

# Optimization of nuclear fuel reloading by the homogenization method

G. Allaire and C. Castro

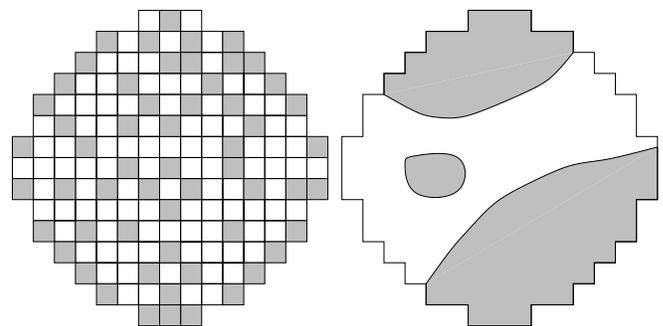
**Abstract** In this paper we apply the homogenization method to the optimization of the position of fuel assemblies in a nuclear reactor core. For this type of problem the state equation is a system of diffusion equations for the neutron flux. Homogenization theory allows us to relax a truly discrete optimization problem into a continuous and well-posed optimization problem. The latter one is solved by using classical methods of optimal control. A discrete admissible distribution of assemblies is recovered by a numerical penalization technique. The main advantage of homogenization is that the resulting reloading pattern is guaranteed to be near optimal whatever the initial guess.

**Key words** homogenization, optimal design, nuclear fuel reloading

## 1 Introduction

This paper is devoted to the application of the homogenization method (see e.g. Allaire 2001; Bendsøe 1995; Cherkaev 2000; Murat and Tartar 1985) to a classical optimization problem in nuclear reactor engineering: the so-called optimal fuel reloading problem. We briefly describe this problem and its physical context (the interested reader can consult e.g. Ho and Rohach 1982; Levine 1986). In most nuclear reactor cores, the nuclear fuel is made of a few hundreds of so-called assemblies, periodically distributed in a plane cross-section of the core

(typically 157 in a 900 MW Pressurized Water Reactor, see the left part of Fig. 1). All assemblies have the same squared cross-section but possibly different physical properties. Each assembly is itself a heterogeneous medium (made by a regular array of uranium fuel pins and control rods immersed in water), but for the sake of simplicity it is modelled as an equivalent homogeneous medium (this is common practice for this type of problem). During the fission process, the fissile isotope of uranium is consumed. This effect, called depletion, progressively decreases the efficiency of the nuclear fuel. Therefore, old assemblies must be changed periodically by new ones (the period, also called a cycle, is about a few months). The difficulty is that the fuel depletion is not spatially uniform in the core. Therefore, only part of the old assemblies (typically one fourth) are removed at the end of each cycle. Furthermore, it is not a good policy to put the new assemblies exactly at the location of the removed ones. It is better, for efficiency reasons, to optimize the position of each type of assemblies. In other words, the fuel re-loading process is not only the replacement of used assemblies by fresh ones but is also a rearrangement of all the assemblies in the core in order to maintain the maximal performance of the reactor. As such, it is a discrete optimization problem which is very difficult for at least three reasons. First, the large number of assemblies yields a huge number of possible combinations. Second, each performance evaluation of an assembly distribution involves the numerical solu-



**Fig. 1** A discrete (left) and a continuous (right) configuration of two types of assemblies in a 900 MW PWR nuclear reactor core (having 157 assemblies)

Received March 26, 2001

G. Allaire<sup>1</sup> and C. Castro<sup>2</sup>

<sup>1</sup> Centre de Mathématiques Appliquées, Ecole Polytechnique, 91128 Palaiseau, France

e-mail: allaire@cmapx.polytechnique.fr

<sup>2</sup> Departamento de Matemática e Informática Aplicadas a la Ingeniería Civil, ETSI Caminos, Canales y Puertos, Universidad Politécnica de Madrid, 28040 Madrid, Spain

e-mail: ccastro@dumbo.caminos.upm.es

tion of a diffusion problem for the neutron flux by using finite elements. Third, this problem lacks any convexity properties.

There are many numerical methods proposed in the literature for solving this discrete optimization problem. Most of them are based on linear programming, simulated annealing, neural networks or genetic algorithms (Ho and Rohach 1982; Kropaczek and Turinsky 1991; Lysenko *et al.* 1999a,b; Maldonado and Turinsky 1995; Parks 1996). However, the huge number of possible permutations, the nonconvexity of the objective function make it a very hard problem to solve and no existing method is fully satisfying. In a previous paper (Allaire and Castro 2001) the authors proposed to apply the homogenization method to this problem. The homogenization method has been very successful in structural optimization (see e.g. Allaire 2001; Bendsøe 1995; Cherkaev 2000; Rozvany *et al.* 1995) and we hope to demonstrate that it can also be very efficient for the fuel re-loading optimization problem. In structural design the homogenization method is regarded as a method for topology optimization, which is not incompatible, but rather complementary, with other classical methods. Likewise in the present setting, our approach should be taken as a topology optimizer, i.e. whatever the starting configuration, it is able to find a quasi-optimal distribution of assemblies, possibly very remote from the starting one. The homogenization method is not a concurrent of other methods, but rather a pre-processor, since its final output could still be refined by these methods.

The main difference to our previous work (Allaire and Castro 2001) is that here we treat the true physical problem which is a system of two coupled diffusion equations (the so-called multi-group neutron diffusion) while (Allaire and Castro 2001) considered the simplified model of one-group neutron diffusion which is a single equation of diffusion. In the present case the mathematical, as well as numerical, difficulties are much more severe. This difference is somehow similar to that between conductivity and elasticity problems in structural optimization. In truth, we do not have a fully explicit relaxation of the two-group diffusion system, and we content ourselves with a partial relaxation. On the contrary the one-group diffusion equation is completely understood, and we refer to Allaire and Castro (2001) for all mathematical details. In order to simplify the exposition we shall not dwell too much on mathematical technicalities and focus rather on the physical and numerical aspects of the problem.

Finally, we conclude this introduction by a brief description of the content of this paper. In Sect. 2, we describe the original discrete optimization problem. Section 3 is devoted to its relaxation which is done in two steps. First, the discrete variables are extended into continuous ones by transforming the original problem into a shape optimization problem (i.e. assemblies can have any shape and size). Second, this continuous shape opti-

mization problem is homogenized by introducing composite designs which are fine mixtures of the original phases. Section 4 is concerned with optimality conditions. Numerical results are finally presented in Sect. 5.

## 2 Description of the problem

In order to give a precise mathematical statement of the optimization problem we are interested in, we first describe the state equation that models the fission process in a nuclear reactor and allows to quantify the efficiency of the assemblies distribution. The power distribution in a nuclear reactor core is usually obtained by solving the so-called criticality eigenvalue problem for a diffusion system of two equations (corresponding to two energy groups of neutrons). Considering more groups, or equivalently a system with more equations, does not increase the difficulty (while a single equation is much simpler, see Allaire and Castro 2001). In a steady-state regime, the criticality problem gives the balance between neutrons produced by fission and neutrons absorbed or diffused by the medium. Denoting by  $\Omega$  the radial section of the core ( $\Omega \subset \mathbb{R}^2$  is a bounded domain in the plane with boundary  $\partial\Omega$ ), the state equation is

$$\begin{cases} -\operatorname{div}(D_1 \nabla u_1) + \Sigma_1 u_1 = \lambda(\sigma_1 u_1 + \sigma_2 u_2) & \text{in } \Omega, \\ -\operatorname{div}(D_2 \nabla u_2) + \Sigma_2 u_2 = \sigma_r u_1 & \text{in } \Omega, \\ u_1 = u_2 = 0 & \text{on } \partial\Omega. \end{cases} \quad (1)$$

Here  $\lambda$  is the first eigenvalue and  $(u_1, u_2)$  the first eigenvector of this system of two coupled equations. The first component  $u_1$  denotes the flux of fast neutrons (with highest kinetic energy), while  $u_2$  is the flux of slow (or thermal) neutrons (with lowest kinetic energy). Apart from the classical diffusion terms with coefficients  $D_1$  and  $D_2$ , the terms  $\Sigma_1 u_1$  and  $\Sigma_2 u_2$  model absorption,  $\sigma_r u_1$  is a collision term (fast neutrons loose kinetic energy during inelastic collisions), and  $\sigma_1 u_1 + \sigma_2 u_2$  is the fission (or production) term (fast neutrons are produced when neutrons hit fissile isotopes). The eigenvalue  $\lambda$  in (1) is therefore interpreted as a balance coefficient between dissipation on the left-hand side and production on the right-hand side (for more details, see e.g. Planchard 1995).

More precisely, the eigenvalue  $\lambda$  measures the criticality of the reactor in a quasistatic limit. If  $\lambda = 1$ , the reactor is said to be critical and can safely be operated: a perfect balance between production and removal of neutrons takes place. If  $\lambda > 1$ , too many neutrons are diffused or absorbed in the core compared to their production by fission: the nuclear chain reaction dies out, and the reactor, being subcritical, cannot operate. If  $\lambda < 1$ , too many neutrons are created by fission, and the reactor, being supercritical, can nevertheless be operated by introducing the control rods (absorbing media) in the core.

Since different types of nuclear fuel are present in the reactor, the coefficients  $D_\alpha$ ,  $\Sigma_\alpha$ ,  $\sigma_\alpha$ ,  $\sigma_r$  ( $\alpha = 1, 2$ ) in (1) are merely bounded and piecewise smooth (but discontinuous) functions. We assume that they satisfy for  $x \in \Omega$

$$\begin{aligned} \Sigma_\alpha(x), \sigma_2(x), \sigma_r(x) &\geq 0, \\ \sigma_1(x) &\geq \sigma_0 > 0, \quad D_\alpha(x) \geq d_0 > 0, \quad \alpha = 1, 2. \end{aligned} \quad (2)$$

Note that (1) is not a self-adjoint system, so the existence of eigenvalues and eigenfunctions is not guaranteed. However, since the coupling of the two equations in (1) is made by zero-order terms only, it satisfies a maximum principle and a Krein–Rutman theorem, i.e. there exists at least one eigenvalue (the smallest one) with a positive eigenfunction. This is a classical result that we recall now (see e.g. Habetler and Martino 1961; Planchard 1995).

**Theorem 1.** *There exists a solution of (1) such that the eigenvalue  $\lambda$  is real, positive, and is the smallest eigenvalue in modulus. Furthermore this eigenvalue is simple and the associated eigenfunction  $(u_1, u_2)$  can be chosen non-negative, i.e.  $u_1, u_2 \geq 0$  in  $\Omega$ , and this is the only eigenfunction which does not change sign.*

*Remark 1.* The only solutions of (1) which have a physical meaning are those for which the eigenfunction  $(u_1, u_2)$  are positive (a necessary feature to be the density functions of neutrons). From now on, we denote by  $(\lambda; u_1, u_2)$  the only solution of (1) with this property (which is called the first eigencouple). Of course,  $(u_1, u_2)$  is unique only up to a multiplicative constant. Thus, (1) gives only the spatial distribution of the neutron flux but not its intensity since the solution is defined up to a multiplicative constant.

In a second step we describe the objective function of the fuel reloading optimization problem. The power distribution is defined as the energy released by fission in the nuclear core: it is therefore proportional to  $\sigma_1 u_1 + \sigma_2 u_2$ . For safety reasons, the power distribution should be as uniform as possible. Indeed, at peak points of the power distribution, the surrounding flow of water could be unable to cool down the fuel pins, yielding a strong increase of the temperature that may eventually cause damage in the assembly. A major issue for safety is thus to have the most uniform power distribution in the core. This can be achieved by minimizing the  $L^r(\Omega)$  norm of  $\sigma_1 u_1 + \sigma_2 u_2$  with  $1 < r < +\infty$  (the largest  $r$ , the closest it is to the maximal value). Since  $(u_1, u_2)$  is defined up to a multiplicative constant, we normalize this  $L^r(\Omega)$  norm by dividing it by the  $L^1(\Omega)$  norm. On the other hand, a reactor can produce energy if its criticality eigenvalue  $\lambda$  is equal to or smaller than 1. However, as time goes by, the fuel depletion has a tendency to increase this eigenvalue. Therefore, at the beginning of a cycle it is highly desirable to have the smallest possible value of  $\lambda$  (or criticality reserve), ensuring that the reactor will be working for the longest possible time. In general these two objectives are contradictory. Therefore, introducing

two positive Lagrange multipliers  $\ell_1, \ell_2 \geq 0$ , our objective function is

$$\min \left\{ \ell_1 \lambda + \ell_2 \frac{\left( \mathcal{M}(|\sigma_1 u_1 + \sigma_2 u_2|^r) \right)^{1/r}}{\mathcal{M}(\sigma_1 u_1 + \sigma_2 u_2)} \right\}, \quad (3)$$

where  $\mathcal{M}$  denotes the average operator in  $\Omega$

$$\mathcal{M}(f) = \frac{1}{\text{vol}(\Omega)} \int_{\Omega} f(x) dx. \quad (4)$$

In practice, there are other constraints and requirements for fuel reloading optimization that we neglect in order to simplify the exposition. In particular, we optimize the assemblies distribution just for one cycle, regardless of what may happen afterwards, and we do not take into account the possibility of rotating the assemblies. We also do not try to minimize the production of undesirable isotopes or species in the fission process. For more information on the actual constraints and objectives, we refer e.g. to Levine (1986).

To finish the mathematical statement of our optimization problem, the third step is to define a space of admissible configurations  $\mathcal{U}_{ad}$  of assemblies in the core. Then, the minimization of the objective function (3) takes place in this space  $\mathcal{U}_{ad}$ . We assume that there are a number  $I$  of different types of assemblies (called phases or components in the sequel) characterized by positive constant coefficients  $(D_\alpha^i, \Sigma_\alpha^i, \sigma_\alpha^i, \sigma_r^i)$  with  $\alpha = 1, 2$  and  $i = 1, 2, \dots, I$ , given in prescribed proportions  $\gamma_i \geq 0$  with

$$\sum_{i=1}^I \gamma_i = \text{vol}(\Omega). \quad (5)$$

A typical value of  $I$  that we use in this paper is  $I = 4$  (the case  $I = 2$  is much simpler but not realistic, while  $I = 4$  is generic and not much easier than any  $I \geq 3$ ). The reason for taking  $I = 4$  is that only one quarter of the assemblies are removed at the end of each cycle. Therefore, there are basically 4 types of assemblies with different ages (or so-called burn up histories). Each of them has thus different coefficients. We make no special assumptions on the ordering of the physical properties of the assemblies, although physically speaking the freshest fuel produce the smallest criticality eigenvalue  $\lambda$ . Finally, since all assemblies have the same size, the core  $\Omega$  contains a finite number of them (see the left part of Fig. 1). Thus,  $\mathcal{U}_{ad}$  is the finite (but very large) set of all possible permutations of these assemblies.

### 3 Relaxation of the problem

As is well-known, integer programming problems are difficult to solve, and a common procedure is to replace integer variables by real ones. This is also our strategy

but we add a new ingredient, namely homogenization. In order to solve the discrete optimization problem (3), we propose to relax it, i.e. to generalize it by transforming it into a continuous problem. This relaxation process is performed in two steps: first, we transform this discrete problem in a continuous one by allowing for any size and shape of assemblies, second, we introduce homogenized designs that are a mixture of the different assembly types in varying proportions.

The first step amounts to change the discrete variables into continuous ones by removing any size and shape constraints on the assemblies which are no longer squares (see the right part of Fig. 1). In other words, we keep the prescribed amount of fuel types (or phases), but it can now be placed in the core as freely as we want, and its repartition does not necessarily follow an assembly pattern. This idea of passing from discrete unknowns to continuous ones is not new, and it has the advantage of being more tractable from a numerical standpoint. In this continuous optimization problem, the unknowns are now the subdomains  $\Omega_i$  of  $\Omega$  occupied by material  $i$  which satisfy the obvious constraints

$$\begin{aligned} \Omega_i \cap \Omega_j &= 0, \quad \text{when } i \neq j, & \cup_{i=1}^I \Omega_i &= \Omega, \\ \text{vol}(\Omega_i) &= \gamma_i, \quad i = 1, \dots, I. \end{aligned} \quad (6)$$

However, the shape of  $\Omega_i$  is totally free.

Introducing the characteristic functions  $(\chi_i)_{i=1\dots I}$  of these subsets  $(\Omega_i)_{i=1\dots I}$ , defined by  $\chi_i(x) = 1$  if  $x \in \Omega_i$  and  $\chi_i(x) = 0$  if  $x \notin \Omega_i$ , the coefficients of (1) are given by

$$\begin{cases} D_\alpha(x) = \sum_{i=1}^I d_\alpha^i \chi_i(x), & \alpha = 1, 2, \\ \Sigma_\alpha(x) = \sum_{i=1}^I \Sigma_\alpha^i \chi_i(x), & \alpha = 1, 2, \\ \sigma_\alpha(x) = \sum_{i=1}^I \sigma_\alpha^i \chi_i(x), & \alpha = 1, 2, r. \end{cases} \quad (7)$$

The space of admissible continuous configurations is thus defined by

$$\begin{aligned} \mathcal{U}_{ad}^c &= \{ \chi = (\chi_i)_{1 \leq i \leq I} \in L^\infty(\Omega; \{0, 1\})^I \text{ such that} \\ &\left. \begin{aligned} &\chi_i \chi_j = 0, \quad i \neq j, \\ &\sum_{i=1}^I \chi_i = 1, \\ &\int_\Omega \chi_i = \gamma_i \end{aligned} \right\}. \end{aligned} \quad (8)$$

Here  $L^\infty(\Omega; \{0, 1\})$  is the space of measurable functions taking only the values 0 or 1. The fuel reloading optimization problem is reduced to find the minimizer of

$$\min_{\chi \in \mathcal{U}_{ad}^c} \left\{ J(\chi) = \ell_1 \lambda + \ell_2 \frac{\left( \mathcal{M}(|\sigma_1 u_1 + \sigma_2 u_2|^r) \right)^{1/r}}{\mathcal{M}(\sigma_1 u_1 + \sigma_2 u_2)} \right\} \quad (9)$$

where  $(\lambda; u_1, u_2)$  is the solution of (1),  $\mathcal{M}$  is the averaging operator in  $\Omega$  defined by (4),  $1 < r < +\infty$ , and the coefficients of (1) are given by (7).

It turns out that the continuous optimization problem (9) is ill-posed in the sense that it does not admit a solution in the space  $\mathcal{U}_{ad}^c$  of all possible continuous distributions of the  $I$  materials (this is a classical difficulty in shape optimization, see Allaire 2001). The reason is that minimizing sequences of almost optimal configurations exhibit very fine mixture of the  $I$  components. On a macroscopic scale these mixtures are composite materials having effective properties different from that of its phase constituents. Their effective or averaged coefficients are found by using homogenization theory.

Therefore, in a second step the continuous optimization problem (9) is further relaxed by enlarging the space of admissible designs, namely by allowing for composite materials obtained by mixing microscopically the  $I$  different fuels. This is the basis of the homogenization method. It has the effect of making the problem well-posed, and to yield very efficient numerical algorithm for computing optimal solutions. We now describe these composite materials in very loose terms: everything can be rigorously justified by homogenization theory and this has been done in this context in our previous work (Allaire and Castro 2001). These composite materials are characterized by the local proportions of each phase, denoted by  $\theta(x) = (\theta_1(x), \dots, \theta_I(x))$ , and by their effective diffusions  $(D_1^*(x), D_2^*(x))$  which depend on their microscopic geometric arrangement. Of course, the proportions satisfy the volume constraints

$$\sum_{i=1}^I \theta_i(x) = 1, \quad \int_\Omega \theta_i(x) dx = \gamma_i, \quad 0 \leq \theta_i(x) \leq 1. \quad (10)$$

It should be emphasized that the  $\theta_i$ 's are usually no longer characteristic functions, but rather densities taking their values in the full range  $[0, 1]$ . Apart from the effective diffusions, the other homogenized coefficients are defined by simple volume averages

$$\begin{aligned} \overline{\Sigma}_\alpha(x) &= \sum_{i=1}^I \theta_i(x) \Sigma_\alpha^i, \quad \alpha = 1, 2, \\ \overline{\sigma}_\alpha(x) &= \sum_{i=1}^I \theta_i(x) \sigma_\alpha^i, \quad \alpha = 1, 2, r. \end{aligned} \quad (11)$$

Therefore, the homogenized problem is

$$\begin{cases} -\text{div}(D_1^* \nabla u_1) + \overline{\Sigma}_1 u_1 = \lambda (\overline{\sigma}_1 u_1 + \overline{\sigma}_2 u_2) & \text{in } \Omega, \\ -\text{div}(D_2^* \nabla u_2) + \overline{\Sigma}_2 u_2 = \overline{\sigma}_r u_1 & \text{in } \Omega, \\ u_1 = u_2 = 0 & \text{on } \partial\Omega, \end{cases} \quad (12)$$

where  $(\lambda; u_1, u_2)$  is the first (positive) eigensolution. Note that Theorem 1 also applies to (12) which therefore admits such a first eigensolution.

It turns out that, although the homogenized cross-sections  $\overline{\Sigma}_\alpha, \overline{\sigma}_\alpha$  are uniquely defined by the limit density  $\theta$ , the homogenized diffusion coefficients  $(D_1^*, D_2^*)$  are not simple volume averages, explicitly characterized by  $\theta$ . Indeed, depending on the geometry of the mixture,  $(D_1^*, D_2^*)$  may be any symmetric positive definite matrix in a set  $G_\theta$ . This is a local constraint defined pointwise in  $\Omega$ . Unfortunately, the set  $G_\theta$  of all possible homogenized diffusion tensors associated to the density  $\theta$  is not explicitly known (except when there are only two phases, i.e.  $I = 2$ , see Cherkaev 2000).

Since the homogenized state system (12) depends on the design parameters  $\theta = (\theta_i)_{1 \leq i \leq I}$  and  $(D_1^*, D_2^*)$ , the set of generalized admissible configuration  $\mathcal{U}_{ad}^h$  is defined by

$$\mathcal{U}_{ad}^h = \left\{ (\theta, D_1^*, D_2^*) \in L^\infty(\Omega) \text{ satisfying (10) and } (D_1^*, D_2^*) \in G_\theta \right\}. \quad (13)$$

Note that we have  $\mathcal{U}_{ad}^c \subset \mathcal{U}_{ad}^h$  if we associate to each characteristic function  $\chi \in \mathcal{U}_{ad}^c$  a diffusion tensor  $D_\alpha = \sum_{i=1}^I d_\alpha^i \chi_i$ .

It remains to characterize the relaxed (or homogenized) objective function. As a consequence of homogenization theory (see Allaire 2001; Allaire and Castro 2001, for details) it is given by

$$J^*(\theta, D_1^*, D_2^*) = \ell_1 \lambda + \ell_2 \frac{(\mathcal{M}(\overline{s}))^{1/r}}{\mathcal{M}(\overline{\sigma}_1 u_1 + \overline{\sigma}_2 u_2)}, \quad (14)$$

where  $(\lambda; u_1, u_2)$  is the first eigensolution of the homogenized problem (12), and  $\overline{s}$  is defined by

$$\overline{s}(x) = \sum_{i=1}^I \theta_i(x) |\sigma_1^i u_1(x) + \sigma_2^i u_2(x)|^r, \quad (15)$$

which is usually different from  $(\overline{\sigma}_1 u_1 + \overline{\sigma}_2 u_2)^r$  for  $r > 1$ . The reason for this seemingly surprising term  $\overline{s}$  is that for characteristic functions  $(\chi_i)_{1 \leq i \leq I}$  we have

$$\begin{aligned} \mathcal{M}(|\sigma_1 u_1 + \sigma_2 u_2|^r) = \\ \frac{1}{\text{vol}(\Omega)} \int_{\Omega} \sum_{i=1}^I \chi_i(x) |\sigma_1^i u_1(x) + \sigma_2^i u_2(x)|^r, \end{aligned}$$

which averages like (15) in the homogenized limit.

The relaxed problem is finally to minimize  $J^*$  over  $\mathcal{U}_{ad}^h$ , i.e.

$$\min_{(\theta, D_1^*, D_2^*) \in \mathcal{U}_{ad}^h} J^*(\theta, D_1^*, D_2^*). \quad (16)$$

As in Allaire and Castro (2001) it can be rigorously justified and the following theorem holds true.

**Theorem 2.** *Assume that  $1 \leq r < +\infty$  in two space dimensions. The relaxation of the continuous optimization problem (9) is (16) in the sense that*

1. *there exists at least one minimizer in  $\mathcal{U}_{ad}^h$  of  $J^*$ ,*
2. *any minimizer  $(\theta, D_1^*, D_2^*)$  of the relaxed problem is the homogenized limit of a minimizing sequence of the continuous problem (9),*
3. *any minimizing sequence of the continuous problem (9) converges, in the sense of homogenization, to a minimizer  $(\theta, D_1^*, D_2^*)$  of the relaxed problem (16).*

The main consequence of Theorem 2 is that relaxation does not change physically the problem but makes it well-posed. In other words, a generalized homogenized design is just a precise and convenient way of characterizing limits of sequences of classical designs. As we already said, the main inconvenient with the relaxed formulation (16) is that we lack an explicit characterization of the set  $G_\theta$  of all homogenized diffusion tensors. Nevertheless, we can restrict ourselves to an explicit subclass of  $G_\theta$  which yields a so-called partial relaxation of the problem (see Allaire 2001). This partial relaxation is then amenable to numerical computations.

We choose to work with the class of simple laminated composite materials which is a (very small) subset of  $G_\theta$ . A simple laminate is obtained by averaging a layered mixture of the  $I$  phases where all slices are orthogonal to a single lamination direction parameterized by an angle  $\gamma$ . In this case, the homogenized diffusion tensors  $D_1^*$  and  $D_2^*$  are fully explicit

$$D_\alpha^* = \begin{pmatrix} \cos \gamma & \sin \gamma \\ -\sin \gamma & \cos \gamma \end{pmatrix} \begin{pmatrix} \mu_\alpha^+ & 0 \\ 0 & \mu_\alpha^- \end{pmatrix} \begin{pmatrix} \cos \gamma & -\sin \gamma \\ \sin \gamma & \cos \gamma \end{pmatrix}, \quad (17)$$

$\alpha = 1, 2,$

where  $\gamma \in [0, \pi)$  is the angle of lamination and  $\mu_\alpha^+, \mu_\alpha^-$  ( $\alpha = 1, 2$ ) are the arithmetic and harmonic averages respectively, i.e.

$$\mu_\alpha^+ = \sum_{i=1}^I \theta_i d_\alpha^i, \quad \frac{1}{\mu_\alpha^-} = \sum_{i=1}^I \frac{\theta_i}{d_\alpha^i}, \quad \alpha = 1, 2. \quad (18)$$

Note that  $\mu_\alpha^+$  and  $\mu_\alpha^-$  are uniquely defined by the density function  $\theta$  and therefore the set of homogenized tensors obtained by simple lamination can be characterized by two parameters: the lamination angle  $\gamma$  and the density  $\theta = (\theta_1, \dots, \theta_I)$ . In the sequel we restrict ourselves to this simpler case and we replace the set of all generalized admissible configurations  $\mathcal{U}_{ad}^h$  by

$$\mathcal{U}_{ad}^l = \left\{ (\theta, \gamma) \in L^\infty(\Omega) \text{ satisfying (10) and } \gamma \in [0, \pi) \right\}, \quad (19)$$

with  $(D_1^*, D_2^*)$  given by (17). From now on, the objective function  $J^*(\theta, D_1^*, D_2^*)$  is equivalently denoted by  $J^*(\theta, \gamma)$ .

Finally, the relaxed minimization problem (13) is simplified and becomes

$$\inf_{(\theta, \gamma) \in \mathcal{U}_{ad}^l} \left\{ J^*(\theta, \gamma) = \ell_1 \lambda + \ell_2 \frac{(\mathcal{M}(\overline{s}))^{1/r}}{\mathcal{M}(\overline{\sigma}_1 u_1 + \overline{\sigma}_2 u_2)} \right\}. \quad (20)$$

The new formulation (20) is fully explicit, but the price to pay is that it may have no minimizer. A possible heuristic justification of working with (20) instead of (13) is twofold. First, it is perfectly legitimate in the one-group diffusion model as proved by Allaire and Castro (2001). Second, it gives very good numerical results in the sense that taking higher order laminates does not improve the results or the convergence.

*Remark 2.* As is usual in the homogenization method, working with a relaxed formulation yields homogenized optimal designs, i.e. a distribution of phases with intermediate densities and not only pure phases. Therefore, for practical applications it must be coupled with a penalization procedure which project an homogenized design onto a classical one. This process is guaranteed to work because of Theorem 2 which states that any optimal composite design is attained as the limit of a sequence of classical designs. This penalization step is purely based on numerical heuristics but it is by now a classical matter although not quite well understood (see e.g. Allaire 2001; Bendsøe 1995).

#### 4 Optimality conditions

One advantage of the relaxed formulation is that it allows us to compute a gradient quite easily. This will be at the root of the numerical algorithm proposed in this paper. This section is therefore devoted to the computation of the gradient of  $J^*$  which, as usual, will be expressed in terms of an adjoint problem. Recall that the relaxed cost functional  $J^*$  is defined by (20). If  $(\delta\theta, \delta\gamma)$  is an admissible increment in  $\mathcal{U}_{ad}^l$ , the directional derivative of  $J^*$  is

$$\begin{aligned} \delta J^* &= \ell_1 \delta\lambda + \ell_2 \frac{\mathcal{M}(\delta\bar{s}) (\mathcal{M}(\bar{s}))^{1/r-1}}{r \mathcal{M}(\bar{\sigma}_1 u_1 + \bar{\sigma}_2 u_2)} - \\ &\ell_2 \frac{(\mathcal{M}(\bar{s}))^{1/r}}{(\mathcal{M}(\bar{\sigma}_1 u_1 + \bar{\sigma}_2 u_2))^2} \mathcal{M}((\bar{\sigma}_1 \delta u_1 + \\ &u_1 \delta \bar{\sigma}_1 + \bar{\sigma}_2 \delta u_2 + u_2 \delta \bar{\sigma}_2)), \end{aligned} \quad (21)$$

where

$$\begin{aligned} \delta \bar{s} &= r \sum_{i=1}^I (\sigma_1^i u_1 + \sigma_2^i u_2)^{r-1} (\sigma_1^i \delta u_1 + \sigma_2^i \delta u_2) \theta_i + \\ &\sum_{i=1}^I (\sigma_1^i u_1 + \sigma_2^i u_2)^r \delta \theta_i. \end{aligned} \quad (22)$$

Here  $\delta\lambda$  is the increment in the first eigenvalue and  $(\delta u_1, \delta u_2)$  is the increment in the first eigenvector. Recall that, since the first eigenvalue of (12) is simple, it is differentiable with respect to the design parameters, as well as the first eigenfunction. In order to obtain an explicit expression of  $\delta J^*$ , let us calculate the corresponding in-

crements. Differentiating (12), we obtain that  $(\delta u_1, \delta u_2)$  is a solution of the system

$$\begin{cases} -\operatorname{div}(D_1^* \nabla \delta u_1) + \bar{\Sigma}_1 \delta u_1 - \lambda(\bar{\sigma}_1 \delta u_1 + \bar{\sigma}_2 \delta u_2) = f_1 \\ \text{in } \Omega, \\ -\operatorname{div}(D_2^* \nabla \delta u_2) + \bar{\Sigma}_2 \delta u_2 - \bar{\sigma}_r \delta u_1 = f_2 \\ \text{in } \Omega, \\ \delta u_1 = \delta u_2 = 0 \\ \text{on } \partial\Omega, \end{cases} \quad (23)$$

where

$$\begin{cases} f_1 = \operatorname{div}(\delta D_1^* \nabla u_1) - \delta \bar{\Sigma}_1 u_1 + \delta \lambda(\bar{\sigma}_1 u_1 + \bar{\sigma}_2 u_2) + \\ \quad \lambda(\delta \bar{\sigma}_1 u_1 + \delta \bar{\sigma}_2 u_2), \\ f_2 = \operatorname{div}(\delta D_2^* \nabla u_2) - \delta \bar{\Sigma}_2 u_2 + \delta \bar{\sigma}_r u_1. \end{cases} \quad (24)$$

Note that (23) is a singular nonhomogeneous system. Therefore, by the Fredholm alternative there exists a solution of (23) if and only if the following condition holds:

$$\int_{\Omega} f_1 v_1 + \lambda \int_{\Omega} f_2 v_2 = 0, \quad (25)$$

where  $(v_1, v_2)$  is the first eigensolution of the adjoint eigenvalue problem

$$\begin{cases} -\operatorname{div}(D_1^* \nabla v_1) + \bar{\Sigma}_1 v_1 - \lambda(\bar{\sigma}_1 v_1 + \bar{\sigma}_r v_2) = 0 & \text{in } \Omega, \\ -\operatorname{div}(D_2^* \nabla v_2) + \bar{\Sigma}_2 v_2 - \bar{\sigma}_2 v_1 = 0 & \text{in } \Omega, \\ v_1 = v_2 = 0 & \text{on } \partial\Omega. \end{cases} \quad (26)$$

Note that the adjoint system (26) admits the same first eigenvalue than the original system (12). Of course, the solution  $(\delta u_1, \delta u_2)$  of (23) is unique only up to the addition of a multiple of the first eigenfunction  $(u_1, u_2)$ . From (24) and (25) we obtain the following expression for  $\delta\lambda$ :

$$\begin{aligned} \delta\lambda &= \\ &\frac{\int_{\Omega} \delta D_1^* \nabla u_1 \nabla v_1 + \int_{\Omega} (\delta \bar{\Sigma}_1 - \lambda \delta \bar{\sigma}_1) u_1 v_1 - \lambda \int_{\Omega} \delta \bar{\sigma}_2 u_2 v_1}{\int_{\Omega} \bar{\sigma}_1 u_1 v_1 + \int_{\Omega} \bar{\sigma}_2 u_2 v_1} + \\ &\lambda \frac{\int_{\Omega} \delta D_2^* \nabla u_2 \nabla v_2 + \int_{\Omega} \delta \bar{\Sigma}_2 u_2 v_2 - \int_{\Omega} \delta \bar{\sigma}_r u_1 v_2}{\int_{\Omega} \bar{\sigma}_1 u_1 v_1 + \int_{\Omega} \bar{\sigma}_2 u_2 v_1}. \end{aligned} \quad (27)$$

We now investigate the last term in formula (21) for  $\delta J^*$ . As usual, to eliminate  $(\delta u_1, \delta u_2)$  an adjoint state  $(q_1, q_2)$  is introduced (see e.g. Lions 1971). It is defined as the solution of

$$\begin{cases} -\operatorname{div}(D_1^* \nabla q_1) + \bar{\Sigma}_1 q_1 - \lambda(\bar{\sigma}_1 q_1 + \bar{\sigma}_r q_2) = g_1 & \text{in } \Omega, \\ -\operatorname{div}(D_2^* \nabla q_2) + \bar{\Sigma}_2 q_2 - \bar{\sigma}_2 q_1 = g_2 & \text{in } \Omega, \\ q_1 = q_2 = 0 & \text{on } \partial\Omega, \end{cases} \quad (28)$$

with

$$\begin{aligned} g_1 &= \frac{(\mathcal{M}(\bar{s}))^{(1-r)/r}}{\mathcal{M}(\bar{\sigma}_1 u_1 + \bar{\sigma}_2 u_2)} \frac{\sum_{i=1}^I (\sigma_1^i u_1 + \sigma_2^i u_2)^{r-1} \sigma_1^i \theta_i}{\operatorname{vol}(\Omega)} - \\ &\frac{(\mathcal{M}(\bar{s}))^{1/r}}{(\mathcal{M}(\bar{\sigma}_1 u_1 + \bar{\sigma}_2 u_2))^2} \frac{\bar{\sigma}_1}{\operatorname{vol}(\Omega)}, \\ g_2 &= \frac{(\mathcal{M}(\bar{s}))^{(1-r)/r}}{\mathcal{M}(\bar{\sigma}_1 u_1 + \bar{\sigma}_2 u_2)} \frac{\sum_{i=1}^I (\sigma_1^i u_1 + \sigma_2^i u_2)^{r-1} \sigma_2^i \theta_i}{\lambda \operatorname{vol}(\Omega)} - \\ &\frac{(\mathcal{M}(\bar{s}))^{1/r}}{(\mathcal{M}(\bar{\sigma}_1 u_1 + \bar{\sigma}_2 u_2))^2} \frac{\bar{\sigma}_2}{\lambda \operatorname{vol}(\Omega)}. \end{aligned} \quad (29)$$

Note that (28) is of the same type as (26) but nonhomogeneous, and the Fredholm alternative implies the existence of  $(q_1, q_2)$  since one can check that

$$\int_{\Omega} g_1 u_1 + \lambda \int_{\Omega} g_2 u_2 = 0. \quad (30)$$

Multiplying the first and second equations in (28) by  $\delta u_1$  and  $\delta u_2$ , respectively, and integrating by parts we obtain

$$\begin{aligned} &\int_{\Omega} D_1^* \nabla q_1 \nabla(\delta u_1) + \int_{\Omega} (\bar{\Sigma}_1 - \lambda \bar{\sigma}_1) q_1 \delta u_1 - \lambda \int_{\Omega} \bar{\sigma}_r q_2 \delta u_1 = \\ &\frac{(\mathcal{M}(\bar{s}))^{(1-r)/r}}{\mathcal{M}(\bar{\sigma}_1 u_1 + \bar{\sigma}_2 u_2)} \mathcal{M} \left( \sum_{i=1}^I (\sigma_1^i u_1 + \sigma_2^i u_2)^{r-1} \sigma_1^i \theta_i \delta u_1 \right) - \\ &\frac{(\mathcal{M}(\bar{s}))^{1/r}}{(\mathcal{M}(\bar{\sigma}_1 u_1 + \bar{\sigma}_2 u_2))^2} \mathcal{M}(\bar{\sigma}_1 \delta u_1), \\ &\int_{\Omega} D_2^* \nabla q_2 \nabla(\delta u_2) + \int_{\Omega} \bar{\Sigma}_2 q_2 \delta u_2 - \int_{\Omega} \bar{\sigma}_2 q_1 \delta u_2 = \\ &\frac{(\mathcal{M}(\bar{s}))^{(1-r)/r}}{\lambda \mathcal{M}(\bar{\sigma}_1 u_1 + \bar{\sigma}_2 u_2)} \mathcal{M} \left( \sum_{i=1}^I (\sigma_1^i u_1 + \sigma_2^i u_2)^{r-1} \sigma_2^i \theta_i \delta u_2 \right) - \\ &\frac{(\mathcal{M}(\bar{s}))^{1/r}}{\lambda (\mathcal{M}(\bar{\sigma}_1 u_1 + \bar{\sigma}_2 u_2))^2} \mathcal{M}(\bar{\sigma}_2 \delta u_2). \end{aligned} \quad (31)$$

Multiplying now the equations in (23) by  $q_1$  and  $q_2$ , respectively, and integrating by parts we have

$$\begin{aligned} &\int_{\Omega} D_1^* \nabla(\delta u_1) \nabla q_1 + \int_{\Omega} (\bar{\Sigma}_1 - \lambda \bar{\sigma}_1) (\delta u_1) q_1 - \lambda \int_{\Omega} \bar{\sigma}_2 \delta u_2 q_1 = \\ &-\int_{\Omega} \delta D_1^* \nabla u_1 \nabla q_1 + \delta \lambda \left( \int_{\Omega} \bar{\sigma}_1 u_1 q_1 + \int_{\Omega} \bar{\sigma}_2 u_2 q_1 \right) + \\ &\lambda \left( \int_{\Omega} \delta \bar{\sigma}_1 u_1 q_1 + \int_{\Omega} \delta \bar{\sigma}_2 u_2 q_1 \right) - \int_{\Omega} \delta \bar{\Sigma}_1 u_1 q_1, \\ &\int_{\Omega} D_2^* \nabla(\delta u_2) \nabla q_2 + \int_{\Omega} \bar{\Sigma}_2 (\delta u_2) q_2 - \int_{\Omega} \bar{\sigma}_r \delta u_1 q_2 = \\ &-\int_{\Omega} \delta D_2^* \nabla u_2 \nabla q_2 - \int_{\Omega} \delta \bar{\Sigma}_2 u_2 q_2 + \int_{\Omega} \delta \bar{\sigma}_r u_1 q_2. \end{aligned} \quad (32)$$

Combining (31) and (32) and introducing

$$t = \sum_{i=1}^I \theta_i (\sigma_1^i u_1 + \sigma_2^i u_2)^{r-1} (\sigma_1^i \delta u_1 + \sigma_2^i \delta u_2),$$

we obtain

$$\begin{aligned} &\frac{(\mathcal{M}(\bar{s}))^{(1-r)/r}}{\mathcal{M}(\bar{\sigma}_1 u_1 + \bar{\sigma}_2 u_2)} \mathcal{M}(t) - \frac{(\mathcal{M}(\bar{s}))^{1/r}}{(\mathcal{M}(\bar{\sigma}_1 u_1 + \bar{\sigma}_2 u_2))^2} \mathcal{M}(\bar{\sigma}_1 \delta u_1 + \bar{\sigma}_2 \delta u_2) = \\ &-\int_{\Omega} \delta D_1^* \nabla u_1 \nabla q_1 - \int_{\Omega} \delta \bar{\Sigma}_1 u_1 q_1 + \\ &\lambda \left( \int_{\Omega} \delta \bar{\sigma}_1 u_1 q_1 + \int_{\Omega} \delta \bar{\sigma}_2 u_2 q_1 \right) + \\ &\delta \lambda \left( \int_{\Omega} \bar{\sigma}_1 u_1 q_1 + \int_{\Omega} \bar{\sigma}_2 u_2 q_1 \right) - \lambda \int_{\Omega} \delta D_2^* \nabla u_2 \nabla q_2 - \\ &-\lambda \int_{\Omega} \delta \bar{\Sigma}_2 u_2 q_2 + \lambda \int_{\Omega} \delta \bar{\sigma}_r u_1 q_2. \end{aligned} \quad (33)$$

Substituting (33) in (21) we obtain the following expression for  $\delta J^*$ :

$$\begin{aligned} \delta J^* &= \delta \lambda \left( \ell_1 + \ell_2 \int_{\Omega} \bar{\sigma}_1 u_1 q_1 + \ell_2 \int_{\Omega} \bar{\sigma}_2 u_2 q_1 \right) - \\ &\ell_2 \int_{\Omega} \delta D_1^* \nabla u_1 \nabla q_1 - \ell_2 \int_{\Omega} \delta \bar{\Sigma}_1 u_1 q_1 + \end{aligned}$$

$$\begin{aligned}
& \ell_2 \lambda \left( \int_{\Omega} \delta \bar{\sigma}_1 u_1 q_1 + \int_{\Omega} \delta \bar{\sigma}_2 u_2 q_1 + \int_{\Omega} \delta \bar{\sigma}_r u_1 q_2 \right) - \\
& \ell_2 \lambda \left( \int_{\Omega} \delta D_2^* \nabla u_2 \nabla q_2 + \int_{\Omega} \delta \bar{\Sigma}_2 u_2 q_2 \right) + \\
& \ell_2 \frac{(\mathcal{M}(\bar{s}))^{(1-r)/r}}{r \mathcal{M}(\bar{\sigma}_1 u_1 + \bar{\sigma}_2 u_2)} \mathcal{M} \left( \sum_{i=1}^I (\sigma_1^i u_1 + \sigma_2^i u_2)^r \delta \theta_i \right) - \\
& \ell_2 \frac{(\mathcal{M}(\bar{s}))^{1/r}}{(\mathcal{M}(\bar{\sigma}_1 u_1 + \bar{\sigma}_2 u_2))^2} \mathcal{M}(\delta \bar{\sigma}_1 u_1 + \delta \bar{\sigma}_2 u_2), \tag{34}
\end{aligned}$$

where

$$\begin{aligned}
\delta \bar{\sigma}_\alpha &= \sum_{i=1}^I \sigma_\alpha^i \delta \theta_i, \quad (\alpha = 1, 2, r), \\
\delta \bar{\Sigma}_i &= \sum_{i=1}^I \Sigma_\alpha^i \delta \theta_i, \quad (\alpha = 1, 2),
\end{aligned}$$

and  $\delta \lambda$  is given by (27). Introducing the combination functions

$$\begin{cases} z_1 = \frac{\ell_1 + \ell_2 \int_{\Omega} (\bar{\sigma}_1 u_1 q_1 + \bar{\sigma}_2 u_2 q_1)}{\int_{\Omega} (\bar{\sigma}_1 u_1 v_1 + \bar{\sigma}_2 u_2 v_1)} v_1 - \ell_2 q_1, \\ z_2 = \lambda \frac{\ell_1 + \ell_2 \int_{\Omega} (\bar{\sigma}_1 u_1 q_1 + \bar{\sigma}_2 u_2 q_1)}{\int_{\Omega} (\bar{\sigma}_1 u_1 v_1 + \bar{\sigma}_2 u_2 v_1)} v_2 - \ell_2 \lambda q_2, \end{cases} \tag{35}$$

the derivative of  $J^*$  reads

$$\begin{aligned}
\delta J^* &= \int_{\Omega} \delta D_1^* \nabla u_1 \nabla z_1 + \int_{\Omega} \delta \bar{\Sigma}_1 u_1 z_1 - \\
& \lambda \left( \int_{\Omega} \delta \bar{\sigma}_1 u_1 z_1 + \int_{\Omega} \delta \bar{\sigma}_2 u_2 z_1 \right) - \int_{\Omega} \delta \bar{\sigma}_r u_1 z_2 + \\
& \int_{\Omega} \delta D_2^* \nabla u_2 \nabla z_2 + \int_{\Omega} \delta \bar{\Sigma}_2 u_2 z_2 + \\
& \ell_2 \frac{(\mathcal{M}(\bar{s}))^{(1-r)/r}}{r \mathcal{M}(\bar{\sigma}_1 u_1 + \bar{\sigma}_2 u_2)} \mathcal{M} \left( \sum_{i=1}^I (\sigma_1^i u_1 + \sigma_2^i u_2)^r \delta \theta_i \right) - \\
& \ell_2 \frac{(\mathcal{M}(\bar{s}))^{1/r}}{(\mathcal{M}(\bar{\sigma}_1 u_1 + \bar{\sigma}_2 u_2))^2} \mathcal{M}(\delta \bar{\sigma}_1 u_1 + \delta \bar{\sigma}_2 u_2). \tag{36}
\end{aligned}$$

As we minimize over the set of simple laminates the variations of the diffusion tensors  $D_\alpha^*$  linearly depend on the increments with respect to the density  $\theta$  and the lamination angle  $\gamma$ , namely

$$\begin{aligned}
\delta D_\alpha^* &= \\
& \begin{pmatrix} \delta \mu_\alpha^+ \cos^2 \gamma + \delta \mu_\alpha^- \sin^2 \gamma & (\delta \mu_\alpha^- - \delta \mu_\alpha^+) \sin \gamma \cos \gamma \\ (\delta \mu_\alpha^- - \delta \mu_\alpha^+) \sin \gamma \cos \gamma & \delta \mu_\alpha^+ \sin^2 \gamma + \delta \mu_\alpha^- \cos^2 \gamma \end{pmatrix} + \\
& (\mu_\alpha^- - \mu_\alpha^+) \begin{pmatrix} \sin 2\gamma & \cos 2\gamma \\ \cos 2\gamma & -\sin 2\gamma \end{pmatrix} \delta \gamma, \tag{37}
\end{aligned}$$

where

$$\delta \mu_\alpha^+ = \sum_{i=1}^I \frac{\partial \mu_\alpha^+}{\partial \theta_i} \delta \theta_i, \quad \delta \mu_\alpha^- = \sum_{i=1}^I \frac{\partial \mu_\alpha^-}{\partial \theta_i} \delta \theta_i, \tag{38}$$

with

$$\frac{\partial \mu_\alpha^+}{\partial \theta_i} = d_\alpha^i, \quad \text{and} \quad \frac{\partial \mu_\alpha^-}{\partial \theta_i} = \frac{-(\mu_\alpha^-)^2}{d_\alpha^i}.$$

Finally the gradient of the objective function  $J^*$  is given by (36), (37), and (38).

According to the structure of  $U_{ad}^l$ , the two design parameters  $\gamma$  and  $\theta$  are independent, and  $J^*$  can be minimized separately with respect to them. We therefore deduce from (36) the partial derivatives of  $J^*$  in the following propositions.

**Proposition 1.** *When  $\delta \theta = 0$ , the partial derivative of  $J^*$  with respect to  $\gamma$  is*

$$\begin{aligned}
& \left\langle \frac{\partial J^*}{\partial \gamma}, \delta \gamma \right\rangle = \\
& \sum_{\alpha=1,2} (\mu_\alpha^- - \mu_\alpha^+) \int_{\Omega} \begin{pmatrix} \sin 2\gamma & \cos 2\gamma \\ \cos 2\gamma & -\sin 2\gamma \end{pmatrix} \nabla u_\alpha \nabla z_\alpha \delta \gamma. \tag{39}
\end{aligned}$$

Therefore, the optimality condition for the angle  $\gamma$  is

$$\begin{aligned}
\tan 2\gamma &= - \frac{\sum_{\alpha=1,2} (\mu_\alpha^- - \mu_\alpha^+) \left( \frac{\partial u_\alpha}{\partial x_2} \frac{\partial z_\alpha}{\partial x_1} + \frac{\partial u_\alpha}{\partial x_1} \frac{\partial z_\alpha}{\partial x_2} \right)}{\sum_{\alpha=1,2} (\mu_\alpha^- - \mu_\alpha^+) \left( \frac{\partial u_\alpha}{\partial x_1} \frac{\partial z_\alpha}{\partial x_1} - \frac{\partial u_\alpha}{\partial x_2} \frac{\partial z_\alpha}{\partial x_2} \right)}. \tag{40}
\end{aligned}$$

**Proposition 2.** *When  $\delta \gamma = 0$ , the partial derivative of  $J^*$  with respect to  $\theta$  is*

$$\left\langle \frac{\partial J^*}{\partial \theta}, \delta \theta \right\rangle = \sum_{i=1}^I \int_{\Omega} \delta \theta_i Q_i(x) dx, \tag{41}$$

where

$$\begin{aligned}
Q_i(x) &= \sum_{\alpha=1,2} \cdot \\
&\left( \frac{\partial \mu_{\alpha}^+}{\partial \theta_i} \cos^2 \gamma + \frac{\partial \mu_{\alpha}^-}{\partial \theta_i} \sin^2 \gamma \quad \left( \frac{\partial \mu_{\alpha}^-}{\partial \theta_i} - \frac{\partial \mu_{\alpha}^+}{\partial \theta_i} \right) \sin \gamma \cos \gamma \right) \\
&\left( \left( \frac{\partial \mu_{\alpha}^-}{\partial \theta_i} - \frac{\partial \mu_{\alpha}^+}{\partial \theta_i} \right) \sin \gamma \cos \gamma \quad \frac{\partial \mu_{\alpha}^+}{\partial \theta_i} \sin^2 \gamma + \frac{\partial \mu_{\alpha}^-}{\partial \theta_i} \cos^2 \gamma \right) \cdot \\
&\nabla u_{\alpha} \nabla z_{\alpha} + \\
&(\Sigma_1^i - \lambda \sigma_1^i) u_1 z_1 - \lambda \sigma_2^i u_2 z_1 + \Sigma_2^i u_2 z_2 - \sigma_r^i u_1 z_2 + \\
&\ell_2 \frac{(\mathcal{M}(\bar{s}))^{(1-r)/r}}{r \mathcal{M}(\bar{\sigma}_1 u_1 + \bar{\sigma}_2 u_2)} \frac{(\sigma_1^i u_1 + \sigma_2^i u_2)^r}{\text{vol}(\Omega)} - \\
&\ell_2 \frac{(\mathcal{M}(\bar{s}))^{1/r}}{(\mathcal{M}(\bar{\sigma}_1 u_1 + \bar{\sigma}_2 u_2))^2} \frac{\sigma_1^i u_1 + \sigma_2^i u_2}{\text{vol}(\Omega)}. \tag{42}
\end{aligned}$$

## 5 Numerical algorithm

This section is devoted to a gradient-type numerical algorithm for solving the proposed relaxed formulation of the re-loading optimization problem (in two space dimensions). It relies on our knowledge of the optimality conditions. The design parameters are the volume fractions  $\theta = (\theta_1, \dots, \theta_I)$  and the rotation angle  $\gamma$ . We use a gradient method for the density  $\theta$ , coupled with a projection step in order to satisfy the admissibility constraints (10). We could do the same for the rotation angle  $\gamma$ , but it is more efficient to use the optimality condition (40). The algorithm is then structured as follows.

1. We initialize the design parameters  $\theta^1 = (\theta_1^1, \dots, \theta_I^1)$  and  $\gamma^1$  (for example, we take a constant angle  $\gamma^1$  and volume fractions  $\theta_i^1$ , which satisfy the volume constraints).
2. Until convergence, for  $n \geq 1$  we iteratively compute the state  $(u_1^n, u_2^n)$  and the adjoint state  $(q_1^n, q_2^n)$ , solutions of (12) and (28), respectively, with the previous design parameters  $(\theta^n, \gamma^n)$ , and then update these parameters by

$$\theta_i^{n+1}(x) = \max \left( 0, \min \left( 1, \theta_i^n(x) - \right. \right.$$

$$\left. \left. t_n \left( \bar{Q}_i^n(x) - C_0^{n+1}(x) - C_i^{n+1} \right) \right) \right)$$

where  $C_i^{n+1}$  are Lagrange multipliers (constant throughout the domain) for the global volume constraints  $\int_{\Omega} \theta_i = \gamma_i$ ,  $C_0^{n+1}(x)$  is the Lagrange multiplier (varying at each point  $x$ ) for the local volume constraint  $\sum_{i=1}^I \theta_i^{n+1}(x) = 1$ , and  $t_n > 0$  is a small step such that

$$J^*(\theta^{n+1}, \gamma^n) < J^*(\theta^n, \gamma^n),$$

and  $\gamma^{n+1}$  is given by the optimality condition

$$\tan 2\gamma^{n+1} = - \frac{\sum_{\alpha=1,2} (\mu_{\alpha}^- - \mu_{\alpha}^+) \left( \frac{\partial u_{\alpha}^n}{\partial x_2} \frac{\partial z_{\alpha}^n}{\partial x_1} + \frac{\partial u_{\alpha}^n}{\partial x_1} \frac{\partial z_{\alpha}^n}{\partial x_2} \right)}{\sum_{\alpha=1,2} (\mu_{\alpha}^- - \mu_{\alpha}^+) \left( \frac{\partial u_{\alpha}^n}{\partial x_1} \frac{\partial z_{\alpha}^n}{\partial x_1} - \frac{\partial u_{\alpha}^n}{\partial x_2} \frac{\partial z_{\alpha}^n}{\partial x_2} \right)}.$$

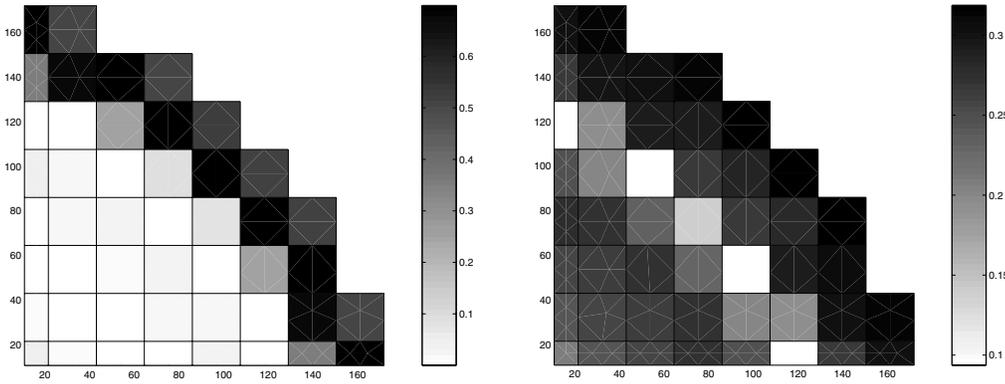
The Lagrange multipliers are iteratively adjusted in an inner loop at each step  $n$  of the above algorithm. This is more delicate for  $I = 4$  phases than for just  $I = 2$  phases (especially during the penalization process). In practice, we made no special efforts to optimize the choice of the step size  $t_n$ , neither did we try to implement a conjugate gradient method or an approximate second-order Newton method (this would be important if CPU time efficiency was our first concern).

**Table 1** Physical constants of the 4 types of assembly

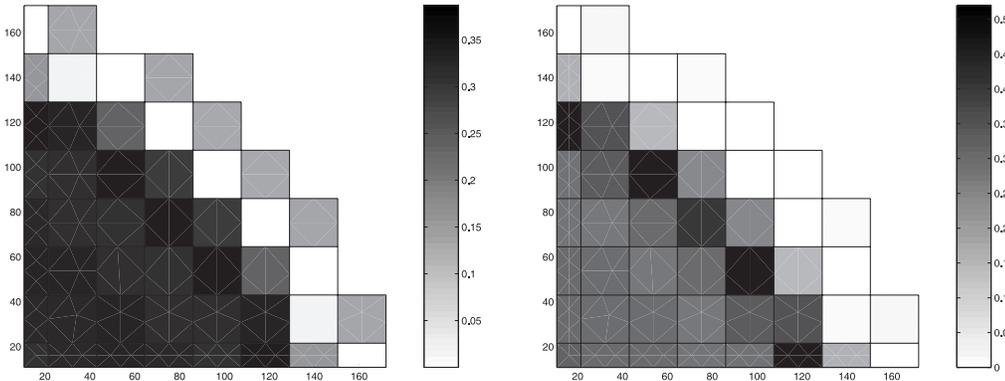
Label of assembly proportion	Diffusion $D_1$ $D_2$	Absorption $\Sigma_1$ $\Sigma_2$	Fission $\sigma_1$ $\sigma_2$	Slackness $\sigma_r$
1	1.340	0.0244	0.0073	0.0149
40/157	0.434	0.103	0.161	
2	1.356	0.0251	0.0063	0.0146
40/157	0.429	0.118	0.173	
3	1.390	0.0256	0.0055	0.0144
40/157	0.428	0.117	0.160	
4	1.410	0.0260	0.0048	0.0143
37/157	0.428	0.114	0.145	

We test our method on a core with 157 squared assemblies (with side length 21.5 cm) of 4 different types with properties given by Table 1 (these data are representative of a 900 MW pressurized water reactor). By symmetry, the computation are performed on one fourth of the geometry using the Matlab software. There are 362  $P1$  finite elements in the mesh and the volume fractions are constant on each assembly. In other words, the spatial discretization is finer for the neutron flux than for the design parameters. The main advantage is that the phase proportions are always constant by assembly, which helps a lot in the penalization process. We choose  $\ell_1 = 0$ ,  $\ell_2 = 1$  and  $r = 10$  in the objective function (other choices work as well). We first compute the optimal solution for the relaxed formulation after 120 iterations. Figures 2 and 3 display the optimal volume fractions, and Fig. 4 the resulting power distribution  $\sigma_1 u_1 + \sigma_2 u_2$ . The convergence is smooth as shown by Fig. 5 and the power peak  $\max(\sigma_1 u_1 + \sigma_2 u_2)$  is globally decreasing (there is no reconstruction of the fine structure of the flux).

In the above example, we started from a previous solution obtained with the one-group diffusion model (see our previous work Allaire and Castro 2001). We checked that,



**Fig. 2** Volume fractions of assembly 1 (left) and 2 (right)



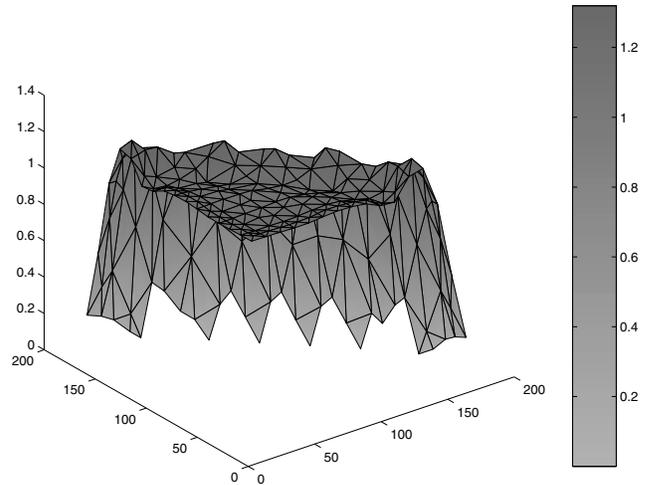
**Fig. 3** Volume fractions of assembly 3 (left) and 4 (right)

if our initial guess is different (typically a random initialization), we converge to the same homogenized solution (we believe we reached a global minimum).

The above relaxed or homogenized optimal solution gives a lower bound on the minimal performance of any discrete distribution of assemblies. More than that, by penalizing the intermediate values of the volume fractions, we can recover a quasi-optimal distribution of assemblies. We introduce a penalized objective function, defined by

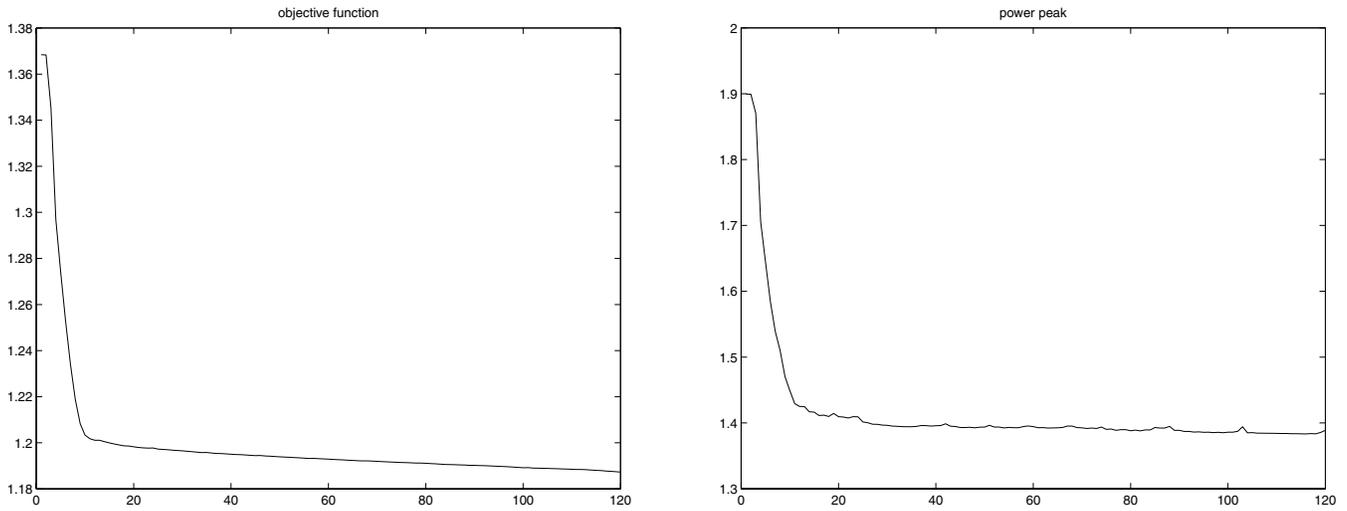
$$J^{pen}(\theta, \gamma) = J^*(\theta, \gamma) + \frac{\eta}{\text{vol}(\Omega)} \int_{\Omega} \sum_{i=1}^I \theta_i (1 - \theta_i) \, dx.$$

For  $\eta = 0$  we recover the relaxed objective function  $J^*$ , while for  $\eta > 0$  we force the volume fractions to take only the values 0 or 1. Starting from the previous relaxed optimal design, we minimize the penalized objective function and increase progressively the value of  $\eta$ . Since by virtue of Theorem 2 any relaxed design is the limit of a sequence of closer and closer classical designs, the penalization process amounts to build such an approximating sequence for which the objective function should not change too much. This procedure is now well-established in structural optimization (see Allaire 2001; Bendsøe 1995; Rozvany *et al.* 1995). Here, we run about 300 iterations with  $\eta$  progressively increasing from 0.01

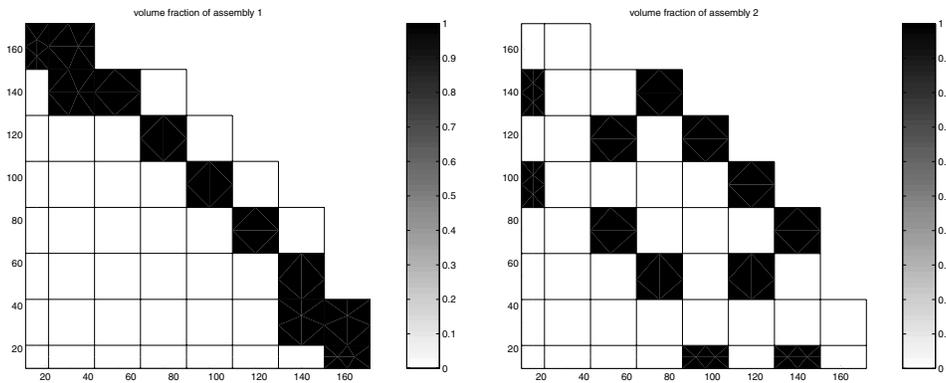


**Fig. 4** Power distribution  $\sigma u$

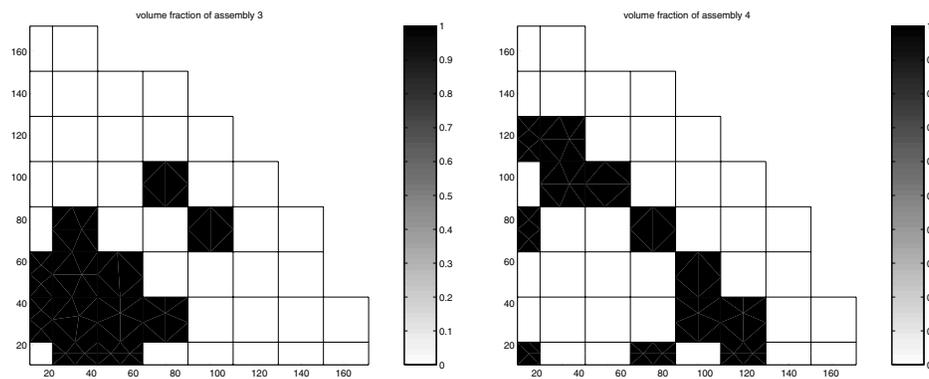
up to 10. This is probably not optimal in terms of CPU time. The reason for this very slow and progressive penalization is that we used the one-quarter symmetry of the core. Indeed, only the assembly of type 4 can be put in the central assembly because of the imposed proportions. Similarly, the half assemblies on the symmetry axes can not be occupied arbitrarily for the same reason of volume constraints.



**Fig. 5** Convergence history: objective function (left) and power peak (right)



**Fig. 6** Distributions of assembly 1 (left) and 2 (right)

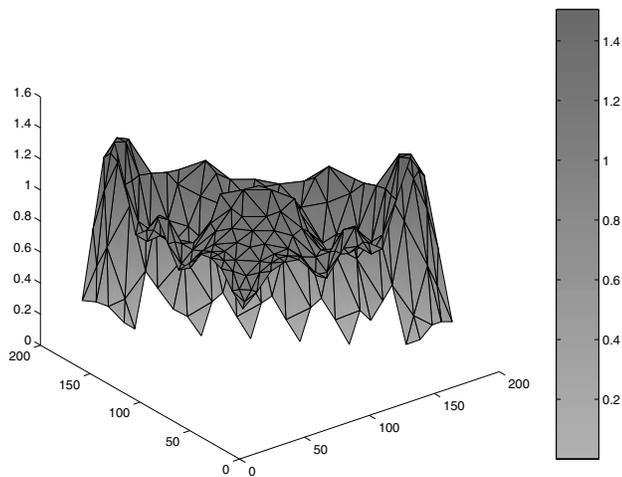


**Fig. 7** Distributions of assembly 3 (left) and 4 (right)

Figures 6 and 7 display the discrete distribution of assemblies, and Fig. 8 the resulting power distribution  $\sigma_1 u_1 + \sigma_2 u_2$ . Note that the obtained pattern is symmetric with respect to the first diagonal. This result is different from that obtained by Allaire and Castro (2001) (where another model was used, namely the one-group diffusion equation).

**Table 2** Comparison between the homogenized and penalized designs

	Objective function	Power peak
homogenized design	1.187	1.388
penalized design	1.225	1.571



**Fig. 8** Power distribution after penalization

In Table 2 we compare the values of the objective function for the relaxed optimal design and for the penalized one (the penalization term  $J^{pen} - J^*$  is almost zero at the end of the penalization process).

## 6 Conclusions

This paper describes the application of the homogenization method for optimizing the fuel assemblies positions in a nuclear reactor core. We believe that this approach is interesting in this context for at least two reasons. First, the homogenized optimal design gives an absolute lower bound to any proposed discrete distribution of assemblies. Therefore, it is a good element of comparison with any other optimization method. Second, the homogenization algorithm is insensitive to the initial guess and the resulting penalized discrete distribution of assemblies is free of any implicit or explicit constraint on its pattern (in structural optimization this is called topology optimization, see e.g. Allaire 2001; Bendsøe 1995). We do not view this method as an alternative to other optimization algorithms but rather as a pre-processing step. Indeed, it gives rise to new patterns that may be different from initial guesses or intuitions, but that can be improved by local optimization using more realistic constraints or objective function. There is still more work to be done in order to treat real industrial problems. Indeed, we have to take into account more realistic constraints such as e.g. multi-cycle optimization, or assembly rotation. Finally, more numerical comparisons with other approaches in the literature are necessary for assessing the potentiality of the homogenization method in this context.

*Acknowledgements* This work has been partially supported by the French Atomic Energy Commission (CEA Saclay, DRN/DMT/SERMA).

## References

- Allaire, G. 2001: *Shape optimization by the homogenization method*. Berlin, Heidelberg, New York: Springer
- Allaire, G.; Castro, C. 2001: A new approach for the optimal distribution of assemblies in a nuclear reactor. *Numerische Mathematik* **89**, 1–29
- Bendsøe, M.P. 1995: *Methods for optimization of structural topology, shape and material*. Berlin, Heidelberg, New York: Springer
- Cherkaev, A. 2000: *Variational methods for structural optimization*. Berlin, Heidelberg, New York: Springer
- Habetler, G.; Martino, M. 1961: Existence theorems and spectral theory for the multigroup diffusion model. In: *Nuclear reactor theory* (Proc. Symp. Appl. Math.), Vol. XI, pp. 127–139, AMS. Providence
- Ho, L.-W.; Rohach, A. 1982: Perturbation theory in nuclear fuel management optimization. *Nucl. Sci. Eng.* **82**, 151–161
- Kropaczek, D.J.; Turinsky, P.J. 1991: In-core nuclear fuel management optimization for pressurized water reactors utilizing simulated annealing. *Nucl. Technol.* **95**, 9
- Levine, S. 1986: In-core fuel management of four reactor types. In: Ronen, Y. (ed.) *Handbook of nuclear reactor calculations*, Vol. II, pp. 87–201. CRC Press
- Lions, J.L. 1971: Optimal control of systems governed by partial differential equations. *Die Grundlehren der mathematischen Wissenschaften* **170**. Berlin, Heidelberg, New York: Springer
- Lysenko, M.G.; Wong, H.I.; Maldonado, G.I. 1999a: Neural network and perturbation theory hybrid models for eigenvalue prediction. *Nucl. Sci. Eng.* **132**
- Lysenko, M.G.; Wong, H.I.; Maldonado, G.I. 1999b: Predicting neutron diffusion eigenvalues with a query-based adaptive neural architecture. *IEEE Trans. Neural Network* **10**
- Maldonado, G.I.; Turinsky, P.J. 1995: Application of nonlinear nodal diffusion generalized perturbation theory to nuclear fuel reload optimization. *Nucl. Technol.* **110**, 198–219
- Murat, F.; Tartar, L. 1985: Calcul des variations et homogénéisation. *Les méthodes de l'homogénéisation: théorie et applications en physique*, Eyrolles, pp. 319–369. English translation in: Cherkaev, A.; Kohn, R. (eds.) *Topics in the mathematical modelling of composite materials*. Progress in nonlinear differential equations and their applications, Vol. 31. Boston: Birkhäuser (1997)
- Parks, G.T. 1996: Multiobjective pressurized water reactor reload core design by nondominated genetic algorithm search. *Nucl. Sci. Eng.* **124**, 178–187
- Planchard, J. 1995: *Méthodes mathématiques en neutronique*. Paris: Eyrolles
- Rozvany, G.I.N.; Bendsøe, M.P.; Kirsch, U. 1995: Layout optimization of structures. *Appl. Mech. Rev.* **48**, 41–118