a^*_\beta, h\rangle_{W_a} + y)^2 \,. \end{aligned} \tag{5.3}$$

We first estimate those terms in (5.3) that are not related to inequality constraints. For this purpose we define

$$\begin{aligned} M(\beta, c)(h, v) = |v|^2_{L^2(H^1_0)} + \beta|h|^2_{W_a} - 2\langle \lambda^*, (-\Delta)^{-1}(hv_x)_x\rangle_{L^2(H^1_0)} \\ + c|(-\Delta)^{-1}[v_t - (a^*_\beta v_x)_x - (hu^*_{\beta,x})_x]|^2_{L^2(H^1_0)} + c|v(0,\cdot)|^2_{L^2} \,. \end{aligned} \tag{5.4}$$

The individual terms in (5.4) are estimated as follows:

$$\begin{aligned} |\langle \lambda^*, (-\Delta)^{-1}(hv_x)_x\rangle|_{L^2(H^1_0)} &= |\langle \nabla\lambda^*, \nabla(-\Delta)^{-1}\nabla(hv_x)\rangle_{L^2(L^2)}| \\ &\leqq |\lambda^*|_{L^2(H^1_0)}|hv_x|_{L^2(L^2)} \leqq k_1|\lambda^*|_{L^2(H^1_0)}|h|_{W_a}|v|_{L^2(H^1_0)} \\ &\leqq \frac{k_1}{\nu}|u^*_\beta - z|_{L^2(H^1_0)}|h|_{W_a}|v|_{L^2(H^1_0)} \,, \end{aligned} \tag{5.5}$$

where in the last term we made use of Lemma 3.3, which implies – compare Lemma 3.2 – that $|\lambda^*|_{L^2(H^1_0)} \leqq \frac{1}{\nu}|u^*_\beta - z|_{L^2(H^1_0)}$. Next we have

$$\begin{aligned} &|(-\Delta)^{-1}[v_t - (a^*_\beta v_x)_x - (hu^*_{\beta,x})_x]|^2_{L^2(H^1_0)} = \\ &\quad \langle v_t - (a^*_\beta v_x)_x - (hu^*_{\beta,x})_x, (-\Delta)^{-1}[v_t - (a^*_\beta v_x)_x - (hu^*_{\beta,x})_x]\rangle \\ &= |v_t|^2_{L^2(H^{-1})} - 2\langle v_t, (-\Delta)^{-1}\nabla(a^*_\beta v_x + hu^*_{\beta,x})\rangle \\ &\quad + \langle \nabla(a^*_\beta v_x + hu^*_{\beta,x}), (-\Delta)^{-1}\nabla(a^*_\beta v_x + hu^*_{\beta,x})\rangle \\ &\geqq |v_t|^2_{L^2(H^{-1})} - 2|v_t|_{L^2(H^{-1})}|(-\Delta)^{-1}\nabla(a^*_\beta v_x + hu^*_{\beta,x})|_{L^2(H^1_0)} \\ &\quad + |(-\Delta)^{-1}\nabla(a^*_\beta v_x + hu^*_{\beta,x})|^2_{L^2(H^1_0)} \\ &\geqq \tfrac{1}{2}|v_t|^2_{L^2(H^{-1})} - 2|a^*_\beta v_x|^2_{L^2} - 2|hu^*_{\beta,x}|^2_{L^2} \\ &\geqq \tfrac{1}{2}|v_t|^2_{L^2(H^{-1})} - 2k_1^2|a^*_\beta|^2_{W_a}|v_x|^2_{L^2} - 2k_1^2|h|^2_{W_a}|u^*_\beta|^2_{L^2(H^1_0)} \\ &\geqq \tfrac{1}{2}|v_t|^2_{L^2(H^{-1})} - 2k_1^2|a^*_\beta|^2_{W_a}|v_x|^2_{L^2} - 2k_1^2\left(\frac{1}{\nu}|\varphi|^2_{L^2} + \frac{1}{\nu^2}|f|^2_{L^2(H^{-1})}\right)|h|^2_{W_a} \,. \end{aligned} \tag{5.6}$$

Combining (5.4)–(5.6) we find

$$(5.7)\qquad M(\beta, c)(h, v) \geqq (1 - c\kappa_1)|v|^2_{L^2(H^1_0)} + (\beta - c\kappa_2)|h|^2_{W_a} - \frac{2k_1}{\nu}|u^*_\beta - z|_{L^2(H^1_0)}|h|_{W_a}|v|_{L^2(H^1_0)} + \frac{c}{2}|v_t|^2_{L^2(H^{-1})},$$

where

$$\kappa_1 = 2k_1^2\gamma^2 \qquad \text{and} \qquad \kappa_2 = 2k_1^2\left(\frac{1}{\nu}|\varphi|^2_{L^2(\Omega)} + \frac{1}{\nu^2}|f|^2_{L^2(H^{-1})}\right).$$

Following [IK2] we set $c = \delta\beta$ with δ to be determined below. Then (5.7) implies for any $\varepsilon \in (0, 1)$

$$\begin{aligned}M(\beta, \delta\beta)(h, v) &\geqq (1 - \delta\beta\kappa_1)|v|^2_{L^2(H^1_0)} + \beta(1 - \delta\kappa_2)|h|^2_{W_a} \\ &\quad - \frac{k_1^2}{\nu^2}\frac{2}{\varepsilon\beta}|u^*_\beta - z|^2_{L^2(H^1_0)}|v|^2_{L^2(H^1_0)} - \frac{\varepsilon}{2}\beta|h|^2_{W_a} + \frac{1}{2}\beta\delta|v_t|^2_{L^2(H^{-1})} \\ &= \left(1 - \delta\beta\kappa_1 - \frac{k_1^2}{\nu^2}\frac{2}{\varepsilon\beta}|u^*_\beta - z|^2_{L^2(H^1_0)}\right)|v|^2_{L^2(H^1_0)} + \frac{1}{2}\beta\delta|v_t|^2_{L^2(H^{-1})} \\ &\quad + \beta\left(1 - \kappa_2\delta - \frac{\varepsilon}{2}\right)|h|^2_{W_a}.\end{aligned}$$

Choosing δ as

$$\delta = \tfrac{1}{2}\max(\kappa_1^{-1}, \kappa_2^{-1})$$

one obtains

$$(5.8)\qquad M(\beta, \delta\beta)(h, v) \geqq \left(1 - \delta\beta\kappa_1 - \frac{2k_1^2}{\varepsilon\beta\nu^2}|u^*_\beta - z|^2_{L^2(H^1_0)}\right)|v|^2_{L^2(H^1_0)} + \tfrac{1}{2}\beta\delta|v_t|^2_{L^2(H^{-1})} + \tfrac{1}{2}\beta(1 - \varepsilon)|h|^2_{W_a}.$$

In view of Proposition 5.1(iv) and (5.1) we have

$$\sup_{U^{\beta_0}}|u^*_{\beta_0} - z|^2_{L^2(H^1_0)} \leqq |u^*_0 - z|^2_{L^2(H^1_0)} + \beta_0\rho(\beta_0) < \beta_0\frac{\nu^2}{2k_1^2}(1 - \beta_0).$$

Since this inequality is strict and since by Proposition 5.1 its left hand side is monotonically increasing with β_0, there exist $\underline{\beta} \in (0, \beta_0)$ and $\varepsilon_1 \in (0, 1)$ such that

$$(5.9)\qquad \sup_{U^\beta}|u^*_\beta - z|^2_{L^2(H^1_0)} \leqq \beta\frac{\nu^2\varepsilon_1}{2k_1^2}(1 - \beta_0)$$

holds for all $\beta \in [\underline{\beta}, \beta_0]$. Inserting (5.9) into (5.8) with $\varepsilon = \varepsilon_1$ leads to

$$(5.10)\qquad M(\beta, \delta\beta)(h, v) \geqq \frac{\beta}{2}|v|^2_{L^2(H^1_0)} + \tfrac{1}{2}\delta\beta|v_t|^2_{L^2(H^{-1})} + \tfrac{1}{2}\beta(1 - \varepsilon_1)|h|^2_{W_a},$$

for all $\beta \in [\underline{\beta}, \beta_0]$. Here we also used the fact that $\delta \leqq \frac{1}{2\kappa_1}$. Returning to (5.3) we find

$$D^2_{(a,u,w)} L^\beta_{\delta\beta}(a^*_\beta, u^*_\beta, w^*_\beta; \Lambda^*)((h, v, y)(h, v, y)) \geqq M(\beta, \delta\beta)(h, v) + \delta\beta(\langle a^*_\beta, h\rangle_{W_a} + y)^2 , \tag{5.11}$$

where for any $\rho \in (0, 1)$

$$(\langle a^*_\beta, h\rangle_{W_a} + y)^2 \geqq \langle a^*_\beta, h\rangle^2 - \frac{1}{\rho^2}\langle a^*_\beta, h\rangle^2_{W_a} - \rho^2 y^2 + y^2 \geqq (1-\rho^2)y^2 - \left(\frac{1}{\rho^2} - 1\right)\gamma^2 |h|^2_{W_a} . \tag{5.12}$$

Hence, together with (5.10) we find

$$D^2_{(a,u,w)} L^\beta_{\delta\beta}(a^*_\beta, u^*_\beta, w^*_\beta; \Lambda^*)((h, v, y)(h, v, y)) \geqq \beta\Big[\tfrac{1}{2}|v|^2_{L^2(H^1_0)} + \tfrac{1}{2}\delta|v_t|^2_{L^2(H^{-1})} + \tfrac{1}{2}(1-\varepsilon_1)|h|^2_{W_a} + \delta(1-\rho^2)y^2 - \delta\gamma^2\left(\frac{1}{\rho^2} - 1\right)|h|^2_{W_a}\Big]$$

$$= \beta\Big[\tfrac{1}{2}|v|^2_{L^2(H^1_0)} + \tfrac{1}{2}\delta|v_t|^2_{L^2(H^{-1})} + \delta(1-\rho^2)y^2 + \left(\tfrac{1}{2}(1-\varepsilon_1) - \delta\gamma^2\left(\frac{1}{\rho^2} - 1\right)\right)|h|^2_{W_a}\Big] .$$

Let ρ_0 be the unique positive root of

$$\delta(1-\rho^2) = \tfrac{1}{2}(1-\varepsilon_1) - \delta\gamma^2\left(\frac{1}{\rho^2} - 1\right)$$

in (0, 1). Then for all $c \geqq c_0 = \delta\beta^0$ the following inequality holds for all $\beta \in [\underline{\beta}, \beta_0]$

$$D^2_{(a,u,w)} L^\beta_c(a^*_\beta, u^*_\beta, w^*_\beta; \Lambda^*)((h, v, y), (h, v, y)) \geqq D^2_{(a,u,w)} L^\beta_{\delta\beta}(a^*_\beta, u^*_\beta, w^*_\beta; \Lambda^*)((h, v, y), (h, v, y)) \geqq \beta\tilde{\sigma}|(h, v, y)|^2_{W_a \times W \times \mathbb{R}} ,$$

where $\tilde{\sigma} = \min(\frac{1}{2}, \frac{1}{2}\delta, \delta(1-\rho_0^2))$. This is the desired estimate (5.2). □

Let us turn to the case $z = u^*_0$. From Proposition 5.1 and (5.8) with $\varepsilon = \frac{1}{4}$ we derive

$$M(\beta, \delta\beta)(h, v) \geqq \left(1 - \frac{\beta}{2} - 8\frac{k_1^2}{\nu^2}\rho(\beta)\right)|v|^2_{L^2(H^1_0)} + \tfrac{1}{2}\beta\delta|v_t|^2_{L^2(H^{-1})} + \tfrac{3}{8}\beta|h|^2_{W_a} .$$

Next we choose β_0 such that

$$\rho(\beta_0) \leqq \frac{\nu^2}{8k_1^2}(1-\beta_0) .$$

Since $\beta \to \rho(\beta)$ is monotonically increasing we have

$$1 - \frac{\beta}{2} - \frac{8k_1^2}{\nu^2}\rho(\beta) \geqq 1 - \frac{\beta_0}{2} - \frac{8k_1^2}{\nu^2}\rho(\beta_0) \geqq \frac{\beta_0}{2} \quad \text{for } \beta \in [0, \beta_0] ,$$

and consequently

$$M(\beta, \delta\beta)(h, v) \geqq \tfrac{1}{2}\beta_0 |v|^2_{L^2(H^1_0)} + \tfrac{1}{2}\beta\delta |v_t|^2_{L^2(H^{-1})} + \tfrac{3}{8}\beta |h|^2_{W_a}$$
$$\geqq \tfrac{1}{2}\beta(|v|^2_{L^2(H^1_0)} + \delta |v_t|^2_{L^2(H^{-1})} + \tfrac{3}{4}|h|^2_{W_a}),$$

for $\beta \in [0, \beta_0]$. We use this estimate in (5.3) and, recalling (5.12), we obtain for $c \geqq \delta\beta_0 = c_0$

$$D^2_{(a,u,w)} L^\beta_c(a^*_\beta, u^*_\beta, w^*_\beta; \Lambda^*)((h, v, y), (h, v, y))$$
$$\geqq \beta\left(\tfrac{1}{2}|v|^2_{L^2(H^1_0)} + \tfrac{1}{2}\delta |v_t|^2_{L^2(H^{-1})} + \delta(1-\rho^2)y^2 + \left(\tfrac{3}{8} - \delta\gamma^2\left(\frac{1}{\rho^2} - 1\right)\right)|h|^2_{W_a}\right).$$

We assign to ρ^2 the value $8\delta\gamma^2(1 + 8\delta\gamma^2)^{-1}$ which ensures that

$$\delta(1-\rho^2) = \frac{\delta}{1 + 8\delta\gamma^2} \quad \text{and} \quad \tfrac{3}{8} - \delta\gamma^2\left(\frac{1}{\rho^2} - 1\right) = \tfrac{1}{4},$$

so that for $c \geqq c_0$ and $\beta \in [0, \beta_0]$

$$D^2_{(a,u,w)} L^\beta_c(a^*_\beta, u^*_\beta, w^*_\beta; \Lambda^*)((h, v, y), (h, v, y)) \geqq \beta(\tfrac{1}{2}|v|^2_{L^2(H^1_0)} + \tfrac{1}{2}\delta |v_t|^2_{L^2(H^{-1})}$$
$$+ \frac{\delta}{1 + 8\delta\gamma^2} y^2 + \tfrac{1}{4}|h|^2_{W_a}) \geqq \beta\tilde{\sigma}|(h, v, y)|^2_{W_a \times W \times R},$$

with $\tilde{\sigma} = \min\left(\tfrac{1}{4}, \tfrac{1}{2}\delta, \dfrac{\delta}{1 + 8\delta\gamma^2}\right)$.

Next we address the special case in which the Lagrange multiplier μ^* is positive. In this case coercivity holds with $\beta = 0$ and the solution (a^*_0, u^*_0) of (P^0) need only be a local one. A sufficient condition for the nonnegative Lagrange multiplier μ^* to be in fact positive is given in [KW]. Of course, due to the complementarity condition, $\mu^* > 0$ can only hold if $|a^*_0|_{W_a} = \gamma$. The precise result is as follows.

Proposition 5.3. *Let (a^*, u^*) denote a local solution of (P^0) and suppose that $\mu^* > 0$ and*

$$|u^*_0 - z|_{L^2(H^1_0)} \leqq \frac{\nu \min(\mu^*, 1)}{4k_1}. \tag{5.13}$$

Then (C) *holds with $\beta = 0$.*

Proof. We use the same notation as in the proof of Proposition 5.2 and find

$$M(0, c)(h, v) + \mu^*|h|^2_{W_a} \geqq |v|^2_{L^2(H^1_0)} - 2\langle \lambda^*, (-\Delta)^{-1}(hv_x)_x\rangle_{L^2(H^1_0)}$$
$$+ c|(-\Delta)^{-1}[v_t - (a^*v_x)_x - (hu^*_x)_x]|^2_{L^2(H^1_0)} + \mu^*|h|^2_{W_a}$$
$$\geqq |v|^2_{L^2(H^1_0)} - 2\frac{k_1}{\nu}|u^* - z|_{L^2(H^1_0)}|h|_{W_a}|v|_{L^2(H^1_0)} + \frac{c}{2}|v_t|^2_{L^2(H^{-1})} - 2ck_1^2\gamma^2|v|^2_{L^2(H^1_0)}$$
$$- 2ck_1^2\left(\frac{1}{\nu}|\varphi|^2_{L^2} + \frac{1}{\nu^2}|f|^2_{L^2(H^{-1})}\right)|h|^2_{W_a} + \mu^*|h|^2_{W_a}$$

$$\geqq \frac{c}{2}|v_t|^2_{L^2(H^{-1})} + \left(1 - 2ck_1^2\gamma^2 - \frac{k_1}{\nu}|u^* - z|_{L^2(H_0^1)}\right)|v|^2_{L^2(H_0^1)}$$

$$+ \left[\mu^* - 2ck_1^2\left(\frac{1}{\nu}|\varphi|^2 + \frac{1}{\nu^2}|f|^2\right) - \frac{k_1}{\nu}|u^* - z|_{L^2(H_0^1)}\right]|h|^2_{W_a}$$

$$= \frac{c}{2}|v_t|^2_{L^2(H^{-1})} + \left(1 - c\kappa_1 - \frac{k_1}{\nu}|u^* - z|_{L^2(H_0^1)}\right)|v|^2_{L^2(H_0^1)}$$

$$+ \left(\mu^* - c\kappa_2 - \frac{k_1}{\nu}|u^* - z|_{L^2(H_0^1)}\right)|h|^2_{W_a}.$$

Inserting this estimate into the second Fréchet derivative of L_c^0 one obtains

$$D^2_{(a,u,w)}L_c^0(a^*, u^*, w^*; \Lambda^*)((h, v, y), (h, v, y)) \geqq \frac{c_0}{2}|v_t|^2_{L^2(H^{-1})}$$

$$+ \left(1 - c\kappa_1 - \frac{k_1}{\nu}|u^* - z|_{L^2(H_0^1)}\right)|v|^2_{L^2(H_0^1)}$$

$$+ \left(\mu^* - c\kappa_2 - \frac{k_1}{\nu}|u^* - z|_{L^2(H_0^1)} - c_0\gamma^2\left(\frac{1}{\rho^2} - 1\right)\right)|h|^2_{W_a} + c_0(1 - \rho^2)y^2$$

for any $c \geqq c_0$. Next we define

$$\delta = \min(\mu^*, 1) \qquad \text{and} \qquad c_0 = \frac{\delta}{4}\max\left(\frac{1}{\kappa_1}, \frac{1}{\kappa_2}\right)$$

and choose $\rho_0 \in (0, 1)$ such that $c_0\gamma^2\left(\frac{1}{\rho_0^2} - 1\right) \leqq \frac{\delta}{4}$. By assumption $\delta > 0$. This gives the estimate

$$D^2_{(a,u,w)}L_c^0(a^*, u^*, w^*; \Lambda^*)((h, v, y), (h, v, y))$$

$$\geqq D^2_{(a,u,w)}L_{c_0}^0(a^*, u^*, w^*; \Lambda^*)((h, v, y), (h, v, y))$$

$$\geqq \frac{c_0}{2}|v_t|^2_{L^2(H^{-1})} + \frac{\delta}{2}|v|^2_{L^2(H_0^1)} + \frac{\delta}{4}|h|^2_{W_a} + c_0(1 - \rho_0^2)y^2 \geqq \sigma|(h, v, y)|^2_{W_a \times W \times \mathbb{R}},$$

for every $c \geqq c_0$, where $\sigma = \min\left(\frac{c_0}{2}, \frac{\delta}{4}, c_0(1 - \rho_0^2)\right)$. This is the desired result. □

As a third case in which one can establish (C) we consider the situation when the optimization problem is finite dimensional. This corresponds to the situation occuring in numerical calculations when the discretization of the variables a and u is fixed. Let W^M and W_a^M be finite dimensional subspaces of W and W_a respectively and consider the unregularized problem

$$(P)^M \qquad \min|u - z|^2_{L^2(H_0^1)}$$

$$\text{subject to } e(a, u) = 0,\ a \in \mathscr{A} \cap W_a^M \quad \text{and} \quad u \in W^M.$$

We assume that there exists a solution (a^*, u^*) of $(P)^M$ and an associated Lagrange multiplier $(\lambda^*, \mu^*, \eta^*)$. Let $\mathscr{P}: L^2(\Omega) \to L^2(\Omega)$ denote the unique continuous extension of $\mathscr{P}: H^1(\Omega) \to L^2(\Omega)$ with $\mathscr{P}_\varphi = \nabla(-\Delta)^{-1}\nabla\varphi$ and observe that $\mathscr{P}$ is an orthogonal projection whose kernel is the set of all constant functions [IK2]. We require moreover that there exists a constant K (depending on W^M, W_a^M and u^*) such that

$$(5.14)\qquad |v_t|^2_{L^2(H^{-1})} + |v(0,\cdot)|^2_{L^2} \leqq K|v|^2_{L^2(H_0^1)}, \quad \text{for all } v \in W^M ,$$

and

$$(5.15)\qquad |\mathscr{P}(hu_x^*)|^2_{L^2(L^2)} \geqq K^{-1}|h|^2_{W_a}, \quad \text{for all } h \in W_a^M .$$

Proposition 5.4. *Let the above assumption on the existence of a solution of* $(P)^M$ *and an associated Lagrange multiplier as well as on the equivalence of the norms according to* (5.14), (5.15) *hold. Then condition* (C) *is satisfied, provided that* $|u^* - z|_{L^2(H_0^1)}$ *is sufficiently small.*

Proof. First we estimate $M(0, c)$ as defined in (5.4) from below:

$$\begin{aligned} M(0,c)(h,v) &= |v|^2_{L^2(H_0^1)} - 2\langle \lambda^*, (-\Delta)^{-1}(hv_x)_x\rangle_{L^2(H_0^1)} \\ &\quad + c|(-\Delta)^{-1}[v_t - (a^*v_x)_x - (hu_x^*)_x]|^2_{L^2(H_0^1)} + c|v(0,\cdot)|^2_{L^2} \\ &\geqq |v|^2_{L^2(H_0^1)} - 2\frac{k_1}{\nu}|u^* - z|_{L^2(H_0^1)}|h|_{W_a}|v|_{L^2(H_0^1)} \\ &\quad + \frac{c}{2}|\mathscr{P}hu_x^*|^2_{L^2(L^2)} - c|(-\Delta)^{-1}(v_t - (a^*v_x)_x)|^2_{L^2(H_0^1)} \\ &\geqq \left(1 - \frac{k_1}{\nu}|u^* - z|_{L^2(H_0^1)} - 2cK - 2ck_1^2\gamma^2\right)|v|^2_{L^2(H_0^1)} \\ &\quad + \left(\frac{c}{2}K^{-1} - \frac{k_1}{\nu}|u^* - z|_{L^2(H_0^1)}\right)|h|^2_{W_a} , \end{aligned}$$

where we have used (5.5), (5.14) and (5.15). In the next estimate use of (5.12) is made

$$\begin{aligned} &D^2_{(a,u,w)}L_c^\beta(a^*, u^*, w^*; \Lambda^*)((h,v,y),\ (h,v,y)) = M(0,c) + \mu^*|h|^2 + c(\langle a^*, h\rangle_{W_a} + y)^2 \\ &\geqq \left(1 - \frac{k_1}{\nu}|u^* - z|_{L^2(H_0^1)} - 2cK - 2ck_1^2\gamma^2\right)|v|^2_{L^2(H_0^1)} \\ &\quad + \left(K^{-1}\frac{c}{2} - \frac{k_1}{\nu}|u^* - z|_{L^2(H_0^1)} - c\gamma^2\left(\frac{1}{\rho^2} - 1\right)\right)|h|^2_{W_a} + c(1-\rho^2)y^2 . \end{aligned}$$

Choosing $\rho^2 = \dfrac{4K\gamma^2}{4K\gamma^2 + 1}$ we arrive at

$$\begin{aligned} &D^2_{(a,u,w)}L_c(a^*, u^*, w^*; \Lambda^*)((h,v,y),(h,v,y)) \\ &\geqq \left(\frac{1}{2} + \frac{1}{2K}\right)\left(1 - \frac{k_1}{\nu}|u^* - z|_{L^2(H_0^1)} - 2cK - 2ck_1^2\gamma^2\right)(|v|^2_{L^2(H_0^1)} + |v|^2_{L^2(H^{-1})}) \\ &\quad + \left(\frac{c}{4}K^{-1} - \frac{k_1}{\nu}|u^* - z|_{L^2(H_0^1)}\right)|h|^2_{W_a} + c(1 + 4K\gamma^2)^{-1}y^2 . \end{aligned}$$

From this estimate it follows that for appropriately defined constants c_0 and σ

$$D^2_{(a,u,w)} L_c(a^*, u^*, w^*; \lambda)((h, v, y),(h, v, y)) \geqq \sigma |(h, v, y)|^2_{W_a \times W \times \mathbb{R}}$$

for all $c \geqq c_0$, provided that $|u^* - z|_{L^1(H_0^1)}$ is sufficiently small.

If $(P)^M$ is considered with a regularization term then (5.15) may be omitted. □

6 A numerical example

The proposed algorithm was tested for numerous numerical examples some of which will be presented together with the relevant numerical specifications in a technical report. Here we give only one specific example which arises from discretizing u with bilinear splines over the grid of 22×38 elements and discretizing a by bilinear splines with half as many elements in each coordinate direction. In particular we consider the estimation of a in (2.1) when $T = 1$,

$$u^*(t, x) = t \sin \pi x \, ,$$

$$a^*(t, x) = 1 + \sin^2 \pi x \sin^2 \pi t \, ,$$

$$f(t, x) = \sin \pi x + \pi^2 t \sin \pi x - \pi^2 t \sin^2 \pi t (2 \sin \pi x \cos^2 \pi x - \sin^3 \pi x) \, ,$$

and $\varphi \equiv 0$. The data are given by

$$z_{ij} = u^*_{ij}(1 + \delta r_{ij})$$

where r_{ij} are uniformly distributed random numbers with values in $[-1, 1]$, $\delta = .01$, $c = 10$, $\beta = 5 \cdot 10^{-6}$ and u^*_{ij} are the values of u^* at the nodal points of the grid. The start up values of λ, a, u were chosen to be $(\lambda^0, a^0, u^0) = (0, 1, z)$. With these specifications the result for a after 5 iterations can be seen in Fig. 2, the optimal parameter is depicted in Fig. 1. We observe that the supremum error is

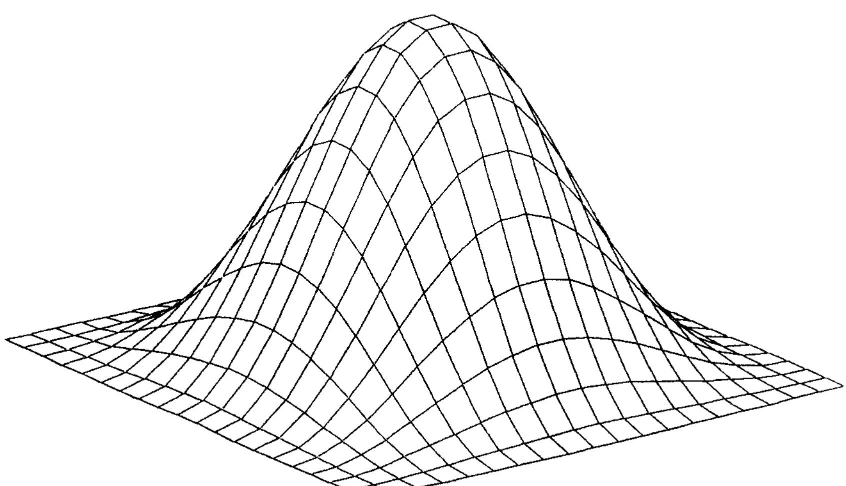

Fig. 1

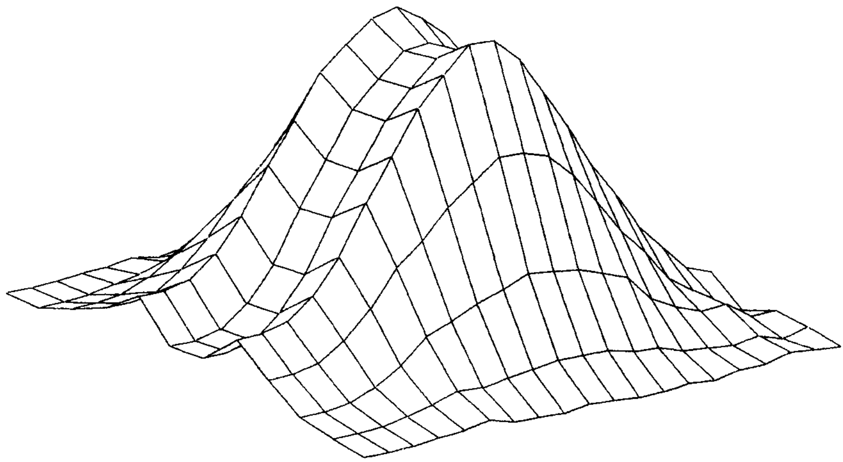

Fig. 2

concentrated along the singular set $\mathscr{S} = \{(t, x) \in Q : u_x^*(t, x) = 0\}$. Numerically the errors are given by

$$|a^{22,38} - a^*|_{L_2} = 0.06$$

$$|a^{22,38} - a^*|_{W_a} = 3.42$$

$$|a^{22,38} - a^*|_\infty = 0.22 \,.$$

In the case of unperturbed data, i.e. $\delta = 0$, the graph of the estimated parameter is indistinguishable from Fig. 1. For example, we found $|a^{38,38} - a^*|_\infty = 0.008$.

Appendix

Proof of Lemma 3.1. Let $(a^n, u^n) \in W_a \times W$ be a minimizing sequence. This implies that $a^n \in \mathscr{A}$ and $u^n = u(a^n)$ by (3.1). Since $\mathscr{A}$ is weakly compact there exists a subsequence (which we denote again by) a^n converging weakly to $a \in \mathscr{A}$. By Lemma 2.3 the corresponding solutions u^n converge strongly in $L^2(H_0^1)$ to $u(a)$. Hence we may infer

$$\begin{aligned}\inf_{\substack{e(a,u)=0\\ a\in\mathscr{A}}} J^\beta(a, u) &= \lim_{n\to\infty} J^\beta(a^n, u^n)\\ &= \frac{1}{2}\lim_{n\to\infty} |u^n - z|^2_{L^2(H_0^1)} + \frac{\beta}{2}\lim_{n\to\infty}\inf |a^n|^2_{W_a}\\ &= \frac{1}{2}|u - z|^2_{L^2(H_0^1)} + \frac{\beta}{2}\lim_{n\to\infty}\inf |a^n|^2_{W_a}\\ &\geqq \frac{1}{2}|u - z|^2_{L^2(H_0^1)} + \frac{\beta}{2}|a|^2_{W_a} = J^\beta(a, u)\,,\end{aligned}$$

where the last inequality is a consequence of the weak (sequential) lower-semicontinuity of the norm.

To show the existence of a Lagrange multiplier at a local solution (a^*, u^*) of (P^β) we refer to the version of the Kuhn–Tucker Theorem in [MZ]. It requires the verification of a regular point condition at (a^*, u^*). Performing the appropriate identifications with the quantities in [MZ] this amounts to the condition (recall the definition of the positive cone in W_a):

$$\begin{aligned}0 \in \operatorname{int}\{(((-\Delta)^{-1}[v_t - (hu_x^*)_x - (a^*v_x)_x], v(0,\cdot)),\\ -\tfrac{1}{2}(|a^*|^2_{W_a} - \gamma^2) - \langle a^*, h\rangle_{W_a} - \rho, a^* - v + h - b):\\ (h, v) \in W_a \times W, \rho \in \mathbb{R}_+, b \in W_a^+\} \subset (L^2(H_0^1) \times L^2(\Omega)) \times \mathbb{R} \times W_a .\end{aligned}$$

Define

$$\mathscr{U} = \{(\tilde{f}, \psi, r, a) \in (L^2(H_0^1) \times L^2(\Omega)) \times \mathbb{R} \times W_a : |r| \leqq \tfrac{1}{4}\gamma(\gamma - v\sqrt{T}), |a|_{L^\infty} \leqq v\sigma\} ,$$

with

$$\sigma = \frac{1}{2}\frac{\gamma - v\sqrt{T}}{\gamma + v\sqrt{T}} .$$

Note that $\mathscr{U}$ is a neighborhood of the origin in $L^2(H_0^1) \times L^2(\Omega) \times \mathbb{R} \times W_a$ (see Lemma 2.1). We show that for any $(\tilde{f}_0, \psi_0, r_0, a_0) \in \mathscr{U}$ there are $(h, v) \in W_a \times W$, $\rho \in \mathbb{R}^+$, $b \in W_a^+$ such that

$$\begin{aligned}&(-\Delta)^{-1}(v_t - (hu_x^*)_x - (a^*v_x)_x) = \tilde{f}_0, \qquad v(0,\cdot) = \psi_0 ,\\ &\tfrac{1}{2}(|a^*|^2_{W_a} - \gamma^2) + \langle a^*, h\rangle_{W_a} + \rho = -r_0 ,\\ &v - a^* - h + b = -a_0 .\end{aligned} \tag{A.1}$$

Choosing

$$h = -\tfrac{1}{2}a^*(1 - \sigma) + \tfrac{1}{2}v(a + \sigma) \in W_a ,$$

we find that b as given by the third equation of (A.1) satisfies

$$b \geqq -v\sigma - \tfrac{1}{2}v(1 - \sigma) + \tfrac{1}{2}a^*(1 + \sigma) \geqq 0 ,$$

and hence $b \in W_a^+$. Inserting h into the second equation of (A.1) together with the estimate

$$|\langle a^*, v\rangle_{W_a}| \leqq v\gamma\sqrt{T} \tag{A.2}$$

implies

$$\begin{aligned}\rho &\geqq -\tfrac{1}{4}\gamma(\gamma - v\sqrt{T}) + \tfrac{1}{2}(1 - \sigma)\gamma^2 - \tfrac{1}{2}\gamma v\sqrt{T}(1 + \sigma)\\ &= \tfrac{1}{4}\gamma(\gamma - v\sqrt{T}) - \tfrac{1}{2}\sigma\gamma(\gamma + v\sqrt{T}) = 0 ,\end{aligned}$$

and therefore $\rho \in \mathbb{R}^+$. The first equation in (A.1) results in

$$\begin{aligned}&v_t - (a^*v_x)_x = (-\Delta)\tilde{f}_0 + (hu_x^*)_x ,\\ &v(0,\cdot) = \psi_0\end{aligned}$$

which has a (unique) solution for any choice of $(\tilde{f}_0, \psi_0)$ in $L^2(H_0^1)\times L^2$. Applying the Kuhn–Tucker theorem in [MZ] we draw the conclusion that there is a multiplier $(\lambda^*, \mu^*, \eta^*)\in(L^2(H_0^1)\times L^2(\Omega))\times \mathbb{R}\times W_a$ such that (a^*, u^*) is a stationary point of $L^\beta(a^*, u^*; \lambda^*, \mu^*, \eta^*)$ and the complementarity conditions hold:

$$\text{(A.3)}\qquad \mu^* g(a^*) = 0 = \langle \eta^*, l(a^*)\rangle_{W_a},$$
$$\mu^* \geqq 0, \qquad \langle \eta^*, a\rangle \geqq 0 \quad \text{for all } a\in W_a^+ .$$ □

For further reference we note that for $(h, v)\in W_a\times W$

(A.4)

$$D_{(a,u)}L^\beta(a^*, u^*; \lambda^*, \eta^*)(h, v) = \langle u^* - z, v\rangle_{L^2(H_0^1)} + \beta\langle a^*, h\rangle_{W_a}$$
$$+ \langle \lambda_1^*, (-\Delta)^{-1}[v_t - (a^* v_x)_x - (hu_x^*)_x]\rangle_{L^2(H_0^1)}$$
$$+ \langle \lambda_2^*, v(0,\cdot)\rangle_{L^2(\Omega)} + \mu^*\langle a^*, h\rangle_{W_a} - \langle \eta^*, h\rangle_{W_a} .$$

Proof of Lemma 3.2. Put $h = 0$ in (A.4) and conclude that in view of the first order necessary optimality condition

(A.5)

$$\langle \lambda_1^*, v_t\rangle_{L^2(L^2)} - \langle (a^*(\lambda_1^*)_x)_x, v\rangle_{L^2(L^2)} + \langle \lambda_2^*, v(0,\cdot)\rangle_{L^2(\Omega)} + \langle (-\Delta)(u^* - z), v\rangle_{L^2(L^2)} = 0$$

holds for all $v\in W$, in particular for $v(t) = \chi(t)w$ where $w\in H_0^1(\Omega)$ and $\chi\in\mathscr{D}(0, T)$. For this choice of v we find

$$\langle \lambda_1^*, v_t\rangle_{L^2(L^2)} = \left\langle \int_0^T \lambda_1^*(t)\dot{\chi}dt, w\right\rangle_{L^2(\Omega)} = -\left\langle \int_0^T D_t\lambda_1^*(t)\chi dt, w\right\rangle_{L^2(\Omega)},$$

where $D_t\lambda_1^*$ denotes the distributional derivative of λ_1^* with respect to t. The remaining terms in (A.5) give

$$\langle(-\Delta)(u^* - z), v\rangle_{L^2(L^2)} - \langle (a^*(\lambda_1^*)_x)_x, v\rangle_{L^2(L^2)} = \left\langle \int_0^T [(-\Delta)(u^* - z) - (a^*(\lambda_1^*)_x)_x]\chi dt, w\right\rangle_{L^2(\Omega)} .$$

Inserting these expressions into (A.5) yields

$$\left\langle \int_0^T [-D_t\lambda_1^*(t) - (a^*(\lambda_1^*)_x)_x - \Delta(u^* - z)]\chi dt, w\right\rangle_{L^2(\Omega)} = 0$$

for all $w\in H_0^1(\Omega)$ and $\chi\in\mathscr{D}(0, T)$. By density of $H_0^1(\Omega)$ in $L^2(\Omega)$ we infer

$$\text{(A.6)}\qquad D_t\lambda_1^* = -(a^*(\lambda_1^*)_x)_x - \Delta(u^* - z),$$

and since the right hand side is an element of $L^2(H^{-1})$ we conclude that $D_t\lambda_1^*\in L^2(H^{-1})$ and consequently $\lambda_1^*\in W$. Hence we may apply (A.6) to the first term in (A.5). Observing (A.6) we find for all $v\in W$

$$0 = \langle -(\lambda_1^*)_t - (a^*(\lambda_1^*)_x)_x + (-\Delta)(u^* - z), v\rangle_{L^2(L^2)}$$
$$+ \int_0^T \frac{d}{dt}\langle \lambda_1^*, v\rangle_{L^2(\Omega)} dt + \langle \lambda_2^*, v(0,\cdot)\rangle_{L^2(\Omega)}$$
$$= \langle \lambda_1^*(T), v(T,\cdot)\rangle_{L^2(\Omega)} - \langle \lambda_1^*(0) - \lambda_2^*, v(0,\cdot)\rangle_{L^2(\Omega)} .$$

Choosing v appropriately we conclude

$$\lambda_1^*(T) = 0 \quad \text{and} \quad \lambda_2^* = \lambda_1^*(0) . \tag{A.7}$$

This establishes the second part of Lemma 3.2 and by uniqueness of the solution of the backward equation (A.6), (A.7) we infer uniqueness of λ^*.

To show uniqueness of (μ^*, η^*) let $(\lambda^*, \mu_i^*, \eta_i^*) \in (L^2(H_0^1) \times L^2(\Omega)) \times \mathbb{R} \times W_a$, $i = 1, 2$ be two pairs of Lagrange multipliers so that

$$D_{(a,u)} L^\beta(a^*, u^*; \lambda^*, \mu_i^*, \eta_i^*)(h, v) = 0$$

is satisfied for all $(h, v) \in W_a \times W$. Set $v = 0$ and subtract the two identities. This gives

$$\delta\mu^* \langle a^*, h \rangle_{W_a} = \langle \delta\eta^*, h \rangle_{W_a} \quad \text{for all } h \in W_a , \tag{A.8}$$

where

$$\delta\mu^* = \mu_1^* - \mu_2^* \quad \text{and} \quad \delta\eta^* = \eta_1^* - \eta_2^* .$$

First assume $a^* \not\equiv v$ and choose $\tilde{h} = a^* - v$. Inserting $\tilde{h}$ into (A.8) and using the complementarity conditions (A.3) one obtains

$$\delta\mu^*(|a^*|_{W_a}^2 - \langle a^*, v \rangle_{W_a}) = (\delta\eta^*, a^* - v \rangle = 0 .$$

Consequently, again referring to (A.3) and (A.2) one finds

$$0 = \delta\mu^*(\gamma^2 - \langle a^*, v \rangle_{W_a}) \geqq \delta\mu^* \gamma(\gamma - v\sqrt{T}) > 0 .$$

This contradiction implies that $\delta\mu^*$ and, by (A.8), $\delta\eta^*$ vanish. A similar argument may be carried out if $a^* \equiv v$. This completes the proof of Lemma 3.2. □

Next we shall use (A.6), (A.7) and (A.3) to derive a priori bounds on the Lagrange multiplier.

Lemma A.1. *There are constants $C_i = C_i(k_1, \gamma, v^{-1}, \beta)$, $i = 1, 2, 3$ such that the following estimates hold:*

(i) $|\lambda^*|_{W \times L^2(\Omega)} \leqq C_1(|\varphi|_{L^2(\Omega)} + |z|_{L^2(H_0^1)} + |f|_{L^2(H^{-1})})$,

(ii) $|\eta^*|_{W_a} \leqq \beta\gamma + C_2(|\varphi|_{L^2(\Omega)}^2 + |z|_{L^2(H_0^1)}^2 + |f|_{L^2(H^{-1})}^2)$,

(iii) $0 \leqq \mu^* \begin{cases} = 0 & \text{if } |a^*|_{W_a} < \gamma , \\ \leqq -\beta + C_3(|\varphi|_{L^2(\Omega)}^2 + |z|_{L^2(H_0^1)}^2 + |f|_{L^2(H^{-1})}^2) & \text{if } |a^*|_{W_a} = \gamma . \end{cases}$

Proof. From the estimates that follow it can be seen that the constants C_i can in principle be calculated explicitly. An argument analogous to that which led to Lemma 2.2 shows that

$$\frac{1}{v} |\lambda_1^*(0)|_{L^2(\Omega)}^2 + |\lambda_1^*|_{L^2(H_0^1)}^2 \leqq \frac{1}{v^2} |\Delta(u^* - z)|_{L^2(H^{-1})}^2 = \frac{1}{v^2} |u^* - z|_{L^2(H_0^1)}^2 ,$$

and

$$|(\lambda_1^*)_t|_{L^2(H^{-1})} \leqq \left[\frac{2k_1^2\gamma^2}{v^2} + 2 \right]^{1/2} (|u^*|_{L^2(H_0^1)} + |z|_{L^2(H_0^1)}) .$$

These estimates combined with Lemma 2.2 and (A.7) imply (i). Evaluating the first order necessary optimality condition (A.4) for $(h, 0) \in W_a \times W$ results in

(A.9)

$$-\langle \lambda_1^*, (-\Delta)^{-1}(hu_x^*)_x \rangle_{L^2(H_0^1)} + \mu^* \langle a^*, h \rangle_{W_a} - \langle \eta^*, h \rangle_{W_a} + \beta \langle a^*, h \rangle_{W_a} = 0, \; h \in W_a .$$

First assume $|a^*|_{W_a} < \gamma$ so that by (A.3) μ^* vanishes. Consequently (A.9) implies

$$|\eta^*|_{W_a} \leqq \beta\gamma + k_1 |\lambda_1^*|_{L^2(H_0^1)} |u^*|_{L^2(H_0^1)} . \tag{A.10}$$

Next we consider $|a^*|_{W_a} = \gamma$. This yields the estimate

$$\langle a^*, a^* - v \rangle_{W_a} \geqq \gamma(\gamma - v\sqrt{T}) > 0 . \tag{A.11}$$

Inserting $h = -a^* + v$ into (A.9) gives in view of (A.3)

$$\mu^* \langle a^*, a^* - v \rangle_{W_a} \leqq \beta \langle a^*, v - a^* \rangle_{W_a} + |\lambda_1^*|_{L^2(H_0^1)} |(v - a^*) u_x^*|_{L^2(L^2)} ,$$

Dividing by $\langle a^*, a^* - v \rangle$ and using (A.11) we arrive at

$$0 \leqq \mu^* \leqq -\beta + |\lambda_1^*|_{L^2(H_0^1)} |u^*|_{L^2(H_0^1)} k_1 \frac{|v - a^*|_{W_a}}{\gamma(\gamma - v\sqrt{T})} .$$

Since $\langle v, a^* \rangle_{W_a} \geqq v^2 T$ we infer

$$|v - a^*|_{W_a}^2 = v^2 T - 2\langle v, a^* \rangle_{W_a} + \gamma^2 \leqq \gamma^2 - v^2 T = (\gamma - v\sqrt{T})(\gamma + v\sqrt{T})$$

and hence

$$0 \leqq \mu^* \leqq -\beta + |\lambda_1^*|_{L^2(H_0^1)} |u^*|_{L^2(H_0^1)} \frac{k_1}{\gamma} \sqrt{\frac{\gamma + v\sqrt{T}}{\gamma - v\sqrt{T}}} .$$

This implies the estimate given in (iii). We return to (A.9) and – still assuming that $|a^*|_{W_a} = \gamma$ – we obtain the estimate

$$|\langle \eta^*, h \rangle_{W_a}| \leqq [(\mu^* + \beta)\gamma + k_1 |\lambda_1^*|_{L^2(H_0^1)} |u^*|_{L^2(H_0^1)}] |h|_{W_a} .$$

Combining the estimate with the one just derived for μ^* results in

$$|\eta^*|_{W_a} \leqq k_1 \left[1 + \frac{1}{\gamma} \sqrt{\frac{\gamma + v\sqrt{T}}{\gamma - v\sqrt{T}}} \right] |\lambda_1^*|_{L^2(H_0^1)} |u^*|_{L^2(H_0^1)} .$$

This inequality together with (A.10) and the estimates for $|\lambda_1^*|_{L^2(H_0^1)}$ and $|u^*|_{L^2(H_0^1)}$ imply the bound for $|\eta^*|_{W_a}$ as shown in (ii). □

Proof of Lemma 3.3. As in the proof of Lemma 3.1 one can show that $(a^*, u^*, -w^*) = (a^*, u^*, -g(a^*))$ is a regular point in the sense of [MZ] which implies the existence of a (unique) Lagrange multiplier that satisfies (i)–(iii) of Lemma 3.3. Instead, we shall verify directly that the quadrupel $(\lambda^*, \mu^*, \mu^*, \eta^*)$ satisfies these conditions if $(\lambda^*, \mu^*, \eta^*)$ is the Lagrange multiplier for (P^β) at (a^*, u^*). Recall (3.8) and observe that

$$\hat{L}_c^\beta(a, u, w; \lambda^*, \mu^*, \mu^*, \eta^*) = L^\beta(a, u; \lambda^*, \mu^*, \eta^*) + \frac{c}{2} |e(a, u)|^2_{L^2(H_0^1) \times L^2(\Omega)}$$
$$+ \frac{c}{2} |g(a) + w|^2 .$$

Therefore Lemma (3.1) implies for $(h, v, y) \in W_a \times W \times \mathbb{R}$

$$D_{(a,u,w)} \hat{L}_c^\beta(a^*, u^*, -g(a^*); \lambda^*, \mu^*, \mu^*, \eta^*)(h, v, y)$$
$$= D_{(a,u)} L^\beta(a^*, u^*; \lambda^*, \mu^*, \eta^*)(h, v) = 0 ,$$

which is (ii) of Lemma 3.3. Conditions (i) and (iii) of Lemma 3.3 coincide with the complementarity conditions (A.3). This completes the proof of Lemma 3.3. □

Proof of Lemma 3.4. We commence the proof with a brief discussion of L_c^β which was defined in (3.6). $L_c^\beta(a, u, w; \lambda, \mu, \eta)$ is three times continuously Fréchet differentiable with respect to (a, u, w). For further reference we note for $(h, v, y) \in W_a \times W \times \mathbb{R}$

(A.12)
$$D^2_{(a,u,w)} L_c^\beta(a, u, w; \lambda^*, \mu^*, \eta^*)((h, v, y), (h, v, y))$$
$$= |v|_{L^2(H_0^1)} + \beta |h|^2_{W_a} - 2\langle \lambda_1^*, (-\Delta)^{-1}(hv_x)_x \rangle_{L^2(H_0^1)}$$
$$+ \mu^* |h|^2_{W_a} + c|(-\Delta)^{-1}[v_t - (av_x)_x - (hu_x)_x]|^2_{L^2(H_0^1)}$$
$$- 2c\langle (-\Delta)^{-1}[u_t - (au_x)_x - f], (-\Delta)^{-1}(hv_x)_x \rangle_{L^2(H_0^1)}$$
$$+ c|v(0, \cdot)|^2_{L^2(\Omega)} + c(\langle a, h \rangle_{W_a} + y)^2 + c(\tfrac{1}{2}(|a|^2_{W_a} - \gamma^2) + w)|h|^2_{W_a} .$$

The third derivative is given by

$$D^3_{(a,u,w)} L_c^\beta(a, u, w; \lambda^*, \mu^*, \eta^*)((h, v, y)^{(3)}) = 3c(\langle a, h \rangle_{W_a} + y)|h|^2_{W_a}$$
$$- 6c\langle (-\Delta)^{-1}[v_t - (av_x)_x - (hu_x)_x], (-\Delta)^{-1}(hv_x)_x \rangle_{L^2(H_0^1)} .$$

This together with the estimate

$$|(-\Delta)^{-1}(av_x)_x|^2_{L^2(H_0^1)} \leqq |av_x|_{L^2(L^2)} |(-\Delta)^{-1}(av_x)_x|_{L^2(H_0^1)}$$
$$\leqq |av_x|^2_{L^2(L^2)} \leqq k_1^2 |a|^2_{W_a} |v|^2_{L^2(H_0^1)}$$

implies the bound

$$|D^3_{(a,u,w)} L_c^\beta(a, u, w; \lambda^*, \mu^*, \eta^*)((h, v, y)^{(3)})|$$
$$\leqq 3c|h|^2_{W_a}(|a|_{W_a}|h|_{W_a} + |y|) + 6ck_1[|v_t|_{L^2(H^{-1})}$$
$$+ k_1|a|_{W_a}|v|_{L^2(H_0^1)} + k_1|h|_{W_a}|u|_{L^2(H_0^1)}]|h|_{W_a}|v|_{L^2(H_0^1)} .$$

Expanding $L_c^\beta(a, u, w; \lambda^*, \mu^*, \eta^*)$ in a neighborhood of a local solution (a^*, u^*, w^*) of (P_c^β) into a Taylor series and taking into account that $L_c^\beta(a, u, w; \lambda^*, \mu^*, \eta^*) = \hat{L}_c^\beta(a, u, w; \lambda^*, \mu^*, \mu^*, \eta^*)$ and that

$$D_{(a,u,w)} \hat{L}_c^\beta(a^*, u^*, -g(a^*); \lambda^*, \mu^*, \mu^*, \eta^*)(h, v, y) = 0 ,$$

we find for $c \geqq c_0$

(A.13)
$$L_c^\beta(a, u, w, \lambda^*, \mu^*, \eta^*) \geqq L_{c_0}^\beta(a, u, w; \lambda^*, \mu^*, \eta^*)$$
$$= J^\beta(a^*, u^*) + D^2_{(a,u,w)} L_{c_0}^\beta(a^*, u^*, w^*; \lambda^*, \mu^*, \eta^*)(\Delta^*, \Delta^*) + R ,$$

where we have set $\Delta^* = (a - a^*, u - u^*, w - w^*)$. It can be shown that the estimate for the third Fréchet derivative of L_c^β derived above implies

(A.14)
$$|R| \leqq \tfrac{1}{6} k_2 c_0 |\Delta^*|^3_{W_a \times W \times \mathbb{R}}$$

with

$$k_2 = 9\gamma + 3 + 6k_1 + 30k_1^2\gamma + 6k_1^2\left(\frac{1}{\nu}|\varphi|^2_{L^2(\Omega)} + \frac{1}{\nu^2}|f|^2_{L^2(H^{-1})}\right)^{1/2} .$$

We further note that (cf. (3.6), (3.9))

$$\begin{aligned}&\mathscr{L}_c^\beta(a, u, w; \lambda^*, \mu^*) - L_c^\beta(a, u, w; \lambda^*, \mu^*, \eta^*)\\ &= \mu^* w - \langle \eta^*, l(a)\rangle_{W_a} \geqq 0 \quad \text{for } w \geqq 0, \quad l(a) \leqq 0 .\end{aligned}$$

The last inequality is a consequence of (A.3). This fact together with

$$\mathscr{L}_c^\beta(a^*, u^*, w^*; \lambda^*, \mu^*) = J^\beta(a^*, u^*), \quad c \in \mathbb{R}^+ ,$$

(A.13), (A.14) and condition (C) justifies the following estimates (c_0 as specified by condition (C))

$$\begin{aligned}\mathscr{L}_c^\beta(a, u, w; \lambda^*, \mu^*) &\geqq \mathscr{L}_{c_0}^\beta(a, u, w; \lambda^*, \mu^*) \geqq L_{c_0}^\beta(a, u, w; \lambda^*, \mu^*, \eta^*)\\ &\geqq J^\beta(a^*, u^*) + \sigma|\Delta^*|^2_{W_a \times W \times \mathbb{R}} - |R|\\ &\geqq \mathscr{L}_c^\beta(a^*, u^*, w^*; \lambda^*, \mu^*) + \frac{\sigma}{2}|\Delta^*|^2_{W_a \times W \times \mathbb{R}} ,\end{aligned}$$

if only

$$\text{(A.15)} \qquad |\Delta^*|^2_{W_a \times W \times \mathbb{R}} \leqq \frac{3\sigma}{k_2 c_0} ,$$

and $w \geqq 0$, $l(a) \leqq 0$. This shows that (a^*, u^*, w^*) is a minimum of $(\tilde{P}_c^\beta)$ in a neighborhood of (a^*, u^*, w^*) which is determined by (A.15). The second assertion of Lemma 3.4 is a consequence of the observation that

$$\mathscr{L}_c^\beta(a, u, w; \lambda^*, \mu^*) = J^\beta(a, u, w) = J^\beta(a, u) ,$$

if (a, u, w) satisfies the constraints of (P_c^β) . □

Next we turn to the discussion of problems (3.3) and (3.4). For this purpose we first consider $(\tilde{P}_c^\beta)$ with (λ^*, μ^*) replaced by $(\lambda, \mu) \in (L^2(H_0^1) \times L^2(\Omega)) \times \mathbb{R}$, i.e.

$$\text{(A.16)} \qquad \min \mathscr{L}_c^\beta(a, u, w; \lambda, \mu) \quad \text{subject to } w \geqq 0, \quad \text{and} \quad l(a) \leqq 0 ,$$

where $\mathscr{L}_c^\beta$ was defined in (3.9). Since (A.16) is not related to $(\tilde{P}_c^\beta)$ anymore, we ensure existence of a solution by introducing the additional constraint $(a, u, w) \in \tilde{B}_\delta$, where for $\delta > 0$,

$$\tilde{B}_\delta = \left\{(a, u, w) \in W_a \times W \times \mathbb{R} : |(a, u) - (a^*, u^*)|_{W_a \times W} \leqq \delta, |w - w^*| \leqq \frac{\mu}{c} + \gamma\delta\right\} .$$

We also recall that $B_\delta = \{(a, u) \in W_a \times W : |(a, u) - (a^*, u^*)|_{W_a \times W} \leqq \delta\}$. Since $\mathscr{L}_c^\beta(a, u, w; \lambda, \mu)$ is weakly lower semicontinuous in (a, u, w) it attains a minimum at a point $(\bar{a}, \bar{u}, \bar{w})$ on the weakly sequentially compact set $\{(a, u, w) \in W_a \times W \times \mathbb{R}: w \geqq 0, l(a) \leqq 0\} \cap \tilde{B}_\delta$. Let us put

$$\text{(A.17)} \qquad w(a) = \max\left(0, -\frac{\mu}{c} - g(a)\right) .$$

Note that for $(a, u) \in W_a \times W$ fixed $w(a)$ minimizes $\mathscr{L}_c^\beta(a, u, \cdot; \lambda, \mu)$ over $w \geqq 0$. Then in view of $w^* = -g(a^*)$ we find

$$|w(a) - w^*| \leqq \frac{\mu}{c} + |g(a) - g(a^*)| ,$$

and hence

$$\text{(A.18)} \qquad |w(a) - w^*| \leqq \frac{\mu}{c} + \frac{1}{2} \big| |a|_{W_a}^2 - |a^*|_{W_a}^2 \big| \leqq \frac{\mu}{c} + \gamma |a - a^*|_{W_a} \leqq \frac{\mu}{c} + \gamma\delta ,$$

provided that $|a - a^*|_{W_a} \leqq \delta$. We also find

$$\mathscr{L}_c^\beta(\bar{a}, \bar{u}, \bar{w}; \lambda, \mu) \geqq \min_{\substack{|w - w^*| \leqq \mu/c + \gamma\delta \\ w \geqq 0}} \mathscr{L}_c^\beta(\bar{a}, \bar{u}, w; \lambda, \mu) = \mathscr{H}_c(\bar{a}, \bar{u}; \lambda, \mu) ,$$

where we used (A.18). Moreover we have

$$\begin{aligned} \min_{\substack{(a,u,w) \in \tilde{B}_\delta \\ w \geqq 0,\, l(a) \leqq 0}} \mathscr{L}(a, u, w; \lambda, \mu) &\leqq \min_{\substack{(a,u) \in \tilde{B}_\delta \\ l(a) \leqq 0}} \; \min_{\substack{|w - w^*| \leqq \mu/c + \gamma\delta \\ w \geqq 0}} \mathscr{L}(a, u, w; \lambda, \mu) \\ &= \min_{\substack{(a,u) \in B_\delta \\ l(a) \leqq 0}} \mathscr{H}_c(a, u; \lambda, \mu) \leqq \mathscr{H}_c(\bar{a}, \bar{u}; \lambda, \mu) , \end{aligned}$$

where again (A.18) was employed. The last two estimates imply

$$\min_{\substack{(a,u,w) \in \tilde{B}_\delta \\ w \geqq 0,\, l(a) \leqq 0}} \mathscr{L}_c^\beta(a, u, w; \lambda, \mu) = \min_{\substack{(a,u) \in B_\delta \\ l(a) \leqq 0}} \mathscr{H}_c(a, u; \lambda, \mu) .$$

The optimization problem on the right hand side of the last equality is problem (3.4), on which our algorithm is based.

The proof of convergence of the augmented Lagrangian algorithm will require the following auxiliary functional. Define

$$\begin{aligned} L^*(a, u; \mu, c) &= J^\beta(a, u) + \langle \lambda^*, e(a, u) \rangle_{L^2(H_0^1) \times L^2(\Omega)} \\ &+ \frac{c_0}{2} |e(a, u)|_{L^2(H_0^1) \times L^2(\Omega)}^2 + \mu^* \hat{g}(a, \mu, c) + \frac{c_0}{2} |\hat{g}(a, \mu, c)|^2 + \langle \eta^*, v - a \rangle_{W_a} , \end{aligned}$$

where c_0 is defined in condition (C) and $(\lambda^*, \mu^*, \eta^*)$ is the Lagrange multiplier associated with the local solution $(a^*, u^*, -g(a^*))$ of (P_c^β). We have the following result

Lemma A.2. *Assume that condition* (C) *holds at a local solution* $(a^*, u^*, -g(a^*))$ *of* (P_c^β). *Choose* $\hat{\mu} \in \mathbb{R}^+$ *arbitrary and define*

$$\delta^* = \frac{1}{2} \sqrt{\frac{3\sigma}{k_2 c_0}} (2 + \gamma)^{-1}, \quad \tilde{c} = \max(\hat{\mu}(\delta^*)^{-1}, 2c_0) ,$$

and $B_{\delta^*} = \{(a, u) : |(a, u) - (a^*, u^*)|_{W_a \times W} < \delta^*\}$. *Then* $(a, u) \in B_{\delta^*}$ *implies*

$$L^*(a, u; \mu, c) \geqq J^\beta(a^*, u^*) + \frac{\sigma}{2} |(a - a^*, u - u^*)|_{W_a \times W}^2 ,$$

for all $c \geqq \tilde{c}$ *and* $\mu \in [0, \hat{\mu}]$.

Since $\hat{g}(a^*, \mu^*, c) = 0$, we have

$$L^*(a^*, u^*; \mu^*, c) = J^\beta(a^*, u^*) \,, \tag{A.19}$$

and thus Lemma A.2 implies in particular that (a^*, u^*) affords an unrestricted minimum to L^* on B_{δ^*}.

Proof of Lemma A.2. Recall the definitions of L_c^β and $w(a)$ ((3.6) and (A.17) respectively) and infer that

$$L_{c_0}^\beta(a, u, w(a); \lambda^*, \mu^*, \eta^*) + \mu^* w(a) = L^*(a, u; \mu, c) \,.$$

Since $\mu^* w(a) \geqq 0$ we conclude that

$$L^*(a, u; \mu, c) \geqq L_{c_0}^\beta(a, u, w(a); \lambda^*, \mu^*, \eta^*) \tag{A.20}$$

In the proof of Lemma 3.4 it was shown that

(A.21)
$$L_{c_0}^\beta(a, u, w; \lambda^*, \mu^*, \eta^*) \geqq J^\beta(a^*, u^*) + \frac{\sigma}{2} |(a - a^*, u - u^*, w - w^*)|^2_{W_a \times W \times \mathbb{R}}$$

provided that

$$|(a - a^*, u - u^*, w - w^*)|^2_{W_a \times W \times \mathbb{R}} \leqq \frac{3\sigma}{k_2 c_0} \,. \tag{A.22}$$

Due to the choice of $\mu \in [0, \hat{\mu}]$, δ^* and $\tilde{c}$ we have

$$\frac{\mu}{c} \leqq \frac{\hat{\mu}}{\tilde{c}} \leqq \delta^* \,.$$

For $(a, u) \in B_{\delta^*}$ it follows from (A.18) that

$$|(a - a^*, u - u^*, w(a) - w^*)|_{W_a \times W \times \mathbb{R}} \leqq |(a - a^*, u - u^*)|_{W_a \times W} + |w(a) - w^*|$$
$$\leqq \delta^* + \frac{\mu}{c} + \gamma\delta^* \leqq (2 + \gamma)\delta^* \leqq \sqrt{\frac{3\sigma}{k_2 c_0}} \,.$$

Therefore $(a, u, w(a))$ satisfies (A.22) if only $(a, u) \in B_{\delta^*}$. Consequently we may combine (A.20) and (A.21) which implies the assertion of Lemma A.2. □

The proof of Theorem 3.1 is similar to the one given in [IK1]. To make the paper selfcontained we present an outline of the proof.

Proof of Theorem 3.1. For every $n \geqq 1$, let (a^n, u^n) be a solution of (P_{c_n}) in B_δ and let (λ^n, μ^n) be determined by the algorithm. Using the update formulas and the fact that the penalty parameters c_n are nondecreasing with $c_1 \geqq 2c_0$, we find for $n = 1, 2, \ldots$

$$(\mu^{n-1} - \mu^*)\hat{g}(a^n, \mu^{n-1}, c_n) + \frac{c_n - c_0}{2} \hat{g}(a^n, \mu^{n-1}, c_n)^2 \tag{A.23}$$
$$\geqq \frac{1}{c_n} [(\mu^n - \mu^*)^2 - (\mu^{n-1} - \mu^*)^2]$$

and

$$\text{(A.24)}\qquad \langle \lambda^{n-1}-\lambda^*, e(a^n,u^n)\rangle_{L^2(H_0^1)\times L^2(\Omega)} + \frac{c_n-c_0}{2}|e(a^n,u^n)|^2_{L^2(H_0^1)\times L^2(\Omega)}$$

$$\geqq \frac{1}{c_n}[|\lambda^n-\lambda^*|^2_{L^2(H_0^1)\times L^2(\Omega)} - |\lambda^{n-1}-\lambda^*|^2_{L^2(H_0^1)\times L^2(\Omega)}]\,.$$

Since $\mu^0 \geqq 0$ it follows from the update formula for μ^n that $\mu^n \geqq 0$ for all $n = 0, 1, \ldots$. This implies that

$$\max(cg(a^*) + \mu^n, 0) \leqq \mu^n \quad \text{for all } n = 0, 1, \ldots\ .$$

This fact is used to establish the estimate

$$\text{(A.25)}\qquad \mathscr{H}_{c_n}(a^n, u^n; \lambda^{n-1}, \mu^{n-1}) \leqq \mathscr{H}_{c_n}(a^*, u^*; \lambda^{n-1}, \mu^{n-1})$$

$$= J^\beta(a^*, u^*) + \frac{1}{2c_n}[(\max(c_n g(a^*) + \mu^{n-1}, 0))^2 - (\mu^{n-1})^2]$$

$$\leqq J^\beta(a^*, u^*)\,.$$

Let us consider the quantity $2(\mu^* + |\lambda^*|_{L^2(H_0^1)\times L^2(\Omega)}) + \mu^0 + |\lambda^0|_{L^2(H_0^1)\times L^2(\Omega)}$. By Lemma A.2 this expression can be bounded by a constant $\bar{c}$ which depends on $(k_1, \gamma, \nu^{-1}, \beta, \varphi, f, z; \mu^0, \lambda^0)$. Henceforth we put $\hat{\mu} = \bar{c}$. Inductively we shall show that $\mu_n \in [0, \hat{\mu}]$. Clearly $\mu_0 \in [0, \hat{\mu}]$ and we now assume that $\mu_{n-1} \in [0, \hat{\mu}]$ as well. Then by (A.25), (A.23), (A.24), Lemma A.2 (which is applicable since $c_n \geqq \tilde{c}$, $\mu^{n-1} \in [0, \hat{\mu}]$ and $(a^n, u^n) \in B_\delta$, with $\delta \leqq \delta^*$) and (A.3) we find

$$J^\beta(a^*, u^*) \geqq \mathscr{H}_{c_n}(a^n, u^n; \lambda^{n-1}, \mu^{n-1})$$

$$= L^*(a^n, u^n; \mu^{n-1}, c_n) + \langle \lambda^{n-1}-\lambda^*, e(a^n,u^n)\rangle_{L^2(H_0^1)\times L^2(\Omega)}$$

$$+ \frac{c_n-c_0}{2}|e(a^n,u^n)|^2_{L^2(H_0^1)\times L^2(\Omega)} + (\mu^{n-1}-\mu^*)\hat{g}(a^n, \mu^{n-1}, c_n)$$

$$+ \frac{c_n-c_0}{2}\hat{g}(a^n, \mu^{n-1}, c_n)^2 - \langle \eta^*, v - a^n\rangle_{W_a}$$

$$\geqq J^\beta(a^*, u^*) + \frac{\sigma}{2}(|a^n - a^*|^2_{W_a} + |u^n - u^*|^2_W)$$

$$+ \frac{1}{c_n}[|\lambda^n-\lambda^*|^2_{L^2(H_0^1)\times L^2(\Omega)} - |\lambda^{n-1}-\lambda^*|^2_{L^2(H_0^1)\times L^2(\Omega)}]$$

$$+ \frac{1}{c_n}[(\mu^n-\mu^*)^2 - (\mu^{n-1}-\mu^*)^2]\,.$$

This estimate implies

$$|a^*-a^n|^2_{W_a} + |u^*-u^n|^2_W + \frac{2}{\sigma c_n}(|\lambda^*-\lambda^n|^2_{L^2(H_0^1)\times L^2(\Omega)} + |\mu^*-\mu^n|^2)$$

$$\leqq \frac{2}{\sigma c_n}(|\lambda^*-\lambda^{n-1}|^2_{L^2(H_0^1)\times L^2(\Omega)} + |\mu^*-\mu^{n-1}|^2)\,.$$

This is (3.7). Further we find

$$|\lambda^* - \lambda^n|^2_{L^2(H_0^1)\times L^2(\Omega)} + |\mu^* - \mu^n|^2 \leqq |\lambda^* - \lambda^{n-1}|^2_{L^2(H_0^1)\times L^2(\Omega)} + |\mu^* - \mu^{n-1}|^2$$
$$\leqq |\lambda^* - \lambda^0|^2_{L^2(H_0^1)\times L^2(\Omega)} + |\mu^* - \mu^0|^2$$

and therefore

$$\max(|\lambda^n|_{L^2(H_0^1)\times L^2(\Omega)}, \mu^n) \leqq 2(|\lambda^*|_{L^2(H_0^1)\times L^2(\Omega)} + \mu^*) + |\lambda^0|_{L^2(H_0^1)\times L^2(\Omega)} + \mu^0 \leqq \hat{\mu}\,.$$

The last estimate determined the choice of $\hat{\mu}$ and it also shows that μ^n and $|\lambda^n|_{L^2(H_0^1)\times L^2(\Omega)}$ are in $[0, \hat{\mu}]$. This ends the proof. □

For the proof of Theorem 3.2 we modify a technique first developed in [IK1, IK2]. It will be assumed henceforth that $(a^n, u^n) \in \operatorname{int} B_\delta$ so that the constraint $(a, u) \in B_{\delta^*}$ is not active. In this case there exists a (unique) Lagrange multiplier $\eta^n \in W_a$ such that for all $(h, v) \in W_a \times W$

$$\text{(A.26)} \qquad D_{(a,u)}\mathscr{H}_{c_n}(a^n, u^n; \lambda^{n-1}, \mu^{n-1})(h, v) - \langle \eta^n, h\rangle = 0\,,$$

and

$$\langle \eta^n, v - a^n\rangle_{W_a} = 0, \qquad \langle \eta^n, a\rangle_{W_a} \geqq 0 \quad \text{for all } a \in W_a^+\,.$$

In the proof of Theorem 3.2 the operator $R: W_a \to W_a \times \mathbb{R}$ given by

$$Rh = (-h, \langle a^*, h\rangle_{W_a})$$

plays a significant role. We summarize some of its properties.

Lemma A.3. *The operator R has closed range and its adjoint is given by*

$$\text{(A.27)} \qquad R^*(a, \alpha) = -a + \alpha a^*\,.$$

The restriction of RR^* to the range of R is a continuous bijection and

$$\|(RR^*)^{-1}R\| = 1\,.$$

Proof. Throughout this proof we use $\langle\cdot,\cdot\rangle$ to denote the inner product in W_a. The verification of the first part of Lemma A.3 is straightforward. In order to calculate the norm of $(RR^*)^{-1}R$, let h be an arbitrary element in W_a. Then there exists $\tilde{h} \in W_a$ such that

$$(RR^*)^{-1}Rh = (RR^*)^{-1}(-h, \langle a^*, h\rangle) = (-\tilde{h}, \langle a^*, \tilde{h}\rangle)$$

is equivalent to

$$(-h, \langle a^*, h\rangle) = (-\tilde{h} - \langle a^*, \tilde{h}\rangle_{W_a} a^*, \langle a^*, \tilde{h}\rangle + \langle a^*, \tilde{h}\rangle |a^*|^2)\,,$$

which gives

$$(RR^*)^{-1}Rh = \left(-h + \frac{\langle a^*, h\rangle}{1 + |a^*|^2}a^*, \frac{\langle a^*, h\rangle}{1 + |a^*|^2}\right).$$

Consequently we have

$$|(RR^*)^{-1}Rh|^2 = |-h + \frac{\langle a^*, h\rangle}{1+|a^*|^2}a^*|^2_{W_a} + \frac{\langle a^*, h\rangle^2}{(1+|a^*|^2)^2}$$

$$= |h|^2 - \frac{\langle a^*, h\rangle^2}{1+|a^*|^2} \leqq |h|^2 .$$

The observation that equality is obtained above for $h \perp a^*$ completes the proof of Lemma A.3. □

Lemma A.4. *Let P_R denote the orthogonal projection of $W_a \times \mathbb{R}$ onto range R. Then the following decomposition holds:*

$$(a^*, 1) = \frac{1+|a^*|^2}{\langle a, a^*\rangle_{W_a}}[(a, 0) - P_R(a, 0)], \quad \textit{provided that } \langle a, a^*\rangle \neq 0 .$$

Proof. Lemma A.3 implies that $\ker R^* = \operatorname{span}\{(a^*, 1)\}$ and that $W_a \times \mathbb{R} = \operatorname{range} R \oplus \ker R^*$. Thus we find

$$(I - P_R)(a, \alpha) = \frac{\langle a, a^*\rangle_{W_a} + \alpha}{1 + |a^*|^2_{W_a}}(a^*, 1), \quad \text{for all } (a, \alpha) \in W_a \times \mathbb{R} ,$$

and the lemma follows. □

Lemma A.5. *Assume that* (C) *holds, that $c_1 \geqq \tilde{c}$, $\delta \leqq \delta^*$, $\sup c_n < \infty$ and that for all n $(a^n, u^n) \in \operatorname{int} B_\delta$. Then there are constants $K_i = K_i(\nu^{-1}, \gamma, k_1, \delta^*, \beta, \lambda^0, \mu^0, c_0, T, \varphi, f, z, \sup c_n)$, $i = 1, 2$, such that*

(i) $|\lambda^* - \lambda^n|_{L^2(H_0^1)\times L^2} \leqq K_1 |(a^* - a^n, u^* - u^n)|_{W_a \times W}$,

(ii) *If* $|a^*| = \gamma$ *then*

$$|(\eta^* - \eta^n, \mu^* - \mu^n)|_{W_a \times \mathbb{R}} \leqq \hat{K}|(a^* - a^n, u^* - u^n)|_{W_a \times W} ,$$

where $\hat{K}$ is a constant which in addition to the arguments of K_1 depends also on $|\langle a^*, a^* - v\rangle_{W_a}|^{-1}$.

(iii) *If $|a^*|_{W_a} < \gamma$, $\delta \in (0, \hat{\delta}]$, $c_1 \geqq \hat{c}$ and $\mu^0 = 0$, then*

$$\mu^n = 0, \quad n = 1, 2, \ldots$$

and

$$|\eta^* - \eta^n|_{W_a} \leqq K_2 |(a^* - a^n, u^* - u^n)|_{W_a \times W}$$

holds for all $n = 1,2, \ldots$.

Proof. In view of the update formula, (A.26) is equivalent to

$$0 = \langle u^n - z, v\rangle_{L^2(H_0^1)} + \beta\langle a^n, h\rangle_{W_a} - \langle \eta^n, h\rangle_{W_a}$$
$$+ \langle \lambda^n, D_{(a,u)}e(a^n, u^n)(h, v)\rangle_{L^2(H_0^1)\times L^2}$$
$$+ \frac{c_n}{2}\langle e(a^n, u^n), D_{(a,u)}e(a^n, u^n)(h, v)\rangle_{L^2(H_0^1)\times L^2}$$
$$+ (\mu^n + \tfrac{1}{2}c_n\hat{g}(a^n, \mu^{n-1}, c_n))\langle a^n, h\rangle_{W_a} .$$

which holds for all $(h, v) \in W_a \times W$. Subtracting the above identity from the first order necessary optimality condition for (a^*, u^*) yields for all $(h, v) \in W_a \times W$

$$
\begin{aligned}
(\text{A.28}) \qquad 0 = {} & \langle u^* - u^n, v\rangle_{L^2(H_0^1)} + \beta\langle a^* - a^n, h\rangle_{W_a} - \langle \eta^* - \eta^n, h\rangle_{W_a} \\
& + \langle \lambda^* - \lambda^n, D_{(a,u)} e(a^n, u^n)(h, v)\rangle_{L^2(H_0^1)\times L^2} \\
& - \langle \lambda^*, D_{(a,u)}(e(a^n, u^n) - e(a^*, u^*))(h, v)\rangle_{L^2(H_0^1)\times L^2} \\
& + \frac{c_n}{2}\langle e(a^*, u^*) - e(a^n, u^n), D_{(a,u)} e(a^n, u^n)(h, v)\rangle_{L^2(H_0^1)\times L^2} \\
& + \mu^*\langle a^*, h\rangle - \mu^n\langle a^n, h\rangle - \frac{c_n}{2}\hat{g}(a^n, \mu^{n-1}, c_n)\langle a^n, h\rangle_{W_a} .
\end{aligned}
$$

Define the operator $M_n : W \to L^2(H_0^1) \times L^2$ by

$$M_n v = D_{(a,u)} e(a^n, u^n)(0, v) = ((-\Delta)^{-1}[v_t - (a^n v_x)_x], v(0, \cdot))$$

and evaluate (A.28) for $(0, v)$, $v \in W$, to obtain

$$
\begin{aligned}
\langle \lambda^* - \lambda^n, M_n v\rangle_{L^2(H_0^1)\times L^2} = {} & -\langle u^* - u^n, v\rangle_{L^2(H_0^1)} \\
+ \langle \lambda_1^*, (-\Delta)^{-1}[(a^* - a^n)v_x)_x]\rangle_{L^2(H_0^1)} & - \frac{c_n}{2}\langle e(a^*, u^*) - e(a^n, u^n), M_n v\rangle_{L^2(H_0^1)\times L^2} .
\end{aligned}
$$

This implies the estimate

$$
\begin{aligned}
|\langle \lambda^* - \lambda^n, M_n v\rangle_{L^2(H_0^1)\times L^2}| \leqq {} & [|u^* - u^n]_{L^2(H_0^1)} + k_1|a^* - a^n|_{W_a}|\lambda_1^*|_{L^2(H_0^1)}]|v|_{L^2(H_0^1)} \\
& + \frac{c_{\max}}{2}|e(a^*, u^*) - e(a^n, u^n)|_{L^2(H_0^1)\times L^2}|M_n v|_{L^2(H_0^1)\times L^2} .
\end{aligned}
$$

The operator M_n is a continuous bijection and by Lemma 2.2

$$|M_n^{-1}(\tilde{f}, \psi)|^2_{L^2(H_0^1)} \leqq \max\left(\frac{1}{\nu}, \frac{1}{\nu^2}\right)|(\tilde{f}, \psi)|^2_{L^2(H_0^1)\times L^2} ,$$

holds for its inverse. Combining these last two estimates we find

$$
\begin{aligned}
(\text{A.29}) \qquad |\lambda^* - \lambda^n|_{L^2(H_0^1)\times L^2} \leqq {} & \left[\max\left(\frac{1}{\nu}, \frac{1}{\nu^2}\right)\right]^{1/2}(|u^* - u^n|_{L^2(H_0^1)} \\
& + k_1|\lambda_1^*|_{L^2(H_0^1)}|a^* - a^n|_{W_a}) + \frac{c_{\max}}{2}|e(a^*, u^*) - e(a^n, u^n)|_{L^2(H_0^1)\times L^2(\Omega)} .
\end{aligned}
$$

A short calculation gives for the last term in (A.29)

$$
\begin{aligned}
(\text{A.30}) \qquad |e(a^*, u^*) - e(a^n, u^n)| \leqq {} & |u^* - u^n|_W + |u^*(0, \cdot) - u^n(0, \cdot)|_{L^2(\Omega)} \\
& + k_1(|u^n|^2_{L^2(H_0^1)} + \gamma^2)^{1/2}|(a^* - a^n, u^* - u^n)|_{W_a\times L^2(H_0^1)} .
\end{aligned}
$$

Moreover we have

$$(\text{A.31}) \qquad |u(0, \cdot)| \leqq \sqrt{\left(1 + \frac{1}{4T}\right)}|u|_W .$$

In fact, integrating

$$|u(0,\cdot)|^2_{L^2(\Omega)} = -\int_0^t \frac{d}{ds}|u(s,\cdot)|^2_{L^2(\Omega)}ds + |u(t,\cdot)|^2_{L^2(\Omega)}$$

with respect to $t\in[0,T]$ and using $|\varphi|^2_{L^2} \leqq \frac{1}{4}|\varphi_x|^2_{L^2}$ for $\varphi\in H^1_0(\Omega)$ gives

$$T|u(0,\cdot)|^2 \leqq T|u|^2_W + \tfrac{1}{4}|u|^2_W$$

which implies (A.31). Inserting (A.30) and (A.31) into (A.29) yields

$$|\lambda^* - \lambda^n|_{L^2(H^1_0)\times L^2} \leqq K_1|(a^* - a^n, u^* - u^n)|_{W_a\times W}\,,$$

where the dependence of the constant K_1 on the parameters specified in the statement of the theorem can be made explicit by Lemma A.1 and Corollary 3.1.

To verify (ii) we evaluate (A.28) for $(h, 0)$, $h\in W_a$ which leads to

$$\text{(A.32)}\qquad \begin{aligned}\langle -(\eta^*-\eta^n) + (\mu^*-\mu^n)a^*, h\rangle_{W_a} &= -\beta\langle a^*-a^n, h\rangle_{W_a}\\ &- \mu^n\langle a^*-a^n, h\rangle_{W_a} + \langle \lambda_1^* - \lambda_1^n, (-\Delta)^{-1}(hu_x^*)_x\rangle_{L^2(H^1_0)}\\ &- \langle \lambda_1^n, (-\Delta)^{-1}[h(u_x^*-u_x^n)_x]\rangle_{L^2(H^1_0)} + \frac{c_n}{2}\hat{g}(a^n, \mu^{n-1}, c_n)\langle a^n, h\rangle_{W_a}\\ &- \frac{c_n}{2}\langle e(a^*,u^*) - e(a^n,u^n), D_{(a,u)}e(a^n,u^n)(h,0)\rangle_{L^2(H^1_0)\times L^2}\,.\end{aligned}$$

The right hand side of (A.32) defines a continuous linear functional $\mathscr{F}$ on W_a. Hence by Riesz' Theorem there is $\tilde{a}\in W_a$, such that $\langle \mathscr{F}, a\rangle_{W_a^*, W_a} = \langle \tilde{a}, a\rangle_{W_a}$ and $|\tilde{a}|_{W_a} = |\mathscr{F}|_{W_a^*}$. Lemma A.3 shows that (A.32) is equivalent to

$$\langle R^*(\eta^*-\eta^n, \mu^*-\mu^n), h\rangle = \langle \tilde{a}, h\rangle_{W_a}, \quad \text{for all } h\in W_a\,,$$

i.e.

$$R^*(\eta^*-\eta^n, \mu^*-\mu^n) = \tilde{a}\,.$$

This implies

$$P_R(\eta^*-\eta^n, \mu^*-\mu^n) = (RR^*)^{-1}R\tilde{a}$$

and, again by Lemma A.3

$$\text{(A.33)}\qquad |P_R(\eta^*-\eta^n, \mu^*-\mu^n)|_{W_a\times R} \leqq |\tilde{a}|_{W_a}\,.$$

We consider two cases: first assume $|a^*| = \gamma$. This implies in particular that $\langle a^*, a^*-v\rangle > 0$. Consequently by Lemma A.4 with $a = a^* - v$ we infer

$$\begin{aligned}|(I-P_R)(\eta^*-\eta^n, \mu^*-\mu^n)|_{W_a\times\mathbb{R}} &= (1+\gamma^2)^{-1/2}|\langle(\eta^*-\eta^n, \mu^*-\mu^n), (a^*, 1)\rangle_{W_a\times\mathbb{R}}|\\ &= \langle a^*, a^*-v\rangle^{-1}(1+\gamma^2)^{1/2}|\langle \eta^*-\eta^n, a^*-v\rangle_{W_a}\\ &\quad - \langle(\eta^*-\eta^n, \mu^*-\mu^n), P_R(a^*-v, 0)\rangle_{W_a\times\mathbb{R}}|\,.\end{aligned}$$

As a consequence of (A.3) and (A.26) one obtains

$$\langle \eta^*-\eta^n, a^*-v\rangle = -\langle \eta^n, a^*-a^n\rangle\,,$$

which together with (A.33) results in

$$|(I - P_R)(\eta^* - \eta^n, \mu^* - \mu^n)| \leqq \frac{(1+\gamma^2)^{1/2}}{\langle a^*, a^* - v\rangle} (|\eta^n|^2_{W_a} + |a^* - v|^2_{W_a})^{1/2} \cdot (|a^* - a^n|^2_{W_a} + |\tilde{a}|^2_{W_a})^{1/2} .$$

and consequently

(A.34) $$|(\eta^* - \eta^n, \mu^* - \mu^n)|_{W_a \times \mathbb{R}} \leqq \hat{K}_1(|a^* - a^n|^2 + 2|\tilde{a}|^2_{W_a})^{1/2}$$

where

$$\hat{K}_1^2 = \frac{1+\gamma^2}{\langle a^*, a^* - v\rangle^2}(|\eta^n|^2_{W_a} + |a^* - v|^2_{W_a}) + 1 .$$

Next we derive an estimate for $\tilde{a}$. Recall that $\mu^n \geqq 0$ for all $n \geqq 0$ and therefore

$$|\hat{g}(a^n, \mu^{n-1}, c^n)| \leqq |g(a^n)| = |g(a^n) - g(a^*)| \leqq \tfrac{1}{2}(\delta^* + 2\gamma)|a^n - a^*|_{W_a} .$$

In view of this estimate and of the first part of Lemma A.5 we obtain the following bound on $|\tilde{a}|$:

$$|\tilde{a}|_{W_a} \leqq [\hat{\mu} + \beta + k_1 K_1 |u^*|_{L^2(H_0^1)} + k_1|\lambda_1^n|_{L^2(H_0^1)} + \frac{c_{\max}}{2} k_1 |u^n|_{L^2(H_0^1)}(1 + k_1(|u^n|^2_{L^2(H_0^1)} + |a^*|^2_{W_a})^{1/2}) + \frac{c_{\max}}{4}(\delta^* + 2\gamma)|a^n|_{W_a}]|(a^* - a^n, u^* - u^n)|_{W_a \times W} .$$

Inserting this estimate into (A.34) and taking into account the bounds implied by Theorem 3.1 and Lemma A.1 we finally arrive at

$$|(\eta^* - \eta^n, \mu^* - \mu^n)|_{W_a \times \mathbb{R}} \leqq \hat{K}|(a^* - a^n, u^* - u^n)|_{W_a \times W} ,$$

where $\hat{K}$ is a constant which in addition to the arguments of K_1 also depends on the angle in W_a between a^* and $a^* - v$.

It remains to consider the case $|a^*| < \gamma$. This implies $g(a^*) < 0$ and consequently $\mu^* = 0$. Define

$$\hat{\delta} = \min(\delta^*, \tfrac{1}{3}(\gamma - |a^*|_{W_a})) \qquad \text{and} \qquad \hat{c} = \max\left(\tilde{c}, \frac{3\hat{\mu}}{(\hat{\delta} + \gamma)(\gamma - |a^*|_{W_a})}\right)$$

and let $(a, u) = B_{\hat{\delta}}$, $\mu \in [0, \hat{\mu}]$ and $c \geqq \hat{c}$. Due to this choice one finds

$$g(a) + \frac{\mu}{c} \leqq \tfrac{1}{2}(\hat{\delta} + |a^*| - \gamma)(\hat{\delta} + |a^*| + \gamma) + \tfrac{1}{3}(\hat{\delta} + \gamma)(\gamma - |a^*|) \leqq \tfrac{1}{3}|a^*|(|a^*| - \gamma) < 0$$

which in turn implies that

$$\mu^n = \mu^{n-1} + \frac{c_n}{2}\hat{g}(a^n, \mu^{n-1}, c_n) = \max\left(\mu^{n-1} + \frac{c_n}{2} g(a^n), \frac{\mu^{n-1}}{2}\right) = \frac{\mu^{n-1}}{2} .$$

Due to the assumption that $\mu^0 = 0$, this establishes that $\mu^n = 0$ and therefore $\hat{g}(a^n, \mu^{n-1}, c_n) = 0$, for all $n = 1, 2, \ldots$. Consequently in the present case no terms involving μ^* and μ^n appear in (A.32). Thus, to obtain a bound for $|\eta^* - \eta^n|$ we may directly refer to (A.32) which gives by arguments similar to those which led to the estimate for $|\tilde{a}|$:

$$|\eta^* - \eta^n|_{W_a} \leqq K_2|(a^* - a^n, u^* - u^n)|_{W_a \times W} .$$

This completes the proof of Lemma A.5. □

Proof of Theorem 3.2. The proof of this theorem follows from an induction argument using Theorem 3.1 and Lemma A.5. □

References

[A] Adams, R.A. (1975) Sobolev Spaces, Academic Press, New York

[AHS] Alt, H.W., Hoffmann, K.-H., Sprekels, J. (1984) A numerical procedure to solve certain identification problems. In: K.H. Hoffmann, W. Krabs (eds.) Optimal Control of Partial Differential Equations. Internat. Series of Numerical Math., Vol. 68. Birkhäuser, Basel, pp. 11–34

[Al] Alt, H.W. (1983) Lipschitzian perturbations of infinite optimization problems. In: A.V. Fiacco (ed.) Mathematical Programming with Data Perturbations II, Lecture Notes in Pure and Applied Mathematics, Vol. 85. Dekker, New York, pp. 7–21

[BKL] Banks, H.T., Kareiva, P.K., Lamm, P.K. (1985) Modelling insect dispersal and estimating parameters when mark-release techniques may cause initial disturbances. J. Math. Biol. **22**, 259–277

[BKM] Banks, H.T., Kareiva, P.M., Murphy, K.A. (1987) Parameter estimation techniques for interaction and redistribution models: a predator-prey example. Oecologia 74, 356–362

[BL] Banks, H.T., Lamm, P.K. (1985) Estimation of variable coefficients in parabolic distributed systems. IEEE Trans. Auto. Control **30**, 386–398

[BS] Baumeister, J., Scondo, W. (1987) Asymptotic embedding methods for parameter estimation. In: Proceedings 26th Conf. Decision & Control, Los Angeles, pp. 170–174

[CK] Colonius, F., Kunisch, K. (1989) Output least squares stability in elliptic systems, Appl. Math. Optim. **19**, 33–63

[EKN] Engl, H.W., Kunisch, K., Neubauer, A. (1989) Tikhonov regularization for the solution of nonlinear illposed problems, Inverse problems **5**, 523–540

[EL] Ewing, R.E., Tao Lin. Parameter identification problems in two-phase flow. Preprint

[IK1] Ito, K., Kunisch, K. (1990) The augmented Lagrangian method for equality and inequality constraints in Hilbert spaces. Math. Prog. **46**, 341–360

[IK2] Ito, K., Kunisch, K. (1990) The augmented Lagrangian method for parameter estimation in elliptic systems. SIAM J. Control Optimization **28**, 113–136

[IKK] Ito, K., Kroller, M., Kunisch, K. (1991) A numerical study of the augmented Lagrangian method for the estimation of parameters in elliptic systems. SIAM J. Sci. Stat. Comput. (to appear)

[KP] Kunisch, K., Peichl, G. (1989) Estimation of a temporally and spatially varying diffusion coefficient in a parabolic system by an augmented Lagrangian technique. Technical Report **146**, University Graz

[KW] Kunisch, K., White, L.W. (1991) Regularity properties and strict complementarity of the output-least-squares approach to parameter estimation in parabolic systems. Inverse Problems (to appear)

[L] Lions, J.L. (1971) Optimal control of systems governed by partial differential equations, Springer, Berlin Heidelberg New York

[LE] Tao Lin, Ewing, R.E. (1986) Parameter estimation for distributed systems arising in fluid flow problems via time series methods. In: Int. Series. Num. Math., Vol. 77. Birkhäuser, Basel

[LM] Lions, J.L., Magenes, E. (1972) Non-homogeneous Boundary Value Problems I, Springer New York

[MZ] Maurer, H., Zowe, J. (1979) First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems, Math. Programming **16**, 98–110

[S] Scondo, W. (1987) Ein Modellabgleichverfahren zur adaptiven Parameteridentifikation in Evolutionsgleichungen, Dissertation, Frankfurt

[W] Wloka, J.W. (1982) Partielle Differentialgleichungen, Teubner, Stuttgart

[Y] Yeh, W.W.-G. (1986) Review of parameter identification procedures in groundwater hydrology: the inverse problem. Water Resources Research **22**, 95–108