Int. J. Appl. Math. Comput. Sci., 2014, Vol. 24, No. 4, 865–886
DOI: 10.2478/amcs-2014-0064

A HYBRID ALGORITHM FOR SOLVING INVERSE PROBLEMS IN ELASTICITY

BARBARA BARABASZ [*], EWA GAJDA-ZAGÓRSKA [*], STANISŁAW MIGÓRSKI [**], MACIEJ PASZYŃSKI [*], ROBERT SCHAEFER [*], MACIEJ SMOŁKA [*]

[*] Department of Computer Science
AGH University of Science and Technology, Mickiewicza 30, 30-059 Kraków, Poland
e-mail: paszynsk@agh.edu.pl

[**] Chair of Optimization and Control Theory
Jagiellonian University, Kraków, Poland

The paper offers a new approach to handling difficult parametric inverse problems in elasticity and thermo-elasticity, formulated as global optimization ones. The proposed strategy is composed of two phases. In the first, global phase, the stochastic *hp*-HGS algorithm recognizes the basins of attraction of various objective minima. In the second phase, the local objective minimizers are closer approached by steepest descent processes executed singly in each basin of attraction. The proposed complex strategy is especially dedicated to ill-posed problems with multimodal objective functionals. The strategy offers comparatively low computational and memory costs resulting from a double-adaptive technique in both forward and inverse problem domains. We provide a result on the Lipschitz continuity of the objective functional composed of the elastic energy and the boundary displacement misfits with respect to the unknown constitutive parameters. It allows common scaling of the accuracy of solving forward and inverse problems, which is the core of the introduced double-adaptive technique. The capability of the proposed method of finding multiple solutions is illustrated by a computational example which consists in restoring all feasible Young modulus distributions minimizing an objective functional in a 3D domain of a photo polymer template obtained during step and flash imprint lithography.

Keywords: inverse problem, hierarchic genetic strategy, hybrid optimization, automatic *hp*-adaptive finite element method.

1. Motivation

Considerable progress has been achieved recently in the field of inverse problems and, as a result, this area is one of the fastest growing domains in applied mathematics and computer science. It is concerned with problems that consist in finding an unknown property of a medium, or an object, from the observation of a response of this medium, or object, to a probing signal. A general framework for inverse problems provides analytical means of estimating constants in mathematical models given appropriate measurements, building mathematical models, and giving insight into the design of experiments (see, for instance, the works of Banks and Kunisch (1989), Isakov (2006), Tarantola (2005) and the references therein). Typically, inverse problems lead to mathematical models which are not well-posed in the sense of Hadamard, i.e., their solution might not be unique and/or might be unstable under data perturbations, and therefore they pose severe numerical difficulties.

The growth of the area of inverse problems has largely been driven by the needs of applications both in sciences and in industry. We refer to Engl *et al.* (2000), Samarski and Vabishchevich (2007) or Tikhonov *et al.* (1995) for a description of inverse problems of different types, including inverse problems in which the equation is not specified completely as some equation coefficients are unknown, boundary inverse problems in which boundary conditions are unknown, and evolutionary inverse problems in which initial conditions are unknown.

In this paper, we deal with a class of inverse problems in linear elasticity in which the coefficients of the partial differential equations are the unknowns. Problems of this class have been frequently formulated and solved in the frame of structural health monitoring (SHM). The general scope of SHM contains a variety of theoretical results, methods and technologies to design,

run and supervise monitoring systems that allow on-line detection of structure damages. The particular SHM tasks are designing and optimizing a system of sensors set on the structure body, verification, statistical analysis and classification of data obtained from sensors, finding weakened structure regions by solving inverse problems properly formulated. The survey of such results may be found in many books (see, e.g., Garibaldi *et al.*, 1999) and scientific journals (see, e.g., *Structural Health Monitoring*, `http://shm.sagepub.com/`).

An interesting characterization of *inverse problems in mechanics* (hereafter referred to as IPMs) appearing in SHM is given by Friswell and Mottershead (2001). A typical IPM is associated here with the discrete forward problem of elastic system vibration with or without viscous damping. The system is assumed to be already discretized, or is discrete itself (e.g., frame, lattice). Usually, the decrement of some stiffness matrix coefficients represents the damage. The IPM consists in finding a weakening coefficient distribution by minimizing the functional being the misfit of the answer of the structure for vibrational forcing or/and the misfit of eigenfrequencies. The damage identification in SHM is sometimes called *model updating*.

The main difficulty in solving such IPMs is the huge dimension of the admissible set. This problem might be partially overcome by an initial selection of the structure part (region, subdomain) in which the damage may occur. In order to speed up the solving process, the objective is frequently linearized and regularized. The other problem is to obtain a discrete system in the case of a massive, lumped structure. Proper FEM or FDM approximation might be helpful, but the results of monitoring are very sensitive to the quality of this approximation. Chase Geoffrey *et al.* (2004) show sample results of the least square regularization method tested on the American Society of Civil Engineers (ASCE) benchmarks of SHM problems.

As far as steepest descent (e.g., gradient) methods are applicable for solving well posed IPMs, other methods replacing or/and supporting these are in frequent use. In the paper of Kirikera *et al.* (2008) a structural artificial neural network (ANN) prototype AI system with a parallel implementation identifying damages in composite materials is presented. An input ANN layer is connected to sensors, while the output layers deliver the data classification information.

The application of machine learning for detecting structural damages using a vibration-based damage identification procedure is presented by Figueiredo *et al.* (2011). Four learning algorithms are compared during the process of learning an auto-associative neural network.

The Bayesian inference approach to solving the inverse problem of locating structural damage based on measured vibrational parameters is presented by Huhtala and Sven (2011). The identified stiffness coefficient decrements were improved by a-posteriori stochastic analysis assuming a known distribution of the measurement errors. Some examples of solving IPMs motivated by SHM by genetic stochastic search are given by Meruane and Heylen (2009) as well as Caicedo and Yun (2011).

Perhaps the most tiring phenomenon appearing when defining and solving an IPM is the multimodality of the objective function, i.e., the existence of multiple local minima. The main causes of multimodality may fall into three groups:

- *Inherent causes*. The mathematical formulation of the problem allows multiple solutions. Sometimes, such a possibility might be formally proven (see, e.g., Cabib *et al.*, 1990) and, more frequently, it is either anticipated from the physical evidence or reflects simply our inability to prove the uniqueness.

- *Uncertainty of the objective function representation*. It appears because of both insufficient knowledge of the problem and the errors in data measurements and representation (see, e.g., Koper *et al.*, 1999; Meruane and Heylen, 2009; Caicedo and Yun, 2011).

- *Uncertainty of methods for problem solving*. Some global optimization strategies (both deterministic and stochastic) may produce artefacts in the form of local objective extrema (see, e.g., Barabasz *et al.*, 2011b). Moreover, an unavoidable error of objective evaluation makes it difficult to distinguish the global extrema among many local ones.

One way of solving ill-posed inverse problems is by using metaheuristics, which are *solution methods that orchestrate an interaction between local improvement procedures and higher level strategies to create a process capable of escaping from local optima and performing a robust search of a solution space* (Osman and Kelly, 1996). There are multiple methods that fulfill the above definition, for example, scatter search, tabu search, simulated annealing, genetic algorithms, ant colony optimization, and many others. An overview of metaheuristics is given, e.g., by Osman and Kelly (1996) or Glover and Kochenberger (2002); for an extensive list of references on this topic, see the work of Osman and Laporte (1996).

Among the existing metaheuristic search strategies, we highlight evolutionary algorithms (EAs), which have been applied to solve inverse problems (see Burczyński and Beluch, 2011; Burczyński *et al.*, 2004). In particular, they deal with multiple material defects in mechanical systems and identification of the shape and the position of a tumour region in a biological tissue domain. Inverse problems in engineering mechanics were the subject of

many international symposia (see, e.g., Tanaka, 2003). EAs were also applied to inverse problems, e.g., by Rocca *et al.* (2009) (inverse scattering problems), Singh *et al.* (2006) (inverse groundwater modelling), Lashin and Likoshvai (2004) (gene networks identification) and Xavier *et al.* (2006) (cardiac bidomain equations), who compare two parallel genetic algorithms; for details concerning this type of algorithms, refer to Cantú-Paz (2000).

Meruane and Heylen (2009) applied a real-coded genetic algorithm to solve the SHM inverse problem formulated as a nonlinear, global optimization one. The main goal of this computationally expensive stochastic approach is to handle the local minima. Fitness was based on operational model data—it considers the initial errors in the numerical model. False damage detection is avoided by using damage penalization. A parallel implementation prevents enormous execution time.

Our particular interest is devoted to EAs, which can handle multimodality by finding all "reasonable" local extrema and then perform proper post-processing in order to distinguish essential ones. Some of them, e.g., niching and sequential niching strategies (see, e.g. Mahfoud, 1997), the hierarchic genetic strategy (HGS) (see Schaefer and Kołodziej, 2003), adaptive, stochastic multi-start (see, e.g., Cabib *et al.*, 1998), are especially robust when handling multimodality in the objective function.

A novel evolutionary algorithm for finding multiple, alternative solutions of the model update of a structural system in the SHM process was presented by Caicedo and Yun (2011). The genetic algorithm identifies global and local minima of the objective function, giving the analyst the option to choose the updated model from a set of plausible ones.

Perhaps the main disadvantage of stochastic global optimization strategies is their enormous computational cost caused by a huge number of objective evaluations. In the case of parametric inverse problems, the objective evaluation requires the solution of a forward problem, which is usually costly itself. The way of reducing the computational cost of stochastic strategies is then to reduce the number of objective evaluations and possibly decrease the computational cost of the forward problem solver.

The main goal of this paper is to propose a two-phase strategy that possesses an asymptotic guarantee of success (see, e.g., Horst and Pardalos, 1995) and can handle objective multimodality. Moreover, the strategy can deliver high final accuracy with a moderate computational cost for solving parametric inverse problems in elastostatics.

The first, global phase is performed by the hierarchic genetic strategy (Schaefer and Kołodziej, 2003) combined with the hp-adaptive finite element method hp-FEM (Demkowicz, 2006; Demkowicz *et al.*, 2007). The hp-HGS develops a tree of dynamically adjustable dependent demes. It starts with a root deme which performs the most chaotic search with low accuracy. After the *metaepoch* (the fixed number of epochs) the best fitted individuals form a seed of a child deme. Sprouting child demes is repeated concurrently for the root and all other demes excluding leaves. It is performed conditionally, if there is room among existing demes at the same level of the hp-HGS tree (the distance of a seed to the centers of existing demes is sufficiently large). Moreover, demes at each level are periodically checked, and redundant demes are reduced. Searches performed by child demes are stopped if no progress in the evolution is observed. The whole strategy is stopped if a sufficient number of well fitted leaves is obtained.

In the second, local phase the satisfactorily fitted individuals (at most one per leaf deme) are improved by a quasi-Newton method, namely, the limited-memory bound-constrained version of the very well-known Broyden–Fletcher–Goldfarb–Shanno (L-BFGS-B) algorithm (Zhu *et al.*, 1997).

The imposed tree structure of demes decreases the total fitness evaluations in comparison with a single population algorithm searching with the maximum accuracy (see, e.g., Schaefer and Kołodziej, 2003; Wierzba *et al.*, 2003). The root deme searches permanently while branch and leaf demes, being much smaller than the root-deme, are invoked only in the promising regions found by their parental demes. Child demes are quickly terminated, just after they stop to search effectively. Two mechanisms, i.e., *redundancy reduction* and *conditional sprouting,* significantly reduce the number of active demes. Finally, well fitted leaf demes concurrently cover separate basins of attractions of local and global minimizers.

The main computational cost reduction is caused by precise accuracy control of the forward self-adaptive hp-FEM solver, which performs misfit evaluation (fitness). This accuracy is adapted to the inverse error at the particular level of the hp-HGS tree, which makes the global phase significantly cheaper. An extensive but necessary mathematical and algorithmic motivation for this procedure is given in Section 2 and Appendix.

Furthermore, only the necessary minimal number of local searches is activated for finding all minimizers with high accuracy. The local gradient searches are expensive in the case of numerical gradient evaluation, which is necessary if, for some reasons, it cannot be obtained analytically (e.g., misfit irregularity or lack of its algebraic formula).

Summing up, in contrast to traditional inversion algorithms producing a unique solution, our two-phase strategy may deliver multiple solutions, which enables an expert in the field to determine the best possible one as

well as its uncertainty level.

The idea of the hp-HGS was introduced by Paszyński *et al.* (2007). Analysis of the asymptotic guarantee of success and computational cost reduction with respect to single- and multi-deme strategies without the common scaling of forward and inverse errors is performed by Schaefer and Barabasz (2008). Barbasz *et al.* (2011b; 2009) show the theory necessary for applying the hp-HGS to inverse, parameter problems in heat flow. They also give computational examples of finding multiple solutions by the hp-HGS method.

The novelty of this paper consists in applying the two-phase strategy combining the hp-HGS and local methods to inverse problems in elastostatics. Moreover, we extend the formal verification of the hp-HGS to the particular class of inverse parametric problems in elastostatics. A crucial role in this analysis is played by the verification of the Lipschitz continuity of the elastic energy functional with respect to the unknown parameters (see Theorem A1 and Corollary A1). Some results of this paper were communicated by Barabasz *et al.* (2011a). We also attach a benchmark example consisting in solving a thermo-elastic IPM associated with the step-and-flash imprint lithography process, which illustrates the advantages coming from the adaptive genetic search and the control of accuracy in the case of ill-posed, inverse problems with the multimodal functional being the composition of elastic energy and boundary displacement misfits.

2. Forward and inverse problems

2.1. Definition of a forward problem. The class of IPMs to be solved is related to the following class of forward problems.

Let an open, bounded, and connected subset $\Omega \subset \mathbb{R}^d, d = 1, 2, 3$ coincide with the undeformed configuration of a linear elastic body. Consider the Sobolev space

$$V = \{\, v \in H^1(\Omega; \mathbb{R}^d) \mid \mathbf{tr}(v) = 0 \text{ on } \Gamma_D \,\},$$

where $\Gamma_D \subseteq \partial\Omega$ with $\mathrm{meas}(\Gamma_D) > 0$ corresponding to the mixed (Dirichlet and Neumann) boundary conditions. Assuming small deformations, the weak equilibrium (variational residual) equation takes the following form: Find $u \in V$ such that

$$a\,(g; u, v) = l\,(v)\,, \quad \forall v \in V, \tag{1}$$

where $g \in L^\infty(\Omega; \mathbb{K}_\gamma)$ is the stiffness tensor parameter. Moreover, $a(g; \cdot, \cdot) : V \times V \to \mathbb{R}$ is bilinear, symmetric, continuous and coercive form with respect to the second and third variables uniformly for all $g \in L^\infty(\Omega; \mathbb{K}_\gamma)$. The right-hand side mapping $l : V \to \mathbb{R}$ is linear and continuous.

The space $L^\infty(\Omega; \mathbb{K}_\gamma), \gamma > 0$ gathers all symmetric bounded tensor functions of fourth order,

$$\mathbb{K}_\gamma = \{\, k \in \mathbb{M}_d \mid k\,\tau \colon \tau \geq \gamma\, \|\tau\|^2_{\mathbb{S}^d} \;\; \forall \tau \in \mathbb{S}^d \,\},$$

and $\sigma \colon \tau$ is the product of tensors given by $\sigma_{ij}\,\tau_{ij}$ for σ, $\tau \in \mathbb{S}^d$. Moreover, for a constant $0 < M_k < +\infty$, we define

$$\mathbb{S}^d \supset \mathbb{M}_d = \{\, k = (k_{ijhl}) \mid k_{ijhl} = k_{jihl} = k_{hlij}, \\ k_{ijhl} \leq M_k,\, i, j, h, l = 1, \ldots, d \,\},$$

where $\mathbb{S}^d$ is the space of symmetric $d \times d$ matrices.

The above conditions are sufficient for the well-posedness of (1) for each $g \in L^\infty(\Omega; \mathbb{K}_\gamma)$ and enable us to rewrite (1) in the equivalent form

$$\arg\min_{u \in V} \; \mathcal{E}\,(g; u)\,, \tag{2}$$

where

$$\mathcal{E}(g; u) = \frac{1}{2}\, a(g; u, u) - l(u) \tag{3}$$

stands for the total energy of the modelled physical system; see, e.g., the work of Ciarlet (1978, Chapter 1), and Appendix with references herein for an explanation of necessary mathematical details.

2.2. Self *hp*–adaptive finite element method. The finite element method (FEM) consists of constructing a subspace $V_{h,p} \subset V$ with a finite basis $\{e^i_{h,p}\}_{i=1,\ldots,N_{h,p}}$, $N_{h,p} < +\infty$. The subspace $V_{h,p}$ is constructed by partitioning the solution domain Ω into a finite number of nonintersecting polyhedra (called elements) and defining basis functions as polynomials satisfying prescribed conditions over their vertices, edges, faces, and interiors. Each non-zero restriction of the basis function $e^i_{h,p}$ to the element is called the *shape function*.

Usually, we restrict the representation of the stiffness tensors (forward problem parameters) to some admissible set $\mathcal{D} \subset L^\infty(\Omega; \mathbb{K}_\gamma)$ for both forward and inverse numerical computations. Typically, $\mathcal{D}$ is a set of stepwise, uniformly bounded functions on the predefined, regular decomposition of the solution domain Ω.

Assuming a value of the parameter $g \in \mathcal{D}$, an approximate solution $u_{h,p}$ to (1) is obtained by using a linear combination of the basis functions

$$u_{h,p} = \sum_{i=1}^{N_{h,p}} u^i_{h,p}\, e^i_{h,p}. \tag{4}$$

Using (4) and a similar representation for $v \in V_{h,p} \subset V$, we can obtain for (1) the system of linear equations

$$\sum_{i=1}^{N_{h,p}} u^i_{h,p}\, a\big(g; e^i_{h,p}, e^j_{h,p}\big) = l\big(e^j_{h,p}\big), \\ j = 1, \ldots, N_{h,p}. \tag{5}$$

The coefficients $u^i_{h,p}$ are called the *degrees of freedom*. A detailed mathematical description of the FEM applied to elliptic variational problems may be found in the work of Ciarlet (1978, Chapter 2), while the convergence of this method for various FEM spaces is discussed by the same author (Ciarlet, 1978, Chapters 3 and 4) or by Descloux (1973).

The accuracy of the approximation depends on the quality of the basis functions. The self-adaptive hp finite element method is an algorithm for automatic construction of the basis function delivering exponential convergence of the accuracy with respect to the mesh size. The algorithm has been formulated by Rachowicz *et al.* (2006), Demkowicz (2006) and Demkowicz *et al.* (2007). It can be summarized in the following steps:

1. Generate an initial basis function $\{e^i_{h,p}\}_{i=1,\ldots,N_{h,p}}$ span over the so-called *initial mesh*. The initial mesh becomes the so-called *coarse mesh* for the first iteration.
2. Solve the coarse mesh problem by computing the degrees of freedom $\{u^i_{h,p}\}_{i=1,\ldots,N_{h,p}}$. After that step, we obtain the coarse mesh approximate solution $u_{h,p} = \sum_{i=1}^{N_{h,p}} u^i_{h,p}\, e^i_{h,p}$.
3. Generate the fine basis $\{e^i_{\frac{h}{2},p+1}\}_{i=1,\ldots,N_{\frac{h}{2},p+1}}$ spanned over the so-called *fine mesh*. The fine mesh is obtained from the coarse mesh by breaking each coarse mesh element into eight elements and incrementing the polynomial order of approximation uniformly by one.
4. Solve the fine mesh problem by computing the degrees of freedom $\{u^i_{\frac{h}{2},p+1}\}_{i=1,\ldots,N_{\frac{h}{2},p+1}}$. After that step, we obtain the fine mesh approximate solution
$$u_{\frac{h}{2},p+1} = \sum_{i=1}^{N_{\frac{h}{2},p+1}} u^i_{\frac{h}{2},p+1}\, e^i_{\frac{h}{2},p+1}.$$
5. Select an optimal refinement strategy for each finite element from the coarse mesh. That should be based on the error estimations $\|e_{\text{rel}}\|^2$, where $e_{\text{rel}} = u_{\frac{h}{2},p+1} - u_{h,p}$, computed by using the coarse and fine mesh solutions (see Demkowicz, 2006). The norm $\|\cdot\|$ is equivalent to the energetic norm on the space V (cf. Section 5 for details). The optimal refinements contain a list of h refinements (requests to break some elements into either two, four, or eight new elements) or p refinements (raising some polynomial orders of approximations by one).
6. Execute all required h refinements.
7. Execute all required p refinements.
8. If the maximum relative error of the solution is greater than the required accuracy, then go to Step 2. The new optimal mesh becomes the coarse mesh for the next iteration.

It has been proven that an appropriate selection of element sizes h and polynomial orders of shape functions p obtained using the above automatic hp-FEM results in exponential convergence of the numerical error with respect to the number of basis functions (degrees of freedom) (Babuška, 1986a; 1986b; Schwab, 19998). No matter how the convergence of hp-FEM holds by $h \to 0$ and $p \to +\infty$, both parameters are dependent. They should satisfy the syntactic rules of element mesh refinement, taking the approximation appropriate for its topology. In the computational practices p rarely exceeds 10, while h is bounded from below by the arithmetic error of the particular implementation.

The self-adaptive hp-FEM algorithm has been also redesigned by using the graph grammar approach (Ryszka *et al.*, 2013; Paszyński, 2009a; Strug *et al.*, 2013; Paszyńska *et al.*, 2008; 2012a; 2012b; 2009) and efficiently implemented on parallel machines (Paszyński, 2009b; Paszyński and Schaefer, 2010; Paszyński and Demkowicz, 2006).

2.3. Rate of *hp*-FEM convergence. In particular, the exponential convergence of the self-adaptive goal oriented hp-FEM is experimentally confirmed as the straight line $y = -ax + b$ in the system of coordinates, where the horizontal axis represents the cube root of the number of degrees of freedom $x = N^{1/3}$ and the vertical axis represents the logarithm of the relative error $y = \log_{10}(\|e_{\text{rel}}\|)$, $\|e_{\text{rel}}\| < 1$, where $e_{\text{rel}} = u_{\frac{h}{2},p+1} - u_{h,p}$ denotes the relative error being the difference between two consecutive approximate solutions, $\|\cdot\|$ denotes the proper norm in the space problem solutions. The constants a and b are positive and problem dependent. This implies the relation $\log_{10}(\|e_{\text{rel}}\|) = -a(N^{1/3}) + b$, which in turn entails $N = -c_1(\log_{10}(c_2\,\|e_{\text{rel}}\|))^3$ for $\|e_{\text{rel}}\| < 1$, where the constants are problem specific $c_1 = a^{-3}, c_2 = 10^{-b} > 0$. The computational cost of the solution of the problem by using a forward solver over the two dimensional mesh depends on the structure of the hp refined mesh. For a regular mesh the cost is of order $O\left(N^{3/2}\right)$. For meshes with pointwise singularities the cost can be reduced down to linear $O\left(N\right)$. Finally,

$$\text{cost} = O\left(-c_1(\log_{10}(c_2\,\|e_{\text{rel}}\|))^{3r}\right), \quad \|e_{\text{rel}}\| < 1, \quad (6)$$

where $r \in [1, 3/2]$, and now $c_1 = a^{-3r}, c_2 = 10^{-b} > 0$.

2.4. Definition of a class of inverse problems. Let us denote hereafter by $u(g) \in V$ and $u_{h,p}(g) \in V_{h,p}$ the unique solutions to (1) and its FEM approximation in the

space $V_{h,p}$ obtained for the particular parameter $g \in \mathcal{D}$, respectively.

We assume that we can observe the total strain energy $\widehat{\mathcal{E}}$ and the boundary displacement $\widehat{\gamma}$ of the deformable elastic body Ω.

The inverse problem under consideration consists in finding (approximating) an unknown physical parameter $\widehat{s} \in L^\infty(\Omega; \mathbb{K}_\gamma)$ resulting in the value of the strain energy $\mathcal{E}(\widehat{s}; u(\widehat{s})) = \widehat{\mathcal{E}}$ (see the formula (3)) and the boundary displacement $\mathbf{tr}(u(\widehat{s})) = \widehat{\gamma}$, where $u(\widehat{s}) \in V$ stands for the exact solution to (2) for the parameter $\widehat{s}$ and $\mathbf{tr}\colon H^1(\Omega; \mathbb{R}^d) \to L^2(\partial\Omega; \mathbb{R}^d)$ is the trace operator (see, e.g., Denkowski *et al.*, 2003b). The exact value of energy $\widehat{\mathcal{E}}$ can be measured in situ or in a laboratory or computed by highly accurate molecular calculations, while the observed value of $\widehat{\gamma}$ can be obtained by laser scanning of the body surface $\partial\Omega$.

We will formulate the misfit functional $J\colon \mathcal{D} \to \mathbb{R}_+$ such as

$$J(g) = \left| \mathcal{E}(g; u(g)) - \widehat{\mathcal{E}} \right| + \omega \, \| \mathbf{tr}(u(g)) - \widehat{\gamma}) \|^2_{L^2(\partial\Omega; \mathbb{R}^d)}, \tag{7}$$

where $\omega \geq 0$ is a scaling factor. Note that the internal strain energy (see (3)) is a square form of the forward solution, so both misfit components are of the same degree.

Since, in general, it is impossible to compute the exact value of (7) for each $\widehat{s} \in L^\infty(\Omega; \mathbb{K}_\gamma)$ effectively (it would be available only if $\widehat{s} \in \mathcal{D}$), we intend to find an approximation $\widehat{g} \in \mathcal{D}$ of the exact parameter $\widehat{s} \in L^\infty(\Omega; \mathbb{K}_\gamma)$ so that for all $g \in \mathcal{D}$ we have

$$\lim_{h \to 0,\, p \to +\infty} J_{h,p}(\widehat{g}) \leq \lim_{h \to 0,\, p \to +\infty} J_{h,p}(g), \tag{8}$$

where

$$J_{h,p}(g) = \left| \mathcal{E}(g; u_{h,p}(g)) - \widehat{\mathcal{E}} \right| + \omega \, \| \mathbf{tr}(u_{h,p}(g)) - \widehat{\gamma}) \|^2_{L^2(\partial\Omega; \mathbb{R}^d)} \tag{9}$$

and $\mathcal{E}(g; u_{h,p}(g))$ stands for the energy of the approximate solution $u_{h,p}(g)$ obtained for the parameter $g \in \mathcal{D}$.

The inverse problem (8) may be classified as a global optimization one because of the frequent ambiguity of its solution (multiple solutions), manifested as objective multimodality caused by its physical nature and/or the uncertainty of the mathematical model, as well as errors in numerical objective evaluation. As a result, not only the global minimizers, but also the local ones with sufficiently low objective values may represent physically meaningful solutions. The other difficulty is often a high computational cost of the approximated evaluation of the misfit (9).

3. *hp*-HGS overview

3.1. Primitives. The HGS produces a tree-structured dynamically changing set of dependent demes (see Fig. 1). The depth of the HGS tree is bounded by $m < +\infty$. The low-order demes (those closer to the root) perform a more chaotic search with lower accuracy, while the high-order demes perform a more accurate, local search.

Each deme, except leaf demes, sprouts a new child deme after a constant number K of genetic epochs called the *metaepoch*. A new child deme is activated in a promising region of the evolutionary landscape surrounding the best fitted individual, distinguished from the parental deme at the end of the metaepoch.

The HGS also implements two mechanisms that allow reduction in search redundancy. The first one, called *conditional sprouting*, disables the sprouting of new demes in regions already occupied or explored by the brother-deme (another child-deme of the same order sprouted by the same parent). The second mechanism, called *branch reduction*, reduces the branches of the same order that perform searches in the common landscape region or in the regions already explored.

The HGS stopping policy is composed of a *local branch stopping condition* that terminates the evolution in leaves and branches and a *global stopping condition* that evaluates the total maturity of the global search. Local stopping conditions monitor the evolution progress in a deme and stop it if it is unsatisfactory. The whole strategy might be stopped if no new demes are sprouted after a sufficiently large number of metaepochs and all active leaves were stopped. The other possibility is to stop the strategy when the satisfactory number of well fitted individuals was already found.

We apply an HGS in order to solve the inverse problem (8). The fitness function for the particular deme should be based on the misfit functional (9). We utilized two type of encodings, obtaining two separate types of strategies.

In the case of the binary encoding, the evolution of each deme is governed by a separate instance of the simple genetic algorithm (SGA) (see Vose, 1999). The search accuracy is dictated by encoding precision and by the length of binary strings used as the genotypes, which are different for different-order demes. The unique deme of the first order (the root) utilizes the shortest genotypes, while the leaf utilizes the longest ones. To obtain search coherency for demes of different orders, a special kind of hierarchical and nested encoding is used. First, the densest mesh of phenotypes in $\mathcal{D}$ for the demes of the m-th order is defined. Next, the meshes for the lower order demes are recursively defined by selecting some nodes from the previous ones. The maximum diameter of mesh δ_j associated with the demes of order j determines the search accuracy at this level of the HGS tree. Obviously,

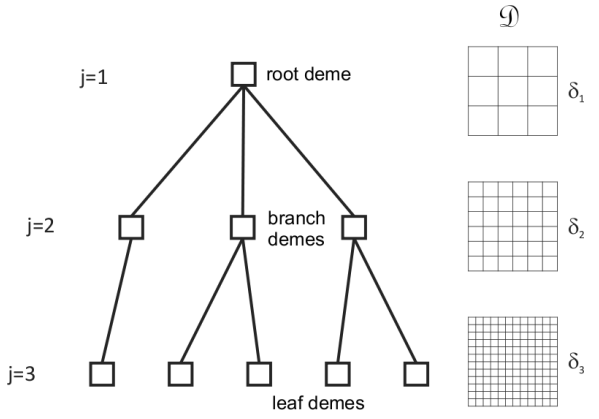

Fig. 1. HGS tree and corresponding two-dimensional meshes.

the mesh parameters satisfy $\delta_m < \cdots < \delta_1$.

In the real-number encoding version of the HGS, a genotype is a vector of floating point numbers. In order to introduce a sequence of increasing genetic spaces for subsequent orders of branches, we use a sequence of scaling coefficients $+\infty > \eta_1 \geq \eta_2 \geq \cdots \geq \eta_m = 1$. Let us denote by $\mathcal{D} = \prod_{i=1}^{N}[a_i, b_i] \subset \mathbb{R}^N$ a search domain, where $a_i, b_i;\ a_i < b_i$ are the lower and upper bounds for the i-th decision variable. The genetic space at the i-th level is defined as $\prod_{i=1}^{N}[0, \frac{b_i - a_i}{\eta_i}] \subset \mathbb{R}^N$. In this way, we obtain genetic spaces that are smaller for lower level branches, closer to the root. The genetic space for leaves $\prod_{i=1}^{N}[0, (b_i - a_i)]$ is of the same size as the admissible domain $\mathcal{D}$, and has the richest numerical representation. If a target search accuracy in leaves equals δ_m, the accuracy in the underlying demes will be reduced to $\delta_j = \eta_j\, \delta_m$, for $j = 1, \ldots, m-1$.

3.2. Dynamic accuracy adjustment. Let us assume that g represents the parameter value decoded from the genotype that appears in the HGS deme of the j-th order, $j \in \{1, \ldots, m\}$. Lemma A6 (see Appendix A.4) delivers the following misfit evaluation:

$$
\begin{aligned}
&J_{\frac{h}{2},p+1}(g) \\
&= \left|\mathcal{E}(g; u_{\frac{h}{2},p+1}(g)) - \widehat{\mathcal{E}}\right| \\
&\quad + \omega \left\| \mathbf{tr}(u_{\frac{h}{2},p+1}(g)) - \widehat{\gamma}) \right\|^2_{L^2(\partial\Omega;\mathbb{R}^3)} \qquad (10) \\
&\leq A_1 \left\| u_{h,p}(g) - u_{\frac{h}{2},p+1}(g) \right\|^2_{H^1(\Omega;\mathbb{R}^d)} \\
&\quad + A_1 \left\| u_{h,p}(g) - u(g) \right\|^2_{H^1(\Omega;\mathbb{R}^d)} \\
&\quad + A_2 \left\| g - \widehat{s} \right\|_{L^\infty(\Omega;\mathbb{K}_\gamma)},
\end{aligned}
$$

where $u(g)$ is the solution to the problem (1), $\widehat{\mathcal{E}} = \mathcal{E}(\widehat{s}; u(\widehat{s}))$ are the real displacement and energy, respectively, and $A_1, A_2 > 0$ are two positive constants.

The first right-hand side component of (10) contains the square of the error decrement $\|e_{\text{rel}}(g)\| = \|u_{\frac{h}{2},p+1}(g) - u_{h,p}(g)\|$ in the single hp-FEM step (see Demkowicz, 2006), while the last component represents the error of the inverse problem solution that characterizes the individuals belonging to the HGS demes of the j-th order.

The main idea of the hp-HGS is to adjust dynamically the accuracy of the misfit evaluation to the particular value of the parameter g encoded in an individual's genotype, as well as to the unavoidable error of solving the inverse problem that characterizes the current HGS branch. This may be obtained by balancing the first and third components of the right-hand side of the formula (10).

Assuming δ_j to be the accuracy of the inverse problem solved by the branch of the j-th order, we will perform the hp-adaptation of the FEM solution of the forward problem while $\|e_{\text{rel}}\|$ is greater than $Ratio\sqrt{\delta_j}$. The value of the strategy parameter $Ratio$ should correspond to the constant $\sqrt{A_2/A_1}$, in view of the rapid diminishing of the middle right-hand side component of (10) as a consequence of the fast hp-FEM convergence.

Corollary A2 verifies the same error scaling rule for the case of the SFIL inverse problem discussed in Section 4.

Algorithm 1. Pseudo-code of the hp-HGS.

```
Initialize the root deme P_root;
AD = {P_root};
while global_stop_condition() = false do
    AW = AD;
    for P ∈ AW do
        Execute in parallel Metaepoch(P);
    end for;
end while;
STOP;
```

3.3. Algorithmic description. A draft of hp-HGS activity is depicted in the pseudo-codes Algorithms 1 and 2. The function $global\_stop\_condition()$ checks whether a satisfactory solution has been found or if the hp-HGS cannot find any more local extremes. The function $branch\_stop\_condition(P)$ returns $true$ if it detects a lack of evolution progress of the current deme P. The generic function $fitness(i)$ computes fitness accordingly to the position of P in the hp-HGS tree.

The *conditional sprouting* mechanism is implemented as follows. The procedure $children\_comparison(x)$ compares the phenotype averages (centroids) of all child demes with the phenotype of the best fitted individual x distinguished from the parental deme P. This procedure returns $true$ if x is sufficiently close to the centroids of the existing child

demes. The generic function $sprout(x, P)$ returns a new child deme surrounding x using proper encoding and sampling, according to the position of the parental deme P in the hp-HGS tree.

Algorithm 2. Pseudo-code of the $Metaepoch$ function.

```
1:  t = 0;
2:  repeat
3:    for (i ∈ P) do
4:      solve the forward problem for g = code(i) on
        the coarse and fine FEM meshes;
5:      compute ‖e_rel(g)‖;
6:      while (‖e_rel(g)‖ > Ratio × sqrt(δ_j)) do
7:        execute one step of hp adaptivity;
8:        solve the problem on the new coarse and fine
          FEM meshes;
9:        compute ‖e_rel(g)‖;
10:     end while;
11:     compute fitness(i) using the FEM mesh finally
        established;
12:   end for
13:   if (P ≠ P_root) then
14:     if (branch_stop_condition(P)) then
15:       AD = AD \ {P};
16:     end if
17:   end if
18:   if ((t = K) ∧ (P is not a leaf)) then
19:     distinguish the best fitted individual x from deme
        P;
20:     if (¬ children_comparison(x)) then
21:       AD = AD ∪ {sprout(x, P)};
22:     end if
23:   end if
24:   perform proportional selection, obtaining a
      multiset of parents;
25:   perform SGA genetic operations on the multiset of
      parents;
26:   establish a new contents of population P after the
      genetic epoch;
27:   t = t + 1;
28: until (t = K)
```

Statements 15 and 21 in Algorithm 2 are mutually exclusive among all instances of $Metaepoch(P)$ function processing in parallel, because the set of active demes AD constitutes common, shared data. A particular implementation-based mechanism of critical section handling is applied. The modifications of the set of alive demes AD imposed by the particular deme P (see Statements 15 and 21 in the $Metaepoch$ routine) do not affect changes performed by other demes, because of their tree structure (see Fig. 1). The *branch reduction* mechanism is omitted in Algorithms 1 and 2 for the sake of simplicity.

The presented general algorithmic description constitutes a basis for various implementations. The serial, trivial one forces to execute the loop 5–7 in Algorithm 1 sequentially.

4. Case study

As a case study, we take step-and-flash imprint lithography (SFIL) being a patterning process utilizing photopolymerization to replicate the topography of a template onto a substrate. During the SFIL process, the shape of the template is transferred into the feature by utilizing the photopolymerization process of the liquid polymer (cf. Paszyński *et al.*, 2010a). Because of different chemical compounds of the liquid polymer, we assume that the resulting feature consists of two layers with different mechanical properties, and the resulting material may be described as the 3D linear elastic body loaded by the thermal stresses (cf. Paszyński *et al.*, 2010a). The inverse problem consists of an identification of all feasible Young modulus distributions minimizing the misfit functional (7) for the two layers of the material, while the Poisson ratio is assumed to be known.

Paszyński *et al.* (2010a) deliver also a reliable value of energy $\widehat{\mathcal{E}}$ obtained by the exhaustive, highly accurate simulation with a nanoscale molecular statics model with different values of interparticle interactions for two layers of the material (see Fig. 2).

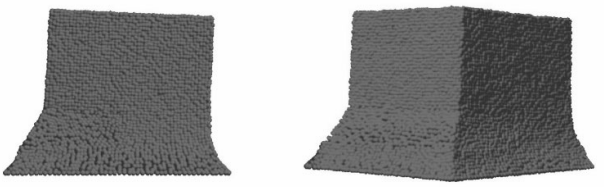

Fig. 2. Results of the nonlinear nano-scale molecular statics model allowing large deformations with quadratic potentials.

4.1. Step-and-flash imprint lithography. Step and flash imprint lithography (SFIL) is a patterning process utilizing photopolymerization used to replicate the topography of a template onto a substrate (cf. Ahopelto and Haatainen, 2002; Colburn *et al.*, 2001; Burns *et al.*, 2004; Paszyński *et al.*, 2005). Nanolithography methods like SFIL are utilized for the production of microprocessors. The SFIL process can be described in the following six steps, as illustrated in Fig. 3.

1. *Dispense*. The SFIL process employs a template/substrate alignment scheme to bring

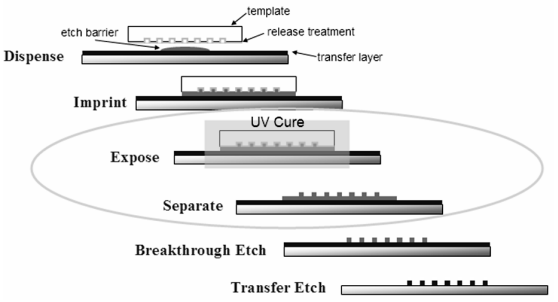

Fig. 3. SFIL process.

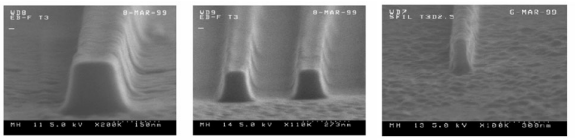

Fig. 4. Example shrinkage of the feature measured after removing the template.

a rigid template and substrate into parallelism, trapping the etch barrier in the relief structure of the template.

2. *Imprint*. The gap is closed until the force that ensures a thin base layer is reached.

3. *Exposure*. The template is then illuminated through the backside to cure the etch barrier.

4. *Separate*. The template is withdrawn, leaving low-aspect ratio, high resolution features in the etch barrier.

5. *Breakthrough etch*. The residual etch barrier (base layer) is etched away with a short halogen plasma etch.

6. *Transfer etch*. The pattern is transferred into the transfer layer with an anisotropic oxygen reactive ion etch, creating high-aspect ratio, high resolution features in the organic transfer layer.

The steps mentioned above include various kinds of physical processing: depositing a low viscosity, silicon containing, photocurable etch barrier on to a substrate; bringing the template into contact with the etch barrier; curing the etch barrier solution through UV exposure; releasing the template, while leaving high-resolution features behind; a short, halogen break-through etch; and finally an anisotropic oxygen reactive ion etch to yield high aspect ratio, high resolution features. Photopolymerization, however, is often accompanied by densification. Densification of the SFIL photopolymer (the etch barrier) may affect both the cross sectional shape of the feature and the placement of relief patterns. Example shrinkage of the feature measured after removing the template is presented in Fig. 4. The linear elasticity model with a thermal expansion coefficient is used to verify the material response of polymerized networks in cured etch-barrier layers that are formed during the exposure and after removal of the template.

We focus on the simulation of the deformation of the feature after removal of the template. It is assumed that the polymer network has been damaged during the removal of the template, and thus the interparticle forces are weaker in one region. The problem has been solved in a 3D cube shape domain, presented in Fig. 5.

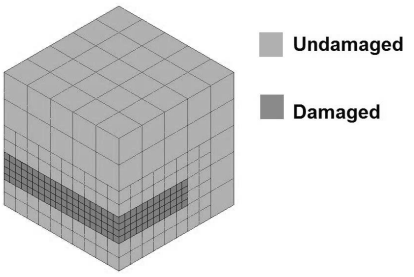

Fig. 5. SFIL problem domain.

4.2. Linear elasticity model with a thermal expansion coefficient. The problem of linear elasticity with a thermal expansion coefficient can be formulated as in the work of Hughes (2000). We use the notation introduced in Section 5. The elastic domain Ω is an open unit cube in $\mathbb{R}^3$. The part of the boundary $\Gamma_D = \{x \in \Gamma \mid x_1 = 0\}$ constitutes the bottom face of Ω and $\Gamma_N = \overline{\Gamma \setminus \Gamma_D}$. Given $\theta \in \mathbb{R}$, $\alpha_{ij} \in \mathbb{R}$ and $f_i \equiv 0$ for $i, j = 1, 2, 3$, the problem under consideration reads as follows: Find the displacement vector field $u \in C^2(\Omega; \mathbb{R}^3) \cap C^1(\bar{\Omega}; \mathbb{R}^3)$ satisfying Eqn. (A1), the boundary conditions

$$\begin{aligned} u_i &= 0 \ \text{ on } \ \Gamma_D, \\ \sigma_{ij} n_j &= 0 \ \text{ on } \ \Gamma_N, \end{aligned} \tag{11}$$

and the constitutive law

$$\sigma_{ij} = k_{ijhl} \left(\varepsilon_{hl}(u) + \theta \alpha_{hl} \right), \quad i, j, h, l = 1, 2, 3, \tag{12}$$

where k_{ijhl} are the elasticity coefficients (known for a given material), θ is the temperature, and α_{hl} are the thermal expansion coefficients.

The solutions of the weak formulation for the SFIL model will be sought in the space $V = \{v \in H^1(\Omega; \mathbb{R}^3) \mid \mathbf{tr}(v) = 0 \, \text{on} \, \Gamma_D\}$. For convenient implementation, we use the notation equivalent to the one introduced in Appendix 5, in which the strain tensor is rearranged into

the following form:

$$\bar{\varepsilon}(u) = \left(\frac{\partial u_1}{\partial x_1}, \frac{\partial u_2}{\partial x_2}, \frac{\partial u_3}{\partial x_3}, \frac{\partial u_2}{\partial x_3} + \frac{\partial u_3}{\partial x_2}, \frac{\partial u_1}{\partial x_3} + \frac{\partial u_3}{\partial x_1}, \frac{\partial u_1}{\partial x_2} + \frac{\partial u_2}{\partial x_1} \right)^\top, \tag{13}$$

and the elasticity coefficients are represented by the second order tensor field of the form

$$\mathrm{D} = \frac{E}{(1+\nu)(1-2\nu)} \times \begin{pmatrix} 1-\nu & \nu & \nu & 0 & 0 & 0 \\ \nu & 1-\nu & \nu & 0 & 0 & 0 \\ \nu & \nu & 1-\nu & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1-2\nu}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1-2\nu}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1-2\nu}{2} \end{pmatrix}. \tag{14}$$

The relation between the coefficients D_{ij} and k_{ijhl} may be found in the work of Hughes (2000). We are looking for $u \in V$ satisfying

$$a(\mathrm{D}; u, w) = l(\mathrm{D}; w) \quad \text{for all } w \in V, \tag{15}$$

where

$$a(\mathrm{D}; u, w) = \int_\Omega \bar{\varepsilon}(w)^\top \, \mathrm{D} \, \bar{\varepsilon}(u) \, \mathrm{d}x$$
$$l(\mathrm{D}; w) = -\theta \int_\Omega \bar{\varepsilon}(w)^\top \, \mathrm{D} \, \alpha \, \mathrm{d}x. \tag{16}$$

In the following numerical simulation, we assume the Poisson ratio $\nu = 0.3$, as provided by Colburn (1978). We also assume $\theta = 1$ (the thermal expansion coefficient expresses the volumetric contraction of the feature when the temperature gradient is equal to 1 °C). Finally, based on the results presented by Paszyński *et al.* (2007), we set $\alpha_{kl} = -0.0615\,\delta_{kl}$. The distribution of the Young modulus E is subject to the inverse analysis.

In the following experiments, there are nine parameters, see Fig. 6, $g \in \mathcal{D} \subset \mathbb{R}^9$. The first three values of the Young modulus $E1, E2, E3$ correspond to the undamaged polymer, while the next six values $E4$ to $E9$ of the Young modulus correspond to the damaged material, on one half of the feature, from 20% to 40% of its height. The top and bottom layers of the polymer are assumed to have a constant value of the Young modulus equal to $E_{\text{top}} = E_{\text{bottom}} = 10^9$. The SFIL problem with damaged material has also been solved by a molecular statics model with weaker interparticle forces assumed in the area of the domain where the damaged material is present (Paszyński, 2005; 2010a). The reference energy $\widehat{\mathcal{E}}$ has been computed based on molecular statics. Displacements obtained in this way are shown in Fig. 2.

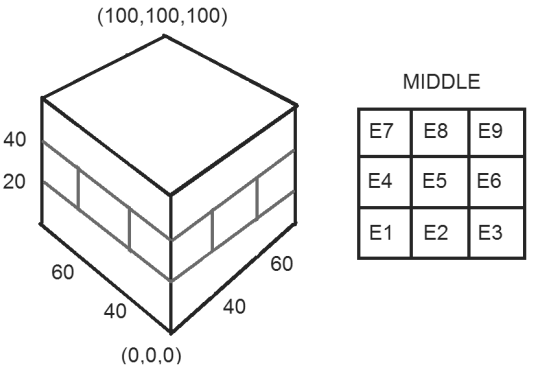

Fig. 6. Location of nine Young modulus being subject to inverse analysis.

4.3. Experiments. Through a computational example we illustrate the advantages of the presented strategy hp-HGS/L-BFGS-B. In particular, we selected a problem with important and practical application and a high computational difficulty. Its difficulty is mainly caused by the irregularity and the shape of the landscape which may result in many solutions, and thus mislocalization of the solution by less accurate methods.

In our experiments we utilized the hybrid strategy for the nine-parameter case, where the hp-HGS algorithm with real-number encoding was used for the location of candidate starting points for a local gradient optimization method. Then, as a second phase, we executed a limited-memory bound-constrained version of the Broyden–Fletcher–Goldfarb–Shanno algorithm (Zhu *et al.*, 1997). It is a quasi-Newton method utilizing a limited-memory approximation of the Hessian matrix.

The hp-HGS instance utilized in the a first phase has two levels. The relative energy error accuracy of the self-adaptive hp-FEM algorithm on particular levels (see Section 2.3) was set to 70 and 30, respectively. We used the real-number encoding of the Young modulus (genotypes) in HGS populations (see Section 3). The parameters of the hp-HGS (see Section 3) are summarized in Table 1. Scaling coefficients η_j shows accuracy degradation at the consecutive HGS levels: the root level and the leaf level. Population sizes on both the levels are small to reduce the number of forward solver calls. The crossing rate is the probability of selecting a genetic individual for reproduction. Analogously, the mutation rate is the probability of selecting an individual for mutation. During the mutation a new (mutated) individual is sampled according to the normal distribution centered in the parent and with the given standard deviation. The sprout distance and the sprout standard deviation parameters are used in procedures $children\_comparison(x)$ and $sprout(x, P)$ of

Table 1. Parameters of the hp-HGS tree.

	Root	Leaves
Scaling coefficients η_j	128.0	1.0
Population size	12	6
Crossing rate	0.5	0.5
Mutation rate	0.1	0.01
Mutation standard deviation	1.0	0.2
Sprout distance	0.5	
Sprout standard deviation	0.1	
Solver accuracy	70	30

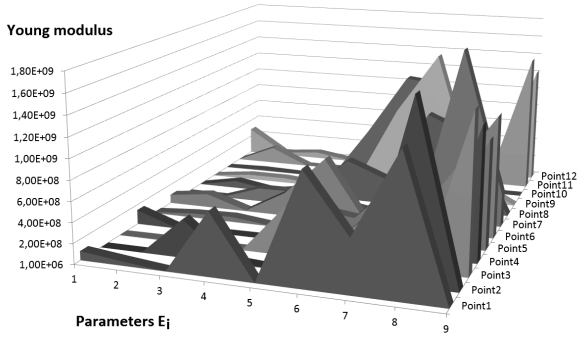

Fig. 7. Individuals found by the first phase (the genetic algorithm). All presented points were taken from the second level of the tree of populations and calculated with accuracy 30.

Algorithm 2.

After several metaepochs we collected individuals from the leaves of the tree of populations. These individuals are presented in Fig. 7. The self-adaptive hp-FEM accuracy for all of these individuals was better than 30. Clearly, all the results have a larger Young modulus in the $E7$-$E9$ area, and the lowest Young modulus in the $E1$-$E6$ area, as expected. The fitness of these 12 individuals is presented in Table 4.3. The absolute fitness values are so high because the parameters of the fitness formula (7) are expressed in nano-scale physical units.

In the second phase, the L-BFGS-B local method was started from the twelve points obtained from the first phase, and converged on five of them. The self-adaptive solver relative accuracy was set to 10 percent, e.g.,

Table 2. Fitness of the individuals found by the first phase.

Point	Fitness	Point	Fitness	Point	Fitness
1	79335	2	82995	3	89526
4	75204	5	34031	6	96797
7	79220	8	89871	9	74177
10	72895	11	48410	12	24984

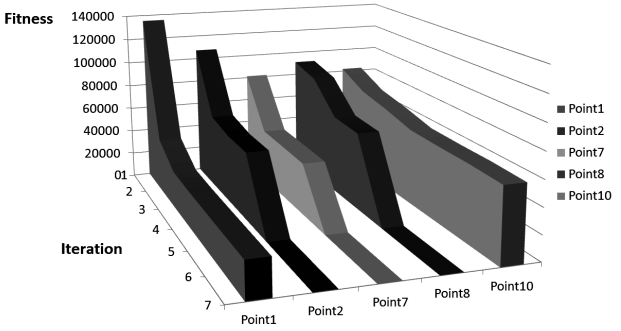

Fig. 8. Convergence of the fitness in the local phase for five individuals found by the first phase.

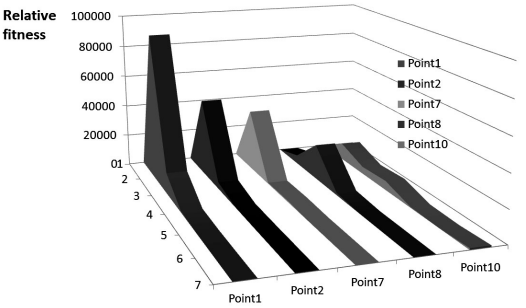

Fig. 9. Relative fitness during convergence in the local phase for five individuals found by the first phase.

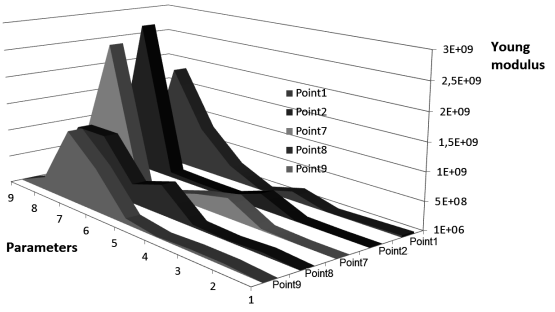

Fig. 10. Results of the second phase.

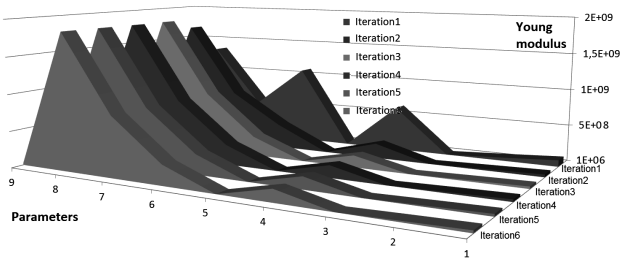

Fig. 11. Second phase convergence history for the first individual found in the first phase.

$100\% * \|u_{\frac{h}{2},p+1} - u_{h,p}\|^2 / \|u_{\frac{h}{2},p+1}\|^2 = 100\% \times \|e_{\text{rel}}\|^2 / \|u_{\frac{h}{2},p+1}\|^2 = 10\%$. This is the percentage of the relative error between the coarse and fine mesh solutions.

The summary of the L-BFGS-B convergence results is presented in Fig. 8, where we plot the convergence of the fitness function as well as in Fig. 9, in which we show the evolution of the relative fitness (the difference of the fitness value between two particular individuals from two consecutive iterations). The plot reflects the fact that we reached the minimum value of the fitness function that can be provided by the hp-FEM algorithm with relative accuracy of 10%. In that sense all five results converged to the lowest fitness value possible to obtain with such accuracy of the hp-FEM solver.

Let us now focus on the analysis of the values of the parameters obtained after the second phase. All five individuals that converged in the second phase are presented in Fig. 10. Their final fitness values are presented in Table 4.3. This fitness represents an integral computed at nano-scale on the boundary of the cubic domain from the difference between the known reference displacement field and the field coming from the numerical experiment.

We may observe that the displacement fields for two sets of parameters computed with the same hp-FEM relative accuracy 10% looks similar (see Figs. 12 and 13). Thus such values of the fitness of the order of 10^4 guarantee small differences between the obtained results and the reference solution. To illustrate the converge process, in Fig. 11 we present exemplary iterations for the first point found at the global stage. We can clearly see that there are three parameters representing the undamaged material and six parameters representing the damaged material. The differences in distribution of small and big values of the Young modulus obtained by these five results illustrate that the problem has many solutions.

4.3.1. Computational cost. We conclude the experimental section with the discussion on the computational cost of the proposed algorithm. The cost of the first phase is related to the number of calls to the self-adaptive hp-FEM solver, as well as to the requested accuracy. The experimental measurements of the computational time for six example individuals are presented in Fig. 14.

Table 3. Fitness of the individuals found by the first phase.

Point	Fitness	Point	Fitness	Point	Fitness
1	28587	2	54730	7	42163
8	61891	9	57880		

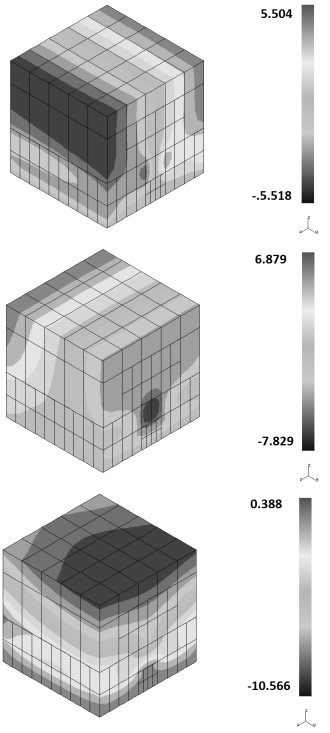

Fig. 12. x, y, z components of the displacement vector field for the first individual.

The hp-HGS algorithm is parallelizable on the level of metaepochs, which means that the total execution time can be reduced to the number of required metaepochs multiplied by the execution time needed to obtain the solution with assumed accuracy. For example, for the accuracy of 30 percent of the relative objective error, the time is of the order of 100 seconds, and the number of metaepochs is around 8. This implies 8×100 seconds which is about 14 minutes. The execution time can be further reduced by using the parallel self-adaptive hp-FEM code (Paszyński *et al.*, 2006; Paszyński and Demkowicz, 2006), delivering 60% efficiency over 16 cores. In other words, the computational time can be further reduced to $14/(16 \times 60\%) = 14/9.6 = 1.5$ minutes.

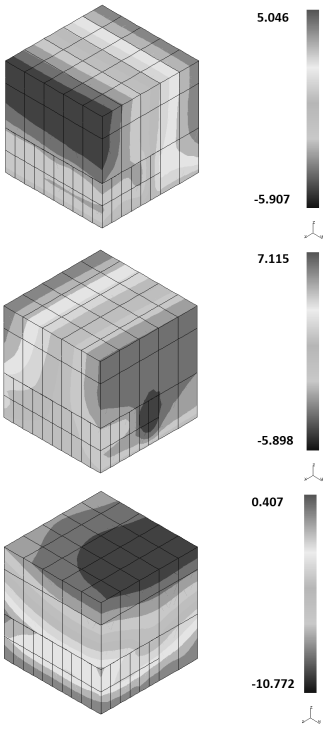

Fig. 13. The x, y, z components of the displacement vector field for the second individual.

Table 4. Number of calls to the self-adaptive hp-FEM solver.

Point	Iterations	Solver calls
1	12	1221
2	4	286
7	4	110
8	4	77
9	2	234

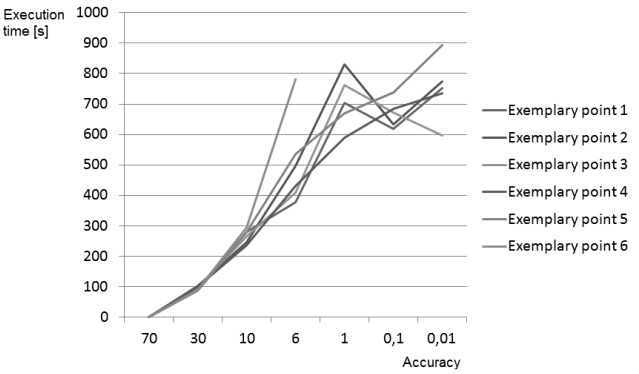

Fig. 14. Computational times for six example points for different accuracies of the self-adaptive hp-FEM solver.

The computational cost of the second phase is different since we use a sequential gradient algorithm working for each starting point separately. We can parallelize the computations on the level of particular starting points, but the solver calls for a given point for each iteration of the local algorithms must be performed in sequential mode.

The number of self-adaptive solver calls is higher than that of iterations, since the L-BFGS-B algorithm utilizes the finite difference method for approximation of the Jacobian in the inverse space. The number of iterations of the L-BFGS-B algorithm and the number of calls for particular individuals in the local phase is summarized in Table 4.3.1. The particular points required between 77–286 solver calls (150 in average), except the expensive first point, which we iterated for 12 iterations in L-BFGS-B without having any significant improvement of the accuracy after the 4th step. Thus, on the average, we can estimate the computational time of the second phase by 150×250 seconds (where 250 is the average time for the solver call with 10 percent accuracy), which is around 10.5 hours. This time can be further reduced by using parallel implementation to $10.5/(16 \times 60\%) = 10.5/9.6 = 1$ hours.

It should be emphasized that we actually performed sequential computations and the above estimates are for the theoretical parallel execution case. The sequential computations took about $8 \times 20 \times 200 = 32000\,\mathrm{s} = 8.8$ h in the first phase (since we had 8 metaepochs with around 20 evaluations each), and $5 \times 150 \times 200 = 150000 = 41.6$ h in the second phase (taking the average number of solver calls equal to 150 and 5 local points being investigated, not including the first point, which was iterated much longer).

5. Conclusions

The proposed strategy can recognize multiple objective minimizers. This allows studying ill-posed inverse problems, which is difficult or even impossible to obtain by other methods. In the case of a finite number of minimizers, all of them can be found. If the set of minimizers is infinite or even continuous, then its volume and shape can be characterized and approximated by the finite number of hp-HGS results.

The hp-HGS minimizes the computational cost in several ways. The number of forward problem solver invocations in the global phase decreases due to the hierarchic structure of demes. The main computational cost reduction is obtained by common scaling of the forward and inverse search accuracy (see Section 3.2). Additional computational cost reduction is obtained by restricting the number of local searches. Only one local search is started in each basin of attraction recognized in the global phase.

The applicability of the hp-HGS is restricted to problems with Lipschitz continuous misfit functionals with respect to the unknown parameters. The current hp-HGS implementation combined with the local search method can be applied to problems with a moderate number of design variables (moderate dimension). The massively parallel implementation will extend the applicability of the method to more complicated problems with a larger number of inverse parameters. A further significant extension might be obtained by economical computing of the misfit gradient and Hessian (Alvarez-Aramberri *et al.*, 2013; Oden and Prudhomme, 2001).

The simple computational example consists in restoring feasible Young modulus distributions in the 3D domain of a photo polymer template obtained during step and flash imprint lithography satisfying the minimum condition (8). The difficulty of this problem is mainly caused by the irregularity and shape of the landscape, which may result in mislocalization of the solution by less accurate methods. The presented approach allows a low-cost landscape analysis, hence a more accurate and expensive search in the most promising regions is available (see Table 2, Fig. 7 and Table 3, Fig. 10).

We showed that the problem of restoring the Young modulus based on the energy and boundary displacement measurements has many solutions, and thus the application of the classical gradient based method may produce misleading results.

The hp-HGS algorithm is parallelizable on the level of metaepochs, and for the accuracy used in the numerical experiments for 8 utilized metaepochs the computational time can be reduced down to 14 minutes. Using the parallel version of the self-adaptive hp-FEM (Paszyński *et al.*, 2006; Paszyński and Demkowicz, 2006), this time can be further reduced down to 1.5 minutes.

The reduction of the computational cost can be obtained by switching to the local gradient search, e.g., by means of the L-BFGS-B algorithm in the areas found in the global phase. In the hybrid case, the second phase executed per individual requires around 10.5 hours. Using the parallel version of the self-adaptive hp-FEM, this time can be also further reduced down to 1 hour per individual.

The hp-HGS can be extended to other cases of global optimization problems in continuous domains in which the computational cost of the objective evaluation depends monotonically on its accuracy.

Acknowledgment

The work of the second author was supported by the Polish National Science Center, grant no. DEC-2012/05/N/ST6/03433. The work of the fourth author was supported by the Polish Ministry of Science and Higher Education, grant no. NN 519 447739. The work of the fifth author was supported by the Polish National Science Center, grant no. DEC-2011/03/B/ST6/01393. The work of the sixth author was supported by the Polish National Science Center, grant no. DEC-2011/01/B/ST6/00674. This research of the third author was supported by the Marie Curie International Research Staff Exchange Scheme Fellowship within the 7th European Community Framework Programme under the grant agreement no. 295118, the National Science Center of Poland, grant no. N N201 604640, an international project co-financed by the Ministry of Science and Higher Education of Republic of Poland under the grant no. W111/7.PR/2012, and the National Science Center of Poland under a Maestro advanced project, DEC-2012/06/A/ST1/00262.

References

Ahopelto J. and Haatainen T. (2002). Step and flash imprint lithography, *in* C.M. Sotomayor Torres (Ed.), *Alternative Lithography. Unleashing the Potentials of Nanotechnology*, Kluwer Academic Publisher, Boston, MA/Dortrecht/London, Chapter 6, pp. 127–142.

Alvarez-Aramberri, J., Pardo, D. and Barucq, H. (2013). Inversion of magnetotelluric measurements using multigoal oriented hp-adaptivity, *Procedia Computer Science* **18**(8): 1564–1573.

Babuška, I. and Guo, B. (1986a). The hp-version of the finite element method, Part I: The basic approximation results, *Computational Mechanics* **1**(1): 21–41.

Babuška, I. and Guo, B. (1986b). The hp-version of the finite element method, Part II: General results and applications, *Computational Mechanics* **1**(3): 203–220.

Banks, H.T. and Kunisch, K. (1989). *Estimation Techniques for Distributed Parameter Systems*, Birkhäuser, Boston, MA.

Barabasz, B., Gajda, E., Migórski, S., Paszyński, M. and Schaefer, R. (2011a). Studying inverse problems in elasticity by hierarchic genetic search, *IPM'2011 Conference Proceedings, Sieniawa, Poland*, pp. 9–10.

Barabasz, B., Migórski, S., Schaefer, R. and Paszyński, M. (2011b). Multi-deme, twin adaptive strategy hp-HGS, *Inverse Problems in Science and Engineering* **19**(1): 3–16.

Barabasz, B., Schaefer, R. and Paszyński, M. (2009). Handling ambiguous inverse problems by the adaptive genetic strategy hp-HGS, *in* G. Allen, J. Nabrzyski, E. Seidel, G.D. van Albada, J. Dongarra and P.M.A. Sloot (Eds.) *Computational Science—ICCS 2009, 9th International Conference, Baton Rouge, LA, USA, May 25–27, 2009, Proceedings, Part II*, Lecture Notes in Computer Sience, Vol. 5545, Springer-Verlag, Berlin/Heidelberg, pp. 904–913.

Burczyński, T. and Beluch, W. (2001). The identification of cracks using boundary elements and evolutionary algorithms, *Engineering Analysis with Boundary Elements* **25**(4–5): 313–322.

Burczyński, T., Kuś, W., Długosz, A. and Orantek, P. (2004). Optimization and defect identification using distributed

evolutionary algorithms, *Engineering Applications of Artificial Intelligence* **17**(4): 337–344.

Burns, R., Johnson, S., Schmid, G., Kim, E., Dickey, D., Meiring, J., Burns, S., Stacey, N., Willson, C., Convey, D., Wei, Y., Fejes, P., Gehoski, K., Mancini, D. Nordquist, K., Dauksher, W.J and Resnick, D.J. (2004). Mesoscale modeling for SFIL simulating polymerization kinetics and densification, *Proceedings of SPIE 2004, Santa Clara, CA, USA*, Vol. 5374, pp. 348–360.

Cabib, E., Davini, C. and Chong-Quing, R. (1990). A problem in the optimal design of networks under transverse loading, *Quarterly of Applied Mathematics* **48**(2): 251–263.

Cabib, E., Schaefer, R. and Telega, H. (1998). A parallel genetic clustering for inverse problems, *in* B. Kagström, J. Dongarra, E. Elmroth and J. Wasniewski (Eds.), *Applied Parallel Computing. Large Scale Scientific and Industrial Problems. 4th International Workshop, PARA'98, Umea, Sweden, June 14–17, Proceedings*, Lecture Notes in Computer Science, Vol. 1541, Springer, Berlin/Heidelberg, pp. 551–556.

Caicedo, J.M. and Yun, G. (2011). A novel evolutionary algorithm for identifying multiple alternative solutions in model updating, *Structural Health Monitoring* **10**(5): 491–501.

Cantú-Paz, E. (2000). *Efficient and Accurate Parallel Genetic Algorithms*, Kluwer Academic Publishers, Norwell, MA.

Chase Geoffrey, J., Barroso, Luciana, R. and Hwank, K.-S. (2004). LMS-based structural health monitoring methods for the ASCE benchmark problem, *Proceedings of the 2004 American Control Conference, Boston, MA, USA,* Vol. 5, pp. 4201–4206.

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

Colburn, M. (1978). *Step and Flash Imprint Lithography: A Low Pressure, Room Temperature Nonoimprint Lithography*, Ph.D. thesis, University of Texas, Austin, TX.

Colburn, M., Suez, I., Choi, B., Meissi, M., Bailey, T., Sreeni-vasan, S., Ekerdt, J. and Willson, C. (2001). Characterization and modeling of volumetric and mechanical properties for SFIL photopoly-mers, *Journal of Vacuum Science and Technology* **19**(6): 2685–2689.

Demkowicz, L. (2006). *Computing with hp-Adaptive Finite Elements, Vol. I: One and Two Dimensional Elliptic and Maxwell Problems*, Chapman and Hall/CRC Applied Mathematics and Nonlinear Science, London.

Demkowicz, L., Kurtz, J., Pardo, P., Paszyński, M., Rachowicz, W. and Zdunek, A. (2007). *Computing with hp-Adaptive Finite Elements, Vol. II: Frontiers: Three-Dimensional Elliptic and Maxwell Problems with Applications*, Chapman and Hall/CRC Applied Mathematics and Nonlinear Science, London.

Denkowski, Z., Migórski, S. and Papageorgiou, N. (2003a). *An Introduction to Nonlinear Analysis: Applications*, Kluwer Academic/Plenum, New York, NY.

Denkowski, Z., Migórski, S. and Papageorgiou, N. (2003b). *An Introduction to Nonlinear Analysis: Theory*, Kluwer Academic/Plenum, New York, NY.

Descloux, J. (1973). *Méthode Des Éléments Finis*, Ecole Polytechnique Fédérale de Lausanne, Lausanne.

Engl, H.W., Hanke, M. and Neubauer, A. (2000). *Regularization of Inverse Problems*, Kluwer, Dordrecht.

Figueiredo, E., Park, G., Farrar, C.R., Worden, K. and Figueiras, J. (2011). Machine learning algorithms for damage detection under operational and environmental variability, *Structural Health Monitoring* **10**(6): 559–572.

Friswell, M.I. and Mottershead, J.E. (2001). Inverse methods in structural health monitoring, *Key Engineering Materials* **204–205**: 201–210.

Garibaldi, L., Surace, C., Holford, K. and Ostachowicz, W.M. (1999). *Damage Assessment of Structures*, Trans Tech Publications, Zürich.

Glover, F. and Kochenberger, G. (2002). *Handbook of Metaheuristics*, Kluwer Academic Publishers, Dordrecht.

Horst, R. and Pardalos, P. (1995). *Handbook of Global Optimization*, Kluwer, Dordrecht.

Hughes, T. (2000). *The Finite Element Method. Linear Statics and Dynamic Finite Element Analysis*, Dover Publications, Mineola, NY.

Huhtala, A. and Sven, B. (2011). A Bayesian approach to vibration based structural health monitoring with experimental verification, *Rakenteiden Mekaniikka (Journal of Structural Mechanics)* **44**(4): 330–344.

Isakov, V. (2006). *Inverse Problems for Partial Differential Equations*, Springer, New York, NY.

Kirikera, Goutham, R., Shinde, V., Schulz, Mark, J., Ghoshal, A., Sundaresan, Mannur, J., Allemang, Randall, J. and Won Lee, J. (2008). A structural neural system for real-time health monitoring of composite materials, *Structural Health Monitoring* **7**(1): 65–83.

Koper, K., Wysession, M. and Wiens, D. (1999). Multimodal function optimization with a niching genetic algorithm: A seismological example, *Bulletin of the Seismological Society of America* **89**(4): 978–988.

Lashin, S. and Likoshvai, V. (2004). Evolutionary algorithms for mathematical models of gene regulatory networks, *Proceedings of the 4th International Conference on Bioinformatics of Genome Regulation and Structure, BGRS 2004, Novosibirsk, Russia,* Vol. 2, pp. 81–84.

Mahfoud, S. (1997). Niching methods, *in* T. Back, D.B. Fogel and Z. Michalewicz (Eds.), *Handbook of Evolutionary Computations*, Oxford University Press, Oxford, Chapter C.6.1, pp. C6.1:1–C6.1:4.

Meruane, V. and Heylen, W. (2009). Damage detection with parallel genetic algorithms and operational modes, *Structural Health Monitoring* **9**(6): 481–496.

Oden, J.T. and Prudhomme, S. (2001). Goal-oriented error estimation and adaptivity for the finite element method, *Computers and Mathematics with Applications* **41**(5): 735–756.

Osman, I. and Kelly, J. (1996). *Meta-Heuristics: Theory and Applications*, Kluwer Academic Publishers, Norwell, MA.

Osman, I.H. and Laporte, G. (1996). Metaheuristics: A bibliography, *Annals of Operations Research* **63**(5): 511–623.

Paszyńska, A., Grabska, E. and Paszyński, M. (2012a). A graph grammar model of the *hp* adaptive three dimensional finite element method, Part I, *Fundamenta Informaticae* **114**(2): 149–182.

Paszyńska, A., Grabska, E. and Paszyński, M. (2012b). A graph grammar model of the *hp* adaptive three dimensional finite element method, Part II, *Fundamenta Informaticae* **114**(2): 183–201.

Paszyńska, A., Paszyński, M. and Grabska, E. (2008). Graph transformations for modeling *hp*-adaptive finite element method with triangular elements, *in* M. Bubak, G.D. van Albada, J. Dongarra and P.M.A. Sloot (Eds.) *Computational Science—ICCS 2008, 8th International Conference, Kraków, Poland, June 23–25, Proceedings, Part III*, Lecture Notes in Computer Science, Vol. 5103, Springer, Berlin, pp. 604–613.

Paszyńska, A., Paszyński, M. and Grabska, E. (2009). Graph transformations for modeling *hp*-adaptive finite element method with mixed triangular and rectangular elements, *in* G. Allen, J. Nabrzyski, E. Seidel, G.D. van Albada, J. Dongarra and P.M.A. Sloot (Eds.), *Computational Science—ICCS 2009, 9th International Conference, Baton Rouge, LA, USA, Proceedings, Part II*, Lecture Notes in Computer Science, Vol. 5545, Springer, Berlin, pp. 875–884.

Paszyński, M. (2009a). On the parallelization of self-adaptive *hp*-finite element methods, Part I: Composite programmable graph grammar model, *Fundamenta Informaticae* **93**(4): 411–434.

Paszyński, M. (2009b). On the parallelization of self-adaptive *hp*-finite element methods, Part II: Partitioning communication agglomeration mapping (PCAM) analysis, *Fundamenta Informaticae* **93**(4): 435–457.

Paszyński, M., Barabasz, B. and Schaefer, R. (2007). Efficient adaptive strategy for solving inverse problems, *in* Y. Shi, G.D. van Albada, J. Dongarra and P.M.A. Sloot (Eds.), *Computational Science—ICCS 2007. 7th International Conference, Beijing China, May 27–30, 2007, Proceedings, Part I,* Lecture Notes in Computer Science, Vol. 4487, Springer, Berlin, pp. 342–349.

Paszyński, M. and Demkowicz, L. (2006). Parallel fully automatic *hp*-adaptive 3D finite element package, *Engineering with Computers* **22**(3–4): 255–276.

Paszyński, M., Kurtz, J. and Demkowicz, L. (2006). Parallel fully automatic *hp*-adaptive 2D finite element package, *Computer Methods in Applied Mechanics and Engineering* **195**(7–8): 711–741.

Paszyński, M., Gurgul, P., Sieniek, M. and Pardo, D. (2010a). Unified modeling language description of the object-oriented multi-scale adaptive finite element method for step-and-flash imprint lithography, *IOP Conference Series: Materials Science and Engineering* **10**(1): 012247.

Paszyński, M., Pardo, D. and Paszyńska, A. (2010b). Parallel multi-frontal solver for *p* adaptive finite element modeling of multi-physics computational problems, *Journal of Computational Science* **1**(1): 48–54.

Paszyński, M., Romkes, A., Collister, E., Meiring, J., Demkowicz, L. and Willson, C. (2005). On the modeling of step-and-flash imprint lithography using molecular statics models, *Technical Report 05-38*, ICES, Austin, TX.

Paszyński, M. and Schaefer, R. (2010). Graph grammar-driven parallel partial differential equation solver, *Concurrency and Computation: Practice and Experience* **22**(9): 1063–1097.

Rachowicz, W., Pardo, D. and Demkowicz, L. (2006). Fully automatic *hp*-adaptivity in three dimensions, *Computer Methods in Applied Mechanics and Engineering* **195**(37–40): 4816–4842.

Rocca, P., Benedetti, M., Donelli, M., Franceschini, D. and Massa, A. (2009). Evolutionary optimization as applied to inverse scattering problems, *Inverse Problems* **25**(12): 123003.

Ryszka, I., Paszyńska, A., Grabska, E. and Paszyński, M. (2013). Graph grammar systems for modeling three dimensional finite element method, *Fundamenta Informaticae*, (submitted).

Samarski, A.A. and Vabishchevich, P.N. (2007). *Numerical Methods for Solving Inverse Problems of Mathematical Physics*, Walter de Gruyter, Berlin.

Schaefer, R. and Barabasz, B. (2008). Asymptotic behavior of *hp*-HGS (*hp*-adaptive finite element method coupled with the hierarchic genetic strategy) by solving inverse problems, *in* M. Bubak, G.D. van Albada, J. Dongarra, and P.M.A. Sloot (Eds.), *Computational Science—ICCS 2008. 8th International Conference, Kraków, Poland, June 23–25, 2008, Proceedings, Part III,* Lecture Notes in Computer Science, Vol. 5103, Springer, Berlin/Heidelberg, pp. 682–692.

Schaefer, R. and Kołodziej, J. (2003). Genetic search reinforced by the population hierarchy, *in* K. DeJong, R. Poli and J. Rowe (Eds.), *Foundations of Genetic Algorithms 7*, Morgan Kaufman, Burlington, MA, pp. 383–399.

Schwab, C. (1998). *p and hp Finite Element Methods*, Oxford University Press, Oxford.

Singh, A., Minsker, B. and Takagi, H. (2006). Interactive genetic algorithms for inverse groundwater modeling: Issues with human fatigue and prediction models, *in* R. Walton (Ed.), *Proceedings of the 2005 World Water and Environmental Resources Congress: Impacts of Global Climate Change, May 15–19, 2005, Anchorage, AK*, Vol. 5, American Society of Civil Engineers, Reston, VA, pp. 3081–3092.

Strug, B., Paszyńska, A., Paszyński, M. and Grabska, E. (2013). Using a graph grammar system in the finite element method, *International Journal of Applied Mathematics and Computer Science* **23**(4): 839–853, DOI:10.2478/amcs-2013-0063.

Tanaka, M. (Ed.) (2003). *Inverse Problems in Engineering Mechanics IV. Proceedings of the International Symposium*

on Inverse Problems in Engineering Mechanics 2003 (ISP 2003), Nagano, Japan, Elsevier, Amsterdam.

Tarantola, A. (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*, SIAM, Philadelphia, PA.

Tikhonov, A.N., Goncharskii, A., Stepanov, V.V. and Yagola, A.G. (1995). *Numerical Methods for the Solution of Ill-Posed Problems*, Mathematics and Its Applications, Vol. 328, Springer-Verlag, Berlin/Heidelberg.

Vose, M. (1999). *The Simple Genetic Algorithm*, MIT Press, Boston, MA.

Wierzba, B., Semczuk, A., Kołodziej, J. and Schaefer, R. (2003). Hierarchical genetic strategy with real number encoding, *Proceedings of the 6th Conference on Evolutionary Algorithms and Global Optimization, Łagów Lubuski, Poland*, pp. 231–237.

Xavier, C., Vieira, V., Martins, D. and Dos Santos, R. (2006). Comparing two parallel genetic algorithms for the inverse problem associated to the cardiac bidomain equations, *Workshop on High Performance Computing in the Life Sciences, HPC LIFE, Ouru Preto, Brazil*.

Zhu, C., Byrd, R. H., Lu, P. and Nocedal, J. (1997). Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization, *ACM Transactions on Mathematical Software* **23**(4): 550–560.

Barbara Barabasz received the M.Sc. degree in 2003 from the Department of Mathematics, Physics and Computer Science, Jagiellonian University, Kraków, Poland. In 2003–2012 she was employed at the Faculty of Metals Engineering and Industrial Computer Science, AGH University of Science and Technology, Kraków. Her interests include solving inverse problems and genetic algorithms.

Ewa Gajda-Zagórska received her M.Sc. degree in computer science at Jagiellonian University, Kraków, Poland, in 2009. Since then, she has been a Ph.D. student of computer science at the AGH University of Science and Technology, Kraków. She is a recipient of the Google Anita Borg Scholarship (2013). Her research interests include evolutionary algorithms and multiobjective optimization.

Stanisław Migórski received his M.Sc. (1985), Ph.D. (1990) and D.Sc. (1997) degrees in mathematics from Jagiellonian University in Kraków, Poland. Currently he is a full professor and the chair of Mathematical Sciences at Jagiellonian University. He has published research work in the field of mathematical analysis and applications (control theory, differential equations, nonlinear functional analysis, calculus of variations and applications) and is the co-author of three monographs on nonlinear analysis by Kluwer/Plenum as well as on hemivariational inequalities and nonlinear inclusions by Springer.

Maciej Paszyński received his Ph.D. (2003) in mathematics with applications to computer science from Jagiellonian University, Kraków, Poland, and his habilitation (2010) in computer science from the AGH University of Science and Technology, Kraków. In 2003–2005 he worked as a postdoctoral fellow at the Institute for Computational Engineering and Sciences (ICES) at the University of Texas at Austin. In 2006 and 2007 he worked as a postdoctoral fellow at the Department of Petroleum and Geosystems Engineering, University of Texas at Austin. Since 2007 he has been a frequent visitor at the ICES, at the King Abdullah University of Science and Technology (Saudi Arabia), at the Basque Center for Applied Mathematics and at the Department of Mathematics, Universidad del Pais Vasco (Bilbao, Spain). Since 2013 he has been a vice-director of the Department of Computer Science, AGH University, and holds a university professor position there. His research interests include the parallel self-adaptive *hp* finite element method, parallel forward solvers, models of concurrency and isogeometric methods.

Robert Schaefer is a full professor at the Department of Computer Science, Faculty of Computer Science Electronics and Telecommunication, AGH University of Science and Technology. The author and a co-author of about 180 books, papers and conference contributions, in particular of the book *Foundation of Genetic Global Optimization* (Springer, 2007). The general chair of the *PPSN 2010* conference and a steering committee member of the *PPSN* series. A PC member and a co-chair of more than 100 scientific conferences in computational sciences and artificial intelligence. Recent research areas: algorithms of oil and gas surveying, genetic and memetic algorithms, computing multi-agent systems, adaptive algorithms of solving forward and inverse problems for PDEs. Former research: modeling of blood flow in arteries, modeling of nonlinear flow in porous media.

Maciej Smołka received his Ph.D. (2000) in the area of shape optimization at Jagiellonian University in Kraków. Until 2012 he worked at the Chair of Optimization and Control Theory at Jagiellonian University. Currently he is a member of the Adaptive Algorithms and Systems Group in the Department of Computer Science, AGH University of Science and Technology, Kraków. His current research interests include stochastic modeling of computational systems, application of evolutionary and hybrid algorithms in solving inverse problems and computational multi-agent systems.

Appendix

Verification of dynamic accuracy adjustment

A.1. Direct problem features. Let us consider a linear elastic body, which in its undeformed state occupies an open, bounded, and connected subset Ω of $\mathbb{R}^d$, $d \geq 1$. Within the framework of linear elasticity theory, assuming small deformations, the equilibrium equation takes the form

$$\frac{\partial \sigma_{ij}}{\partial x_j} + f_i = 0 \ \text{ in } \Omega, \tag{A1}$$

where $f = (f_i)$ is the volume force. The

strain–displacement relationship can be written as

$$\varepsilon_{ij}(u) = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right)$$

in Ω, where $u = (u_i)$ denotes the displacement. The constitutive equation is given by the Hooke law,

$$\sigma_{ij}(u) = k_{ijhl}\,\varepsilon_{hl}(u) \text{ in } \Omega, \tag{A2}$$

where $k = (k_{ijhl})$, $i, j, h, l = 1, \ldots, d$ is the elasticity tensor. Let $\mathbb{S}^d$ be the space of symmetric $d \times d$ matrices and

$$\begin{aligned}\mathbb{M}_d = \{\, k = (k_{ijhl}) \mid\ & k_{ijhl} = k_{jihl} = k_{hlij},\\ & k_{ijhl} \le M_k,\ i,j,h,l = 1,\ldots,d\,\},\end{aligned} \tag{A3}$$

with entries uniformly bounded by a constant $0 < M_k < +\infty$.

Given $\gamma > 0$, we introduce the space $L^\infty(\Omega;\mathbb{K}_\gamma)$ of admissible fourth order tensors $k = k(x)$ with

$$\mathbb{K}_\gamma = \{\, k \in \mathbb{M}_d \mid k\,\tau\colon \tau \ge \gamma\,\|\tau\|^2_{\mathbb{S}^d} \text{ for all } \tau \in \mathbb{S}^d\,\}, \tag{A4}$$

where the product of tensors is given by $\sigma\colon \tau = \sigma_{ij}\,\tau_{ij}$ for $\sigma,\ \tau \in \mathbb{S}^d$. Consider the space $V = \{\, v \in H^1(\Omega;\mathbb{R}^d) \mid \mathbf{tr}(v) = 0 \text{ on } \Gamma_D\,\}$, where $\Gamma_D \subseteq \partial\Omega$ with $\mathrm{meas}(\Gamma_D) > 0$. This space corresponds to the mixed boundary conditions in the elasticity problem. Since $\mathrm{meas}(\Gamma_D) > 0$, the following Korn inequality holds:

$$\|v\|_{H^1(\Omega;\mathbb{R}^d)} \le c\,\|\varepsilon(v)\|_{L^2(\Omega;\mathbb{S}^d)}, \tag{A5}$$

where $c > 0$ depends only on Ω and Γ. This implies that the norm $\|\cdot\| = \|\varepsilon(\cdot)\|_{L^2(\Omega;\mathbb{S}^d)}$ is equivalent on V to the norm $\|\cdot\|_{H^1(\Omega;\mathbb{R}^d)}$. For future convenience, let us denote the norm equivalence constants as $0 < \kappa_0 < \kappa_1 < +\infty$, so that

$$\kappa_0\|v\|_{H^1(\Omega;\mathbb{R}^d)} \le \|v\| \le \kappa_1\|v\|_{H^1(\Omega;\mathbb{R}^d)} \text{ for all } v \in V. \tag{A6}$$

The weak formulation of the problem is as follows: Find $u \in V$ such that $a(k;u,v) = l(v)$ for all $v \in V$, where $a\colon L^\infty(\Omega;\mathbb{K}_\gamma) \times V \times V \to \mathbb{R}$ and $l\colon V \to \mathbb{R}$ are given by

$$\begin{aligned}a(k;u,v) &= \int_\Omega k(x)\,\varepsilon(u)\colon \varepsilon(v)\,\mathrm{d}x\\ &\text{for } k \in L^\infty(\Omega;\mathbb{K}_\gamma),\ u,v \in V,\end{aligned} \tag{A7}$$

$$l(v) = \int_\Omega f \cdot v\,\mathrm{d}x \text{ for } v \in V, \tag{A8}$$

where $f \in V^*$ and $f \cdot v = f_i v_i$. Then it is clear that, for all $k \in L^\infty(\Omega;\mathbb{K}_\gamma)$, the form $a(k;\cdot,\cdot)$ is bilinear, continuous on $V \times V$ (i.e., $|a(k;u,v)| \le M\,\|u\|\,\|v\|$ for $u,\ v \in V$ with $M > 0$), coercive uniformly with respect to k (i.e., $a(k;v,v) \ge \gamma\|v\|^2$ for $v \in V$) and the linear functional l is continuous (i.e., $l \in V^*$). Therefore, by application of the Lax–Milgram lemma (cf. Denkowski *et al.*, 2003a, Chapter 2.7), for every $f \in V^*$ and $k \in L^\infty(\Omega;\mathbb{K}_\gamma)$ there exists $u = u(k) \in V$ a unique solution to the problem,

$$a(k;u,v) = l(v) \quad \text{for all } v \in V \tag{A9}$$

and

$$\|u\| \le \frac{\|f\|_{V^*}}{\gamma}. \tag{A10}$$

Moreover, since for all $k \in L^\infty(\Omega;\mathbb{K}_\gamma)$ the bilinear form $a(k;\cdot,\cdot)$ is symmetric on $V \times V$ (i.e. $a(k;u,v) = a(k;v,u)$ for $u,\ v \in V$) and coercive uniformly with respect to k, it is well known (cf. Denkowski *et al.*, 2003a, Chapter 2.7) that for all $k \in L^\infty(\Omega;\mathbb{K}_\gamma)$ the problem (A9) is equivalent to the following minimization one:

$$\mathcal{E}(k;u) \le \mathcal{E}(k;v) \quad \text{for all} \quad v \in V,$$

with the elasticity energy functional $\mathcal{E}\colon L^\infty(\Omega;\mathbb{K}_\gamma) \times V \to \mathbb{R}$ given by

$$\begin{aligned}\mathcal{E}(k;v) = \frac{1}{2}\,a(k;v,v) - l(v)\\ \text{for } k \in L^\infty(\Omega;\mathbb{K}_\gamma),\ v \in V.\end{aligned} \tag{A11}$$

Let us observe that because both constants M_k and γ in the formulas (A3) and (A4) defining the set of admissible parameter functions are universal (Descloux, 1973), we may select the constants γ and M uniform for the whole family of bilinear forms $\{a(k;\cdot,\cdot)\}$, $k \in L^\infty(\Omega;\mathbb{K}_\gamma)$ such that for all $k \in L^\infty(\Omega;\mathbb{K}_\gamma)$ we have

$$\gamma\,\|v\|^2 \le a(k;v,v) \le M\,\|v\|^2 \text{ for all } v \in V. \tag{A12}$$

A.2. Lipschitz continuity of the elastic energy functional.

Theorem A1. *Let the energy functional $\mathcal{E}$ be of the form* (A11) *with a and l defined by* (A7) *and* (A8)*, respectively. Then there exists a constant $L > 0$ such that for all k_1, $k_2 \in L^\infty(\Omega;\mathbb{K}_\gamma)$, we have*

$$\begin{aligned}|\,\mathcal{E}(k_1;u(k_1)) - \mathcal{E}(k_2;u(k_2))|\\ \le L\,\|k_1 - k_2\|_{L^\infty(\Omega;\mathbb{K}_\gamma)},\end{aligned}$$

where $u(k_i) \in V$ is the unique solution to (A9) *corresponding to k_i for $i = 1, 2$.*

Proof. Let $k_1, k_2 \in L^\infty(\Omega;\mathbb{K}_\gamma)$ and $u_1 = u(k_1)$, $u_2 = u(k_2)$ be the unique solutions to (A9) which correspond to k_1 and k_2, respectively.

Since $u_1,\ u_2 \in V$ are solutions to (A9), we have $a(k_1;u_1,v) = l(v)$ and $a(k_2;u_2,v) = l(v)$ for all $v \in V$. Hence $a(k_1;u_1 - u_2,v) = -a(k_1 - k_2;u_2,v)$ for all

$v \in V$. From the latter, using the coercivity of the form a, (A10) and the Hölder inequality, we obtain

$$\begin{aligned}\gamma \|u_1 - u_2\|^2 &\le |\, a(k_1 - k_2; u_2, u_1 - u_2)| \\ &\le \|k_1 - k_2\|_{L^\infty(\Omega;\mathbb{K}_\gamma)} \|u_2\| \, \|u_1 - u_2\| \\ &\le \frac{\|f\|_{V^*}}{\gamma} \|k_1 - k_2\|_{L^\infty(\Omega;\mathbb{K}_\gamma)} \|u_1 - u_2\|,\end{aligned}$$

which implies the following lemma.

Lemma A1. *The solution of the forward problem (A9) is Lipschitz continuous with respect to the parameter*

$$\|u_1 - u_2\| \le C_0 \, \|k_1 - k_2\|_{L^\infty(\Omega;\mathbb{K}_\gamma)}, \tag{A13}$$

where $C_0 = \gamma^{-2}\|f\|_{V^*}$ *and* $\|\cdot\| = \|\varepsilon(\cdot)\|_{L^2(\Omega;\mathbb{S}^d)}$.

Subsequently, using (A10), we have

$$\begin{aligned}&|a(k_1; u_1, u_1) - a(k_2; u_1, u_1)| \\ &\quad = |\int_\Omega (k_1 - k_2)\, \varepsilon(u_1) \colon \varepsilon(u_1)\, \mathrm{d}x| \\ &\quad \le \|k_1 - k_2\|_{L^\infty(\Omega;\mathbb{K}_\gamma)} \int_\Omega \|\varepsilon(u_1)\|^2_{\mathbb{S}^d}\, \mathrm{d}x \\ &\quad = \|k_1 - k_2\|_{L^\infty(\Omega;\mathbb{K}_\gamma)} \|u_1\|^2 \\ &\quad \le \frac{\|f\|^2_{V^*}}{\gamma^2} \|k_1 - k_2\|_{L^\infty(\Omega;\mathbb{K}_\gamma)}.\end{aligned} \tag{A14}$$

From (A13), the triangle inequality for the norm and the elementary inequality $\|\xi\|^2_{\mathbb{S}^d} - \|\eta\|^2_{\mathbb{S}^d} \le \|\xi - \eta\|^2_{\mathbb{S}^d}$ for all $\xi, \eta \in \mathbb{S}^d$, we deduce

$$\begin{aligned}&|a(k_2; u_1, u_1) - a(k_2; u_2, u_2)| \\ &\quad = |\int_\Omega k_2 \left(\|\varepsilon(u_1)\|^2_{\mathbb{S}^d} - \|\varepsilon(u_2)\|^2_{\mathbb{S}^d}\right) \mathrm{d}x| \\ &\quad \le \|k_2\|_{L^\infty(\Omega;\mathbb{K}_\gamma)} \int_\Omega \|\varepsilon(u_1 - u_2)\|^2_{\mathbb{S}^d}\, \mathrm{d}x \\ &\quad = \|k_2\|_{L^\infty(\Omega;\mathbb{K}_\gamma)} \|u_1 - u_2\|^2 \\ &\quad \le C^2 \, \|k_2\|_{L^\infty(\Omega;\mathbb{K}_\gamma)} \|k_1 - k_2\|^2_{L^\infty(\Omega;\mathbb{K}_\gamma)} \\ &\quad \le C^2 \, \|k_2\|_{L^\infty(\Omega;\mathbb{K}_\gamma)} \left(\|k_1\|_{L^\infty(\Omega;\mathbb{K}_\gamma)}\right. \\ &\qquad \left. + \|k_2\|_{L^\infty(\Omega;\mathbb{K}_\gamma)}\right) \|k_1 - k_2\|_{L^\infty(\Omega;\mathbb{K}_\gamma)} \\ &\quad \le 2\,\beta^2\, C^2 \, \|k_1 - k_2\|_{L^\infty(\Omega;\mathbb{K}_\gamma)},\end{aligned}$$

where

$$\beta = \max\{\|k\|_{L^\infty(\Omega;\mathbb{K}_\gamma)} \mid k \in L^\infty(\Omega;\mathbb{K}_\gamma)\}.$$

Similarly, from (A13), it follows that

$$\begin{aligned}|\, l(u_2 - u_1)| &\le \|l\|_{V^*} \, \|u_2 - u_1\| \\ &\le C \, \|f\|_{V^*} \, \|k_1 - k_2\|_{L^\infty(\Omega;\mathbb{K}_\gamma)}.\end{aligned} \tag{A15}$$

Finally, by the inequalities (A14) and (A15), we have

$$\begin{aligned}&|\, \mathcal{E}(k_1; u_1) - \mathcal{E}(k_2; u_2)| \\ &= \frac{1}{2}|a(k_1; u_1, u_1) - a(k_2; u_1, u_1)| \\ &\quad + \frac{1}{2}|a(k_2; u_1, u_1) - a(k_2; u_2, u_2)| + |l(u_2 - u_1)| \\ &\le \frac{\|f\|^2_{V^*}}{2\gamma^2} \|k_1 - k_2\|_{L^\infty(\Omega;\mathbb{K}_\gamma)} \\ &\quad + \beta^2 \, \frac{\|f\|^2_{V^*}}{\gamma^4} \|k_1 - k_2\|_{L^\infty(\Omega;\mathbb{K}_\gamma)} \\ &\quad + \frac{\|f\|^2_{V^*}}{\gamma^2} \|k_1 - k_2\|_{L^\infty(\Omega;\mathbb{K}_\gamma)} \\ &= \left(\frac{\|f\|_{V^*}}{\gamma}\right)^2 \left(\frac{3}{2} + \left(\frac{\beta}{\gamma}\right)^2\right) \|k_1 - k_2\|_{L^\infty(\Omega;\mathbb{K}_\gamma)}\end{aligned}$$

which completes the proof. ■

A.3. Lipschitz continuity of the energy functional for the SFIL model. Under the hypotheses that θ and ν are known positive constants, and α is a known positive diagonal matrix, both forms a and l can be written as follows:

$$\begin{aligned}a(E; u, w) &= \int_\Omega E\, \bar{\varepsilon}(w)^\top \, \mathrm{P}\, \bar{\varepsilon}(u)\, \mathrm{d}x \\ l(E; w) &= \int_\Omega E\, \bar{\varepsilon}(w)^\top \, (-\theta\, \mathrm{P}\, \alpha)\, \mathrm{d}x,\end{aligned} \tag{A16}$$

where P is a constant matrix

$$\mathrm{P} = \frac{1}{(1+\nu)(1-2\nu)} \times \begin{pmatrix} 1-\nu & \nu & \nu & 0 & 0 & 0 \\ \nu & 1-\nu & \nu & 0 & 0 & 0 \\ \nu & \nu & 1-\nu & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1-2\nu}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1-2\nu}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1-2\nu}{2} \end{pmatrix}. \tag{A17}$$

We are looking for a Young modulus E that belongs to the space $L^\infty(\Omega; [\underline{E}, \overline{E}])$, where $0 < \underline{E} < \overline{E} < +\infty$ are some prescribed positive constants.

Both forms given by (A16) inherit all features analogously to the proof in Section 5, i.e., $a(E; \cdot, \cdot)$ is bilinear and continuous on $V \times V$, and coercive with the constant $\gamma > 0$, and $l(E; \cdot)$ is continuous on V for all $E \in L^\infty(\Omega; [\underline{E}, \overline{E}])$. Therefore, by the Lax–Milgram lemma, there exists $u(E) \in V$ a unique solution to the problem

$$a(E; u, v) = l(E; v) \quad \text{for all } \ v \in V, \tag{A18}$$

for all $E \in L^\infty(\Omega; [\underline{E}, \overline{E}])$, where

$$V = \{v \in H^1(\Omega; \mathbb{R}^3) \mid \mathbf{tr}(v) = 0 \text{ on } \Gamma_D\}. \tag{A19}$$

Furthermore, we have the following uniform estimate: There is a constant $C_l > 0$ such that

$$\|u(E)\| \leq \frac{1}{\gamma} \max_{H \in L^\infty(\Omega;[\underline{E},\overline{E}])} \|l(H;\cdot)\|_{V^*} = C_l. \quad \text{(A20)}$$

The next important property of the functional l is its Lipschitz continuity with respect to the Young modulus, i.e. there is a constant $L_l > 0$ such that for all E_1, $E_2 \in L^\infty(\Omega;[\underline{E},\overline{E}])$ and all $v \in V$ we have

$$\begin{aligned} &|\, l(E_1;v) - l(E_2;v)| \\ &\quad = |\, l(E_1 - E_2;v)| \\ &\quad \leq L_l \, \|E_1 - E_2\|_{L^\infty(\Omega;[\underline{E},\overline{E}])} \, \|v\|. \end{aligned} \quad \text{(A21)}$$

From the arguments used in the proof of Theorem A1, we have the following result:

Corollary A1. *Let the energy functional for the SFIL model be of the form*

$$\mathcal{E}(E;v) = \frac{1}{2}\, a(E;v,v) - l(E;v) \quad \text{for } E \in L^\infty(\Omega;[\underline{E},\overline{E}]), \quad v \in V, \quad \text{(A22)}$$

where a and l are given by (A16), (A17) and V is defined by (A19). Then there exists a constant $L > 0$ such that for all E_1, $E_2 \in L^\infty(\Omega;[\underline{E},\overline{E}])$, we have

$$\begin{aligned} &|\, \mathcal{E}(E_1;u(E_1)) - \mathcal{E}(E_2;u(E_2))| \\ &\quad \leq L \, \|E_1 - E_2\|_{L^\infty(\Omega;[\underline{E},\overline{E}])}, \end{aligned} \quad \text{(A23)}$$

where $u(E_i) \in V$ is the unique solution of (A18) corresponding to E_i for $i = 1, 2$.

Proof. It is analogous to the proof of Theorem A1, and thus we indicate only the main differences with that result. First, we observe that for all E_1, $E_2 \in L^\infty(\Omega;[\underline{E},\overline{E}])$ and $v \in V$ we have $a(E_1;u_1-u_2,v) = -a(E_1-E_2;u_2,v) - l(E_2 - E_1;v)$. Using (A20) and (A21), we obtain

$$\begin{aligned} &\gamma \, \|u_1 - u_2\|^2 \\ &\leq a(E_1;u_1 - u_2, u_1 - u_2) \\ &\leq |\, a(E_1 - E_2;u_2,u_1-u_2)| \\ &\quad + |\, l(E_1 - E_2;u_1 - u_2)| \\ &\leq (C_l + L_l)\|E_1 - E_2\|_{L^\infty(\Omega;[\underline{E},\overline{E}])}\|u_1 - u_2\| \end{aligned}$$

so the following inequality, analogous to (A13) in the proof of Theorem A1, remains valid:

$$\|u_1 - u_2\| \leq \frac{C_l + L_l}{\gamma} \, \|E_1 - E_2\|_{L^\infty(\Omega;[\underline{E},\overline{E}])}. \quad \text{(A24)}$$

Next, from (A20) and (A24), we deduce

$$\begin{aligned} &|\, a(E_1;u_1,u_1) - a(E_2;u_1,u_1)| \\ &\quad \leq C_l^2 \, \|E_1 - E_2\|_{L^\infty(\Omega;[\underline{E},\overline{E}])}, \end{aligned} \quad \text{(A25)}$$

$$\begin{aligned} &|\, a(E_2;u_1,u_1) - a(E_2;u_2,u_2)| \\ &\quad \leq 2\,\beta^2 \left(\frac{C_l + L_l}{\gamma}\right)^2 \|E_1 - E_2\|_{L^\infty(\Omega;[\underline{E},\overline{E}])}, \end{aligned} \quad \text{(A26)}$$

respectively. Moreover, from (A24), it follows that

$$\begin{aligned} &|\, l(E_2;u_2 - u_1)| \\ &\quad = \Big|\int_\Omega E\,\bar{\varepsilon}(u_2 - u_1)^\top \, (-\theta\,\mathrm{P}\,\alpha)\,\mathrm{d}x\Big| \\ &\quad \leq \theta\,\overline{E}\,|\, r(u_2 - u_1)| \leq \theta\,\overline{E}\,\|r\|_{V^*}\|u_2 - u_1\| \\ &\quad \leq \theta\,\overline{E}\,\|r\|_{V^*}\left(\frac{C_l + L_l}{\gamma}\right) \|E_1 - E_2\|_{L^\infty(\Omega;[\underline{E},\overline{E}])}, \end{aligned} \quad \text{(A27)}$$

where $r(v) = \int_\Omega \bar{\varepsilon}(v)^\top \,\mathrm{P}\,\alpha\,\mathrm{d}x$. From the inequalities (A21), (A25)–(A27), we have

$$\begin{aligned} &|\, \mathcal{E}(E_1;u_1) - \mathcal{E}(E_2;u_2)| \\ &\leq \frac{1}{2}|\, a(E_1;u_1,u_1) - a(E_2;u_1,u_1)| \\ &\quad + \frac{1}{2}|\, a(E_2;u_1,u_1) - a(E_2;u_2,u_2)| \\ &\quad + |\, l(E_2;u_1 - u_2)| + |\, l(E_1 - E_2;u_1)| \\ &\leq \Big(\frac{C_l^2}{2} + \beta^2 \left(\frac{C_l + L_l}{\gamma}\right)^2 \\ &\quad + \theta\overline{E}\,\|r\|_{V^*}\left(\frac{C_l + L_l}{\gamma}\right) + C_l\,L_l\Big) \\ &\quad \times \|E_1 - E_2\|_{L^\infty(\Omega;[\underline{E},\overline{E}])}, \end{aligned}$$

which completes the proof of the corollary. ■

A.4. Misfit estimation. Let us denote by $u_{h,p}(k) \in V_{h,p}$ and $u_{\frac{h}{2},p+1}(k) \in V_{\frac{h}{2},p+1}$ two consecutive hp-FEM solutions obtained in the hp-adaptation process (see Section 2.1), where $k \in L^\infty(\Omega;\mathbb{K}_\gamma)$ is an admissible parameter function. Notice that hp-adaptation rule forces $V_{h,p} \subset V_{\frac{h}{2},p+1} \subset V$ (see, e.g., Demkowicz, 2006).

Lemma A2. *For all $k \in L^\infty(\Omega;\mathbb{K}_\gamma)$ we have*

$$\begin{aligned} &2\,|\mathcal{E}\,(k;u_{h,p}(k)) - \mathcal{E}\,(k;u(k))| \\ &\quad \leq\ M\,\|u_{h,p}(k) - u(k)\|^2, \end{aligned} \quad \text{(A28)}$$

where $u(k)$ is the solution to the exact problem (A9) and

$$\begin{aligned} &2\,\left|\mathcal{E}\,(k;u_{h,p}(k)) - \mathcal{E}\left(k;u_{\frac{h}{2},p+1}(k)\right)\right| \\ &\quad \leq\ M\,\|u_{h,p}(k) - u_{\frac{h}{2},p+1}(k)\|^2, \end{aligned} \quad \text{(A29)}$$

where M is the uniform continuity constant (see (A12)).

Proof. We start the proof with a simple lemma.

Lemma A3. (Descloux, 1973, Theorem I.1.1.) *Let U be a linear space over $\mathbb{R}$ and $b : U \times U \to \mathbb{R}$ a bilinear, continuous, symmetric and positively defined form, whereas $\phi : U \to \mathbb{R}$ is a linear, continuous functional. Moreover, let $F : U \to \mathbb{R}$ be the quadratic functional given by $F(v) = \frac{1}{2}b(v,v) - \phi(v)$ for $v \in U$.*

If $u \in U$ is a solution of the variational equation

$$b(u,v) = \phi(v) \ \text{ for all } \quad v \in U,$$

then

$$F(v) - F(u) = \frac{1}{2}\, b(v-u, v-u) \ \text{ for all } \quad v \in U. \tag{A30}$$

Taking into account that $V_{h,p} \subset V$, Lemma A3, together with (A12), allows rewriting the left-hand side of (A28) in the form

$$\begin{aligned} &2\,|\mathcal{E}\,(k; u_{h,p}(k)) - \mathcal{E}\,(k; u(k))| \\ &\quad = a(k; u_{h,p}(k) - u(k), u_{h,p}(k) - u(k)) \\ &\quad \le M \|u_{h,p}(k) - u(k)\|^2. \end{aligned}$$

Analogously, since $u_{\frac{h}{2},p+1}$ is a solution of (A9) on $V_{\frac{h}{2},p+1}$ and since $V_{h,p} \subset V_{\frac{h}{2},p+1}$, we have

$$\begin{aligned} &2\left|\mathcal{E}\left(k; u_{\frac{h}{2},p+1}(k)\right) - \mathcal{E}\,(k; u_{h,p}(k))\right| \\ &\quad = a(k; u_{\frac{h}{2},p+1}(k) - u_{h,p}(k), u_{\frac{h}{2},p+1}(k) - u_{h,p}(k)) \\ &\quad \le M \|u_{\frac{h}{2},p+1}(k) - u_{h,p}(k)\|^2, \end{aligned}$$

which concludes the proof of Lemma A2. ∎

Lemma A4. *For all $k, g \in L^\infty(\Omega; \mathbb{K}_\gamma)$ we have*

$$\begin{aligned} &\left|\mathcal{E}\left(g; u_{\frac{h}{2},p+1}(g)\right) - \mathcal{E}\,(k; u(k))\right| \\ &\quad \le M' \|u_{h,p}(g) - u_{\frac{h}{2},p+1}(g)\|^2_{H^1(\Omega;\mathbb{R}^d)} \\ &\qquad + M' \|u_{h,p}(g) - u(g)\|^2_{H^1(\Omega;\mathbb{R}^d)} \\ &\qquad + L \|g - k\|_{L^\infty(\Omega;\mathbb{K}_\gamma)}, \end{aligned} \tag{A31}$$

where $u(g)$ is the solution to the problem (A9) and L the Lipschitz continuity constant (see Theorem A1). Moreover, $M' = \frac{1}{2}M\kappa_1^2$, where M is the uniform continuity constant (see (A12)) and κ_1 norm equivalence constant (see (A6)).

Proof. The left-hand side of (A31) might be rewritten in the form

$$\begin{aligned} &|\,\mathcal{E}\left(g; u_{\frac{h}{2},p+1}(g)\right) - \mathcal{E}\,(g; u_{h,p}(g)) \\ &\qquad + \mathcal{E}\,(g; u_{h,p}(g)) - \mathcal{E}\,(g; u(g)) \\ &\qquad + \mathcal{E}\,(g; u(g)) - \mathcal{E}\,(k; u(k))\,|. \end{aligned} \tag{A32}$$

It is enough to apply Lemma A2 to the first and second difference, the norm equivalence condition in $H^1(\Omega; \mathbb{R}^d)$ to the second one and Theorem A1 to the last difference. So, using (A6) we obtain $M' = \frac{1}{2}M^2\kappa_1^2$. ∎

In the sequel we derive a similar evaluation for the boundary part of the misfit functional.

Lemma A5. *There are constants $C_1, C_2 > 0$ such that for all $k, g \in L^\infty(\Omega; \mathbb{K}_\gamma)$ we have*

$$\begin{aligned} &\|\mathbf{tr}(u_{\frac{h}{2},p+1}(g)) - \mathbf{tr}(u(k))\|^2_{L^2(\partial\Omega;\mathbb{R}^d)} \\ &\quad \le C_1 \|u_{h,p}(g) - u_{\frac{h}{2},p+1}(g)\|^2_{H^1(\Omega;\mathbb{R}^d)} \\ &\qquad + C_1 \|u_{h,p}(g) - u(g)\|^2_{H^1(\Omega;\mathbb{R}^d)} \\ &\qquad + C_2 \|g - k\|_{L^\infty(\Omega;\mathbb{K}_\gamma)}, \end{aligned} \tag{A33}$$

where $u(g)$ is the solution to the problem (A9).

Proof. Because the trace operator $\mathbf{tr}\colon L^2(\partial\Omega; \mathbb{R}^d) \to H^1(\Omega; \mathbb{R}^d)$ is linear and continuous, there exists $C_{\text{tr}} > 0$ such that

$$\begin{aligned} &\|\mathbf{tr}(u_{\frac{h}{2},p+1}(g)) - \mathbf{tr}(u(k))\|_{L^2(\partial\Omega;\mathbb{R}^d)} \\ &\quad \le C_{\text{tr}} \|u_{\frac{h}{2},p+1}(g) - u(k)\|_{H^1(\Omega;\mathbb{R}^d)}. \end{aligned}$$

Now, by the Lemma A1, we obtain

$$\begin{aligned} &\|\mathbf{tr}(u_{\frac{h}{2},p+1}(g)) - \mathbf{tr}(u(k))\|_{L^2(\partial\Omega;\mathbb{R}^d)} \\ &\le C_{\text{tr}} \|u_{h,p}(g) - u_{\frac{h}{2},p+1}(g)\|_{H^1(\Omega;\mathbb{R}^d)} \\ &\quad + C_{\text{tr}} \|u_{h,p}(g) - u(g)\|_{H^1(\Omega;\mathbb{R}^d)} \\ &\quad + C_0 \|g - k\|_{L^\infty(\Omega;\mathbb{K}_\gamma)}. \end{aligned} \tag{A34}$$

Applying twice the elementary inequality $(a+b)^2 \le 2a^2 + 2b^2$ to the right-hand side of the formula (A34), we obtain

$$\begin{aligned} &\|\mathbf{tr}(u_{\frac{h}{2},p+1}(g)) - \mathbf{tr}(u(k))\|^2_{L^2(\partial\Omega;\mathbb{R}^d)} \\ &\le 4C_{\text{tr}}^2 \|u_{h,p}(g) - u_{\frac{h}{2},p+1}(g)\|^2 \\ &\quad + 4C_{\text{tr}}^2 \|u_{h,p}(g) - u(g)\|^2_{H^1(\Omega;\mathbb{R}^d)} \\ &\quad + 2C_0^2 \|g - k\|^2_{L^\infty(\Omega;\mathbb{K}_\gamma)}. \end{aligned} \tag{A35}$$

Finally, it is enough to observe that there exists a constant $C' > 0$ such that $\|g-k\|^2_{L^\infty(\Omega;\mathbb{K}_\gamma)} \le C' \|g-k\|_{L^\infty(\Omega;\mathbb{K}_\gamma)}$ so we have $C_1 = 4C_{tr}^2$ and $C_2 = 2C'C_0^2$. ∎

Summing up the above, we may state the following inequality, which may be useful in the misfit functional evaluation (9).

Lemma A6. *For all $k, g \in L^\infty(\Omega; \mathbb{K}_\gamma)$ we have*

$$\begin{aligned} &\left|\mathcal{E}\left(g; u_{\frac{h}{2},p+1}(g)\right) - \mathcal{E}\,(k; u(k))\right| \\ &\quad + \omega \|\mathbf{tr}(u_{\frac{h}{2},p+1}(g)) - \mathbf{tr}(u(k))\|_{L^2(\partial\Omega;\mathbb{R}^d)} \\ &\le A_1 \|u_{h,p}(g) - u_{\frac{h}{2},p+1}(g)\|^2_{H^1(\Omega;\mathbb{R}^d)} \\ &\quad + A_1 \|u_{h,p}(g) - u(g)\|^2_{H^1(\Omega;\mathbb{R}^d)} \\ &\quad + A_2 \|g - k\|_{L^\infty(\Omega;\mathbb{K}_\gamma)}, \end{aligned} \tag{A36}$$

where $u(g)$ is the solution to the problem (A9), $\omega > 0$ an arbitrary positive scaling parameter and $A_1 = M' + \omega C_1$, $A_2 = L + \omega C_2$ are two positive constants (see Lemmas A4 and A5).

Using a similar approach, we may obtain the misfit evaluation for the case study associated with the forward SFIL problem (see Section 4). The energy functional in the SFIL problem (A22) satisfies the same conditions as (A11) (see Corollary A1) while the space $L^\infty(\Omega; [\underline{E}, \overline{E}]$ is a special case of the space $L^\infty(\Omega; \mathbb{K}_\gamma)$.

Corollary A2. *For all E_1, $E_2 \in L^\infty(\Omega; [\underline{E}, \overline{E}])$ it follows that*

$$\begin{aligned}
&\left|\mathcal{E}\left(E_1; u_{\frac{h}{2},p+1}(E_1)\right) - \mathcal{E}\left(E_2; u(E_2)\right)\right| \\
&\quad + \omega \left\|\mathbf{tr}(u_{\frac{h}{2},p+1}(E_1)) - \mathbf{tr}(u(E_2))\right\|_{L^2(\partial\Omega;\mathbb{R}^3)} \\
&\leq A_1 \left\|u_{h,p}(E_1) - u_{\frac{h}{2},p+1}(E_1)\right\|^2_{H^1(\Omega;\mathbb{R}^3)} \\
&\quad + A_1 \left\|u_{h,p}(E_1) - u(E_1)\right\|^2_{H^1(\Omega;\mathbb{R}^3)} \\
&\quad + A_2 \left\|E_1 - E_2\right\|_{L^\infty(\Omega;[\underline{E},\overline{E}])},
\end{aligned} \tag{A37}$$

where $u(E_i) \in V$ is the unique solution of (A18) corresponding to E_i for $i = 1, 2$ and A_1, A_2 are two positive constants.

Received: 21 August 2013
Revised: 29 March 2014
Re-revised: 29 May 2014