Journal of Computational and Applied Mathematics 310 (2017) 129–142

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

Identification problems with given material interfaces

J. Haslinger, R. Blaheta, R. Hrtus *

Institute of Geonics AS CR, Studentska 1768, Ostrava, Czech Republic

ARTICLE INFO

Article history:
Received 30 December 2015
Received in revised form 16 June 2016

Keywords:
Identification of conductivity coefficients
Control of systems by coefficients of PDE
Algebraic sensitivity analysis
Scalarization technique
Pumping test

ABSTRACT

This paper is devoted to the identification of coefficients in scalar elliptic partial differential equations using an optimal control approach. The coefficients usually can be interpreted as material characteristics and play the role of the control variable. The paper focuses on processes in domains which can be split into a finite number of materially homogeneous subdomains, i.e. the coefficients to be identified are piecewise constant on them. In addition, we assume that the material interfaces are a priori known. We prove the existence of at least one solution of the optimal control problem for a large class of cost functionals and show that solutions can be obtained as limits of solutions to the problems which are governed by finite element discretizations of the state equations. Further, the unified algebraic sensitivity analysis of the first and the second order for several least squares type cost functions is investigated. Finally, a model problem of the identification of coefficients characterizing hydraulic conductivity by pumping tests in groundwater flow modeling is numerically solved using the sequential approach and the scalarization technique.

1. Introduction

Many processes which are important in engineering, health care, science, etc. are described by partial differential equations (PDEs). Mathematical modeling for predicting the behavior of systems requires the selection of proper control variables and equations describing the process. Control variables may have different nature. This paper deals with the identification of coefficients in the 2nd order scalar elliptic equations. The coefficients of PDEs mostly represent material properties of a structure and their proper determination can be based on the knowledge of the state problem solution for a given input data. To this end we use an optimal control approach whose abstract form reads as follows:

$$\text{find } a^* \in \operatorname*{argmin}_{a \in \mathcal{U}_{ad}} J(a, u(a)) \tag{1.1}$$

where J is a cost functional depending on: control variables a (coefficient of PDE), that are required to be elements of a set of admissible controls $\mathcal{U}_{ad}$ and the solution $u(a)$ to the state problem corresponding to $a \in \mathcal{U}_{ad}$. A particular choice of J depends on the goals we want to achieve. One of the most frequently used functionals in practical applications is a least squares type

$$J(a, u(a)) = \frac{1}{2} \|R\, u(a) - z_d\|^2 . \tag{1.2}$$

Here R stands for the selection (observation) operator, z_d is a target and $\|\bullet\|$ is an appropriate norm. The straightforward approach to the identification problem (1.1) combines a suitable minimization method with a procedure for computing

* Corresponding author.
E-mail addresses: hasling@karlin.mff.cuni.cz (J. Haslinger), blaheta@ugn.cas.cz (R. Blaheta), hrtus@ugn.cas.cz (R. Hrtus).

http://dx.doi.org/10.1016/j.cam.2016.06.023
0377-0427/

the solution to the state problem. If a gradient type minimization method is used then an extra procedure for the efficient computation of derivatives of the objective functional has to be involved. In many applications (and this is just our case) the set $\mathcal{U}_{ad}$ is (isomorphic to) a compact subset of the Euclidean space $\mathbb{R}^r$.

This fact simplifies a lot the mathematical analysis since no additional constraints ensuring compactness of $\mathcal{U}_{ad}$ are needed. The simple case is introduced in Section 2: the coefficients of the PDEs are piecewise constant in a fixed number of subdomains $\{\Omega_i\}_{i=1}^r$, that define a partition of a domain of definition Ω of the state problem with a priori known interfaces. Such situations arise in many applications when the decomposition of Ω into materially homogeneous sub-domains Ω_i can be detected either by various types of tomography (e.g. [1]) or less rigorously just by an engineering guess (e.g. [2]). The obtained results remain valid when besides the (piecewise constant) coefficients in $\{\Omega_i\}_{i=1}^r$ also interfaces among them are the object of the identification, provided that the shapes of all Ω_i are described by a finite number of parameters. The existence of a solution to (1.1) is based on compactness and continuity arguments. To prove continuity of the composed functional $\mathcal{J}(a) := J(a, u(a))$, Section 2 analyzes continuity of the control-to-state mapping $\varphi : \mathcal{U}_{ad} \mapsto H^1(\Omega)$, $\varphi(a) = u(a)$ and also continuity of $\tilde{\varphi} : \mathcal{U}_{ad} \mapsto H^2(\Omega')$, $\tilde{\varphi}(a) = u(a)_{|_{\Omega'}}$, where Ω' is a sub-domain lying strictly inside of Ω_i for some i. The latter result is important in the case when J uses pointwise targets. Moreover we proved that any accumulation point of the sequence of solutions to the discretized identification problems when the discretization parameter tends to zero is a solution to the original continuous setting. Section 3 deals with computing the first and second derivatives of the discretized least squares type cost function of the form

$$\boldsymbol{\mathcal{J}}(\boldsymbol{a}) = \frac{1}{2}\left(\boldsymbol{B}(\boldsymbol{a})\left(\boldsymbol{R}\boldsymbol{u}(\boldsymbol{a}) - \boldsymbol{z_d}\right), \boldsymbol{R}\boldsymbol{u}(\boldsymbol{a}) - \boldsymbol{z_d}\right)_s, \tag{1.3}$$

where $\boldsymbol{a}$ is the vector representation of $a \in \mathcal{U}_{ad}$, $\boldsymbol{B}(\boldsymbol{a})$ is a symmetric, positive definite matrix, $\boldsymbol{R}$ is the matrix representation of the observation mapping R, $\boldsymbol{z_d}$ is a target vector, $\boldsymbol{u}(\boldsymbol{a})$ is the solution of a linear algebraic system $\boldsymbol{K}(\boldsymbol{a})\boldsymbol{u}(\boldsymbol{a}) = \boldsymbol{b}$, and $(\bullet, \bullet)_s$ stands for the Euclidean inner product in $\mathbb{R}^s$. This form of $\boldsymbol{\mathcal{J}}$ is quite general. According to the choice of the matrices $\boldsymbol{B}(\boldsymbol{a})$ and $\boldsymbol{R}$ one gets the Euclidean, energy and equation error functions, respectively. The expressions for the corresponding gradient, Hessian as well as the Gauss–Newton approximations to the Hessian are derived. The use of the Euclidean error functional is standard, see e.g. [3], the use of the energy and equation error function was inspired by [4–6] and the references therein. One of the advantages of the last two mentioned error functions is the fact that in the case of targets distributed in the whole Ω one can easily derive the expression for the full Hessian $\mathbb{D}^2\boldsymbol{\mathcal{J}}$ and show that $\mathbb{D}^2\boldsymbol{\mathcal{J}}$ is positive semi-definite, i.e. the minimized functions are convex. Section 4 is devoted to the optimization techniques which are used in the last section. First, we recall the Levenberg–Marquardt method frequently used in identification problems. Due to the demanding character of measurements the only way how to get more data is to repeat them for different input data (multi-response). This leads to an identification problem involving more cost functions. For solving such a type of problems two approaches are proposed: a sequential method and the weighted sum method. In Section 5 the identification problem for coefficients characterizing hydraulic conductivity by using pumping tests is formulated. This model example is a modification of the one in [7]. It has a structure which is strongly different from the standard setting introduced in Section 2. The characteristic features of this example are the following: (i) only pointwise targets are available and their number is very small (only eight in our case), (ii) conductivities are piecewise constant. These are identified only in a small domain $\omega \subset \Omega$ on the boundary of which all measurement points are located, (iii) the right hand side of the state problem is given by two Dirac distributions, (iv) the multi-response character of gaining data: eight measurement points are split into four disjoint pairs. With any such pair one experiment which provides the input data in the remaining 6 points is associated. In this way four least squares type functions are generated.

Finally, in Section 6 we use the sequential and weighted sum method with the Euclidean error function for solving the model example considering exact and noisy data.

The paper addresses inverse identification problems characterized by piecewise constant (material) coefficients, some a priori knowledge of domains, where coefficients can be supposed as constant, mostly pointwise (not distributed) measurements and utilization of multi-response data. Such problems frequently appear in engineering applications, see e.g. [8].

The contribution of the presented paper can be seen in

- simple and straightforward existence and convergence analysis for a class of identification problems which also involves cost functions defined by pointwise measurements;
- unified sensitivity analysis: computations of the gradients and Hessians or Gauss–Newton approximations to the Hessians for a large class of least squares type functions including Euclidean, energy and equation error cost functions;
- efficient numerical realization of a model multi-response problem taken from hydrogeology.

2. Identification problems with given material interfaces—the scalar case

The aim of this section is to give the mathematical justification of problems arising in identification of materials separated by interfaces whose position is known. To this end we use an optimal control approach in which coefficients of partial differential equations play the role of control variables. We restrict ourselves to the identification of coefficients in a *scalar*

elliptic equation. First we prove that the respective minimization problem has a solution. The second part of this section deals with a discretization of this type of problems and convergence analysis.

Let us consider a state problem

$$\begin{cases} -\mathrm{div}\,(a\nabla u) = f & \text{in } \Omega \subset \mathbb{R}^d, \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{2.1}$$

where the coefficient a represents a *control* variable. We suppose that $\Omega \subset \mathbb{R}^d$, $d = 2, 3$ is a domain, decomposed into mutually disjoint subdomains Ω_i, $i = 1, \ldots, r$, i.e. $\bar{\Omega} = \bigcup_{i=1}^{r} \bar{\Omega}_i$, $\Omega_i \cap \Omega_j = \emptyset$ for $i \neq j$.

The set $\mathcal{U}_{ad}$ of admissible controls is defined as follows:

$$\mathcal{U}_{ad} = \left\{ a \in L^\infty(\Omega) \mid a_{\min} \leq a \leq a_{\max} \text{ in } \Omega,\, a_{|_{\Omega_i}} \in P_0(\Omega_i),\, i = 1, \ldots, r \right\},$$

i.e. $\mathcal{U}_{ad}$ is the set of piecewise constant functions on the partition $\{\Omega_i\}_{i=1}^{r}$ and $0 < a_{\min} < a_{\max}$ are given.

The weak formulation of (2.1) reads as follows: given $a \in \mathcal{U}_{ad}$,

$$(\mathcal{P}(a)) \quad \begin{cases} \text{find } u := u(a) \in H_0^1(\Omega) \text{ such that} \\ (a\nabla u, \nabla v)_{0,\Omega} = (f, v)_{0,\Omega} \quad \forall v \in H_0^1(\Omega), \end{cases}$$

where the right hand side $f \in L^2(\Omega)$.

Let $J : \mathcal{U}_{ad} \times H_0^1(\Omega) \mapsto \mathbb{R}^1$ be a functional and denote

$$\mathcal{J}(a) := J(a, u(a)), \tag{2.2}$$

with $u(a) \in H_0^1(\Omega)$ being a unique solution to $(\mathcal{P}(a))$, $a \in \mathcal{U}_{ad}$. We consider the following optimal control problem:

$$(\mathbb{P}) \quad \begin{cases} \text{find } a^* \in \mathcal{U}_{ad} \text{ such that} \\ \mathcal{J}(a^*) \leq \mathcal{J}(a) \quad \forall a \in \mathcal{U}_{ad}. \end{cases}$$

Next we suppose that J is *lower semicontinuous*:

$$\left.\begin{array}{ll} a_n \underset{n\to\infty}{\longrightarrow} a \text{ in } L^\infty(\Omega), & a_n, a \in \mathcal{U}_{ad} \\ y_n \underset{n\to\infty}{\longrightarrow} y \text{ in } H^1(\Omega), & y_n, y \in H_0^1(\Omega) \end{array}\right\} \Rightarrow \liminf_{n\to\infty} J(a_n, y_n) \geq J(a, y). \tag{2.3}$$

Under this assumption the following existence result holds.

Theorem 2.1. *Let (2.3) be satisfied. Then* $(\mathbb{P})$ *has a solution.*

To prove this theorem we show that the *control-to-state* mapping $u : \mathcal{U}_{ad} \mapsto H_0^1(\Omega)$, $a \mapsto u(a)$, where $u(a)$ solves $(\mathcal{P}(a))$ is continuous.

Proposition 2.1. *Let* $\{a_n\}$, $a_n \in \mathcal{U}_{ad}$ *be a sequence such that* $a_n \underset{n\to\infty}{\longrightarrow} a \in \mathcal{U}_{ad}$ *in* $L^\infty(\Omega)$ *and* $u_n := u(a_n) \in H_0^1(\Omega)$ *be the solution to* $(\mathcal{P}(a_n))$, $n = 1, 2, \ldots$. *Then*

$$u_n \underset{n\to\infty}{\longrightarrow} u \quad \text{in } H_0^1(\Omega) \tag{2.4}$$

and $u := u(a)$ *solves* $(\mathcal{P}(a))$.

Proof. From Friedrich's inequality and the definition of $\mathcal{U}_{ad}$ it follows that $\{u_n\}$ is bounded in $H_0^1(\Omega)$:

$$\exists c = const. > 0 : \quad \|u_n\|_{1,\Omega} \leq c \;\forall n \in \mathbb{N}.$$

Thus there exists a subsequence $\{u_{n_j}\}$ such that

$$u_{n_j} \underset{j\to\infty}{\rightharpoonup} u \quad \text{in } H_0^1(\Omega). \tag{2.5}$$

We prove that u solves $(\mathcal{P}(a))$. The weak formulation of $\left(\mathcal{P}\left(a_{n_j}\right)\right)$ yields:

$$\left(a_{n_j}\nabla u_{n_j}, \nabla v\right)_{0,\Omega} = (f, v)_{0,\Omega} \quad \forall v \in H_0^1(\Omega).$$

Letting $j \to \infty$ and using (2.5) we obtain:

$$(a\nabla u, \nabla v)_{0,\Omega} = (f, v)_{0,\Omega} \quad \forall v \in H_0^1(\Omega),$$

i.e. u solves $(\mathcal{P}(a))$. Since the solution of $(\mathcal{P}(a))$ is unique, (2.5) holds for the whole sequence. To prove strong convergence of $\{u_n\}$ to u we proceed as follows:

$$\begin{aligned}\|a\nabla u_n\|^2_{0,\Omega} &= \int_\Omega a\,|\nabla u_n|^2\,dx = \int_\Omega (a-a_n)\,|\nabla u_n|^2\,dx + \int_\Omega a_n\,|\nabla u_n|^2\,dx \\ &= \int_\Omega (a-a_n)\,|\nabla u_n|^2\,dx + \int_\Omega f u_n dx \underset{n\to\infty}{\longrightarrow} \int_\Omega f u dx = \|a\nabla u\|^2_{0,\Omega}\,.\end{aligned}$$

Here we used the fact that $a_n \underset{n\to\infty}{\longrightarrow} a$ in $L^\infty(\Omega)$ and the definition of $(\mathcal{P}(a))$. Since $\|a\nabla \bullet\|_{0,\Omega}$ is a norm on $H^1_0(\Omega)$ we obtain (2.4). □

Proof of Theorem 2.1. Let $\{(a_n, u_n)\}$ be a minimizing sequence of J, i.e.

$$q := \inf_{a\in\mathcal{U}_{ad}} J(a, u(a)) = \lim_{n\to\infty} J(a_n, u_n)\,, \tag{2.6}$$

where $a_n \in \mathcal{U}_{ad}$ and $u_n := u(a_n)$ is a solution to $(\mathcal{P}(a_n))$. Since $\mathcal{U}_{ad}$ is compact in the $L^\infty(\Omega)$-norm, there exists a subsequence $\{a_{n_j}\}$ and $a^* \in \mathcal{U}_{ad}$ such that

$$\left.\begin{aligned} a_{n_j} &\underset{j\to\infty}{\longrightarrow} a^* \quad \text{in } L^\infty(\Omega) \\ u_{n_j} &\underset{j\to\infty}{\longrightarrow} u(a^*) \quad \text{in } H^1(\Omega)\end{aligned}\right\} \tag{2.7}$$

by Proposition 2.1. From (2.3), (2.6) and (2.7) we obtain

$$q = \lim_{n\to\infty} J(a_n, u_n) = \lim_{j\to\infty} J\left(a_{n_j}, u_{n_j}\right) \geq J\left(a^*, u\left(a^*\right)\right),$$

i.e. a^* is a solution to $(\mathbb{P})$ and $u(a^*)$ solves $(\mathcal{P}(a^*))$. □

Remark 2.1. There exists a large class of functionals J satisfying (2.3). For example, the following standard least square functionals satisfy this assumption:

$$J(y) = \frac{1}{2}\|y - z\|^2_{0,\Omega}\,, \quad z \in L^2(\Omega) \text{ given}, \tag{2.8}$$

or

$$J(y) = \frac{1}{2}\|\nabla y - \bar{z}\|^2_{0,\Omega}\,, \quad \bar{z} \in \left(L^2(\Omega)\right)^d \text{ given}. \tag{2.9}$$

In addition, they do not depend explicitly on $a \in \mathcal{U}_{ad}$. Another class of functionals which will be used in computations is of the form

$$J(a, y) = \int_D (a\nabla(y - z_d), \nabla(y - z_d))\,dx, \quad z_d \in \left(H^1(\Omega)\right)^d \text{ given}, \tag{2.10}$$

where $D \subseteq \Omega$. If $D = \Omega$ then (2.10) defines the square of the energy norm of $(y - z_d)$. Observe that all of these cost functionals are even continuous in $L^\infty(\Omega) \times H^1_0(\Omega)$.

Remark 2.2. Sometimes we have at our disposal pointwise measurements at a finite number of points $\{x_i\}_{i=1}^m$, $x_i \in \Omega\ \forall i$ and J is defined by

$$J(y) = \frac{1}{2}\sum_{i=1}^m (y(x_i) - z_i)^2\,, \quad \boldsymbol{z} = (z_1, \ldots, z_m) \in \mathbb{R}^m \text{ given}. \tag{2.11}$$

To evaluate such J we need the argument y to be represented by a continuous function, but this is not the case of functions belonging to $H^1_0(\Omega)$ if $d \geq 2$. Fortunately it is known (see [9]) that for any subdomain $\Omega' \subset \Omega_i$, $i = 1, \ldots, r$ and any $a \in \mathcal{U}_{ad}$ the restriction $u(a)_{|_{\Omega'}} \in H^2(\Omega')$. Moreover, there exists a constant $c > 0$ such that

$$\|u(a)\|_{2,\Omega'} \leq c \quad \forall a \in \mathcal{U}_{ad}. \tag{2.12}$$

Next, we suppose that $x_i \in \Omega \setminus \cup_{j=1}^r \partial\Omega_j$, $i = 1, \ldots, m$. Since $H^2(\Omega') \subset C(\bar{\Omega}')$ for $d \leq 3$, the function (2.11) is well-defined. Then $(\mathbb{P})$ with the cost functional (2.11) has a solution, too. Indeed, if $a_n \underset{n\to\infty}{\longrightarrow} a$ in $L^\infty(\Omega)$ it follows from Proposition 2.1 that

$u_n \underset{n\to\infty}{\longrightarrow} u\,(a)$ in $H_0^1(\Omega)$, where $u_n := u(a_n)$, $u(a)$ is a solution of $(\mathcal{P}(a_n))$ and $(\mathcal{P}(a))$, respectively. At the same time we see from (2.12) that $\{u_n\}$ is bounded in $H^2(\Omega')$ so that there exists a subsequence $\{u_{n_j}\}$ such that

$$u_{n_j} \underset{j\to\infty}{\rightharpoonup} u \quad \text{in } H^2(\Omega'),$$

implying

$$u_{n_j} \underset{j\to\infty}{\rightrightarrows} u \quad \text{(uniformly) in } \bar{\Omega}'.$$

Therefore the functional $\mathcal{J}(a) := J(u(a))$, where J is defined by (2.11) is continuous and consequently $(\mathbb{P})$ has a solution.

Now we pass to a discretization of $(\mathbb{P})$. Since $\mathcal{U}_{ad}$ is already a discrete set it remains to discretize the state space $H_0^1(\Omega)$. To this end we suppose that the domains $\Omega_i,\ i = 1,\dots,r$ defining the decomposition of Ω are *polygonal*. Let $\{\mathcal{T}_h\},\ h \to 0+$ be a regular family of triangulations of $\bar{\Omega}$ such that $\mathcal{T}_{h|_{\Omega_i}}$ is a triangulation of $\bar{\Omega}_i,\ i = 1,\dots,r\ \forall h > 0$. With any $\mathcal{T}_h$ we associate the space of piecewise linear functions on $\mathcal{T}_h$:

$$V_h = \left\{v_h \in C(\bar{\Omega}) \mid v_{h|_T} \in P_1(T)\ \forall T \in \mathcal{T}_h,\ v_h = 0 \text{ on } \partial\Omega\right\}.$$

The state problem $(\mathcal{P}(a))$, $a \in \mathcal{U}_{ad}$, will be discretized by the standard Galerkin method on V_h: given $a \in \mathcal{U}_{ad}$

$$(\mathcal{P}(a))_h \quad \begin{cases} \text{find } u_h := u_h(a) \in V_h \text{ such that} \\ (a\nabla u_h, \nabla v_h)_{0,\Omega} = (f, v_h)_{0,\Omega} \quad \forall v_h \in V_h. \end{cases}$$

The discretization of $(\mathbb{P})$ reads as follows:

$$(\mathbb{P})_h \quad \begin{cases} \text{find } a^{h*} \in \mathcal{U}_{ad} \text{ such that} \\ \mathcal{J}_h(a^{h*}) \le \mathcal{J}_h(a^h) \quad \forall a^h \in \mathcal{U}_{ad}, \end{cases}$$

where $\mathcal{J}_h(a^h) := J(a^h, u_h(a^h))$, and $u_h(a^h) \in V_h$ is the solution to $(\mathcal{P}(a^h))_h$.[1] Using the classical continuity and compactness arguments it is easy to show that $(\mathbb{P})_h$ has a solution for any $h > 0$.

Next we study the mutual relation between solutions to $(\mathbb{P})$ and $(\mathbb{P})_h$ for $h \to 0+$. To this end we need a stronger *continuity* assumption on J. Suppose that J is continuous in the following sense:

$$\left.\begin{aligned} &a^h \underset{h\to 0}{\longrightarrow} a \text{ in } L^\infty(\Omega), \quad a^h, a \in \mathcal{U}_{ad} \\ &y_h \underset{h\to 0}{\longrightarrow} y \text{ in } H^1(\Omega), \quad y_h \in V_h, y \in H_0^1(\Omega) \end{aligned}\right\} \Rightarrow \lim_{h\to 0} J(a^h, y_h) = J(a, y). \tag{2.13}$$

We need the following parallel of Proposition 2.1.

Proposition 2.2. *Let $\{a^h\}$, $a^h \in \mathcal{U}_{ad}$ be a sequence such that $a^h \underset{h\to 0+}{\longrightarrow} a \in \mathcal{U}_{ad}$ in $L^\infty(\Omega)$ and $u_h := u_h(a^h) \in V_h$ be the solution to $(\mathcal{P}(a^h))_h$. Then*

$$u_h \underset{h\to 0+}{\longrightarrow} u \quad \text{in } H_0^1(\Omega),$$

and $u := u(a)$ is the solution to $(\mathcal{P}(a))$.

Proof. It is readily seen that the sequence $\{u_h\}$ is bounded in $H_0^1(\Omega)$ so that there exists a subsequence $\{u_{h_j}\}$ such that

$$u_{h_j} \underset{j\to\infty}{\rightharpoonup} u \quad \text{in } H_0^1(\Omega). \tag{2.14}$$

Let $\bar{v} \in H_0^1(\Omega)$ be arbitrary. Since the system $\{V_h\}$ is dense in $H_0^1(\Omega)$, one can find a sequence $\{\bar{v}_h\}$, $\bar{v}_h \in V_h$ such that

$$\bar{v}_h \underset{h\to 0+}{\longrightarrow} \bar{v} \quad \text{in } H_0^1(\Omega). \tag{2.15}$$

The weak formulation of $(\mathcal{P}(a^{h_j}))_{h_j}$ yields:

$$\left(a^{h_j}\nabla u_{h_j}, \nabla \bar{v}_{h_j}\right)_{0,\Omega} = \left(f, \bar{v}_{h_j}\right)_{0,\Omega}.$$

Passing to the limit with $j \to \infty$ and using (2.14), (2.15) we obtain:

$$(a\nabla u, \nabla \bar{v})_{0,\Omega} = (f, \bar{v})_{0,\Omega} \quad \forall \bar{v} \in V,$$

i.e. u solves $(\mathcal{P}(a))$. Since $(\mathcal{P}(a))$ has a unique solution, (2.14) holds for the whole sequence. Strong convergence can be shown exactly as in Proposition 2.1. □

[1] The upper index h in a^h does not denote a discretization of a since $\mathcal{U}_{ad}$ is discrete itself. It means only that $a^h \in \mathcal{U}_{ad}$ is used in $(\mathcal{P}(a))_h$ with $a := a^h$.

If $a^* \in \mathcal{U}_{ad}$ is a solution to $(\mathbb{P})$ and u^* solves $(\mathcal{P}(a^*))$ then the pair (a^*, u^*) is termed to be *an optimal pair* of $(\mathbb{P})$ (and the same for $(\mathbb{P})_h$). The main result of this section is the following convergence theorem.

Theorem 2.2. *Let $\{(a^{h*}, u_h^*)\}$ be a sequence of optimal pairs of $(\mathbb{P})_h$, $h \to 0+$. Then there exists a subsequence $\{(a^{h_j*}, u_{h_j}^*)\}$ and a pair $(a^*, u^*) \in \mathcal{U}_{ad} \times H_0^1(\Omega)$ such that*

$$a^{h_j*} \underset{j\to\infty}{\longrightarrow} a^* \quad \text{in } L^\infty(\Omega), \tag{2.16}$$

$$u_{h_j}^* \underset{j\to\infty}{\longrightarrow} u^* \quad \text{in } H_0^1(\Omega). \tag{2.17}$$

Moreover, (a^*, u^*) is an optimal pair of $(\mathbb{P})$. Any accumulation point of $\left\{\left(a^{h*}, u_h^*\right)\right\}$ in the sense of (2.16) and (2.17) has this property.

Proof. The existence of a subsequence $\{(a^{h_j*}, u_{h_j}^*)\}$ satisfying (2.16) and (2.17) is a consequence of the compactness of $\mathcal{U}_{ad}$ and (2.17). From the definition of $(\mathbb{P})_{h_j}$ we have

$$\mathcal{J}_{h_j}(a^{h_j*}) \le \mathcal{J}_{h_j}\left(a^{h_j}\right) \quad \forall a^{h_j} \in \mathcal{U}_{ad}, \ \forall j.$$

Passing to the limit with $j \to \infty$, using (2.16), (2.17) and the fact that also

$$u_h(a) \underset{h\to 0+}{\longrightarrow} u(a) \quad \text{in } H_0^1(\Omega),$$

for any $a \in \mathcal{U}_{ad}$, where $u(a)$ solves $(\mathcal{P}(a))$, we arrive at the assertion. □

Remark 2.3. The proof of the counterpart of Theorem 2.2 is much more involved since the solution to the state problem is not globally regular.

3. Discrete sensitivity analysis

The aim of this section is to show how to compute the first and second order derivatives of the cost functionals presented in Section 2. To this end we pass to the algebraic formulation of $(\mathbb{P})_h$.

Each $a \in \mathcal{U}_{ad}$ can be identified with the vector $\boldsymbol{a} = (a_1, \dots, a_r) \in \mathbb{R}^r$, where $a_i = a_{|\Omega_i}$, $i = 1, \dots, r$ and the admissible set $\mathcal{U}_{ad}$ with a compact subset $\mathcal{U}$ of $\mathbb{R}^r$:

$$\mathcal{U} = \left\{\boldsymbol{a} = (a_1, \dots, a_r) \in \mathbb{R}^r \,|\, a_{\min} \le a_i \le a_{\max}, \ i = 1, \dots, r\right\}. \tag{3.1}$$

Let the discretization parameter $h > 0$ be fixed. The discrete state problem $(\mathcal{P}(a))_h$ leads to a system of linear equations in $\mathbb{R}^n$:

$$(\boldsymbol{P}(\boldsymbol{a})) \quad \boldsymbol{K}(\boldsymbol{a})\boldsymbol{u}(\boldsymbol{a}) = \boldsymbol{f},$$

where $\boldsymbol{K}(\boldsymbol{a}) = \left(k_{ij}(\boldsymbol{a})\right)_{i,j=1}^n$ is a *symmetric*, *positive definite* matrix for any $\boldsymbol{a} \in \mathcal{U}$, $\boldsymbol{u}(\boldsymbol{a}) \in \mathbb{R}^n$ is the vector of nodal values of $u_h(a)$ and $\boldsymbol{f} = (f_i)_{i=1}^n \in \mathbb{R}^n$. Recall that

$$\left.\begin{aligned} k_{ij}(\boldsymbol{a}) &= \int_\Omega a \nabla\varphi_i \nabla\varphi_j dx, \\ f_i &= \int_\Omega f\varphi_i dx, \quad i,j = 1, \dots, n, \end{aligned}\right\} \tag{3.2}$$

where $\{\varphi_i\}_{i=1}^n$ is the Courant basis of V_h.

Let $\mathcal{J} : \mathcal{U} \mapsto \mathbb{R}^1$ be defined by

$$\mathcal{J}(\boldsymbol{a}) := J(a, u_h(a)), \quad a \in \mathcal{U}_{ad},$$

where $u_h(a) \in V_h$ solves $(\mathcal{P}(a))_h$.

The algebraic form of $(\mathbb{P})_h$ reads as follows:

$$(\boldsymbol{P}) \quad \begin{cases} \text{find } \boldsymbol{a}^* \in \mathcal{U} \text{ such that} \\ \mathcal{J}\left(\boldsymbol{a}^*\right) \le \mathcal{J}(\boldsymbol{a}) \quad \forall \boldsymbol{a} \in \mathcal{U}. \end{cases}$$

Next we suppose that $\mathcal{J}$ is a least square type function of the following form:

$$\mathcal{J}(\boldsymbol{a}) = \frac{1}{2}\left(\boldsymbol{B}(\boldsymbol{a})\left(\boldsymbol{R}\boldsymbol{u}(\boldsymbol{a}) - \boldsymbol{z}_{\boldsymbol{d}}\right), \boldsymbol{R}\boldsymbol{u}(\boldsymbol{a}) - \boldsymbol{z}_{\boldsymbol{d}}\right)_s, \tag{3.3}$$

where $\boldsymbol{u}(\boldsymbol{a})$ is the solution of $(\boldsymbol{P}(\boldsymbol{a}))$, $\boldsymbol{B}(\boldsymbol{a}) \in \mathbb{R}^{s\times s}$ is a *symmetric*, *positive definite* $(s \times s)$ matrix for $\forall \boldsymbol{a} \in \mathcal{U}$. Further $\boldsymbol{R} \in \mathbb{R}^{s\times n}$ is a matrix representing the restriction mapping of $\mathbb{R}^n$ onto $\mathbb{R}^s$, i.e. if $\mathcal{I} \subset \{1, \dots, n\}$ is an index set of cardinality s then $\boldsymbol{R}\boldsymbol{x} = \boldsymbol{x}_{\mathcal{I}}$, where $\boldsymbol{x} = (x_i)_{i=1}^n$, $\boldsymbol{x}_{\mathcal{I}} = (x_i)_{i\in\mathcal{I}}$. Finally, $\boldsymbol{z}_{\boldsymbol{d}} \in \mathbb{R}^s$ is given.

Remark 3.1. This form of $\mathcal{J}$ covers all cost functionals from Remarks 2.1 and 2.2. Indeed,

(i) (*full energy norm*) If J is defined by (2.10) with $D = \Omega$, then $s = n, \boldsymbol{B}(\boldsymbol{a}) = \boldsymbol{K}(\boldsymbol{a})\ \forall \boldsymbol{a} \in \mathcal{U}$ and $\boldsymbol{R} = \boldsymbol{id} \in \mathbb{R}^{n\times n}$;

(ii) (*partial energy norm*) If $D \subsetneq \Omega$ in (2.10) then we suppose that D is a *polygonal* domain such that $\mathcal{T}_{h|_D}$ defines a standard triangulation of $\bar{D}$, where $\mathcal{T}_h$ is a triangulation used for the construction of V_h. Let $\mathcal{I} \subset \{1, \dots, n\}$ be the set of indices of all nodes of $\mathcal{T}_{h|_{\bar{D}}}$ and *card* $\mathcal{I} = p$. Then $s = p, \boldsymbol{R} \in \mathbb{R}^{p\times n}$ is the restriction of $\mathbb{R}^n$ onto $\mathbb{R}^p$, and $\boldsymbol{B}(\boldsymbol{a}) = \boldsymbol{K}_D(\boldsymbol{a}) \in \mathbb{R}^{p\times p}$, $\boldsymbol{K}_D(\boldsymbol{a}) = \left(k_{ij}(\boldsymbol{a})\right)_{i,j\in\mathcal{I}}$, where the entries are defined by $(3.2)_1$;

(iii) (*Euclidean norm*) If J is defined by (2.11) we suppose that the measurement points $\{x_i\}, i \in \mathcal{I}$ belong to the nodes of $\mathcal{T}_h$. Then $s = m$, where $m = card\ \mathcal{I}, \boldsymbol{B}(\boldsymbol{a}) = \boldsymbol{id} \in \mathbb{R}^{m\times m}$ and $\boldsymbol{R}$ is the restriction of $\mathbb{R}^n$ onto $\mathbb{R}^m$.

Next we derive the expression for the gradient $\mathbb{D}\mathcal{J}(\boldsymbol{a})$ and Hessian $\mathbb{D}^2\mathcal{J}(\boldsymbol{a})$ of $\mathcal{J}$ defined by (3.3) at $\boldsymbol{a} \in \mathcal{U}$. To this end we suppose that the matrix function $\boldsymbol{B} : \mathcal{U} \mapsto \mathbb{R}^{s\times s}$ is *continuously differentiable* on $\mathcal{U}$.

Let $\boldsymbol{a} \in \mathcal{U}$ and $\boldsymbol{b} \in \mathbb{R}^r$ be given. The *directional derivative* of $\mathcal{J}$ at $\boldsymbol{a}$ and the direction $\boldsymbol{b}$ is defined by

$$\mathcal{J}'(\boldsymbol{a}, \boldsymbol{b}) = \lim_{t\to 0_+} \frac{\mathcal{J}(\boldsymbol{a} + t\boldsymbol{b}) - \mathcal{J}(\boldsymbol{a})}{t}.$$

It is known that if the mapping $\boldsymbol{a} \mapsto \mathcal{J}'(\boldsymbol{a}, \boldsymbol{b})$ is continuous at $\boldsymbol{a}$ for any $\boldsymbol{b} \in \mathbb{R}^r$ then choosing $\boldsymbol{b} = \boldsymbol{e}_1, \dots, \boldsymbol{e}_r$ (the canonical basis of $\mathbb{R}^r$) we obtain $\mathbb{D}\mathcal{J}(a)$. Analogously we define the directional derivatives of vector and matrix functions. To simplify notation we write $\mathcal{J}'(\boldsymbol{a})$ instead of $\mathcal{J}'(\boldsymbol{a}, \boldsymbol{b})$ if the direction $\boldsymbol{b}$ is apparent from the context.

From $(3.2)_1$, it follows:

$$k_{ij}(\boldsymbol{a}) = \sum_{k=1}^{r} a_k c_{ij}^{(k)},$$

where $\boldsymbol{a} = (a_1, \dots, a_r)$, $c_{ij}^{(k)} = \int_{\Omega_k} \nabla\varphi_i \nabla\varphi_j dx$, and $\{\Omega_k\}_{k=1}^r$ is the decomposition of Ω, $i, j = 1, \dots, n$. Therefore

$$\mathbb{D}k_{ij}(\boldsymbol{a}) = (c_{ij}^{(1)}, \dots, c_{ij}^{(r)})$$

and

$$\mathbb{D}^2 k_{ij}(\boldsymbol{a}) = \boldsymbol{0}.$$

Let $\boldsymbol{a} \in \mathcal{U}, \boldsymbol{b} \in \mathbb{R}^r$ be given and $\boldsymbol{u}(\boldsymbol{a})$ be the solution to $(\mathcal{P}(\boldsymbol{a}))$. The directional derivative $\boldsymbol{u}'(\boldsymbol{a}) \in \mathbb{R}^n$ at $\boldsymbol{a}$ and the direction $\boldsymbol{b}$ solves the following system of algebraic equations:

$$\boldsymbol{K}(\boldsymbol{a})\boldsymbol{u}'(\boldsymbol{a}) = -\boldsymbol{K}'(\boldsymbol{a})\boldsymbol{u}(\boldsymbol{a}), \tag{3.4}$$

where $\boldsymbol{K}'(\boldsymbol{a}) = \left(k'_{ij}(\boldsymbol{a})\right)_{i,j=1}^n \in \mathbb{R}^{n\times n}$.

Then from (3.3) and (3.4) we obtain:

$$\mathcal{J}'(\boldsymbol{a}) := \mathcal{J}'(\boldsymbol{a}, \boldsymbol{b}) = \frac{1}{2}\left(\boldsymbol{B}'(\boldsymbol{a})(\boldsymbol{R}\boldsymbol{u}(\boldsymbol{a}) - \boldsymbol{z}_d), \boldsymbol{R}\boldsymbol{u}(\boldsymbol{a}) - \boldsymbol{z}_d\right)_s + (\boldsymbol{R}^T\boldsymbol{B}(\boldsymbol{a})(\boldsymbol{R}\boldsymbol{u}(\boldsymbol{a}) - \boldsymbol{z}_d), \boldsymbol{u}'(\boldsymbol{a}))_n \tag{3.5}$$

making use of the symmetry of $\boldsymbol{B}(\boldsymbol{a})$. To eliminate $\boldsymbol{u}'(\boldsymbol{a})$ we use the *adjoint state equation*:

$$\boldsymbol{K}(\boldsymbol{a})\boldsymbol{p}(\boldsymbol{a}) = \boldsymbol{R}^T\boldsymbol{B}(\boldsymbol{a})(\boldsymbol{R}\boldsymbol{u}(\boldsymbol{a}) - \boldsymbol{z}_d). \tag{3.6}$$

The vector $\boldsymbol{p}(\boldsymbol{a}) \in \mathbb{R}^n$ is termed the *adjoint state*.

We multiply (3.6) by $\boldsymbol{u}'(\boldsymbol{a})$ and use (3.4):

$$\begin{aligned}\left(\boldsymbol{R}^T\boldsymbol{B}(\boldsymbol{a})(\boldsymbol{R}\boldsymbol{u}(\boldsymbol{a}) - \boldsymbol{z}_d), \boldsymbol{u}'(\boldsymbol{a})\right)_n &= \left(\boldsymbol{K}(\boldsymbol{a})\boldsymbol{p}(\boldsymbol{a}), \boldsymbol{u}'(\boldsymbol{a})\right)_n \\ &= \left(\boldsymbol{p}(\boldsymbol{a}), \boldsymbol{K}(\boldsymbol{a})\boldsymbol{u}'(\boldsymbol{a})\right)_n = -\left(\boldsymbol{p}(\boldsymbol{a}), \boldsymbol{K}'(\boldsymbol{a})\boldsymbol{u}(\boldsymbol{a})\right)_n.\end{aligned}$$

From this and (3.5) we arrive at the following expression for the derivative of $\mathcal{J}$ at $\boldsymbol{a} \in \mathcal{U}$ and the direction $\boldsymbol{b} \in \mathbb{R}^r$:

$$\mathcal{J}'(\boldsymbol{a}) = \frac{1}{2}\left(\boldsymbol{B}'(\boldsymbol{a})(\boldsymbol{R}\boldsymbol{u}(\boldsymbol{a}) - \boldsymbol{z}_d), \boldsymbol{R}\boldsymbol{u}(\boldsymbol{a}) - \boldsymbol{z}_d\right)_s - \left(\boldsymbol{p}(\boldsymbol{a}), \boldsymbol{K}'(\boldsymbol{a})\boldsymbol{u}(\boldsymbol{a})\right)_n, \tag{3.7}$$

where $\boldsymbol{p}(\boldsymbol{a}) \in \mathbb{R}^n$ solves (3.6).

On the basis of (3.7) we obtain the directional derivatives of the functions $\mathcal{J}$ introduced in Remark 3.1:

(ad i) (*full energy norm*) $\boldsymbol{B}(\boldsymbol{a}) = \boldsymbol{K}(\boldsymbol{a})\ \forall \boldsymbol{a} \in \mathcal{U}$ and $\boldsymbol{R} = \boldsymbol{id} \in \mathbb{R}^{n\times n}, s = n$. Then

$$\mathcal{J}(\boldsymbol{a}) = \frac{1}{2}(\boldsymbol{K}(\boldsymbol{a})(\boldsymbol{u}(\boldsymbol{a}) - \boldsymbol{z}_d), \boldsymbol{u}(\boldsymbol{a}) - \boldsymbol{z}_d)_n.$$

The adjoint equation (3.6) reads:

$$\boldsymbol{K}(\boldsymbol{a})\boldsymbol{p}(\boldsymbol{a}) = \boldsymbol{K}(\boldsymbol{a})(\boldsymbol{u}(\boldsymbol{a}) - \boldsymbol{z}_d)$$

so that

$$\boldsymbol{p}(\boldsymbol{a}) = \boldsymbol{u}(\boldsymbol{a}) - \boldsymbol{z_d}.$$

Hence

$$\mathcal{J}'(\boldsymbol{a}) = -\frac{1}{2}\left(\boldsymbol{K}'(\boldsymbol{a})\left(\boldsymbol{u}(\boldsymbol{a}) + \boldsymbol{z_d}\right), \boldsymbol{u}(\boldsymbol{a}) - \boldsymbol{z_d}\right)_n. \tag{3.8}$$

(ad ii) (*partial energy norm*) $\boldsymbol{B}(\boldsymbol{a}) = \boldsymbol{K}_D(\boldsymbol{a}) \in \mathbb{R}^{p\times p}$, $\forall \boldsymbol{a} \in \mathcal{U}$, $\boldsymbol{R} \in \mathbb{R}^{p\times n}$, $\boldsymbol{Ru}(\boldsymbol{a}) = \boldsymbol{u}_{\mathcal{I}}(\boldsymbol{a})$, where $\boldsymbol{u}_{\mathcal{I}}(\boldsymbol{a})$ is the vector of the nodal values of $\boldsymbol{u}(\boldsymbol{a})$ at the nodes of $\mathcal{T}_h$ belonging to $\bar{D}$. Then

$$\mathcal{J}(\boldsymbol{a}) = \frac{1}{2}\left(\boldsymbol{K}_D(\boldsymbol{a})\left(\boldsymbol{u}_{\mathcal{I}}(\boldsymbol{a}) - \boldsymbol{z_d}\right), \boldsymbol{u}_{\mathcal{I}}(\boldsymbol{a}) - \boldsymbol{z_d}\right)_p$$

and

$$\mathcal{J}'(\boldsymbol{a}) = -\frac{1}{2}\left(\boldsymbol{K}_D'(\boldsymbol{a})\left(\boldsymbol{u}_{\mathcal{I}}(\boldsymbol{a}) - \boldsymbol{z_d}\right), \boldsymbol{u}_{\mathcal{I}}(\boldsymbol{a}) - \boldsymbol{z_d}\right)_p - \left(\boldsymbol{p}(\boldsymbol{a}), \boldsymbol{K}'(\boldsymbol{a})\boldsymbol{u}(\boldsymbol{a})\right)_n, \tag{3.9}$$

where $\boldsymbol{p}(\boldsymbol{a})$ solves the adjoint equation

$$\boldsymbol{K}(\boldsymbol{a})\boldsymbol{p}(\boldsymbol{a}) = \boldsymbol{R}^T\boldsymbol{K}_D(\boldsymbol{a})\left(\boldsymbol{u}_{\mathcal{I}}(\boldsymbol{a}) - \boldsymbol{z_d}\right).$$

(ad iii) (*Euclidean norm*) $\boldsymbol{B}(\boldsymbol{a}) = \boldsymbol{id} \in \mathbb{R}^{m\times m}$, $\boldsymbol{R} \in \mathbb{R}^{m\times n}$, $\boldsymbol{Ru}(\boldsymbol{a}) = \boldsymbol{u}_{\mathcal{I}}(\boldsymbol{a})$, where $\mathcal{I}$ is the set containing the indices of the nodes of $\mathcal{T}_h$, where m measurements are available. Then

$$\mathcal{J}(\boldsymbol{a}) = \frac{1}{2}\left(\boldsymbol{u}_{\mathcal{I}}(\boldsymbol{a}) - \boldsymbol{z_d}, \boldsymbol{u}_{\mathcal{I}}(\boldsymbol{a}) - \boldsymbol{z_d}\right)_m$$

and

$$\mathcal{J}'(\boldsymbol{a}) = -\left(\boldsymbol{p}(\boldsymbol{a}), \boldsymbol{K}'(\boldsymbol{a})\boldsymbol{u}(\boldsymbol{a})\right)_n \tag{3.10}$$

using that $\boldsymbol{B}'(\boldsymbol{a}) = \boldsymbol{0}\ \forall \boldsymbol{a} \in \mathcal{U}$. The adjoint state $\boldsymbol{p}(\boldsymbol{a})$ solves:

$$\boldsymbol{K}(\boldsymbol{a})\boldsymbol{p}(\boldsymbol{a}) = \boldsymbol{R}^T\left(\boldsymbol{u}_{\mathcal{I}}(\boldsymbol{a}) - \boldsymbol{z_d}\right).$$

From (3.8), (3.9), (3.10) we obtain the following expressions for the partial derivatives $\frac{\partial \mathcal{J}(\boldsymbol{a})}{\partial a_k}$, $k = 1, \ldots, r$:

$$\frac{\partial \mathcal{J}(\boldsymbol{a})}{\partial a_k} = -\frac{1}{2}\left(\frac{\partial \boldsymbol{K}(\boldsymbol{a})}{\partial a_k}\left(\boldsymbol{u}(\boldsymbol{a}) + \boldsymbol{z_d}\right), \boldsymbol{u}(\boldsymbol{a}) - \boldsymbol{z_d}\right)_n \tag{3.11}$$

$$\frac{\partial \mathcal{J}(\boldsymbol{a})}{\partial a_k} = -\frac{1}{2}\left(\frac{\partial \boldsymbol{K}_D(\boldsymbol{a})}{\partial a_k}\left(\boldsymbol{u}_{\mathcal{I}}(\boldsymbol{a}) - \boldsymbol{z_d}\right), \boldsymbol{u}_{\mathcal{I}}(\boldsymbol{a}) - \boldsymbol{z_d}\right)_p - \left(\boldsymbol{p}(\boldsymbol{a}), \frac{\partial \boldsymbol{K}(\boldsymbol{a})}{\partial a_k}\boldsymbol{u}(\boldsymbol{a})\right)_n \tag{3.12}$$

$$\frac{\partial \mathcal{J}(\boldsymbol{a})}{\partial a_k} = -\left(\boldsymbol{p}(\boldsymbol{a}), \frac{\partial \boldsymbol{K}(\boldsymbol{a})}{\partial a_k}\boldsymbol{u}(\boldsymbol{a})\right)_n, \tag{3.13}$$

respectively, where $\boldsymbol{p}(\boldsymbol{a})$ in (3.12) and (3.13) is the respective adjoint state.

Now we pass to the computation of the Hessian $\mathbb{D}^2\mathcal{J}(\boldsymbol{a})$ of $\mathcal{J}$ at $\boldsymbol{a}$ resulting from the full energy norm. From (3.11) we easily obtain:

$$\mathbb{D}^2\mathcal{J}(\boldsymbol{a}) = \left[\left(\boldsymbol{K}(\boldsymbol{a})\frac{\partial \boldsymbol{u}(\boldsymbol{a})}{\partial a_k}, \frac{\partial \boldsymbol{u}(\boldsymbol{a})}{\partial a_l}\right)_n\right]_{k,l=1}^r \tag{3.14}$$

making use of the symmetry of $\boldsymbol{K}(\boldsymbol{a})$, (3.4) and the fact that the second derivative of $\boldsymbol{K}(\boldsymbol{a})$ is the null matrix for all $\boldsymbol{a} \in \mathcal{U}$. Since $\boldsymbol{K}(\boldsymbol{a})$ is positive definite for any $\boldsymbol{a} \in \mathcal{U}$, the Hessian $\mathbb{D}^2\mathcal{J}(\boldsymbol{a})$ is positive semidefinite and consequently the function $\mathcal{J}$ is *convex* in $\mathcal{U}$. Indeed, let $\boldsymbol{b} \in \mathbb{R}^r$ be arbitrary. Then

$$\left(\mathbb{D}^2\mathcal{J}(\boldsymbol{a})\boldsymbol{b}, \boldsymbol{b}\right)_n = \sum_{k,l=1}^r \left(\boldsymbol{K}(\boldsymbol{a})\frac{\partial \boldsymbol{u}(\boldsymbol{a})}{\partial a_k}, \frac{\partial \boldsymbol{u}(\boldsymbol{a})}{\partial a_l}\right)_n b_k b_l \geq 0 \quad \forall \boldsymbol{b} \in \mathbb{R}^n.$$

If $\mathcal{J}$ is defined by the partial energy or Euclidean norm, we use the Gauss–Newton approximation of $\mathbb{D}^2\mathcal{J}(\boldsymbol{a})$ whose construction will be now described. Let

$$\boldsymbol{\Phi}(\boldsymbol{a}) = \frac{1}{2}\left(\boldsymbol{A}(\boldsymbol{a})\left(\boldsymbol{x}(\boldsymbol{a}) - \boldsymbol{z_d}\right), \boldsymbol{x}(\boldsymbol{a}) - \boldsymbol{z_d}\right)_s, \quad \boldsymbol{z_d} \in \mathbb{R}^s \text{ given}$$

be a quadratic function such that

– $\boldsymbol{A}(\boldsymbol{a}) \in \mathbb{R}^{s\times s}$ is a symmetric, positive definite matrix for any $\boldsymbol{a} \in \mathcal{U}$; (3.15)

– $\boldsymbol{A} : \mathcal{U} \mapsto \mathbb{R}^{s\times s}, \boldsymbol{a} \mapsto \boldsymbol{A}(\boldsymbol{a})$ is a linear function of $\boldsymbol{a} \in \mathcal{U}$; (3.16)

– $\boldsymbol{x} : \mathcal{U} \mapsto \mathbb{R}^{s}, \boldsymbol{a} \mapsto \boldsymbol{x}(\boldsymbol{a}), \boldsymbol{a} \in \mathcal{U}$ is a continuously differentiable vector function in $\mathcal{U}$. (3.17)

A direct computation yields:

$$\frac{\partial^2 \boldsymbol{\Phi}(\boldsymbol{a})}{\partial a_k \partial a_l} = \left(\frac{\partial \boldsymbol{A}(\boldsymbol{a})}{\partial a_k}\frac{\partial \boldsymbol{x}(\boldsymbol{a})}{\partial a_l}, \boldsymbol{x}(\boldsymbol{a}) - \boldsymbol{z_d}\right)_s + \left(\frac{\partial \boldsymbol{A}(\boldsymbol{a})}{\partial a_l}\frac{\partial \boldsymbol{x}(\boldsymbol{a})}{\partial a_k}, \boldsymbol{x}(\boldsymbol{a}) - \boldsymbol{z_d}\right)_s + \left(\boldsymbol{A}(\boldsymbol{a})\frac{\partial^2 \boldsymbol{x}(\boldsymbol{a})}{\partial a_k \partial a_l}, \boldsymbol{x}(\boldsymbol{a}) - \boldsymbol{z_d}\right)_s + \left(\boldsymbol{A}(\boldsymbol{a})\frac{\partial \boldsymbol{x}(\boldsymbol{a})}{\partial a_k}, \frac{\partial \boldsymbol{x}(\boldsymbol{a})}{\partial a_l}\right)_s, \quad k, l = 1, \ldots, r$$

making use of (3.15)–(3.17). By the *Gauss–Newton approximation* of the Hessian $\mathbb{D}^2\boldsymbol{\Phi}(\boldsymbol{a})$ we call a matrix

$$\mathbb{D}^2_{GN}\boldsymbol{\Phi}(\boldsymbol{a}) = \left[\left(\boldsymbol{A}(\boldsymbol{a})\frac{\partial \boldsymbol{x}(\boldsymbol{a})}{\partial a_l}, \frac{\partial \boldsymbol{x}(\boldsymbol{a})}{\partial a_k}\right)_s\right]_{k,l=1}^{r}.$$

As before one can show that the matrix $\mathbb{D}_{GN}\boldsymbol{\Phi}(\boldsymbol{a}) \in \mathbb{R}^{r\times r}$ is *positive semidefinite* for any $\boldsymbol{a} \in \mathcal{U}$.

The Gauss–Newton approximation of $\mathbb{D}^2\boldsymbol{\mathcal{J}}(\boldsymbol{a})$ at $\boldsymbol{a} \in \mathcal{U}$, where $\boldsymbol{\mathcal{J}}$ is defined by the partial energy and Euclidean norm is given by

$$\mathbb{D}^2_{GN}\boldsymbol{\mathcal{J}}(\boldsymbol{a}) = \left[\left(\boldsymbol{K}_D(\boldsymbol{a})\frac{\partial \boldsymbol{u}_I(\boldsymbol{a})}{\partial a_k}, \frac{\partial \boldsymbol{u}_I(\boldsymbol{a})}{\partial a_l}\right)_p\right]_{k,l=1}^{r} \tag{3.18}$$

and

$$\mathbb{D}^2_{GN}\boldsymbol{\mathcal{J}}(\boldsymbol{a}) = \left[\left(\frac{\partial \boldsymbol{u}_I(\boldsymbol{a})}{\partial a_k}, \frac{\partial \boldsymbol{u}_I(\boldsymbol{a})}{\partial a_l}\right)_m\right]_{k,l=1}^{r}, \tag{3.19}$$

respectively.

At the end of this section we mention the discrete form of the so-called *equation error method* [4,6]. The cost function $\boldsymbol{\mathcal{J}}$ results from (3.3) with $\boldsymbol{B}(\boldsymbol{a}) = \boldsymbol{K}^2(\boldsymbol{a})\ \forall \boldsymbol{a} \in \mathcal{U}, \boldsymbol{R} = \boldsymbol{id} \in \mathbb{R}^{n\times n}, \boldsymbol{z_d} \in \mathbb{R}^n, s = n$ using also the definition of the state problem $(\mathcal{P}(\boldsymbol{a}))$:

$$\begin{aligned}\boldsymbol{\mathcal{J}}(\boldsymbol{a}) &= \frac{1}{2}\left(\boldsymbol{K}^2(\boldsymbol{a})(\boldsymbol{u}(\boldsymbol{a}) - \boldsymbol{z_d}), \boldsymbol{u}(\boldsymbol{a}) - \boldsymbol{z_d}\right)_n \\ &= \frac{1}{2}\|\boldsymbol{K}(\boldsymbol{a})(\boldsymbol{u}(\boldsymbol{a}) - \boldsymbol{z_d})\|_n^2 = \frac{1}{2}\|\boldsymbol{f} - \boldsymbol{K}(\boldsymbol{a})\boldsymbol{z_d}\|_n^2.\end{aligned} \tag{3.20}$$

It is worth noticing that to evaluate $\boldsymbol{\mathcal{J}}(\boldsymbol{a}), \boldsymbol{a} \in \mathcal{U}$ no solution to the state problem is needed. In addition, $\boldsymbol{\mathcal{J}}$ is *quadratic* in $\mathbb{R}^r$ since $\boldsymbol{K}$ depends linearly on $\boldsymbol{a}$. Thus the directional derivative of $\boldsymbol{\mathcal{J}}$ at $\boldsymbol{a} \in \mathcal{U}$ and a direction $\boldsymbol{b} \in \mathbb{R}^r$ is

$$\boldsymbol{\mathcal{J}}'(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{K}(\boldsymbol{a})\boldsymbol{z_d} - \boldsymbol{f}, \boldsymbol{K}(\boldsymbol{b})\boldsymbol{z_d})_n. \tag{3.21}$$

Let $\boldsymbol{L} := \boldsymbol{L}(\boldsymbol{z_d}) \in \mathbb{R}^{r\times n}$ be the matrix defined by

$$\boldsymbol{K}(\boldsymbol{b})\boldsymbol{z_d} = \boldsymbol{L}\boldsymbol{b} \quad \forall \boldsymbol{b} \in \mathbb{R}^r.$$

Then (3.21) can be written as follows:

$$\boldsymbol{\mathcal{J}}'(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{L}\boldsymbol{a} - \boldsymbol{f}, \boldsymbol{L}\boldsymbol{b})_n = (\boldsymbol{L}^T(\boldsymbol{L}\boldsymbol{a} - \boldsymbol{f}), \boldsymbol{b})_r. \tag{3.22}$$

Choosing $\boldsymbol{b} = \boldsymbol{e}_k$ we get from (3.22):

$$\frac{\partial \boldsymbol{\mathcal{J}}(\boldsymbol{a})}{\partial a_k} = ((\boldsymbol{L}^T\boldsymbol{L})_k, \boldsymbol{a})_r - ((\boldsymbol{L}^T)_k, \boldsymbol{f})_n$$

and

$$\frac{\partial^2 \boldsymbol{\mathcal{J}}(\boldsymbol{a})}{\partial a_k \partial a_l} = (\boldsymbol{L}^T\boldsymbol{L})_{k,l}, \quad k, l = 1, \ldots, r,$$

where $(\bullet)_k$, $(\bullet)_{k,l}$ denotes the kth row and (k, l)-element of a matrix $\bullet$, respectively. Another advantage of this method is the fact that the unconstrained minimization of $\boldsymbol{\mathcal{J}}$ given by (3.20) leads to a normal system of equations. Indeed, from (3.22) we see that

$$\nabla \boldsymbol{\mathcal{J}}(\boldsymbol{a}) = \boldsymbol{0} \Leftrightarrow \boldsymbol{L}^T\boldsymbol{L}\boldsymbol{a} = \boldsymbol{L}^T\boldsymbol{f}.$$

4. Optimization techniques

The minimization of non-linear least squares functionals $\mathcal{J}$ is frequently performed by Levenberg–Marquardt methods belonging to a class of quasi-Newton methods [10,11]. If $\boldsymbol{a}^k$ is an approximation of the minimum of $\mathcal{J}$, the next iteration is searched in the form

$$\boldsymbol{a}^{k+1} = \boldsymbol{a}^k - \alpha_k \boldsymbol{d}_k,$$

where $\alpha_k > 0$ is a damping parameter and $\boldsymbol{d}_k$ is a trial step being the solution to the following linear system:

$$\left[\mathbb{D}^2_{GN}\mathcal{J}\left(\boldsymbol{a}^k\right) + \nu_k \boldsymbol{D}(\boldsymbol{a}^k)\right]\boldsymbol{d}^k = \mathbb{D}\mathcal{J}(\boldsymbol{a}^k) \tag{4.1}$$

where $\mathbb{D}^2_{GN}\mathcal{J}(\boldsymbol{a}^k)$ stands for the Gauss–Newton approximation of the Hessian of $\mathcal{J}$ at $\boldsymbol{a}^k$, $\boldsymbol{D}$ is a diagonal regularization matrix function with positive entries and $\nu_k > 0$ is a regularization parameter which can be adaptively changed at each iteration. Two possible choices of the regularization (scaling) matrix are $\boldsymbol{D}(\boldsymbol{a}^k) := \boldsymbol{D} = \boldsymbol{id}$ and $\boldsymbol{D}(\boldsymbol{a}^k) = \mathrm{diag}(\mathbb{D}^2_{GN}\mathcal{J}(\boldsymbol{a}^k))$. The latter one is used in the sequel since the former one turned out not to be suitable for identification of parameters which differ by several orders of magnitude. Note that for $\boldsymbol{D}(\boldsymbol{a}^k) = \mathrm{diag}(\mathbb{D}^2_{GN}\mathcal{J}(\boldsymbol{a}^k))$ the diagonal entries are $D_{ii}(\boldsymbol{a}^k) = \|\boldsymbol{c}_i\|^2$, where $\boldsymbol{c}_i$ is the ith column of $\mathbb{D}\mathcal{J}(\boldsymbol{a}^k)$. A possible zero diagonal element can be replaced by a positive number $\delta > 0$, but in the tests reported in Section 6 such measure was not necessary. The damping parameter α_k is mainly used to ensure global convergence of the method. To ensure possible local quadratic convergence of $\boldsymbol{a}^k$, the parameters ν_k decrease when $\boldsymbol{a}^k$ approaches the minimum of $\mathcal{J}$ and $\min \mathcal{J}$ is small. This trend is respected when e.g. $\nu_k = c\sqrt{\mathcal{J}\left(\boldsymbol{a}^k\right)}$ with a positive constant $c > 0$ determined experimentally, see the choice in Section 6. Note that the relation of the Levenberg–Marquardt method to the trust region technique and a subsequent adaptive procedure for determination of ν_k can be found e.g. in [12]. The set $\mathcal{U}$ of admissible controls involves constraints on control variables. One possibility how to treat such constraints is to augment $\mathcal{J}$ by adding a penalty functional, another possibility is to use constrained variants of Levenberg–Marquardt methods [13]. In our case we used the unconstrained variant (4.1). The constraints were chosen in such a way that none of them became active.

In many engineering identification problems, the (insufficient) amount of data can be increased by measurements under different input conditions (multi-response data) generating cost functionals $\mathcal{J}_1, \ldots, \mathcal{J}_L$. Let $(\boldsymbol{P})_k$ denote the optimization problem for single $\mathcal{J}_k$, $k = 1, \ldots, L$. Since the individual problems $(\boldsymbol{P})_k$ are not usually able to give satisfactory results, they have to be mutually interlocked in some way. We propose two solution procedures how to do that, namely the *sequential* approach and the *weighted sum method*. The idea of the sequential approach is very simple: each $(\boldsymbol{P})_k$ is (approximately) solved by an iterative method, a found solution is used as the initial approximation in $(\boldsymbol{P})_{k+1}$. The whole cycle is repeated until a termination criteria is satisfied. Note that in our computations we use a *one shot* variant of the sequential approach, i.e. we perform just one iteration of (4.1) and proceed to the next problem. The weighted sum method is based on the scalarization approach: we define the function $\mathcal{J}_{sc}$ as the linear combination of $\mathcal{J}_1, \ldots, \mathcal{J}_L$:

$$\mathcal{J}_{sc} = \sum_{k=1}^{L} w_k \mathcal{J}_k, \tag{4.2}$$

where $w_k > 0$ are appropriate weights. Subsequent minimization of $\mathcal{J}_{sc}$ can be carried out by (4.1), e.g.

5. Numerical example

This section deals with a steady state groundwater flow problem in a confined aquifer in 2D. The model problem represents a pumping test [7] used to determine hydraulic conductivity in an aquifer. The state problem is the Darcy flow equation (2.1) with the coefficient a being hydraulic conductivity.

The computational domain Ω is a square of size 1200×1200 (in m) with another square ω of size 100×100 (in m) which is located in the middle of Ω. We assume that ω is split into eight subdomains $\{\Omega_i\}_{i=1}^8$, which are materially homogeneous and whose position is a priori known. Two of the subdomains contain the same material (violet), i.e. the number of different materials in ω to be identified (distinguished by different colors in Fig. 1) is seven. The material in Ω_0 (the complement of ω) is known. Thus our model example belongs to the class of problems with a priori known material interfaces.

To reveal the conductivities in $\{\Omega_i\}_{i=1}^8$, we use a pumping test based on a system of eight wells placed on $\partial\omega$ and numbered by $k \in \mathcal{I} = \{1, \ldots, 8\}$, see Fig. 2. The wells are grouped into four pairs, namely $\mathcal{I}_1 = \{1, 8\}$, $\mathcal{I}_2 = \{2, 7\}$, $\mathcal{I}_3 = \{3, 6\}$, $\mathcal{I}_4 = \{5, 4\}$. The pumping test consists of four individual tests, each associated with one pair $\mathcal{I}_k = \{k_1, k_2\}$, when the fluid is injected into Ω through the well k_1 and drained out from k_2. The remaining wells are used to measure the hydraulic head response.

With any $\mathcal{I}_l$, $l = 1, \ldots, 4$ we associate one identification problem:

$$(\mathbb{P})_l \quad \begin{cases} \text{find } a_l^* \in \mathcal{U}_{ad} \text{ such that} \\ \mathcal{J}_l\left(a_l^*\right) \le \mathcal{J}_l(a) \quad \forall a \in \mathcal{U}_{ad}, \end{cases} \tag{5.1}$$

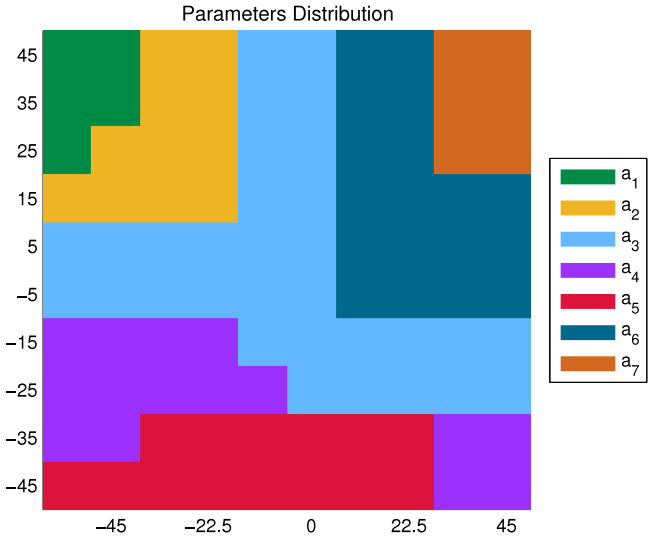

Fig. 1. Distribution of materials in ω. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

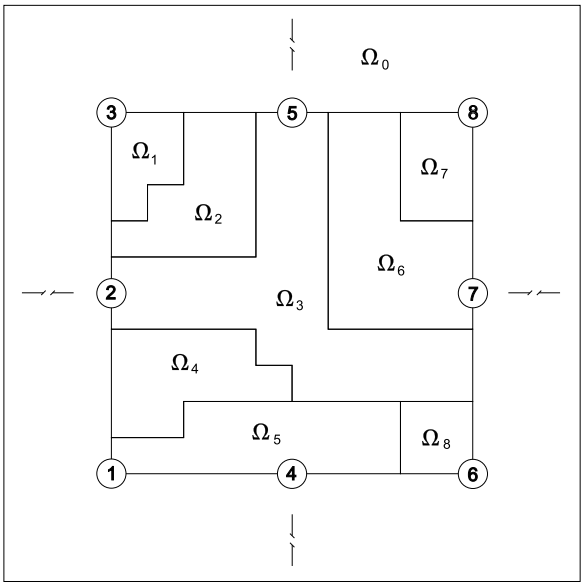

Fig. 2. Position of the wells.

where $\mathcal{J}_l(a) := J_l(a, u_l(a))$ is a functional, $u_l(a)$ is the solution to the corresponding state problem

$$(\mathcal{P}_l(a)) \quad \begin{cases} -\operatorname{div}(a\nabla u_l) = f_l & \text{in } \Omega, \\ u_l = 0 & \text{on } \partial\Omega. \end{cases} \tag{5.2}$$

The right hand side f_l in $(\mathcal{P}_l(a))$ characterizes the action of injection and drainage. It is given by two Dirac distributions concentrated at the wells.

For the discretization of $(\mathcal{P}_l(a))$, $l = 1, \ldots, 4$, $a \in \mathcal{U}_{ad}$ we use piecewise linear functions on a uniform triangulation $\mathcal{T}_h$ of $\bar{\Omega}$ having 460 800 elements. The triangulation is aligned with the subdomains Ω_i, $\forall i$ with 160, 352, 992, 352, 448, 576, 192 and 128 elements located in $\Omega_1, \ldots, \Omega_8$, respectively. Any coefficient $a \in \mathcal{U}_{ad}$ will be represented by a vector $\boldsymbol{a} = (a_1, \ldots, a_7) \in \mathbb{R}^7$.

Denote by $\boldsymbol{u}_l(\boldsymbol{a})$ the solution of the discrete state problem

$$(\boldsymbol{P}_l(\boldsymbol{a})) \quad \boldsymbol{K}(\boldsymbol{a})\,\boldsymbol{u}_l(\boldsymbol{a}) = \boldsymbol{f}_l,$$

where $\boldsymbol{K}$ is assembled with the aid of an efficient MATLAB procedure, see [10], and the discrete load vector $\boldsymbol{f}_l$ has only two non-zero components determined by injection 0.001 m^3/s and drainage -0.001 m^3/s.

Hydraulic conductivity of semi-pervious rocks varies from 10^{-10} to 10^{-6} order of magnitude. In our example we use scaled values of the coefficients $a := 10^{10}a$, and correspondingly we set the bounds $a_{\min} = 10^0$ and $a_{\max} = 9 \cdot 10^4$ in the definition of $\mathcal{U}_{ad}$.

By $\mathcal{J}_l$, $\mathcal{J}_{sc}$ we denote the algebraic form of the functionals $\mathcal{J}_l$, $l = 1, \dots, 4$ and scalarized functional $\mathcal{J}_{sc}$, respectively. According to the number of computational experiments, we finally decided to use the *Euclidean norm* to define all $\mathcal{J}_l$. Thus

$$\mathcal{J}_l(\boldsymbol{a}) = \frac{1}{2}\|\boldsymbol{u}_l(\boldsymbol{a})_{|\mathcal{D}_l} - \boldsymbol{z}_{\boldsymbol{d}}^l\|^2_{\mathbb{R}^6},$$

where $\boldsymbol{u}_l(\boldsymbol{a})_{|\mathcal{D}_l} \in \mathbb{R}^6$ is the restriction of $\boldsymbol{u}_l(\boldsymbol{a})$ on the nodes of $\mathcal{D}_l = \mathcal{N}\backslash\mathcal{N}_l$. The target vector $\boldsymbol{z}_{\boldsymbol{d}}^l$ is defined as follows: $\boldsymbol{z}_{\boldsymbol{d}}^l = \boldsymbol{u}_l(\boldsymbol{a}^*)_{|\mathcal{D}_l}$, $l = 1, \dots, 4$, where $\boldsymbol{a}^* = (21\,000,\ 2000,\ 670,\ 380,\ 60,\ 142,\ 9)$. Our aim is to recover $\boldsymbol{a}^*$.

We also compute examples with data polluted by η % noise:

$$\boldsymbol{z}_{\boldsymbol{d}}^l := \boldsymbol{z}_{\boldsymbol{d}}^l + \frac{\eta}{100}\boldsymbol{\varepsilon} \odot \boldsymbol{z}_{\boldsymbol{d}}^l, \tag{5.3}$$

where $\boldsymbol{\varepsilon} \odot \boldsymbol{z}_{\boldsymbol{d}}^l = (\varepsilon_1 z_1, \dots, \varepsilon_6 z_6)$ with ε_i randomly generated in $(-1, 1)$ and z_i denoting the components of $\boldsymbol{z}_{\boldsymbol{d}}^l$.

The quality of computed parameters $\boldsymbol{a}^k$ after performing k Levenberg–Marquardt iterations is measured by the mean square error:

$$MSE^k = \frac{1}{7}\sum_{i=1}^{7}(a_i^k - a_i^*)^2, \quad k = 0, 1, 2, \dots,$$

where a_i^k, a_i^* is the ith component of the vector $\boldsymbol{a}^k$, and $\boldsymbol{a}^*$, respectively.

Remark 5.1. The sequential approach is a purely heuristic method. It is easy to be implemented regardless the number of the cost functions $\mathcal{J}_l$. On the other hand no convergence results are at our disposal.

The weighted sum method transforms the problem into the standard form introduced in Section 2. However, owing to the specific character of the problem, mentioned in the introduction, the application of the theoretical results is not straightforward. First of all, the Dirac distributions which define the right hand sides f_l have to be approximated by appropriate L^2 regularizations, otherwise the classical concept of weak solutions cannot be used. If it is so and, in addition, if the targets z_d^l are distributed in the whole ω and the cost functionals $\mathcal{J}_l$ have the form (2.8) or (2.9) with ω in place of Ω then the results of Theorems 2.1 and 2.2 can be applied.

For pointwise measurements the situation is more involved. In order to define the cost functionals $\mathcal{J}_l$, the measurement points have to be located in the interior of Ω_i, $i = 1, \dots, 8$ (see Remark 2.2). Since all the wells are located on $\partial\omega$, this condition is not fulfilled. To satisfy it, the wells have to be slightly moved from $\partial\omega$ into the interior of the adjacent Ω_i. Then Theorem 2.1 and Remark 2.2 guarantee the existence of a solution to the continuous setting of the weighted sum method.

6. Numerical results

This section presents the numerical results of the identification of coefficients in the model problem introduced in the previous section. The minimization of $\mathcal{J}_l$, $l = 1, \dots, 4$ and $\mathcal{J}_{sc}$ is carried out by the Levenberg–Marquardt method, see Section 4. Remind that the Euclidean norm is used to define $\mathcal{J}_l$.

As we have already mentioned, the regularization parameter ν is adaptively modified during computations. Here we use similar strategy to the one proposed in [12], namely $\nu = c\ \min\{1, \|\boldsymbol{u}_l(\boldsymbol{a})_{|\mathcal{D}_l} - \boldsymbol{z}_{\boldsymbol{d}}^l\|_{\mathbb{R}^6}\}$, where $c = 1$ while in our computations we use $c = const. > 0$.

In computations the following values of α from (4.1) are used: $\alpha = 0.4$ for the sequential approach and $\alpha = 1$ for the weighted sum method. The initial iteration $\boldsymbol{a}^0 = (17\,800,\ 1700,\ 810,\ 540,\ 73,\ 120,\ 6)$ is the same in all numerical experiments.

Table 1 shows results of the identification based on four isolated pumping tests $(\mathbb{P})_l$, $l = 1, \dots, 4$. One can see that the individual problems are not able to identify all parameters with the same quality, especially the results for the lowest transmissivity are identified only roughly. In all tables we present rounded values with the rounding error on the second to fourth decimal place. The regularization multiplier c in the definition of ν is $c = 10^7$ in this case. The parameter ν has to be tuned carefully in particular in the presence of big jumps among the components of $\boldsymbol{a}$ which is just the case. The value of c is tuned experimentally to attain reasonable convergence and prevent oscillations of the coefficients during the iterative identification process. For small values of c the coefficients to be identified oscillate while for c too large convergence of the iterations becomes too slow. Note that the optimization is terminated after 40 iterations since the values of coefficients stagnate.

Table 2 displays results for the sequential and the weighted sum approach involving all four pumping tests. We see that both methods recover the coefficients accurately and efficiently. Especially the weighted sum approach (all weights are equal to one) is very efficient. Practically the exact value $\boldsymbol{a}^*$ was found using only 20 Levenberg–Marquardt iterations. In addition, convergence of parameters computed by the weighted sum approach turned out to be monotone and robust. The

Table 1
Single problems, zero noise.

Problem	MSE^{40}	$\boldsymbol{a}^{40}$						
$\mathbb{P}_1$	44 994	20 441	2009	674	417	55	129	28
$\mathbb{P}_2$	1 947	21 114	1979	671	386	61	145	1.7
$\mathbb{P}_3$	7 277	20 811	1918	755	376	57	109	21
$\mathbb{P}_4$	81 720	20 244	2024	677	364	63	131	20
$\boldsymbol{a}^*$	0	21 000	2000	670	380	60	142	9

Table 2
Multi-response approach, zero noise.

	$\boldsymbol{a}^{it.}$						
Sequential a.—96 it. $MSE^{96} = 7.9 \cdot 10^{-5}$	21 000	2000	670	380	60	142	9
Weighted sum a.—20 it. $MSE^{20} = 8.5 \cdot 10^{-5}$	21 000	2000	670	380	60	142	9
$\boldsymbol{a}^*$	21 000	2000	670	380	60	142	9

Table 3
The weighted sum approach with noise.

	Noise	$\boldsymbol{a}^{it.}$						
Weighted sum a.—67 it. $MSE^{67} = 4.2 \cdot 10^{4}$	1%	21 523	1875	708	363	67	117	13
Weighted sum a.—54 it. $MSE^{54} = 1.7 \cdot 10^{4}$	0.5%	21 341	1933	689	371	64	130	10

sequential approach is less stable and suffered from parameter oscillations and therefore it needed stronger regularization and more iterative steps. The regularization multiplier is $c = 10^3$, and $c = 10^7$ for the weighted sum and the sequential approach, respectively.

The identification reported so far used noiseless data input. Now we will suppose that input data are perturbed by a random 0.5% and 1% noise, see (2.3). The regularization multiplier is $c = 3 \cdot 10^7$ for 1% noise and $c = 6 \cdot 10^7$ for 0.5% noise. The weighted sum approach, which was found to be more efficient is used. The results are reported in Table 3. They show a substantial change of coefficients, especially for low conductivities and a higher noise in the measurements. The sequential approach is not used for noisy data due to its lower stability.

7. Conclusions

The paper is devoted to identification of piecewise constant coefficients of a scalar elliptic PDE (state problem). The coefficients are constant on a finite number of materially homogeneous subdomains, locations of which are known a priori. We present a mathematically simple and straightforward existence and convergence analysis in this type of problems. The exposition illustrates and mathematically supports the identification procedure frequently used in engineering applications. A specific point is that the presented analysis covers the case of cost functionals with not only distributed but also pointwise measurements.

Further we perform sensitivity analysis for the discretized problems, namely computations of the gradients and Hessians or Gauss–Newton approximations to the Hessians for a large class of least squares type functions including Euclidean, energy and equation error cost functions. The energy and equation error cost functions are suitable in the case of distributed measurements. We show a way of handling these functions also for pointwise measurements. For this reason the numerical results presented in Section 5 are computed by means of the classical Euclidean error function.

In Section 4 we describe the Levenberg–Marquardt optimization procedure, which is used for numerical experiments reported in Section 6. Note that the regularized variant of the Gauss–Newton approximation to the Hessian by adding a multiple of its diagonal is efficient also in the case when the spatial variation of the coefficients is of several orders of magnitude.

The lack of measurements can be improved by multi-response analysis. In Section 4 we introduce two procedures for solving multi-response identification problems. The Section 5 describes a model problem taken from hydrogeology. Section 6 contains numerical results showing high efficiency of the multi-response analysis in both computational effort and accuracy of the computed coefficients. The numerical experiments cover also the case of noisy measurements.

Acknowledgments

This work was supported by the European Regional Development Fund in the IT4Innovations Centre of Excellence project; contract/grant number: CZ.1.05/1.1.00/02.0070 and by The Ministry of Education, Youth and Sports of Tze Czech Republic from the National Programme of Sustainability (NPU II) project "IT4Innovations excellence in science", project number: LQ1602.

References

[1] R. Blaheta, R. Kohut, A. Kolcun, K. Souček, L. Staš, L. Vavro, Digital image based numerical micromechanics of geocomposites with application to chemical grouting, Int. J. Rock Mech. Min. Sci. 77 (2015) 77–88.
[2] R. Blaheta, P. Byczanski, M. Čermák, R. Hrtus, R. Kohut, A. Kolcun, J. Malík, S. Sysala, Analysis of Äspö pillar stability experiment: Continuous thermo-mechanical model development and calibration, J. Rock Mech. Geotech. Eng. 5 (2013) 124–135.
[3] C.R. Vogel, Computational Methods for Inverse Problems, Society for Industrial and Applied Mathematics, Philadelphia, 1999.
[4] T. Karkainen, An equation error method to recover diffusion from the distributed observation, Inverse Problems 13 (1997) 1033–1051.
[5] M.S. Gockenbach, A.A. Khan, An abstract framework for elliptic inverse problems: Part 1. An output least-squares approach, Math. Mech. Solids 12 (2007) 259–276.
[6] M.S. Gockenbach, B. Jadamba, A.A. Khan, A comparative numerical study of optimization approaches for elliptic inverse problems, JMI Int. J. Math. Sci. 1 (2010) 35–56.
[7] M. Cardiff, P.K. Kitanidis, Efficient solution of nonlinear, underdetermined inverse problems with a generalized PDE model, Comput. Geosci. 34 (2008) 1480–1491.
[8] F.D. Moura Neto, A.J. da Silva Neto, An Introduction to Inverse Problems with Applications, Springer-Verlag, Berlin, Heidelberg, 2013.
[9] J. Nečas, Direct Methods in the Theory of Elliptic Equations, Springer-Verlag, Berlin, Heidelberg, 2012.
[10] T. Rahman, J. Valdman, Fast MATLAB assembly of FEM matrices in 2D and 3D: Nodal elements, Appl. Math. Comput. 219 (2013) 7151–7158.
[11] D. Marquardt, An algorithm for least squares estimation of nonlinear parameters, J. Soc. Ind. Appl. Math. 11 (1963) 431–441.
[12] C.T. Kelley, Iterative Methods for Optimization, in: Frontiers in Applied Mathematics, SIAM, Philadelphia, 1999.
[13] C. Kanzow, N. Yamashita, M. Fukushima, Levenberg–Marquardt methods with strong local convergence properties for solving nonlinear equations with convex constraints, J. Comput. Appl. Math. 172 (2004) 375–397.