ELSEVIER

Comput. Methods Appl. Mech. Engrg. 190 (2001) 6577–6604

Computer methods in applied mechanics and engineering

www.elsevier.com/locate/cma

Design of multiphysics actuators using topology optimization – Part I: One-material structures

O. Sigmund

Department of Mechanical Engineering, Solid Mechanics, The Technical University of Denmark, Nils Koppels Allé 404, DK-2800 Lyngby, Denmark

Received 29 February 2000

Abstract

This is the first part of a two-paper description of the topology optimization method applied to the design of multiphysics actuators and electrothermomechanical systems in particular. The first paper is focussed on one-material structures, the second on two-material structures. The extensions of the topology optimization method in this first part include coupled and non-linear finite element analyses, constitutive modelling of elements with intermediate densities, adjoint sensitivity analyses and formulation of optimization problems with multiple constraints. The application in mind is the design of thermally and electrothermally driven micro actuators for use in microelectromechanical systems (MEMS). MEMS are microscopic mechanical systems coupled with electrical circuits. MEMS are fabricated using techniques known from the semi-conductor industry. Examples include design of thermal and electrothermal micro actuators with one, two and three degrees of freedom (d.o.f.). The differences between modelling and optimizing the actuators using linear and non-linear finite element analyses are discussed in detail. Ways for systematic interpretation and transfer of the topology optimized designs to the micromachining process are also discussed and finally, some test results prove the validity of the method.

Keywords: Topology optimization; Non-linear finite element analysis; Coupled problems; Optimization; Microelectromechanical systems (MEMS)

1. Introduction

Topology optimization methods have, up to now, mostly been used in solving linear, 'single physics' problems with single constraints, such as in stiffness optimization of mechanical structures with constraint on structural weight. The recent developments in theory, computational speed and large-scale optimization algorithms however, allow for extensions of the topology optimization method to problems involving multiphysics, multiple constraints, large numbers of elements and non-linear modelling. Compared to the vast amount of literature on general topology optimization (see [1] for an overview), examples of topology optimization methods applied to multiphysics problems are relatively few and are limited to thermally loaded structures [2], thermal actuators [3], material micro structures with extreme thermoelastic properties [4,5], piezo-electric material micro structures [6–9], three-dimensional micro structures with extremal stiffness and bounds on conductivity [10] and preliminary results for electrothermally actuated micro devices [11,12]. Preferably, actuators should undergo large deformations and therefore the usual assumptions on small displacements are invalid. Solving topology optimization problems involving geometrical non-linearities has so far only been done in a small number of papers for stiffness optimization [13–16] and for compliant mechanism design [16–19].

E-mail address: sigmund@mek.dtu.dk (O. Sigmund).

0045-7825/01/$ - see front matter
PII: S0045-7825(01)00251-1

The possible applications of topology optimization in multiphysics problems are numerous but one very interesting application is in the systematic design of microelectromechanical systems (MEMS) [20]. MEMS are microscopic (below 1 mm^3) mechanical systems coupled with electronic circuits. MEMS are built using etching and micromachining processes known from the semi-conductor industry. Examples of MEMS applications are medical instruments for in-body surgery, hearing aids, air-bag sensors, micro pumps and optics and tilting mirrors for projection devices. Because of the small scale, hinges and bearings as seen in conventional mechanisms are impossible or difficult to manufacture. Thus MEMS typically work as compliant actuators which gain their mobility from elastic deformation. Several actuation principles for MEMS have appeared. Examples are electrostatics, thermal, electromagnetic, shape-memory and other effects. Here, electrothermomechanical actuation based on Joule heating is investigated.

Manufacturing and processing techniques for MEMS have reached a high level of maturity and new devices can be built in a matter of days in labs and foundries. By contrast, modelling, and in particular, development of *systematic* design methods for MEMS is still in its infancy. Due to the lack of existing systematic design methods for MEMS, many devices are designed using intuition, experience and trial and error approaches. Furthermore, many devices are built from rectangular sub-elements ordered in 'Manhattan'-like horizontal/vertical grids. Obviously, systematic design methods such as topology optimization should be able to improve existing designs considerably and come up with entirely new or more efficient devices or devices with increased functionality. The topology optimization method has recently shown great promise in this area (e.g., [3,11,12,21–26]). The reader can also visit the web-site http://www.topopt.dtu.dk for an interactive topology optimization program for compliant mechanism synthesis relevant to MEMS design.

Some of the challenges in applying topology optimization to non-linear multiphysics and MEMS problems are discussed in the following.

1.1. Material interpolations

For simple compliance optimization problems, the micro structure that optimizes the local stiffness is known and can be used to find optimal homogenized material properties for areas with intermediate densities. This was used in the original works on the homogenization approach to topology optimization [27] and in many following works. For non-compliance problems and coupled field problems the optimal micro structures are not known. Therefore, either sub-optimal micro structures or other interpolation schemes must be used. Here, the solid isotropic material with penalization (SIMP) approach [28–30] is used and in this case, a material property to density relation must be developed for each physical domain. The SIMP approach has also been called the 'fictitious material' approach since it was believed that no material would attain the material properties modelled by the power-law for intermediate densities. However, a recent paper [31] has proved that the SIMP approach is physically permissible as long as a simple condition on the power is fulfilled (e.g. the power must be greater or equal to three for Poisson's ratio equal to a third). The present paper discusses and extends the results of [31] to interpolation problems appearing in multiphysics.

1.2. Optimization algorithms

Optimization problems for systems modelled in multiple physical domains often require several constraints and many thousand variables as will be seen in the examples. Advanced mathematical programming algorithms must therefore be applied to solve these problems, as opposed to the case of simple compliance optimization problems which can be solved using optimality criteria or intuition-based algorithms.

1.3. Linear vs. non-linear modelling

In this paper, it is assumed that material properties are linear in the temperature range of interest (below 500°C). In reality, many material parameters may be non-linear for even smaller temperatures which will

change the performance of the devices and the optimized topologies. However, from lab experiments it is concluded that geometrical non-linearities are more important than material non-linearities.

When considering non-linear displacements (and material behaviours for that sake) in topology optimization, a whole new world of problems and possibilities arises. The problems may consist in the finite element modelling, buckling behaviour and the need for introducing additional constraints to prevent ill behaviour and the new possibilities may include path-generating mechanisms, bi-stable mechanisms and other complex mechanisms. The earlier mentioned papers considering non-linear modelling have opened our eyes for some of these new aspects. In the present paper, some new ideas for geometrically non-linear topology optimization with applications to compliant actuator design are introduced.

One of the problems arising when dealing with large-displacement modelling is that element stiffness matrices for low-density elements may become excessively distorted resulting in negative definite stiffness matrices. Above-mentioned papers deal with the problem in different ways. Kemmler et al. [14] ignore the stress-induced part of the non-linear element matrices when solving the stability problem, Bruns and Tortorelli [16] suggest to use a hyper-elastic material law, Bruns and Tortorelli [17] suggest to remove low-density elements and Buhl et al. [15] suggest to ignore nodes surrounded by low-density elements in the convergence analysis. In this paper, the latter approach is adopted.

1.4. The paper

In this paper, methods and results from three proceedings papers [3,11,12] are reviewed and extended. The extensions include new material interpolation schemes, more discussion on objective functions and sensitivity analysis, geometrically non-linear modelling and improved examples. Since the main application of the developed software will be MEMS-design and MEMS usually are planar structures, only two-dimensional problems with in-plane deflections will be discussed in this paper. However, the basic ideas also hold for out-of-plane and three-dimensional problems.

The first part of this paper is restricted to one-material structures and the second part of the paper considers the extensions to multi-material design problems.

This first part of this paper is organized as follows. In Section 2, the governing equations for an electrothermomechanical system are formulated and in Section 3 these equations are reformulated in finite element notation. In Section 4 the material interpolation schemes are discussed. Then the formulation of the optimization problem is discussed in Section 5 and the sensitivity analysis is discussed in Section 6, followed by discussions of the computational procedure in Section 7. Finally, the multiphysics topology optimization algorithm is applied to various actuator design problems appearing in MEMS in Section 8 and the conclusions are drawn in Section 9.

2. Governing equations

A quite general problem in microscopic actuator or MEMS design is to convert an electrical input to a mechanical output. Many energy conversion principles such as piezoelectrics, electrostatics and electromagnetics a.o. can solve this task. Here, the principle of electrothermomechanical actuation will be focussed on. In the electrothermomechanical actuation principle an electrical current is converted to heat by Joule's heating and the heat then causes thermal strain, which in turn causes structural deformation. How to design an efficient electrothermomechanical actuator is obviously a complicated task. This section discusses the governing equations for a general electrothermomechanically loaded structure. For manufacturing reasons, the structure must be two-dimensional. However everything discussed in the following can easily be extended to the three-dimensional case.

The actuators designed in this paper undergo rather large displacements, on the other hand, experimental results at moderate temperatures have shown no sign of plastic deformation of the structures which would have been seen as hysteresis or failure in returning the output point to the original position after actuation. Therefore, we can assume geometrical non-linearity but small strains and linear material behaviour in the modelling of the actuators.

Assuming that the geometrical changes do not influence the convection or conduction properties (e.g., no modelling of contact between parts), the small strains allow that the electrical and thermal fields can be modelled using linear modelling and thus the system is only weakly coupled in the sense that the heat equations do not depend on the elasticity equations and that the electric field equations are independent of the heat equations.

The structure to be analysed is a sub-domain of the domain Ω. The boundaries are composed of six, pairwise disjoint regions $\Gamma = \overline{\Gamma_{u_0} \cup \Gamma_{t_0}} = \overline{\Gamma_{u_1} \cup \Gamma_{t_1}} = \overline{\Gamma_{u_2} \cup \Gamma_{t_2}}$ where the indices 0, 1 and 2 refer to electrical, thermal and elastic boundary conditions, respectively, and subscripts u and t refer to displacement and traction boundary conditions, respectively (see Fig. 1). Regions Γ_{u_i} can coincide, partly overlap or be fully disjoint, which holds for regions Γ_{t_i} as well. The domain Ω contains a linear thermoelectroelastic material. The governing equations for the electrical, the thermal and the elastic fields, respectively, are:

$$\begin{aligned} &\sigma_0 u_{0,ii} + b_0 = 0 \quad \text{in } \Omega, \\ &u_0 = \overline{u}_0 \quad \text{on } \Gamma_{u_0}, \\ &-\sigma_0 u_{0,n} = \overline{t}_0 \quad \text{on } \Gamma_{t_0}, \end{aligned} \tag{1}$$

$$\begin{aligned} &\sigma_1 u_{1,ii} + b_1 = 0 \quad \text{in } \Omega, \\ &b_1 = -\sigma_0 u_{0,i}^2 - h(\overline{u}_1 - u_1), \\ &u_1 = \overline{u}_1 \quad \text{on } \Gamma_{u_1}, \\ &-\sigma_1 u_{1,n} = \overline{t}_1 \quad \text{on } \Gamma_{t_1}, \end{aligned} \tag{2}$$

$$\begin{aligned} &s_{ij,j} + b_{2(i)} = 0 \quad \text{in } \Omega, \\ &s_{ij} = E_{ijkl}(\eta_{kl} - \alpha_{kl} u_1) \quad \text{in } \Omega, \\ &\eta_{ij} = \tfrac{1}{2}\left[u_{2(i,j)} + u_{2(j,i)} + u_{2(k,i)} u_{2(k,j)}\right] \quad \text{in } \Omega, \\ &u_{2(i)} = \overline{u}_{2(i)} \quad \text{on } \Gamma_{u_2}, \\ &s_{ij} n_j = \overline{t}_{2(i)} \quad \text{on } \Gamma_{t_2} \end{aligned} \tag{3}$$

where the u's are state variables with u_0 being voltage, u_1 temperature change and $u_{2(i)}$ displacement component in direction i, the notation $_{,i}$ means differentiation with respect to coordinate direction i, b_0 is the prescribed internal current source per unit volume, h is the convection coefficient, b_1 models the internal heat generation per unit volume and the convection from the top surfaces, $b_{2(i)}$ are body force components, σ_0 and σ_1 are the (isotropic) electrical and thermal conductivities, respectively, E_{ijkl} is the elastic stiffness tensor, $\overline{u}_0$

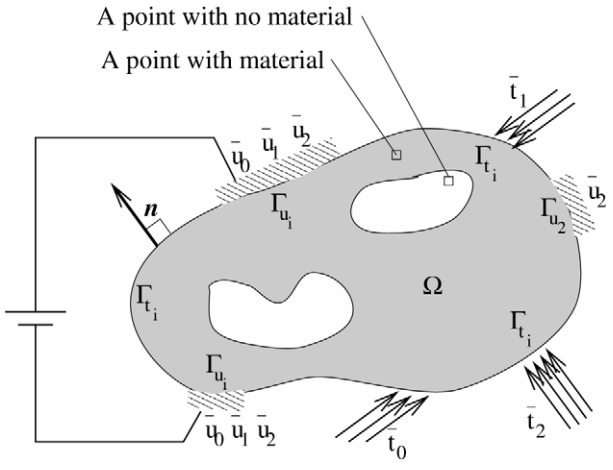

Fig. 1. Loadings and boundary conditions for the general electrothermomechanical load problem.

and $\overline{u}_1$ are prescribed (or neutral) voltage and temperature, respectively, $\overline{u}_{2(i)}$ are the prescribed displacement components, $\overline{t}_0$ and $\overline{t}_1$ are prescribed current and heat flux, respectively, $\overline{t}_{2(i)}$ are traction force components, η_{ij}, s_{kl} and α_{ij} are the Green–Lagrange strain, second Piola Kirchoff stress and thermal expansion tensors, respectively, n_i are the components of the unit normal vector to the surface Γ. Note that the internal heat generation b_1 is a function of the electric field u_0 and that the stress s_{ij} is a function of the temperature field u_1.

In two special cases, the governing equations (1)–(3) must be modified. In case of geometrically linear modelling, the last term in the third row of (3) cancels out. In the case of pure thermal loading, the temperature u_1 is constant throughout the domain and thus only the elasticity equations (3) must be solved in order to determine the response of the structure.

3. Finite element formulation

To simplify the equations, it is assumed that the body force and the internal electric current source are zero. The two-dimensional design domain is assumed to be rectangular and is discretized in a regular way such that the shapes of all the 4-node quadrilateral finite elements are equal and rectangular. The number of elements used in the discretization is N and the number of nodes is m. The structure is modelled using the plane stress assumption. To begin with, we consider small displacement (linear) theory, but at the end of the subsection we extend the theory to geometrically non-linear modelling.

Since the goal of the topology optimization is to optimize material distribution by controlling material density in each finite element, we let material properties depend on the relative density ρ^e of material in element e. An N-vector containing the design variables is denoted $\boldsymbol{\rho}$. We will return to the exact relations between material properties and ρ^e later.

The discretized electrothermomechanical system can be modelled by three finite element problems:

$$\mathbf{R}_0(\boldsymbol{\rho}) = \mathbf{P}_0 - \mathbf{K}_0(\boldsymbol{\rho})\mathbf{U}_0(\boldsymbol{\rho}) = \mathbf{0}, \tag{4}$$

$$\mathbf{R}_1(\mathbf{U}_0(\boldsymbol{\rho}), \boldsymbol{\rho}) = \mathbf{P}_1(\mathbf{U}_0(\boldsymbol{\rho}), \boldsymbol{\rho}) - \mathbf{K}_1(\boldsymbol{\rho})\mathbf{U}_1(\boldsymbol{\rho}) = \mathbf{0}, \tag{5}$$

$$\mathbf{R}_2(\mathbf{U}_1(\boldsymbol{\rho}), \boldsymbol{\rho}) = \mathbf{P}_2(\mathbf{U}_1(\boldsymbol{\rho}), \boldsymbol{\rho}) - \mathbf{K}_2(\boldsymbol{\rho})\mathbf{U}_2(\boldsymbol{\rho}) = \mathbf{0}. \tag{6}$$

As before, index 0 refers to electrical, 1 to thermal and 2 to mechanical modelling. As discussed in Section 2, the electrical field results in thermal heating which again results in thermal strain. Consequently the finite element problems (4)–(6), must be solved in the order they are listed. For the electrical and thermal problems, $\mathbf{K}_0$ and $\mathbf{K}_1$ are $(m \times m)$ global conductivity matrices whereas the global stiffness matrix $\mathbf{K}_2$ is a $(2m \times 2m)$ matrix (two displacement degrees of freedom (d.o.f.) per node). $\mathbf{U}_0$ and $\mathbf{U}_1$ are the voltage and temperature m-vectors, $\mathbf{U}_2$ is the $2m$ displacement vector, $\mathbf{P}_0$ and $\mathbf{P}_1$ are the electrical and thermal load vectors (m) and $\mathbf{P}_2$ is the structural load vector $(2m)$. All the system matrices and vectors are dependent on the element relative density vector $\boldsymbol{\rho}$.

The system matrices $\mathbf{K}_i$ are assembled the usual way as sums over elements matrices

$$\mathbf{K}_0(\boldsymbol{\rho}) = \sum_{e=1}^{N} \mathbf{k}_0^e(\rho^e), \quad \mathbf{K}_1(\boldsymbol{\rho}) = \sum_{e=1}^{N} (\mathbf{k}_1^e(\rho^e) + \mathbf{h}^e) \quad \text{and} \quad \mathbf{K}_2(\boldsymbol{\rho}) = \sum_{e=1}^{N} \mathbf{k}_2^e(\rho^e), \tag{7}$$

where $\mathbf{k}_0^e$ (4×4) is the element electrical conductivity matrix, $\mathbf{k}_1^e$ and $\mathbf{h}$ (4×4) are the element thermal conductivity and convection matrices, respectively, and $\mathbf{k}_2^e$ (8×8) is the element stiffness matrix.

The element matrices $\mathbf{k}_i^e$ are functions of the material properties in the element, whereas the element convection matrix $\mathbf{h}$ is assumed to be independent of the element density (otherwise low-density elements would not be cooled). Defining $\sigma_0(\rho^e)$ as the electrical conductivity, $\sigma_1(\rho^e)$ as the thermal conductivity, h as the convection coefficient, $\mathbf{E}(\rho^e)$ as the elasticity matrix of element e, $\mathbf{N}_0$ as the 4-vector of shape functions, $\mathbf{B}_0$ as the voltage- or temperature-gradient matrix, $\mathbf{B}_2$ as the strain–displacement matrix and V^e as the volume of element e, the element system matrices can be written as

$$\mathbf{k}_0^e(\rho^e) = \sigma_0(\rho^e) \int_{V^e} \mathbf{B}_0^{\mathrm{T}} \mathbf{B}_0 \, \mathrm{d}V, \tag{8}$$

$$\mathbf{k}_1^e(\rho^e) = \sigma_1(\rho^e)\int_{V^e}\mathbf{B}_0^{\mathrm{T}}\mathbf{B}_0\,\mathrm{d}V, \tag{9}$$

$$\mathbf{h}^e = h\int_{S^e}\mathbf{N}_0^{\mathrm{T}}\mathbf{N}_0\,\mathrm{d}S, \tag{10}$$

$$\mathbf{k}_2^e(\rho^e) = \int_{V^e}\mathbf{B}_2^{\mathrm{T}}\mathbf{E}(\rho^e)\mathbf{B}_2\,\mathrm{d}V. \tag{11}$$

For the case of geometrically non-linear modelling, we write the Hooke's law in vector form as $\mathbf{s} = \mathbf{E}\boldsymbol{\eta}$ and redefine the finite element matrix $\mathbf{B}_2$ as the (8×3) matrix that transforms a change in element nodal displacements $\mathrm{d}\mathbf{U}_2$ into a change in strain $\mathrm{d}\boldsymbol{\eta}$, i.e.,

$$\mathbf{B}_2(\mathbf{U}_2) = \frac{\mathrm{d}\boldsymbol{\eta}}{\mathrm{d}\mathbf{U}_2}. \tag{12}$$

Now, the residual for the elastic problem is redefined as the error in obtaining the elastic equilibrium

$$\mathbf{R}_2(\mathbf{U}_2,\mathbf{U}_1,\boldsymbol{\rho}) = \mathbf{P}_2(\mathbf{U}_2,\mathbf{U}_1,\boldsymbol{\rho}) - \int_{\Omega}\mathbf{B}_2^{\mathrm{T}}(\mathbf{U}_2)\mathbf{s}(\mathbf{U}_2,\boldsymbol{\rho})\,\mathrm{d}\Omega = \mathbf{0}, \tag{13}$$

where $\mathbf{P}_2$ is the structural $(2m)$ load vector. Using a total Lagrangian formulation, the integration is performed over the undeformed volume. The equilibrium is reached when the residual vector is equal to the zero vector.

There are several ways to find the equilibrium of a non-linear structure (e.g., [32] or [33]), one of the most popular being the Newton–Raphson (NR) iterative scheme. To find the equilibrium using NR-iterations a relationship between the error increment $\mathrm{d}\mathbf{R}_2$ and the displacement increment $\mathrm{d}\mathbf{U}_2$ must be defined

$$\mathrm{d}\mathbf{R}_2 = -\left[\int_{\Omega}\mathbf{B}_2^{\mathrm{T}}(\mathbf{U}_2)\frac{\partial\mathbf{s}(\mathbf{U}_2,\boldsymbol{\rho})}{\partial\mathbf{U}_2}\,\mathrm{d}\Omega - \int_{\Omega}\frac{\partial\mathbf{B}_2^{\mathrm{T}}(\mathbf{U}_2)}{\partial\mathbf{U}_2}\mathbf{s}(\mathbf{U}_2,\boldsymbol{\rho})\,\mathrm{d}\Omega\right]\mathrm{d}\mathbf{U}_2 = \mathbf{K}_{2\mathrm{T}}\,\mathrm{d}\mathbf{U}_2 \tag{14}$$

with $\mathbf{K}_{2\mathrm{T}}$ being the tangent stiffness matrix. More details on determining the tangent stiffness matrix and implementation of the NR iterative scheme can be found in standard books on non-linear finite element theory (e.g., [32] or [33]).

Since the topology optimization procedure may require the solving of several thousand non-linear finite element problems, we have chosen to solve the equilibrium equation (6) in one step (only one increment), i.e., without an incremental scheme. The implications of this decision are discussed in further detail in Section 7.

4. Material interpolation scheme

Solving topology optimization problems using mathematical programming algorithms, require the design variables to be continuous thus we need a scheme for the interpolation of material properties between solid $\rho^e = 1$ and void $\rho^e = 0$. Here, we use the so-called power-law approach or SIMP where the Young's modulus depends on some power of the density and the Poisson's ratio is constant (e.g. [28–30]). Until recently, the power-law approach was believed to result in 'artificial materials' that violated theoretical bounds on material behaviour [34]. However, in a recent paper [31] it was proved that the power-law is thermodynamically admissible if the power is selected according to a simple rule (see later). The power-law approach has the advantage that the material properties between solid and void are interpolated with a smooth continuous function which only depends on the material density.

The power-law interpolation of the Young's modulus E can be written as

$$E(\rho^e) = (\rho^e)^p E^{(0)}, \tag{15}$$

where $E^{(0)}$ is the Young's modulus of solid material and p is a penalization power typically chosen larger than 3. High values of the penalization power p decreases the stiffness of intermediate density elements thus making the use of such element uneconomical if the volume is constrained. In this way, high values of the power result in 'black and white' or '0/1' design which are easily manufacturable.

As discussed above, the power-law approach was earlier called the 'fictitious' or 'artificial' material approach since it was believed that it was impossible to build materials with properties corresponding to the power-law interpolation (i.e., Young's modulus given by Eq. (15) and constant Poisson's ratio $v = v^{(0)}$). However, it has been proved [31] that the power-law is thermodynamically admissible if the power p is selected according to the rule

$$p \geqslant \max\left(\frac{2}{1-v^{(0)}}, \frac{4}{1+v^{(0)}}\right). \tag{16}$$

For reasons that will become apparent in the second part of the paper, we will here choose to write the interpolation scheme in an alternative way formulated in terms of material bulk and shear moduli. Instead of using the interpolation on the Young's modulus (15) with constant Poisson's ratio, it is here chosen to interpolate the bulk and shear modulus of the material separately, i.e.,

$$\kappa(\rho^e) = (\rho^e)^p \kappa^{(0)} \quad \text{and} \quad \mu(\rho^e) = (\rho^e)^p \mu^{(0)}, \tag{17}$$

where bulk and shear moduli for the base material can be written in terms of the Young's modulus and the Poisson's ratio

$$\kappa^{(0)} = \frac{E^{(0)}}{2(1-v^{(0)})} \quad \text{and} \quad \mu^{(0)} = \frac{E^{(0)}}{2(1+v^{(0)})}. \tag{18}$$

Similar to the condition on the power p written in terms of Young's modulus and Poisson's ratio of the base material in Eq. (16), one can write the condition in terms of bulk and shear moduli as

$$p \geqslant \max\left(\frac{\kappa^{(0)}+\mu^{(0)}}{\mu^{(0)}}, \frac{\kappa^{(0)}+\mu^{(0)}}{\kappa^{(0)}}\right). \tag{19}$$

The inequality in Eq. (19) is (likewise (16)) calculated under the condition that the Hashin–Shtrikman bounds for bulk and shear moduli must be satisfied for any density ρ^e.

The electrical and thermal conductivities for elements with intermediate densities are also interpolated using the power-law approach

$$\sigma_0(\rho^e) = (\rho^e)^p \sigma_0^{(0)} \quad \text{and} \quad \sigma_1(\rho^e) = (\rho^e)^p \sigma_1^{(0)}. \tag{20}$$

To satisfy the Hashin–Shtrikman bounds, the power p for the conductivities must satisfy $p \geqslant 2$.

The thermal expansion coefficient is independent of the density and the convection coefficient is assumed independent of the element density.

In principle, the power p could take different values for each physical property. However, for simplicity and to avoid having to choose multiple parameters, the same power is selected for all material interpolations, which means that p is governed solely by Eq. (16) or (19).

5. Optimization problem

Design problems for a thermally and an electrothermally actuated structure are sketched in Fig. 2. The design problems consist in distributing a given amount of material in the planar design domains such that the deformation u_{out} of a workpiece of given stiffness k_s is maximized. At the same time, the displacement component perpendicular to the output (spring) direction $\hat{u}_{\text{out}}$ should be constrained. Some areas of the design domain may be restricted to be void (white areas) and some may be restricted to be solid material (black areas). The optimal actuator topology will depend on the stiffness of the workpiece. For a soft workpiece a small force is needed to deform the workpiece, whereas for a stiff workpiece a large output force is needed to deform the workpiece. In this way, one can design force-oriented or displacement-oriented actuators by varying the stiffness of the workpiece.

Preliminary tests of the algorithm showed that the electrothermally actuated topologies often become almost short-circuited with very high temperatures and resultingly large deflections, but also with a very

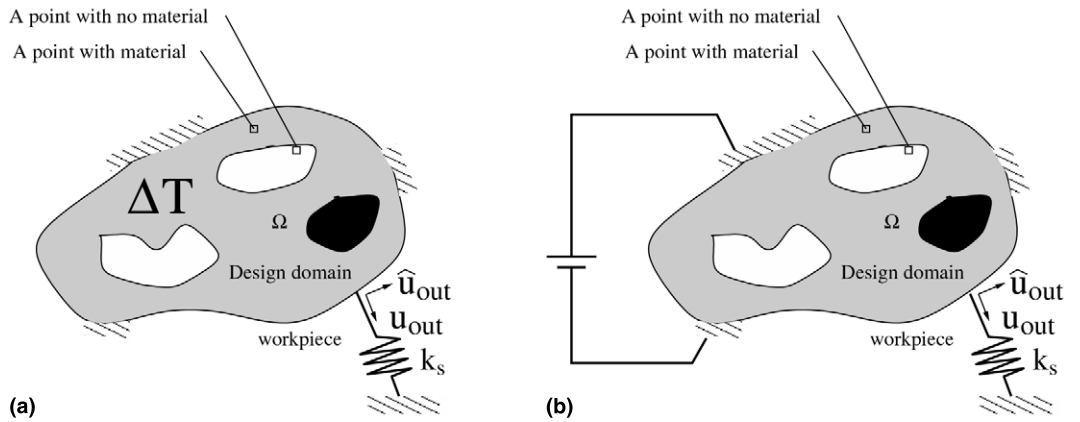

Fig. 2. Design problems for thermal actuation (a) and electrothermomechanical actuation (b).

high current running through the structures. To prevent this short-circuiting, an upper bound constraint on the current is added to the design problem.

Based on above arguments, the optimization problem can be written as:

$$\begin{aligned} \max_{\boldsymbol{\rho}} \quad & u_{\text{out}}(\boldsymbol{\rho}) \\ \text{s.t.} \quad & \sum_{e=1}^{N} \rho^e V^e \leqslant V^*, \\ & I(\boldsymbol{\rho}) \leqslant I^*, \\ & \frac{(\hat{u}_{\text{out}}(\boldsymbol{\rho}))^2}{(u_{\text{out}}(\boldsymbol{\rho}))^2} \leqslant \varepsilon^*, \\ & \mathbf{0} < \boldsymbol{\rho}_{\min} \leqslant \boldsymbol{\rho} \leqslant \mathbf{1}, \\ & \text{equilibrium equations,} \end{aligned} \tag{21}$$

where V^e is the element volume, V^* is the constraint on material volume, $I(\boldsymbol{\rho})$ is the electrical current, I^* is the upper bound on the current, $\boldsymbol{\rho}_{\min}$ is the lower bound vector on relative element densities ($\boldsymbol{\rho}_{\min} = 10^{-3}$) and ε^* is a small positive number (usually $\varepsilon^* = 0.01$).

The introduction of a volume constraint in the optimization of micro actuators may seem unnecessary both because material cost for such small devices is low and also because the superfluous material may have to be etched away from the structure. However, the volume constraint is mainly introduced in order to produce narrow structures with thin bars and members that can be under-etched in a reasonable amount of time. Furthermore, the volume constraint ensures that we end up with black and white structures that make efficient use of the available material. Without the volume constraint, islands of material that are disconnected from the actual mechanism may appear. Furthermore, these superfluous islands may prevent the real structure in growing up in those areas. Another problem is, that without the volume constraint, the whole idea of the power-law approach is lost and there is a risk of ending up with structures with large 'grey' (intermediate density) regions. This is because the basic idea of the power-law approach is to make intermediate density areas 'uneconomical'. Then, if there is no volume constraint, intermediate density areas will be just as 'economical' as full density areas. It should be noted that volume constraint always will be active. The reason for this is that the objective function requires the output force to be maximized and this requires maximum amount of material in the actuator. This question is discussed in further detail in Section 8.1.1.

5.1. Multiple inputs and outputs

In order to solve problems with multiple inputs and outputs, the single input/output optimization problem (21) must be extended. The selection of additional constraints and objective functions may vary

from problem to problem. Here, the formulation for a multiple input/output actuator is demonstrated by an XY-actuator.

The design problem is sketched in Fig. 15(a). The goal is to come up with a device that moves the output point (center right edge) horizontally when an electric potential is applied over the top terminal and the lower left terminal and vertically when an electric potential is applied over the top terminal and the lower right terminal. The horizontal and vertical displacements for the two input currents should have equal magnitudes and at the same time there should be no cross-sensitivity between the outputs. Since the relation between heating and applied current is non-linear, the output point does not necessarily move in a 45° direction when the two input potentials are applied simultaneously. Therefore, a third load case with corresponding constraints is added to the optimization problem to ensure a higher linearity of the output behaviour. To solve this design problem several new constraints are added to the original optimization problem (21). The more complex optimization problem is written as:

$$\begin{aligned}
\max_{\boldsymbol{\rho}} \quad & u_{\text{out},1}(\boldsymbol{\rho}) + u_{\text{out},2}(\boldsymbol{\rho}) \\
\text{s.t.} \quad & \sum_{e=1}^{N} \rho^e V^e \leqslant V^*, \\
& I_i(\boldsymbol{\rho}) \geqslant I^*, \quad i = 1,2, \\
& \frac{(\hat{u}_{\text{out},i}(\boldsymbol{\rho}))^2}{(u_{\text{out},i}(\boldsymbol{\rho}))^2} \leqslant \varepsilon^*, \quad i = 1,3, \\
& \frac{(u_{\text{out},1}(\boldsymbol{\rho}) - u_{\text{out},2}(\boldsymbol{\rho}))^2}{(u_{\text{out},1}(\boldsymbol{\rho}) + u_{\text{out},2}(\boldsymbol{\rho}))^2} \leqslant \varepsilon^*, \\
& \mathbf{0} < \boldsymbol{\rho}_{\min} \leqslant \boldsymbol{\rho} \leqslant \mathbf{1}, \\
& \text{equilibrium equations},
\end{aligned} \tag{22}$$

where $u_{\text{out},i}$ are the deflections of the output port when the left, right and both electrical inputs are applied, respectively, and $\hat{u}_{\text{out},i}$ are the corresponding cross-sensitivities (displacements perpendicular to the output direction). The horizontal and the vertical deflections for outputs 1 and 2 should have similar values which is looked after by the fifth line in the optimization problem.

6. Sensitivity analysis

To solve the optimization problems (21) or (22), the sensitivities of the structural response to changes in the design variables must be determined. The complete sensitivity analysis will not be performed here, instead it is shown how the sensitivity of the linear output displacement of the actuator with respect to change in a design variable $\mathrm{d}u_{\text{out}}/\mathrm{d}\rho^e$ can be determined using the adjoint sensitivity analysis method (e.g., [35]).

The displacement of the output point u_{out} can be written as

$$u_{\text{out}} = \mathbf{L}^{\mathrm{T}}\mathbf{U}_2, \tag{23}$$

where $\mathbf{L}^{\mathrm{T}}$ is a vector consisting of zeroes except for the position i, corresponding to the d.o.f. of the output direction, where its value is one.

Assuming that the finite element problems (4)–(6) have been solved, nothing is changed by rewriting (23) as

$$u_{\text{out}} = \mathbf{L}^{\mathrm{T}}\mathbf{U}_2(\boldsymbol{\rho}) + \boldsymbol{\lambda}_0^{\mathrm{T}}\mathbf{R}_0(\mathbf{U}_0) + \boldsymbol{\lambda}_1^{\mathrm{T}}\mathbf{R}_1(\mathbf{U}_1,\mathbf{U}_0) + \boldsymbol{\lambda}_2^{\mathrm{T}}\mathbf{R}_2(\mathbf{U}_2,\mathbf{U}_1), \tag{24}$$

where $\boldsymbol{\lambda}_0^{\mathrm{T}}$, $\boldsymbol{\lambda}_1^{\mathrm{T}}$ (m-vectors) and $\boldsymbol{\lambda}_2^{\mathrm{T}}$ ($2m$-vector) are adjoint vectors. The sensitivity of the output displacement u_{out} with respect to design variable ρ^e can then be calculated as

$$\begin{aligned}\frac{\mathrm{d}u_{\mathrm{out}}}{\mathrm{d}\rho^e} &= \mathbf{L}^{\mathrm{T}}\frac{\mathrm{d}\mathbf{U}_2}{\mathrm{d}\rho^e} + \boldsymbol{\lambda}_0^{\mathrm{T}}\left(\frac{\partial \mathbf{R}_0}{\partial \mathbf{U}_0}\frac{\mathrm{d}\mathbf{U}_0}{\mathrm{d}\rho^e} + \frac{\partial \mathbf{R}_0}{\partial \rho^e}\right) + \boldsymbol{\lambda}_1^{\mathrm{T}}\left(\frac{\partial \mathbf{R}_1}{\partial \mathbf{U}_1}\frac{\mathrm{d}\mathbf{U}_1}{\mathrm{d}\rho^e} + \frac{\partial \mathbf{R}_1}{\partial \rho^e}\right) + \boldsymbol{\lambda}_2^{\mathrm{T}}\left(\frac{\partial \mathbf{R}_2}{\partial \mathbf{U}_2}\frac{\mathrm{d}\mathbf{U}_2}{\mathrm{d}\rho^e} + \frac{\partial \mathbf{R}_2}{\partial \rho^e}\right)\\ &= \mathbf{L}^{\mathrm{T}}\frac{\mathrm{d}\mathbf{U}_2}{\mathrm{d}\rho^e} + \boldsymbol{\lambda}_0^{\mathrm{T}}\left(\frac{\mathrm{d}\mathbf{P}_0}{\mathrm{d}\rho^e} - \frac{\mathrm{d}\mathbf{K}_0}{\mathrm{d}\rho^e}\mathbf{U}_0\right) + \boldsymbol{\lambda}_1^{\mathrm{T}}\left(\frac{\partial \mathbf{P}_1}{\partial \rho^e} - \frac{\mathrm{d}\mathbf{K}_1}{\mathrm{d}\rho^e}\mathbf{U}_1\right) + \boldsymbol{\lambda}_2^{\mathrm{T}}\left(\frac{\partial \mathbf{P}_2}{\partial \rho^e} - \frac{\mathrm{d}\mathbf{K}_2}{\mathrm{d}\rho^e}\mathbf{U}_2\right)\\ &\quad - \boldsymbol{\lambda}_0^{\mathrm{T}}\mathbf{K}_0\frac{\mathrm{d}\mathbf{U}_0}{\mathrm{d}\rho^e} - \boldsymbol{\lambda}_1^{\mathrm{T}}\mathbf{K}_1\frac{\mathrm{d}\mathbf{U}_1}{\mathrm{d}\rho^e} - \boldsymbol{\lambda}_2^{\mathrm{T}}\mathbf{K}_2\frac{\mathrm{d}\mathbf{U}_2}{\mathrm{d}\rho^e} + \boldsymbol{\lambda}_1^{\mathrm{T}}\frac{\partial \mathbf{P}_1}{\partial \mathbf{U}_0}\frac{\mathrm{d}\mathbf{U}_0}{\mathrm{d}\rho^e} + \boldsymbol{\lambda}_2^{\mathrm{T}}\frac{\partial \mathbf{P}_2}{\partial \mathbf{U}_1}\frac{\mathrm{d}\mathbf{U}_1}{\mathrm{d}\rho^e}. \end{aligned} \tag{25}$$

In order to get rid of the terms including sensitivities of the displacement vectors, the terms including $\mathrm{d}\mathbf{U}_i/\mathrm{d}\rho^e$ should be zero, i.e.

$$\left(\boldsymbol{\lambda}_1^{\mathrm{T}}\frac{\partial \mathbf{P}_1}{\partial \mathbf{U}_0} - \boldsymbol{\lambda}_0^{\mathrm{T}}\mathbf{K}_0\right)\frac{\mathrm{d}\mathbf{U}_0}{\mathrm{d}\rho^e} = 0, \quad \left(\boldsymbol{\lambda}_2^{\mathrm{T}}\frac{\partial \mathbf{P}_2}{\partial \mathbf{U}_1} - \boldsymbol{\lambda}_1^{\mathrm{T}}\mathbf{K}_1\right)\frac{\mathrm{d}\mathbf{U}_1}{\mathrm{d}\rho^e} = 0 \quad \text{and} \quad \left(\mathbf{L} - \boldsymbol{\lambda}_2^{\mathrm{T}}\mathbf{K}_2\right)\frac{\mathrm{d}\mathbf{U}_2}{\mathrm{d}\rho^e} = 0. \tag{26}$$

This can be obtained by selecting the adjoint vectors as the solutions to the following three adjoint load finite element problems:

$$\mathbf{K}_0\boldsymbol{\lambda}_0 = \frac{\partial \mathbf{P}_1}{\partial \mathbf{U}_0}\boldsymbol{\lambda}_1, \quad \mathbf{K}_1\boldsymbol{\lambda}_1 = \frac{\partial \mathbf{P}_2}{\partial \mathbf{U}_1}\boldsymbol{\lambda}_2 \quad \text{and} \quad \mathbf{K}_2\boldsymbol{\lambda}_2 = \mathbf{L}. \tag{27}$$

These three adjoint load problems correspond to solving the same three finite element problems as before ((4)–(6)) but with other right-hand sides and in the opposite order. This is a computationally cheap operation since the stiffness matrices already have been factorized.

Having solved these three adjoint load cases for the adjoint vectors $\boldsymbol{\lambda}_0^{\mathrm{T}}$, $\boldsymbol{\lambda}_1^{\mathrm{T}}$ and $\boldsymbol{\lambda}_2^{\mathrm{T}}$, the sensitivity (25) can be simplified to

$$\frac{\mathrm{d}u_{\mathrm{out}}}{\mathrm{d}\rho^e} = \boldsymbol{\lambda}_0^{\mathrm{T}}\left(\frac{\mathrm{d}\mathbf{P}_0}{\mathrm{d}\rho^e} - \frac{\mathrm{d}\mathbf{K}_0}{\mathrm{d}\rho^e}\mathbf{U}_0\right) + \boldsymbol{\lambda}_1^{\mathrm{T}}\left(\frac{\partial \mathbf{P}_1}{\partial \rho^e} - \frac{\mathrm{d}\mathbf{K}_1}{\mathrm{d}\rho^e}\mathbf{U}_1\right) + \boldsymbol{\lambda}_2^{\mathrm{T}}\left(\frac{\partial \mathbf{P}_2}{\partial \rho^e} - \frac{\mathrm{d}\mathbf{K}_2}{\mathrm{d}\rho^e}\mathbf{U}_2\right). \tag{28}$$

The sensitivity of the current I is also found using adjoint sensitivity analysis involving the solving of one extra electrical load-case.

For multiple input/output actuators, the three finite element problems (4)–(6) and their corresponding adjoint load-cases (27) must be solved for each output. Determining a cross-displacement sensitivity $\mathrm{d}\,\hat{u}_{\mathrm{out}}/\mathrm{d}\rho^e$ requires one extra elastic adjoint load case and determining the electric current requires one extra electrical adjoint load case. In total, a number of three finite element problems and a total of eight load-cases must be solved for each input/output set. This means that the XY-actuator design problem requires the solution of nine finite element problems and 24 load cases in order to determine the structural response and to perform the sensitivity analysis. This may sound like a lot, but in reality the electrical and thermal problems are solved in a fraction of the time of the elastic problems since the number of d.o.f. and the bandwidth of these problems are half that of the elastic problem. Therefore the main part of computational time is spent on the solving of the three elastic problems. Using a customized finite element code, these problems can also be solved quite efficiently.

When considering purely thermal actuators, there is no need to solve the electrical and thermal fields. The temperature is assumed constant in the domain and thus the sensitivity analysis only involves the solving of two extra elastic load-case (one for the output direction and one for the cross-displacement).

For geometrical non-linear modelling, the sensitivity analysis is performed using the adjoint method as described above, with the added complexity of having to derive the sensitivities of the residual (13). For the non-linear case, solving the adjoint problems is comparatively cheap, since the tangent stiffness matrix $\mathbf{K}_{2\mathrm{T}}$ already has been found during the solving of the real displacements. Further aspects of adjoint sensitivity analysis for non-linear topology optimization can be found in [19,36].

7. Numerical implementation

The design algorithm is implemented as a custom-made code written in the FORTRAN90 programming language. A simple in-data file allows for easy definition of design problems, material data and constraints. A graphical output showing electric, thermal and displacement fields for each load case together with design changes allows the user to follow the progress in the design. Animations of the design process or actuator responses for various inputs can easily be generated.

In the following subsections, a number of implementation aspects and ways to solve numerical problems are discussed in more detail.

7.1. Mathematical programming method

The optimization problems (21) and (22) are non-linear and must be solved by an iterative procedure where linearized sub-problems are solved in each step.

Topology optimization problems with only one constraint (on volume) can be solved by numerous types of more or less theoretically well-founded optimization algorithms. Examples are optimality criteria, evolutionary algorithms, hard-kill/soft-kill methods and many others seen in early and recent literature on topology optimization. Many of these methods are based on intuition and make little or no use of sensitivity analysis information that can be obtained computationally efficiently. When considering more complicated objective functions and multiple constraints, however, it is the author's belief that none of those methods will be able to do a good job. Advanced topology optimization problems must be solved with theoretically well-founded mathematical programming methods making use of sensitivity analysis approximation concepts and, possibly, information from prior iterations. Examples of such algorithms are sequential linear programming (SLP), CONLIN [37] or the method of moving asymptotes (MMA) [38].

The iterative optimization scheme used here is MMA by Svanberg [38]. The MMA-scheme solves a sequence of linearized sub-problems and uses information from prior iterations to improve the approximation of the design space. The method is quite stable and computationally efficient. Solving topology optimization problem with multiple (e.g., seven for the XY-actuator) non-trivial constraints puts a heavy load on the optimization algorithm. In the author's experience, MMA is the most reliable and stable optimization code suited for topology optimization problems that is currently available.

7.2. Computational procedure

The procedure is initiated by specifying material constants, boundary conditions, loadings, workpiece stiffnesses, geometry of design domain, etc. Then the design domain is discretized by a large number of finite elements and the available material is distributed evenly throughout the domain. Now the iterations start. Each iteration consists of the following three steps: (1) solving three finite element problems for the electric, the thermal and the elastic field, respectively; (2) performing an analytical sensitivity analysis (adjoint method) involving the solving of five additional load cases to the finite element problems and (3) doing an optimal material redistribution based on the MMA-scheme. For multiple input/output structures, steps 1 and 2 must be repeated for each input/output relation. The procedure has converged when the changes in design variables from iteration to iteration get below a certain value (typically 10^{-3}). Usually, one can get an idea of the final topology in less than 50 iterations but at least 300 iterations are needed to converge to a good design and this number may be significantly exceeded for problems involving many constraints. The number of design iterations is large compared to traditional stiffness design. This can be explained by the fact that the stress field in stiffness design (self-adjoint problem) is more or less known from iteration number 1 (at least for statically determinate structures) whereas the electric, thermal and elastic fields in the present design problem change drastically during the design process. Depending on the number of elements and constraints and the complexity of the design domain, the whole design process may take from a few minutes to a night on a powerful workstation. To save computational time, the examples are often solved using a coarse finite element discretization. Upon convergence, the resulting design is re-discretized using a finer mesh and the optimization is continued.

7.3. Newton–Raphson procedure

As discussed earlier, we solve the non-linear elasticity problem in one step without an incremental scheme. The advantage of this is the saving in computational time although there is a risk that the method may fail to converge or fail to converge to the right equilibrium state for larger displacements. However, several numerical tests where an incremental scheme was compared with the one-step scheme showed no differences in the resulting topologies except for a drastically increased computational time for the incremental scheme.

7.4. Numerical problems

Various problems occur when solving the topology optimization problem as described.

7.4.1. Checkerboards and mesh-dependencies

Often one encounters regions with alternating solid and void elements, referred to as checkerboards and a strong mesh dependency of the solutions. These problems are described in the literature on topology optimization (e.g., [1] or in a recent survey paper [39]). Here, a mesh-independency algorithm [24,40], that both eliminates the checkerboard problem and the mesh-dependency problem is applied. The method works by substituting the element-wise sensitivities with a weighted average of the sensitivities of their neighbours within a given radius.

7.4.2. Continuation method

Since the topology optimization problem is highly non-convex, one can never be sure to converge to the global optimum. However, by using a continuation approach, the algorithm seems to produce consistent designs independent on starting guesses. The continuation approach makes use of the filtering technique described above. In the beginning of the design process, a large-sized filter is used to convexify the problem. During the design process, the filter size is gradually reduced. Typical filter size in the beginning of the design process is 10% of the smaller dimension of the design domain and it is decreased to 1.2 times the element size at the end of the design process. No other material or optimization parameters need to be adjusted during the design process.

7.4.3. Finite element convergence problems

When the finite element analysis is based on the Green–Lagrange strain measure, large displacements may cause the tangent stiffness matrix to become indefinite or even negative definite. This phenomenon is observed frequently during the topology optimization process and results in non-convergence of the equilibrium iterations. Numerical experiments show that the problem occurs in low-density elements with minimum or close to minimum stiffness. The problem is 'artificial' since the elements with minimum stiffness represent void and therefore their behaviour should not influence the structural response. It should be noted that this problem occurs both for the one-step and the incremental schemes.

As mentioned in Section 1, the problem is solved by relaxing the convergence criterion for the equilibrium iterations [15,19]. Usually, the NR iterative scheme is stopped when the changes in nodal displacements get below a certain value. In the topology optimization case we experience non-convergence of the algorithm since displacements in nodal points surrounded by 'void' (low density) elements keep oscillating. Since such nodes should have no structural importance, nodes surrounded by void elements are eliminated from the convergence criterion. This solution to the problem has proven very efficient and we seldom experience convergence problems.

A way to save computing time for the optimization algorithm is to use the displacements from the previous design iteration as a starting guess for the new equilibrium iterations. However, our experience is that this is a bad strategy, since the algorithm may get stuck in a wrong equilibrium. This was frequently seen during experiments, especially in low density regions that could get stuck in solutions that implied 'folding' of elements. Therefore, we always reset the displacement vectors before a new NR-iteration.

7.4.4. Computational time

Solving the equilibrium equations using NR-iteration implies that the system finite element matrix must be assembled and solved several times for each load case. It is our experience that the number of NR-iterations increases with increasing displacements/non-linearity. For example, finding the equilibrium of an actuator that has to perform work on a very stiff output spring may require three to five NR-iterations, whereas the number of NR-iterations may exceed 10 for problems with very soft output springs. In rare cases, the number of NR-iterations exceeds 40 which usually indicates that one or more bars have become unstable (buckling).

Since the total computational time for linear systems is mainly governed by the time spent on the finite element analysis, the above-mentioned numbers of NR-iterations also correspond to the factor of increase in computational time when considering non-linear systems compared to linear systems.

8. Examples

The examples are grouped into thermal and electrothermal actuators. For the thermal actuators it is possible to establish a theoretical bound for their performance. This makes it possible to test the performance of the proposed design algorithm. For the electrothermal actuators, which involve coupled physics combined with design-dependent heat-losses due to convection and conduction, no theoretical bound for their performance has been established.

8.1. Thermal actuators

The first example considers the design of a thermal micro actuator. Ideally, the actuator would consist of rigid body parts connected with moment free hinges. Unfortunately, it is impossible to manufacture efficient hinges and the small scale of MEMS does not allow the usual assembly processes. Therefore, the actuator must be built as a *compliant mechanism*. A compliant mechanism gains its mobility from flexibility of its members [41]. A compliant mechanism will always perform worse than its rigid body counterpart since some of the input energy will be stored as elastic energy in its members. To obtain a measure of the efficiency of a compliant thermal actuator, a theoretical bound on the work performed by the thermal actuator is derived in the next subsection. In this part of the paper, we derive the bound using geometrically linear modelling. In part II of this paper, some interesting non-linear effects due to stress stiffening are discussed in further detail.

8.1.1. Derivation of a bound for thermal actuator performance

An 'ideal' thermal actuator is shown in Fig. 3(a) and consists of a rod of length L and area A performing work on an output spring with stiffness k_s. The Young's modulus of the rod is $E^{(0)}$ and the thermal expansion coefficient is $\alpha^{(0)}$. The work performed by the actuator on a given spring can be optimized by

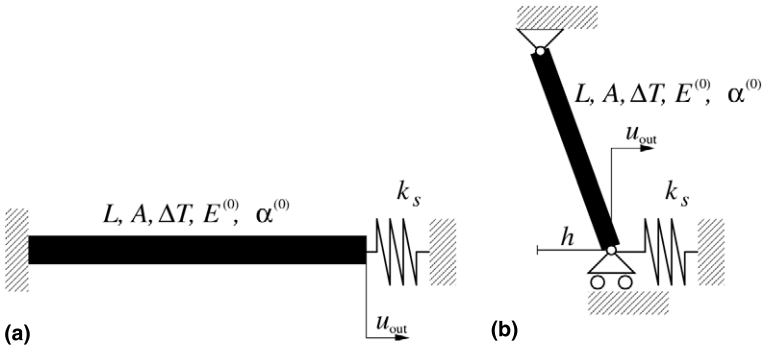

Fig. 3. Ideal thermal actuators: (a) A straight rod performing work on a spring and (b) a displacement amplification actuator performing work on a spring.

finding the length of the actuator that maximizes the displacement. It is assumed that the volume of the actuator is constant, i.e. $V^* = AL$, and that the transfer of force between the actuator and the output spring is loss-free.

Force equilibrium of the actuator spring system can be written as

$$E^{(0)}A\left(\varepsilon - \alpha^{(0)}\Delta T\right) + k_s u_{out} = 0, \tag{29}$$

where the engineering strain is defined as $\varepsilon = u_{out}/L$.

The output displacement u_{out} of the actuator spring system is found as the solution to the equilibrium equation (29).

$$u_{out} = \frac{E^{(0)}A\alpha^{(0)}\Delta T}{(E^{(0)}A/L) + k_s} = \frac{LE^{(0)}V^*\alpha^{(0)}\Delta T}{E^{(0)}V^* + L^2 k_s}. \tag{30}$$

The maximum displacement is obtained for

$$L = L_{max} = \sqrt{\frac{E^{(0)}V^*}{k_s}}. \tag{31}$$

By substituting L_{max} into (30) we find the maximum work performed on the spring as

$$W_{max} = \frac{1}{2}k_s u_{max}^2 = \frac{1}{8}\Delta T^2\left(\alpha^{(0)}\right)^2 E^{(0)}V^*. \tag{32}$$

The maximum work is seen to be independent of the stiffness of the spring whereas the length of the optimal actuator L_{max} does depend on the spring stiffness. The maximum work is proportional to the actuator volume V^* and the Young's modulus $E^{(0)}$ and proportional to the square of the temperature change ΔT and the thermal expansion coefficient $\alpha^{(0)}$. After some algebra we get that the stiffness k_α of the optimal actuator is equal to the spring stiffness, i.e., $k_\alpha = k_s$. This equality was expected and is often referred to as 'impedance matching' when describing mechanical systems with electrical analogies.

The free displacement (without the output spring) of the actuator is

$$u_\alpha = L\alpha^{(0)}\Delta T, \tag{33}$$

and the actuator blocking force is

$$F_\alpha = \frac{E^{(0)}V^*\alpha^{(0)}\Delta T}{L}. \tag{34}$$

Exactly the same results can be obtained for the actuator shown in Fig. 3(b). This actuator consists of a rod of length L, simply supported at one end and supported by a roller and attached to the output spring at the other end. The height h determines the displacement magnification of the actuator. A small h gives a large displacement and a large h gives a large force.

The height h_{max} that maximizes the work performed on an output spring of given stiffness k_s is found as

$$h_{max} = L^2\sqrt{\frac{k_s}{E^{(0)}V^*}}. \tag{35}$$

With height equal to h_{max}, the actuator performs exactly as well as the simple rod-actuator from Fig. 3(a). The advantage of the actuator in Fig. 3(b) is that it can be made more compact than the one in Fig. 3(a) due to the displacement magnification.

For optimization purposes it is interesting to note that the best actuator is obtained if the volume is large. Therefore, a volume constraint in the optimization problem will always be active.

8.1.2. Optimal design of thermally actuated compliant micro actuator

In this subsection, the output work of the ideal actuator (32) is used as a theoretical limit for a compliant actuator design example. The design domain is sketched in Fig. 4 and consists of a square domain with side

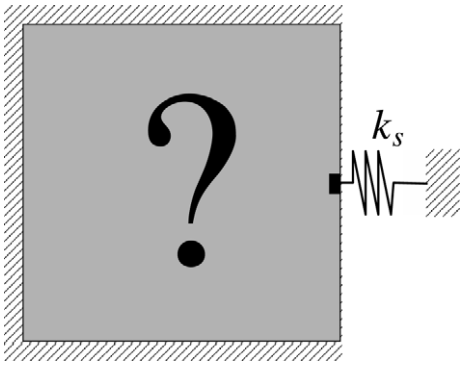

Fig. 4. Design domain for a compliant thermal actuator mechanism.

lengths 500 μm and a thickness of 15 μm. 25% of the design domain may be filled with Nickel which has the thermal expansion coefficient $\alpha^{(0)} = 15 \times 10^{-6}$ K^{-1}, the Young's modulus $E^{(0)} = 200$ GPa and the Poisson's ratio $v = 0.31$. For a temperature change of 100 K, the maximum work performed by the ideal actuator, is 26.5 nJ (from (32)). For output spring stiffnesses of 2000, 200 and 20 N/m, the 'optimal' lengths of the ideal rod-like actuators would be 7.5, 23.7 and 75 mm, respectively. The goal of the optimization is therefore to find three compliant actuator topologies with performances close to the ideal ones but with a size restriction of 500 by 500 μm^2. Due to symmetry, only half of the design domain is discretized using 5000 4-node finite elements.

First, we solve the design problem using linear modelling. At the end of the subsection we analyse the resulting structures using non-linear modelling and compare them with structures optimized using non-linear modelling.

8.1.2.1. Linear modelling. The topology optimized actuators for output spring stiffnesses of 2000, 200 and 20 N/m are shown in Fig. 5. The work performed by the three actuators is 70%, 64% and 59% of the theoretical

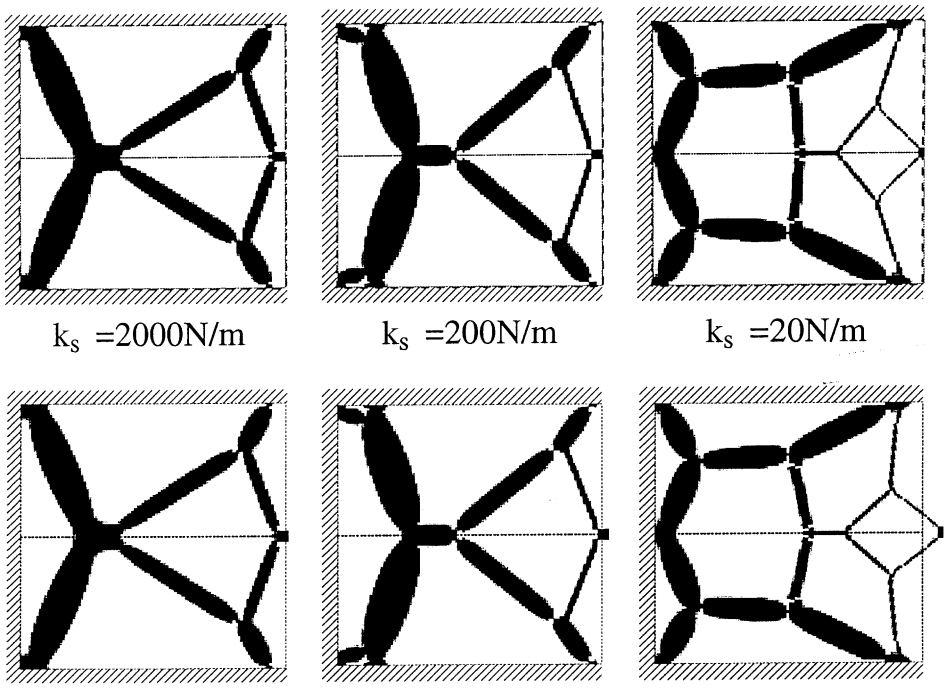

Fig. 5. Design of compliant thermal actuator mechanism. Top row: optimized topologies for output spring stiffnesses of 2000, 200 and 20 N/m, respectively, and linear modelling. Bottom row: (linear) displacements patterns of the optimized actuators.

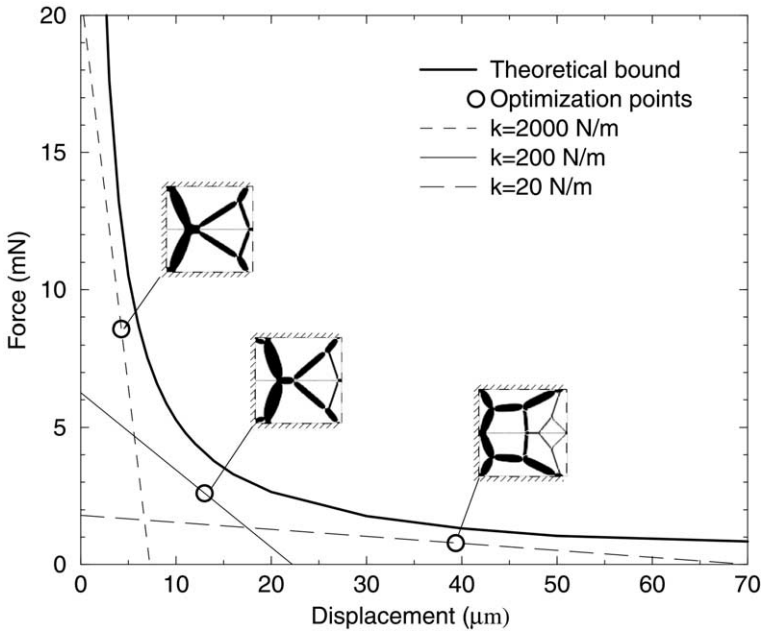

Fig. 6. Force–displacement diagram for the thermal actuators from Fig. 5 (using linear modelling).

limit, respectively. We see that the optimal topology changes with varying k_s. For a high k_s the actuator has thick 'hinges' and bars, whereas for a low value of k_s, the optimal topology shows thin hinges and bars which allow for bending and large displacements. However, the large displacement is obtained on the cost of low actuator force. The force/displacement curves of the three actuators for linear analysis are plotted together with the theoretical iso-work curve in Fig. 6. The obtained values of reaction force and output displacement are shown with small circles in the figure. The work-lines of the individual actuators are shown as straight lines. The free displacement of an actuator (without the workpiece) is the value where the line crosses the displacement axis and the blocking force is the force value where the work-line crosses the force axis. The individual actuators are seen to have the best performance around their design point. It is also seen that the actuator optimized for spring stiffness $k_s = 20$ N/m has a maximum (unloaded) displacement of 65 μm and a blocking force just below 2 mN whereas the actuator optimized for spring stiffness $k_s = 2000$ N/m has a maximum (unloaded) displacement of 7 μm and a blocking force just above 20 mN, thus, as expected, there is an inverse relation between the output force and the output displacement of an optimized actuator. The stiffnesses of the optimized actuators are 2870, 282 and 25 N/m, respectively. The obtained stiffnesses are not exactly equal to the applied spring stiffnesses as was expected from the analytical study. This is due to the compliance of the actuators. The ideal actuator would have matching impedance but the compliant actuator that deforms elastically must be stiffer than the workpiece in order to compensate for the energy lost in the hinges.

8.1.2.2. Non-linear modelling. First, the actuators *optimized* using linear modelling above are *analysed* using non-linear analysis. For the actuators optimized for output spring stiffnesses of 2000 and 200 N/m there is practically no difference in the performance when using linear or non-linear analysis. For the actuator optimized for spring stiffness 20 N/m however, the beams leading to the output point buckle and the efficiency is drastically reduced (from 59% using linear analysis to 22% using non-linear analysis). The non-linear displacement patterns of the actuators optimized using linear modelling for spring stiffnesses of 200 and 20 N/m are shown in Fig. 7(a) and (c), respectively.

Figs. 7(b) and (d) show *optimized* actuators obtained using non-linear modelling and spring stiffnesses of 200 and 20 N/m, respectively. It is seen that the resulting topologies are significantly different from the ones obtained using linear modelling and the efficiencies are 87% and 68% (compared to 62% and 22% for the ones optimized using linear modelling). It is interesting to note that the efficiencies of the actuators

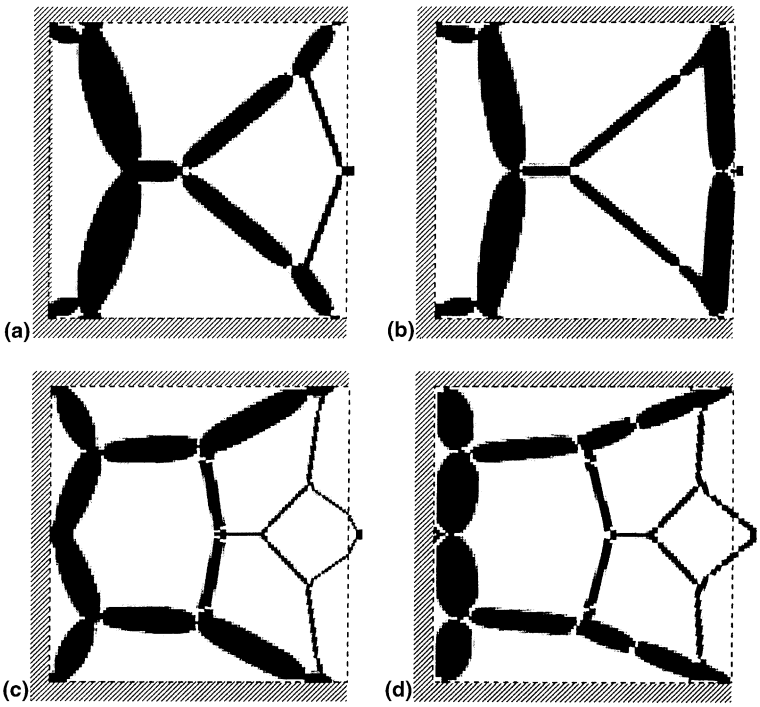

Fig. 7. Comparison of actuators optimized using linear (a) and (c) and non-linear (b) and (d) finite element analysis. (a) Non-linear displacement pattern of actuator optimized using linear FE and $k_s = 200$ N/m (efficiency is 62% (64%)); (b) non-linear displacement pattern of actuator optimized using non-linear FE and $k_s = 200$ N/m (efficiency is 87% (1%)); (c) non-linear displacement pattern of actuator optimized using linear FE and $k_s = 20$ N/m (efficiency is 22% (59%)); (d) non-linear displacement pattern of actuator optimized using non-linear FE and $k_s = 20$ N/m (efficiency is 68% (9%)). Numbers in parentheses are the efficiencies obtained for linear modelling.

optimized using non-linear modelling are only 1% and 9%, respectively, when they are modelled using linear analysis.

From this example, it can be concluded that the algorithm produces compliant thermal actuators with performances close to the theoretical limit. However, since there is no guarantee that the global optimum is found, it can be that there exist actuators with performances closer to the theoretical limit. The optimized structures shown here are the best ones obtained for varying starting guesses.

The example also proves that the use of geometrically non-linear modelling is extremely important for compliant actuator design and analysis when the displacements exceed approximately 2% of the design domain or if buckling appears in parts of the structures. Part II of this paper has a more detailed discussion on non-linear effects in thermal actuator design.

Manufacturing and tests of various thermal micro actuators designed by the algorithm have been reported in [42,43].

8.2. Electrothermal actuators

In the following, we consider the design of electrothermomechanical actuators. For manufacturing reasons, the base material is again chosen to be electro-plated Nickel (Ni) on a Silicon wafer and the thickness of the devices is 15 μm. Material properties of electro-plated Nickel are taken as: Young's modulus $E^{(0)} = 200$ GPa, thermal expansion coefficient $\alpha^{(0)} = 15 \times 10^{-6}$ K^{-1}, thermal conductivity $\sigma_1^0 = 90.7$ W/(K m), Poisson's ratio 0.31 and the electric conductivity is $\sigma_0^0 = 6.4 \times 10^6 (\Omega\ \mathrm{m})^{-1}$. These material properties are not exactly equal to the bulk values of Nickel but are found from experiments (e.g., [42,43]). The thermal conductivity of air is $k_{\mathrm{air}} = 3.13$ W/(K m). The convection from the surfaces of the structure is governed by the Nusselt number which is dependent on Reynolds number, geometry and a

characteristic length scale. For the small structures considered here, the Reynolds number is close to zero and therefore the Nusselt number can be estimated to be constant $Nu = 0.36$. A simplification was made in assuming that the individual parts of the structure are horizontal cylinders with diameter given by the characteristic length scale $D = 20$ μm. The convection coefficient from a surface can then be written as $h = k_{\text{air}}Nu/D = 56 \times 10^3$ W/(m^2 K). However, during practical tests [44] a representative convection coefficient was measured to be $h = k_{\text{air}}Nu/D = 18.7 \times 10^3$ W/(m^2 K), and therefore we use the latter value in the examples. For computational simplicity, the convection is assumed to happen only from the top surfaces. Another assumption is that radiation is ignored. A hand-calculation for a simple micro-beam shows that the heat-loss due to radiation is negligible compared to convection and conduction losses when the temperature is below 700°C.

8.2.1. Conventional actuator

As an example of design of an electrically actuated micro actuator, the problem sketched in Fig. 8(a) is considered. The design domain is rectangular with dimensions 500×200 μm. 30% of the design domain can be filled with material. The left edge of the design domain has two terminals (shown in black) where a prescribed voltage of 0.3 V is applied. A small rectangular domain (shown in white) is fixed to be void. The optimization problem consists in maximizing the force exerted on a vertical spring of stiffness 100 N/m, which is attached to the structure at the center of the right edge.

An intuitive solution to this design problem is known from the literature [45] and is shown in Fig. 9(a). This thermal actuator has become popular in the MEMS-community and is included in some standard device libraries. In the following, the electrothermal actuator in Fig. 9 will be referred to as the 'conventional' actuator. The principle of actuation in the conventional actuator is simple to explain. Due to higher

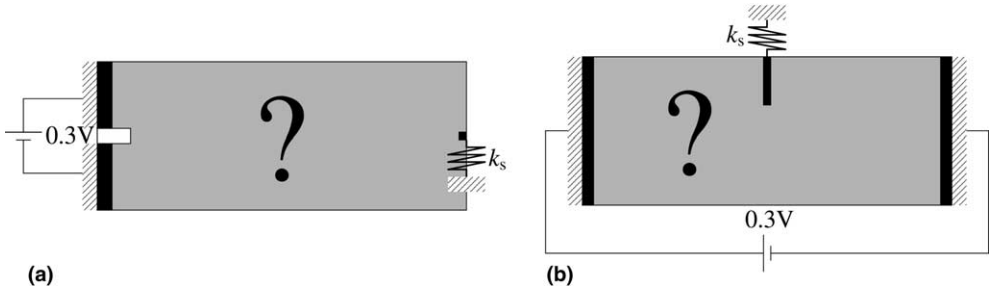

Fig. 8. Definition of the electrothermomechanical actuator design problem. (a) Both terminal on left edge and (b) terminal at opposite sides.

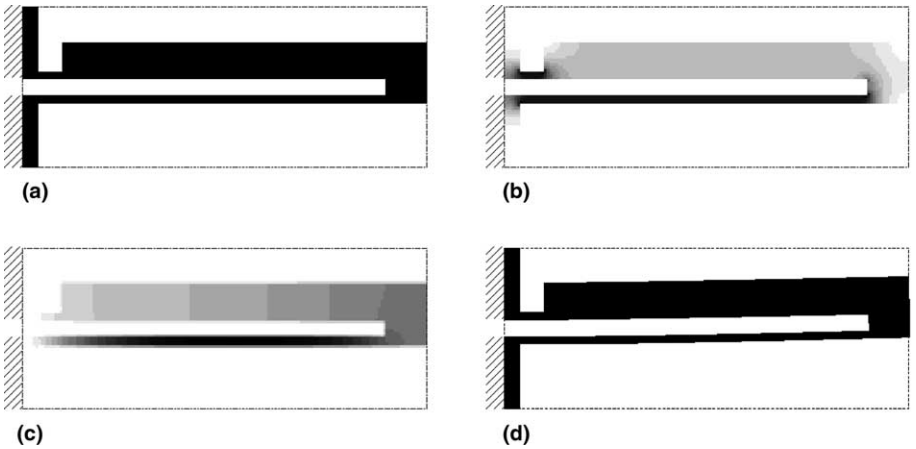

Fig. 9. Conventional electrothermal micro actuator. (a) Topology; (b) current intensity; (c) temperature distribution ($\Delta T_{\max} = 468$ K); (d) displacement field.

resistance in the lower part of the actuator, the current intensity will be higher in that part of the structure as seen in Fig. 9(b). The high current intensity in the lower part of the structure will result in increased temperature due to Joule heating as seen in Fig. 9(c). Therefore, the thin lower part will expand more than the top part and cause an upward deflection as seen in Fig. 9(d). The work performed on the spring by the 'conventional' actuator is 3.5 nJ and the resistance of the device is 0.6 Ω. Other data for the conventional actuator are listed in Table 1.

To check if a better design than the 'conventional' actuator can be obtained, the topology optimization method developed in this paper is used to solve the same problem but now allowing the material to be distributed freely in the design domain. The topology optimized actuator is shown in Fig. 10(a). It is seen that the optimal structure is almost short-circuited (Fig. 10(b)) resulting in a large temperature increase near the terminals (Fig. 10(c)). The deformation mechanism (Fig. 10(d)) is rather complex but consists mainly in a very hot beam in the left part of the structure that pushes the rest of the structure upwards by a levering effect. The work performed on the spring is 32.8 nJ or nine times more than before but the resistance is very low – namely 0.025 Ω and the maximum temperature increase is $\Delta T_{\max} = 730$ K compared to $\Delta T_{\max} = 468$ K for the conventional actuator (see also Table 1).

The very low resistance and the very high temperature of the new design (Fig. 10) makes it impractical and therefore a maximum constraint on the current is imposed on the optimization problem. Constraining the current to be less or equal to the current running through the 'conventional' actuator (Fig. 9) the optimized actuator topology shown in Fig. 11(a) is obtained. The basic functionality of the new design is the same as for the conventional, with the main difference being that the hinge of the upper part has moved to the center of the actuator. In this way, more heat is conducted away from the upper (cold) part of the structure to the supports. The output work performed on the spring by the new actuator is 6.3nJ corresponding to a more than 70% improvement.

Table 1
Data for the electrothermomechanical actuator design problem defined in Fig. 8

Figure	Example	u_{out} (μm)	F_{out} (mN)	W (nJ)	P (W)	W/P (ns)	$\Delta T_{\max}$ (K)	u_{free} (μm)	F_{block} (mN)	k_a (N/m)
Fig. 9(a)	Conventional	8.4	0.84	3.5	0.15	24.2	402	13.3	2.3	169
Fig. 10(a)		25.6	2.56	*32.8*	3.54	9.3	730	44.4	6.0	135
Fig. 11(a)	$R^* = 0.6\ \Omega$	11.2	1.12	*6.3*	*0.15*	41.8	468	17.1	3.2	187
Fig. 11(c)	max W/P	9.5	0.95	4.5	0.08	*54.0*	363	13.9	3.0	216
Fig. 14(a)		18.4	1.84	*16.8*	0.63	27.4	472	19.2	40.0	2073
–	$R^* = 0.6\ \Omega$	12.7	1.27	*8.1*	*0.15*	54.2	359	15.5	7.3	471
Fig. 14(c)	max W/P	16.3	1.63	13.3	0.23	*59.3*	338	18.6	13.2	711

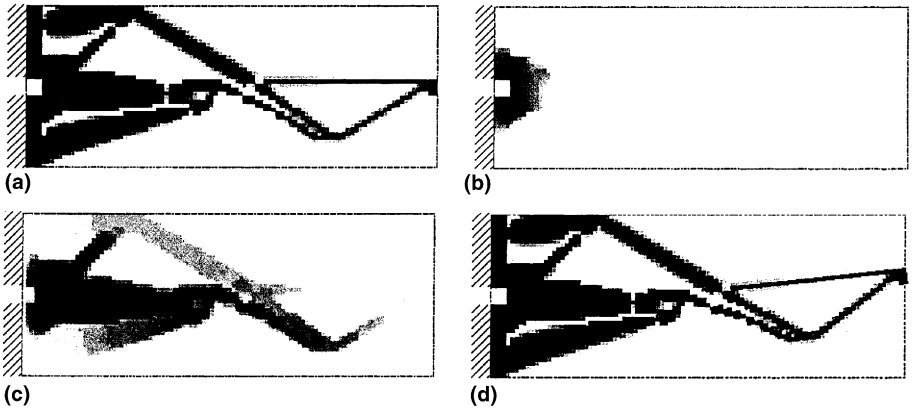

Fig. 10. (a) Optimized topology without a constraint on resistance; (b) current intensity; (c) temperature distribution; (d) displacement field.

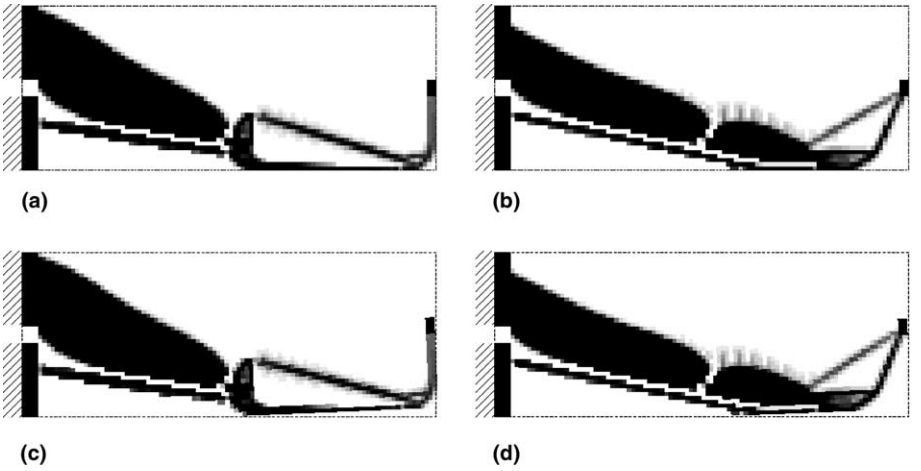

Fig. 11. (a) Optimized actuator for a constraint on resistance ($R^* = 0.6\ \Omega$); (b) displacement field; (c) actuator topology with maximum work to power ratio and (d) displacement field.

Instead of maximizing the output work with a constraint on the resistance, one can maximize the ratio between the output work and the driving power in order to get the electromechanically most efficient structure. The resulting topology for this objective is shown in Fig. 11(c). The work/power (W/P) ratio for this actuator is 54 ns compared to 24.2, 9.3 and 41.8 ns for the previous examples (see Table 1). The optimized topology, Fig. 11(c), is seen to have a long thin bar which causes a high resistance and thereby a smaller power consumption. It is also interesting to note that the optimized topology displays some 'cooling ribs' in the right part of the structure.

8.2.1.1. Non-linear effects. Although the 1 d.o.f. electrothermal actuators discussed above were optimized for rather high spring stiffnesses and had resultingly low displacements, there may be a big change in their performance when they are modelled using non-linear finite element analysis. The reason is instability or buckling of parts of the actuators.

Fig. 12(a) shows the actuator from Fig. 10 which was optimized using linear analysis. Fig. 12(b) shows the same actuator modelled using non-linear analysis. It is seen that a part of the actuator is buckled. A force/voltage diagram is shown in Fig. 13. The thin solid line shows the response of the actuator using linear modelling and the thick solid line shows the response of the actuator when using non-linear modelling. It is

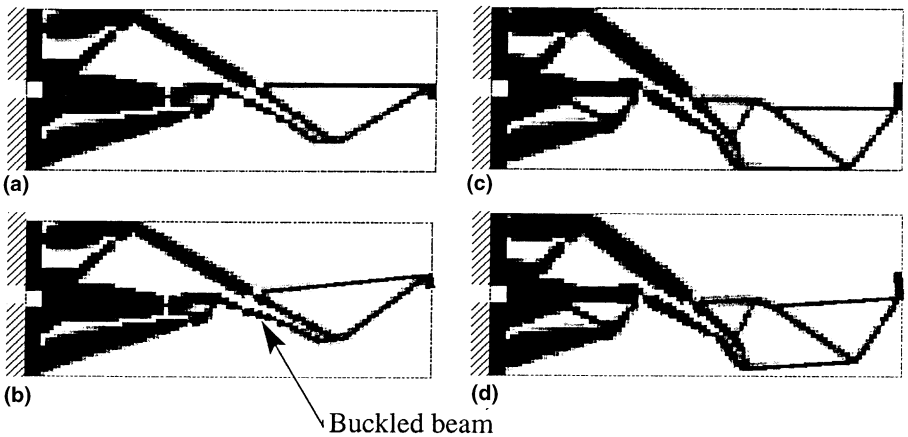

Fig. 12. Electrothermal actuators optimized using (a) linear modelling and (c) non-linear modelling. Displacement patterns for the same topologies using non-linear modelling are shown in (b) and (d), respectively.

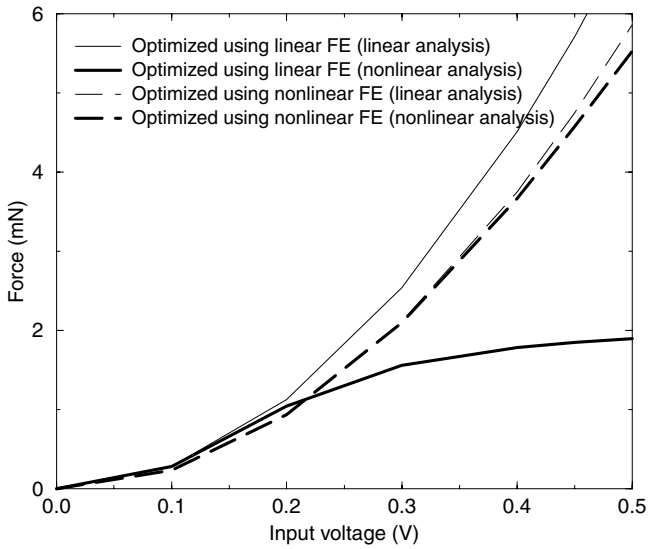

Fig. 13. Force/voltage diagram showing response of actuators optimized for linear and non-linear modelling from Figs. 12(a) and (c).

seen that the linear and non-linear output forces are equal for input voltages up to 0.2 V, but for higher voltages, the curve obtained using non-linear modelling flattens out, resulting in a maximum force of 2 mN for 0.5 V. It can be concluded that the buckling mode is non-catastrophic but simply results in decreased performance (more than three times lower than predicted by the linear analysis).

The actuator is now re-optimized using non-linear analysis. In order to ensure a stable actuator, a blocking force constraint is added to the optimization problem (see part II for the definition of blocking force constraints). The blocking force is required to be four times higher than the work force (on the output spring of stiffness 100 N/m). Otherwise all parameters are the same as before. The resulting actuator is shown in Fig. 12(c) and the non-linear displacement pattern is shown in Fig. 12(d). The force as a function of the input voltage is shown as a thick dashed line in Fig. 13 which also shows the linear response as a thin dashed line. The new actuator does not have problems with buckling and it is seen that the non-linear behaviour is very close to the linear behaviour of the actuator for input voltages less than 0.5 V.

It is once again concluded that non-linear analysis is very important in topology design of compliant actuators if large displacements or buckling effects are present. Similar conclusions hold for the other electrothermal actuators. In practice, it is costly to perform non-linear analysis for all design problems. Therefore, we usually solve a new design problem using linear analysis first. If the resulting design has large displacements or if the difference in objective functions for linear and non-linear analysis is significant, the design process is redone using non-linear modelling.

8.2.2. Electrothermal actuator with alternative terminal placement

In the previous example, the two electrical input terminals were placed on the left edge of the design domain. In this example the terminals are placed at each end of the device and the output point is at the center of the top edge as sketched in Fig. 8(b).

The optimized topologies for free minimum resistance and maximum work to power ratio are shown in Figs. 14(a) and (c), respectively. The output work of the first design is five times higher than that for the conventional actuator (Fig. 9) but only half that of the design in Fig. 10(a). However, the temperature is only 472 K compared to 730 K. The work/power ratio of the actuator in Fig. 14(c) is 59.3 ns which is approximately 10% better than the design in Fig. 11(c). From this example it is concluded that not only the topology of the devices but also the placement of the terminals influence actuator performance. Other criteria, however, may put a constraint on the placement of the terminals, but in principle one ought to include the placement of terminals in the formulation of the optimization problem.

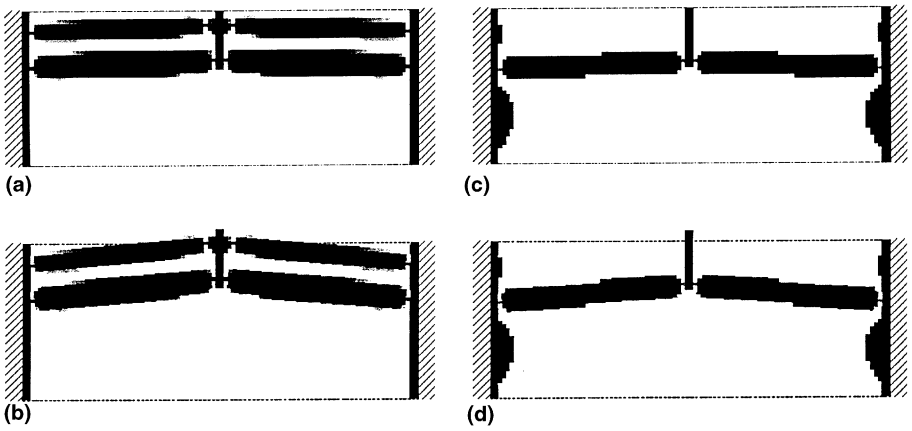

Fig. 14. Optimized topologies for opposite terminal placement and (a) maximization of output work and (c) maximization of work/power ratio. (b) and (d) show the corresponding displacement fields.

8.2.3. XY-scanning device

The next example demonstrates how actuators with multiple inputs and outputs can be designed using the proposed topology optimization method. The design problem is sketched in Fig. 15(a). The goal is to come up with a device that moves the output point (center right edge) horizontally when an electric potential is applied to the lower left terminal and vertically when an electric field is applied to the lower right terminal. The definition of the design problem was discussed in Sections 5.1 and (22) and includes seven constraints and a multi-objective function. The stiffness of the output springs are 40 N/m and the material volume to be distributed is limited to 22% of the design domain. The input voltages are 0.3 V and the currents are constraint to $I^* = 2$ A. The design domain has the dimension 500 by 400 μm^2 and is discretized using 8000 elements. It took more than 24 h to solve the (linear) problem on a fast workstation.

8.2.3.1. Linear modelling. The optimized design for linear modelling is shown in Fig. 15(b) and the (linear) displacement patterns subject to the three electrical inputs are shown in Fig. 15(e). The principle of actuation is relatively simple. For an electrical current over the left terminals, the actuator works like the 1 d.o.f. actuator from Fig. 14: the left part pushes the right part in the horizontal direction. For an electrical current over the right terminals, the two 'actuation bars' are offset resulting in an upward rotation of the output point. Even though the solution seems intuitively simple, it would be very hard, if not impossible, to come up with a solution based on simple intuition. The output displacements of the 2 d.o.f. actuator are 18, 16 and 24 μm, for the left, right and simultaneous electrical inputs, respectively.

Output displacements for input voltages that should result in the output point following a square path followed by a diagonal path are shown as a solid line in the XY-displacement diagram in Fig. 16(a). It is seen that the output point follows a straight line for proportional horizontal, vertical and diagonal outputs whereas the path is not perfectly straight for fixed (maximum) X-displacement and varying Y-displacement and vice versa. This non-linearity is due to the nonlinearity of the heating (proportional to the current intensity squared). It is also seen that the constraint on the cross-sensitivity is fulfilled – the Y-displacement for X-actuation is less than 10% of the X-displacement.

8.2.3.2. Non-linear modelling. To check whether the XY-actuator works as expected also for geometrically non-linear finite element modelling, the non-linear output path is shown as a dashed line in Fig. 16(a). Although the non-linear path does follow the linear path for very small displacements, it is seen that it does not follow at all for larger displacements. In other words, the small displacement assumption is totally invalid for this example.

The non-linear behaviour of the actuator obtained using linear modelling (Figs. 15(b) and 16(a)) is far from satisfactory although it performs very well using linear modelling. The optimization problem is

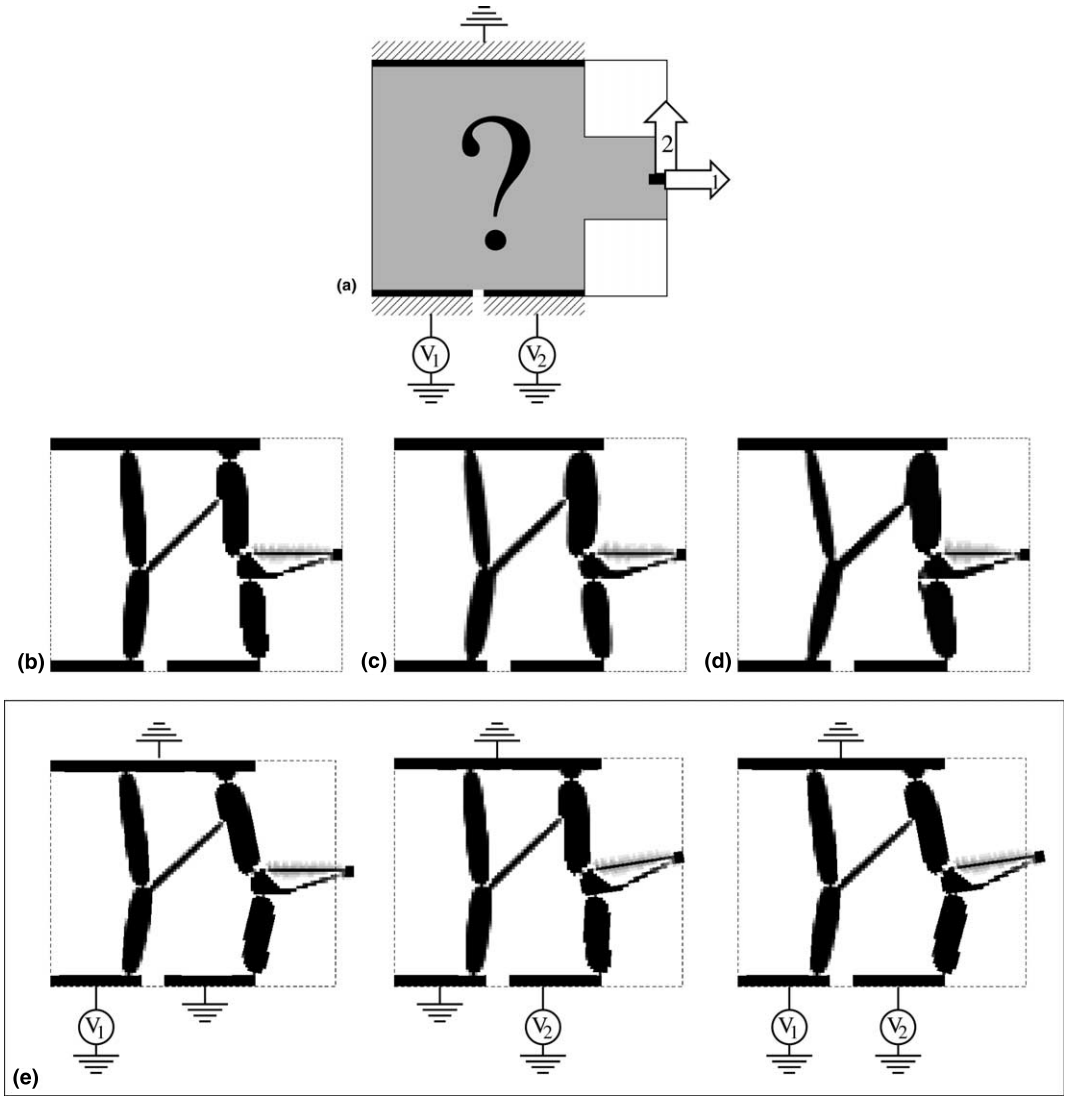

Fig. 15. Design of an *XY*-scanner where the output point moves vertically for one electrical input and horizontally for another electrical input. (a) Design problem; (b) solution for linear analysis; (c) solution for non-linear analysis; (d) solution for non-linear analysis with extra precision points; (e) horizontal, vertical and diagonal displacement modes for 0.4 V input fields. The design domain is discretized by 8000 elements.

therefore repeated using non-linear modelling. The resulting topology is shown in Fig. 15(c). At first sight, the new topology does not seem to be significantly different from the actuator optimized using linear modelling (Fig. 15(b)). However, the angles and thicknesses of bars as well as the hinge positions have changed, resulting in a higher linearity and a much lower cross-sensitivity (shown as a solid line in Fig. 16(b)).

Unfortunately, the non-linear response of the new actuator is still not fully satisfactory. The goal of the original optimization problem was to produce a 2 d.o.f. actuator with a maximum cross-sensitivity of 10%. In Fig. 16(b) it is seen that the actuator does fulfill the cross-sensitivity constraint in the corner points (precision points) of the work area (which was to be expected since the cross-sensitivity constraints were imposed only at these points). However, the cross-sensitivity exceeds 10% for lower input voltages seen as a curved response for increasing *X*-actuation. To circumvent the problem, three additional constraints (precision points) were imposed on the cross-sensitivities for *XY*-input voltages of (0 V, 0.2 V), (0.2 V, 0.2 V) and (0.2 V, 0 V). The resulting topology is shown in Fig. 15(d) and its response is shown as a solid line in

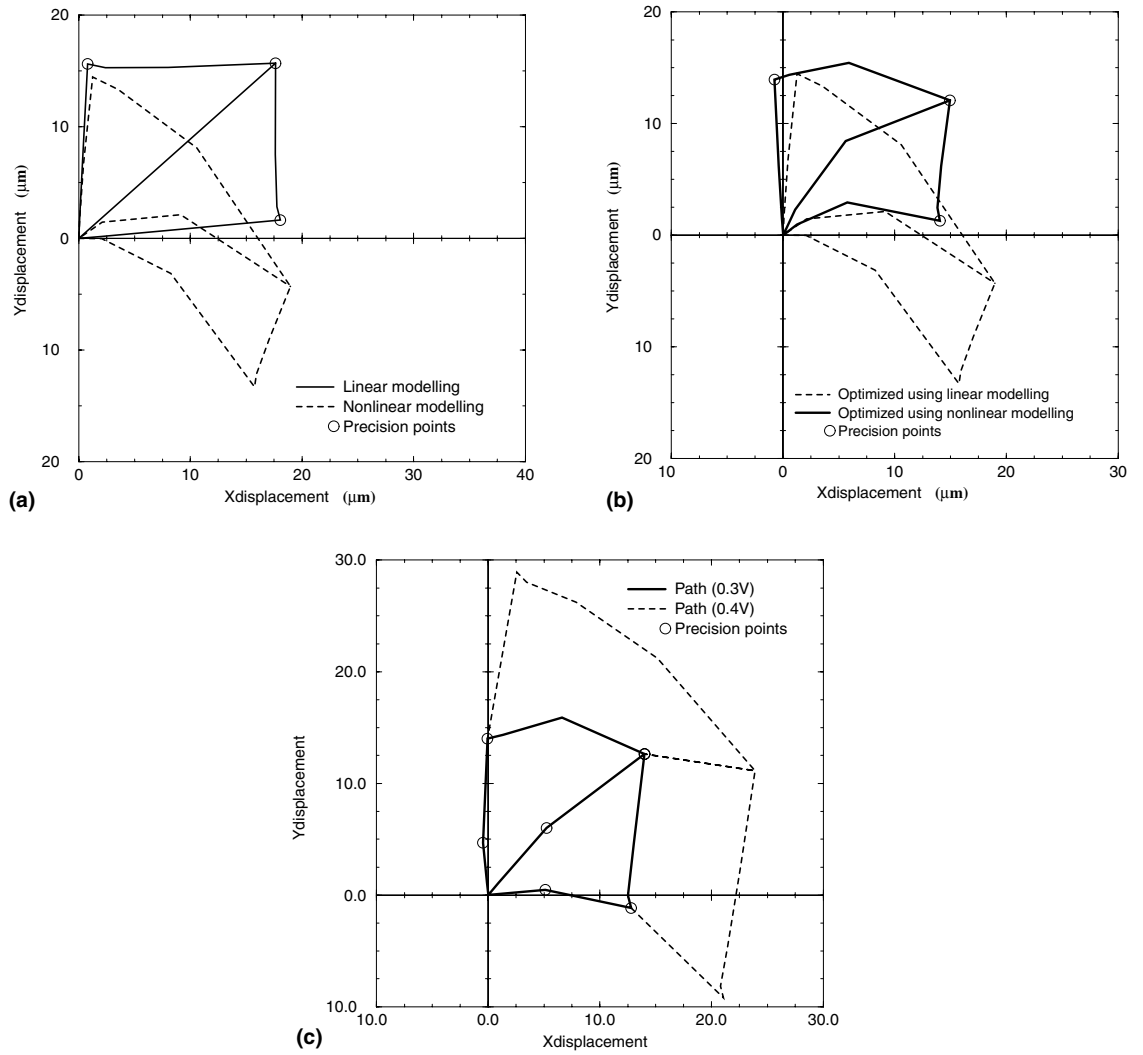

Fig. 16. Horizontal and vertical displacement of the *XY*-actuators from Fig. 15. (a) Solid line: *XY*-actuator optimized and modelled using linear modelling (Fig. 15(b)) and dashed line: the same *XY*-actuator modelled using non-linear analysis. (b) Solid line: *XY*-actuator optimized using non-linear modelling with three precision points (Fig. 15(c)) and dashed line: *XY*-actuator optimized using linear modelling (Fig. 15(b)). (c) Solid line, *XY*-actuator optimized using non-linear modelling with six precision points (Fig. 15(d)) driven with 0.3 V and dashed line: same actuator driven with 0.4 V.

Fig. 16(c). Again the main topological change consists in change of bar-areas and hinge positions and the response now fulfills the cross-sensitivity requirement also in-between the precision points.

The topology from Fig. 15(c) was optimized for input voltages up to 0.3 V. To see what happens for higher input voltages, the trace of the actuator for input voltages up to 0.4 V is shown as a dashed line in Fig. 16(c). It is seen that the cross-sensitivity for these input voltages greatly exceeds the cross-sensitivity constraints. The conclusion is that a large number of constraints (precision points) must be imposed if a linear response over a large voltage/displacement range is required.

8.2.4. Three d.o.f. device

The last example demonstrates the design of a 3 d.o.f. gripping mechanism. The design problem is sketched in Fig. 17(a) and the design goal is to move the gripping head in the horizontal direction for the first electric input, in the vertical direction for the second electric input and to open the jaws for the third electrical input. A total of 14 constraints make sure that the output motion follows the desired directions

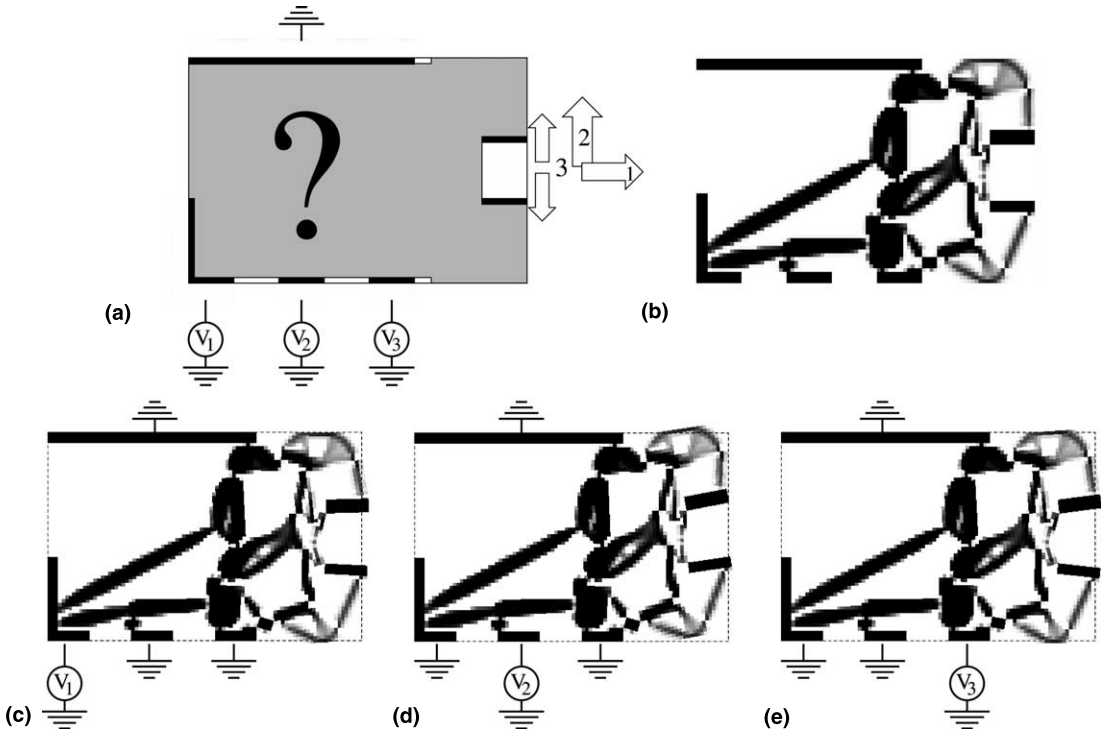

Fig. 17. Design of a 3 d.o.f. gripping mechanism: (a) design domain; (b) optimized topology and (c)–(e) displacement fields for input voltages of 0.4 V.

with small cross-sensitivities and that the current for all three electric inputs is limited to 3 A. The input voltages are 0.3 V and the spring stiffnesses are 100 N/m. A total of nine finite element problems and 24 adjoint load cases are used in the analysis. The optimized topology is shown in Fig. 17(b) and the displacement fields for three electric inputs of 0.4 V are shown in Figs. 17(c)–(e).

The resulting displacements for the three electrical inputs are 6, 11 and 8 μm (for each gripper), respectively. These displacements are quite small compared to the single d.o.f. structures discussed earlier. In general, it can be said that the maximum displacement of the actuators decreases with increasing number of constraints and load cases which was to be expected. Adding more constraints to the problem limits the ability of the actuator to move freely.

9. Manufacturing

The proposed devices can be built in micro scale using several different techniques. In order to ensure rapid prototyping, Mikroelektronik Centret (MIC) at the Technical University of Denmark has developed a laser-micromachining system that allows for the prototyping of the structures in a few days [46]. A laser beam controlled etching process with a resolution of one micrometer is used to produce a negative of the actuator on top of a silicon chip. Using different electro-plating techniques and selective etching processes, an actuator built in Nickel attached to the top of a Silicon wafer can be manufactured. For more details on this process, the reader is referred to [42–44].

The laser-micromachining requires a bitmap of the actuator with a resolution of one micrometer per. bit. Since it is impossible to do topology optimization with such a fine resolution with the current computing power (the 2 d.o.f. actuator would be discretized by $500 \times 400 = 200.000$ elements), the coarse grey-scale bitmap resulting from the topology optimization algorithm must be converted to a black and white, finer resolution picture for the lasermicromaching process. The best fine-scale bitmap is obtained by doing a contour plot of the grey-scale picture where the contour-line thresholds at relative density 0.6. Such a

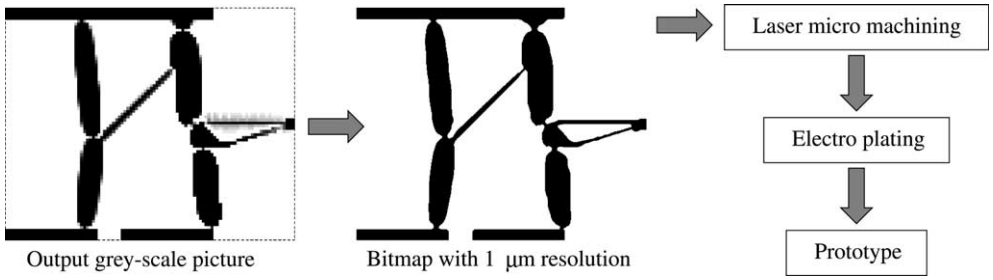

Fig. 18. Conversion of the resulting grey-scale image from the topology optimization algorithm to a 0/1 cif-bitmap image which is used as input to the laser micromachining software.

post-processed bitmap of the 2 d.o.f. actuator is shown in Fig. 18. The original grey-scale picture has a resolution of 100 by 80 whereas the post-processed 0/1 picture has a resolution of 500 by 400 which can be directly transferred to the laser micromachining control software. To verify that this post-processing does not significantly change the output behaviour of the device, a medium resolution picture (e.g., 250 by 200 elements) can be reanalysed with the software. Usually the error introduced by the post-processing is less than 10% meaning that the change in absolute output displacements compared to the displacements using the coarse grid stays within a reasonable margin.

9.1. Tests

Fig. 19(a) shows a laser-micromachined 2 d.o.f. actuator. The topology of this actuator is somewhat different from the 2 d.o.f. actuator optimized in Section 8.2.3, however, the basic functionality is the same. The actuator was tested by applying electrical inputs over the left, right and both terminal sets simultaneously and the output displacements were measured using a video-based frame grapper and image processing techniques [43,44]. A graph of the output displacements as results of various electrical input voltages is shown in Fig. 19(b). It is seen that the output point moves horizontally and vertically with less than 10% cross-sensitivity for X- and Y-actuation, respectively. However, the amplitude of the X-actuation is only half that of the Y-actuation although the difference was restricted to 10% in the optimization algorithm. The discrepancy can be explained in many different ways and is most likely due to a combination of several different effects. For example, an error is introduced in converting the grey-scale topology optimized image to the black and white fine-resolution picture. Also non-linearity effects play a role as seen in Fig. 16 and there may be errors in the assumptions on conductivity, convection and material properties. On-going simulations and experiments are used to clarify the error sources. For a more detailed discussion, the reader is referred to the thesis by Jonsmann [43].

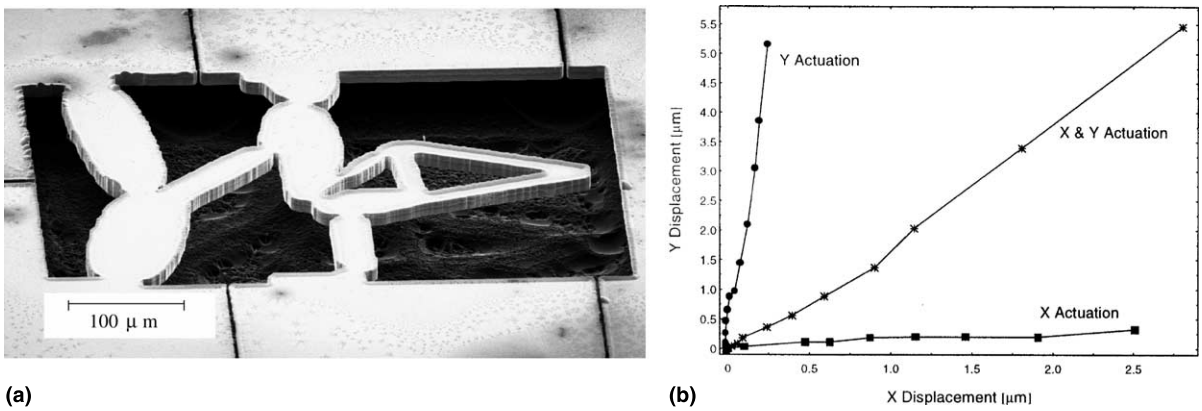

Fig. 19. (a) Micromachined 2 d.o.f. electrothermomechanical micro actuator fabricated by J. Jonsmann and S. Bouwstra at Mikroelektronik Centret (MIC), DTU, Denmark. (b) Test results for a 2 d.o.f. actuator.

10. Conclusions

The topology optimization method has been extended to be used as a systematic tool for MEMS design. The extensions include geometrically non-linear finite element analyses in multiple physical domains, adjoint sensitivity analysis and the formulation and solution of optimization problems with multiple constraints.

All the examples were done using geometrically linear and non-linear analysis. It was concluded that non-linear analysis is extremely important for cases with large displacements (exceeding 2% of the design domain) and for cases with buckling problems. In practice it is expensive to run optimizations using non-linear modelling and therefore, new problems are first solved using linear analysis. If the displacements of the resulting structures modelled using non-linear analysis are large or if the difference in objective functions for linear and non-linear modelling is significant, the design process is redone using non-linear analysis.

In order to measure the performance of the algorithm, a theoretical bound for the work of an ideal thermal actuators was established. The compliant topology optimized actuators all had performances above 60% of the theoretical limit. The discrepancy is mainly due to elastic energy taken up in the hinges of the compliant mechanisms but may also be explained by convergence to local minima.

A way to improve performance of the actuators could be to introduce a second material in the design domain. The second material could have a different thermal expansion coefficient which would allow for bimorph actuators or it could have a lower conductivity that would allow to isolate different pairs of the structure from each other. These possibilities are explored in part II of this paper.

Further conclusions and perspectives will also be given in part II of this paper.

Acknowledgements

The author is grateful to Krister Svanberg for supplying the MMA-optimization subroutines, to Thomas Buhl and Claus B. W. Pedersen for providing subroutines for non-linear analysis and to Martin P. Bendsøe, Pauli Pedersen and Daniel A. Tortorelli for helpful discussions. The author also gratefully acknowledges the work of Jacques Jonsmann and Siebe Bouwstra at Mikroelektronik Centret (MIC), Technical University of Denmark, who produced the micromachined devices. The work presented in this paper received support from the Danish Technical Research Council through the THOR/Talent-programme: Design of MicroElectroMechanical Systems (MEMS).

References

[1] M.P. Bendsøe, Optimization of Structural Topology, Shape and Material, Springer, Berlin, 1995.
[2] H. Rodriques, P. Fernandes, A material based model for topology optimization of thermoelastic structures, Int. J. Numer. Methods Engrg. 38 (1995) 1951–1965.
[3] O. Sigmund, Design of thermomechanical actuators using topology optimization, in: W. Gutkowski, Z. Mroz (Eds.), Second World Congress on Structural and Multidiciplinary Optimization, Zakopane, Poland, May, IFTR, 1997, pp. 393–398.
[4] O. Sigmund, S. Torquato, Composites with extremal thermal expansion coefficients, Appl. Phys. Lett. 69 (21) (1996) 3203–3205.
[5] O. Sigmund, S. Torquato, Design of materials with extreme thermal expansion using a three-phase topology optimization method, J. Mech. Phys. Solids 45 (6) (1997) 1037–1067.
[6] E.C.N. Silva, J.S.O. Fonseca, N. Kikuchi, Optimal design of piezoelectric microstructures, Comput. Mech. 19 (5) (1997) 397–410.
[7] E.C.N. Silva, J.S.O. Fonseca, N. Kikuchi, Optimal design of periodic piezocomposites, Comput. Methods Appl. Engrg. 159 (1998) 49–77.
[8] O. Sigmund, S. Torquato, I.A. Aksay, On the design of 1–3 piezocomposites using topology optimization, J. Mater. Res. 13 (4) (1998) 1038–1048.
[9] O. Sigmund, S. Torquato, L.V. Gibiansky, I.A. Aksay, On the design of hydrophones made as 1–3 piezoelectrics, in: Y.A. Bahei-El Din, G.J. Dvorak (Eds.), IUTAM Symposium on Transformation Problems in Composite and Active Materials, IUTAM, Kluwer Academic Publishers, Dordrecht, 1998, pp. 147–159.
[10] O. Sigmund, On the optimality of bone microstructure, in: P. Pedersen, M.P. Bendsøe (Eds.), Synthesis in Bio Solid Mechanics, IUTAM, Kluwer Academic Publishers, Dordrecht, 1999, pp. 221–234.
[11] O. Sigmund, Systematic design of electrothermomechanical microactuators using topology optimization, in: Modelling and Simulation of Microsystems, Semiconductors, Sensors and Actuators, MSM98, Santa Clara, CA, 1998, pp. 350–355.

[12] O. Sigmund, Systematic design of micro actuators using topology optimization, in: V.K. Varadan, P.J. McWhorter, R.A. Singer, M.J. Vellekoop (Eds.), SPIE's 5th Annual International Symposium on Smart Structures and Materials, Smart Electronics and MEMS, San Diego, CA, vol. 3328, March 1998, SPIE, pp. 23–31.
[13] C.S. Jog, Distributed-parameter optimization and topology design for nonlinear thermoelasticity, Comput. Methods Appl. Mech. Engrg. 132 (1-2) (1997) 117–134.
[14] R. Kemmler, S. Schwarz, E. Ramm, Topology optimization including geometrically nonlinear response, in: 3rd WCSMO, Niagara Falls, May 1999.
[15] T. Buhl, C.B.W. Pedersen, O. Sigmund, Stiffness design of geometrically non-linear structures using topology optimization, Struct. Multidisciplinary Optim. 19 (2) (2000) 93–104.
[16] T.E. Bruns, D.A. Tortorelli, Topology optimization of nonlinear elastic structures and compliant mechanisms, Comput. Methods Appl. Mech. Engrg. 190 (2001) 3443–3459.
[17] T.E. Bruns, D.A. Tortorelli, An element removal and reintroduction strategy for the topology optimization of structures and compliant mechanisms, Int. J. Numer. Methods Engrg., accepted for publication.
[18] A. Saxena, G.K. Ananthasuresh, Topology synthesis of compliant mechanisms for nonlinear force-deflection and curved path specification, Trans. ASME – R – J. Mech. Design 123 (1) (2001) 33–42.
[19] C.B.W. Pedersen, T. Buhl, O. Sigmund, Topology synthesis of large-displacement compliant mechanisms, Int. J. Numer. Methods Engrg. 50 (2001) 2683–2705.
[20] K.E. Petersen, Silicon as a mechanical material, Proc. IEEE 70 (5) (1982) 420–457.
[21] G.K. Ananthasuresh, S. Kota, Y. Gianchandani, A methodical approach to the design of compliant micromechanisms, in: Solid-state Sensor and Actuator Workshop, 1994, pp. 189–192.
[22] U.D. Larsen, O. Sigmund, S. Bouwstra, Design and fabrication of compliant mechanisms and material structures with negative Poisson's ratio, J. Microelectromech. Syst. 6 (2) (1997) 99–106.
[23] O. Sigmund, Some inverse problems in topology design of materials and mechanisms, in: D. Bestle, W. Schielen (Eds.), Symposium on Optimization of Mechanical Systems, IUTAM, Kluwer Academic Publishers, Dordrecht, 1996, pp. 277–284.
[24] O. Sigmund, On the design of compliant mechanisms using topology optimization, Mech. Struct. Mach. 25 (4) (1997) 495–526.
[25] N.L. Pedersen, Design of cantilever probes for atomic force microscopy (AFM), Engrg. Optim. 32 (2000) 373–392.
[26] N.L. Pedersen, Maximization of eigenvalues using topology optimization, Struct. Multidisciplinary Optim. 20 (1) (2000) 2–11.
[27] M.P. Bendsøe, N. Kikuchi, Generating optimal topologies in optimal design using a homogenization method, Comput. Methods Appl. Mech. Engrg. 71 (1988) 197–224.
[28] M.P. Bendsøe, Optimal shape design as a material distribution problem, Struct. Optim. 1 (1989) 193–202.
[29] M. Zhou, G.I.N. Rozvany, The COC algorithm, Part II: topological, geometry and generalized shape optimization, Comput. Methods Appl. Mech. Engrg. 89 (1991) 197–224.
[30] H.P. Mlejnek, Some aspects of the genesis of structures, Struct. Optim. 5 (1992) 64–69.
[31] M.P. Bendsøe, O. Sigmund, Material interpolations in topology optimization, Arch. Appl. Mech. 69 (1999) 635–654.
[32] M.A. Crisfield, Non-linear Finite Element Analysis of Solids and Structure, vols. 1–2, Wiley, England, 1997.
[33] O.C. Zienkiewicz, The Finite Element Method, McGraw-Hill, New York, 1972.
[34] Z. Hashin, S. Shtrikman, A variational approach to the theory of the elastic behaviour of multiphase materials, J. Mech. Phys. Solids (1963) 127–140.
[35] P. Michaleris, D.A. Tortorelli, C. Vidal, Tangent operators and design sensitivity formulations for transient non-linear coupled problems with applications in elastoplasticity, Int. J. Numer. Methods Engrg. 37 (14) (1994) 2471–2499.
[36] M. Kleiber, H. Antunez, T.D. Hien, P. Kowalczyk, Parameter Sensitivity in Nonlinear Mechanics: Theory and Finite Element Computations, Wiley, Sussex, UK, 1997.
[37] C. Fleury, CONLIN: an efficient dual optimizer based on convex approximation concepts, Struct. Optim. 1 (1989) 81–89.
[38] K. Svanberg, The method of moving asymptotes – a new method for structural optimization, Int. J. Numer. Methods Engrg. 24 (1987) 359–373.
[39] O. Sigmund, J. Petersson, Numerical instabilities in topology optimization: a survey on procedures dealing with checkerboards, mesh-dependencies and local minima, Struct. Optim. 16 (1) (1998) 68–75.
[40] O. Sigmund, Design of material structures using topology optimization, Ph.D. thesis, Department of Solid Mechanics, Technical University of Denmark, December 1994.
[41] N.M Sevak, C.W. McLarnan, Optimal synthesis of flexible link mechanisms with large static deflection, ASME Paper No. 74-Det-83, 1974.
[42] J. Jonsmann, O. Sigmund, S. Bouwstra, Compliant thermal microactuators, Sensors Actuators 76 (1999) 463–469.
[43] J. Jonsmann, Technology development for topology optimized thermal microactuators, Ph.D. thesis, Technical University of Denmark, Mikroelektronik Centret (MIC), December 1999.
[44] J. Jonsmann, O. Sigmund, S. Bouwstra, Compliant electro-thermal microactuators, in: MEMS'99, 1999, pp. 588–593.
[45] J.H. Comtois, V. M. Bright, Thermal microactuators for surface-micromachining processes, in: Proceedings of SPIE, vol. 2642, 1995, pp. 10–21.
[46] M. Müllenborn, M. Heschel, U.D. Larsen, H. Dirac, S. Bouwstra, Laser direct etching of silicon on oxide for rapid prototyping, J. Micromech. Microengrg. 6 (1996) 49–51.