



**HAL**  
open science

# Optimization of dispersive coefficients in the homogenization of the wave equation in periodic structures

Grégoire Allaire, T Yamada

► **To cite this version:**

Grégoire Allaire, T Yamada. Optimization of dispersive coefficients in the homogenization of the wave equation in periodic structures. *Numerische Mathematik*, 2018, 140 (2), pp.265-326. hal-01341082

**HAL Id: hal-01341082**

**<https://hal.science/hal-01341082v1>**

Submitted on 4 Jul 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Optimization of dispersive coefficients in the homogenization of the wave equation in periodic structures

G. Allaire\*

T. Yamada†

July 2, 2016

## Abstract

We study dispersive effects of wave propagation in periodic media, which can be modelled by adding a fourth-order term in the homogenized equation. The corresponding fourth-order dispersive tensor is called Burnett tensor and we numerically optimize its values in order to minimize or maximize dispersion. More precisely, we consider the case of a two-phase composite medium with an 8-fold symmetry assumption of the periodicity cell in two space dimensions. We obtain upper and lower bound for the dispersive properties, along with optimal microgeometries.

**Keywords:** periodic structure, homogenization, Bloch waves, dispersion, shape optimization

**Mathematics Subject Classification:** 35B27, 49K20

## 1 Introduction

Wave propagation in periodic heterogeneous media is ubiquitous in engineering and science. Denoting by  $\varepsilon$  the small ratio between the period size and a characteristic lengthscale, it can be modeled by the following scalar wave equation

$$\begin{cases} \frac{\partial^2 u_\varepsilon}{\partial t^2} - \operatorname{div}(a_\varepsilon \nabla u_\varepsilon) = f, \\ u_\varepsilon(0, x) = u^{\text{init}}(x), \quad \frac{\partial u_\varepsilon}{\partial t}(0, x) = v^{\text{init}}(x), \end{cases} \quad (1.1)$$

with periodic coefficients  $a_\varepsilon(x) := a\left(\frac{x}{\varepsilon}\right)$ , a right hand side  $f(t, x)$  and initial data  $u^{\text{init}}(x), v^{\text{init}}(x)$ . For simplicity, we assume that the domain of propagation is the full space  $\mathbb{R}^d$ . It does not change much our results to consider another domain but the case of the full space avoids to discuss the boundary conditions, as well as the issue of boundary layers in the homogenization process. Here,  $a(y)$  a  $Y$ -periodic symmetric tensor and  $Y$  is the unit cube  $[0, 1]^d$ . We assume that, for  $0 < \alpha \leq \beta$ ,

$$\alpha I_d \leq a(y) \leq \beta I_d \quad \text{for a.e. } y \in Y. \quad (1.2)$$

For very small values of  $\varepsilon$ , problem (1.1) can be studied by means of the homogenization theory [10], [12], [28], [45], [51]. The result of homogenization theory is that the solution  $u_\varepsilon$  of (1.1) can be well approximated by the solution  $u$  of the following homogenized wave equation

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} - \operatorname{div}(a^* \nabla u) = f, \\ u(0, x) = u^{\text{init}}(x), \quad \frac{\partial u}{\partial t}(0, x) = v^{\text{init}}(x), \end{cases} \quad (1.3)$$

where, in the periodic homogenization setting,  $a^*$  is a constant effective tensor given by an explicit formula involving cell problems (see Section 2 for details).

Although it is less classical, it is known that the homogenized equation can be improved by adding a small fourth-order operator and modifying the source term. This is the concept of “high order homogenized equation”

---

\*CMAP, École Polytechnique, Palaiseau, FRANCE – gregoire.allaire@polytechnique.fr

†Department of Mechanical Engineering and Science, Kyoto University, Kyoto, JAPAN – takayuki@me.kyoto-u.ac.jp

that goes back to [10], [46] and has been studied by many authors [1], [2], [4], [21], [22], [30], [48]. In the present setting it reads

$$\frac{\partial^2 v_\varepsilon}{\partial t^2} - \operatorname{div}(a^* \nabla v_\varepsilon) + \varepsilon^2 \mathbb{D}^* \nabla^4 v_\varepsilon = f + \varepsilon^2 \operatorname{div}(d^* \nabla f). \quad (1.4)$$

where  $\mathbb{D}^*$  is a fourth-order tensor, called Burnett tensor and studied in [17], [19], [20], and  $d^*$  is some second-order tensor (see Section 2 for details). Equations (1.3) and (1.4) can be established by two different methods: two-scale asymptotic expansions (Section 2), and Bloch wave expansions (Section 3). The interpretation of  $\mathbb{D}^*$  is that it plays the role of a dispersion tensor. This is explained and numerically illustrated in Section 4. In particular, for long times of order up to  $\varepsilon^{-2}$ , (1.4) is a better approximation of the wave equation (1.1) than (1.3) (this key observation was first made by [46] and further discussed in the references, just quoted above).

Dispersion is classically defined as the phenomenon by which waves with different wavelengths propagate with different velocities. In practice, it induces severe deformations of the profile of the propagating waves in the long time limit. Here we focus exclusively on dispersion induced by homogenization and not by the more classical dispersion effects arising in the high frequency limit (or geometric optics, see e.g. chapter 3 in [42]). In such a homogenization setting, dispersion can occur only in heterogeneous media. Since composite materials (which are of course heterogeneous) are ubiquitous in engineering, it is therefore very important to study its associated dispersion properties. Dispersion can be a good thing or a bad thing, depending on the type of applications which we have in mind. Clearly, dispersion is a nasty effect if one is interested in preserving the profile of a wave or signal during its propagation. On the opposite, dispersion could be beneficial if one wants to spread and thus diminish the intensity of, say, a sound wave. In any case, it makes sense to optimize the periodic structure, namely the coefficient  $a(y)$  in (1.1), in order to achieve minimal or maximal dispersion. There are a few rigorous bounds on the dispersive properties of periodic structures [19], [20] but no systematic numerical study. The goal of the present paper is to make a first numerical investigation in the optimization of these dispersive properties. We restrict our attention to two-phase composite materials with isotropic constituents. In this setting there is an extensive literature on the precise characterization of the set of all possible values of the homogenized tensor  $a^*$  (see [3], [51] and references therein). However, to our knowledge, nothing is known about the Burnett tensor  $\mathbb{D}^*$  (except the few bounds in [19], [20]). Therefore, we use shape optimization techniques in order to optimize the values of  $\mathbb{D}^*$  for a two-phase composite with prescribed volume fractions. Since  $\mathbb{D}^*$  is a fourth-order tensor, to simplify the analysis, we restrict ourselves to a plane 2-d setting and to a geometric 8-fold symmetry in the unit cell  $Y$ , which yields a kind of isotropy for  $\mathbb{D}^*$ .

A key feature of the dispersion tensor  $\mathbb{D}^*$  is that, contrary to the homogenized tensor  $a^*$ , it is not scale invariant. More precisely, we prove in Lemmas 6.1 and 6.3 that if the periodicity cell is scaled by a factor  $\kappa$  then the dispersion tensor  $\mathbb{D}^*$  is scaled by  $\kappa^2$ . Actually, even if the periodicity cell is fixed to be the unit cube  $Y$ , then the microgeometry can be periodically repeated  $k$  times in  $Y$  (with  $k \geq 1$  any integer) and the dispersion tensor  $\mathbb{D}^*$  is thus divided by  $k^2$ . In other words, by considering finer details in the unit periodicity cell, the dispersion tensor  $\mathbb{D}^*$  can be made as small as we want (in norm). As a consequence, the minimization of (norm of)  $\mathbb{D}^*$  is an ill-posed problem (except if geometric constraint are added) while one can expect that the maximization problem is meaningful.

More specifically, we consider coefficients defined by

$$a(y) = a^A \mathbf{1}_{Y_A}(y) + a^B \mathbf{1}_{Y_B}(y),$$

where  $a^A, a^B > 0$  are two constant real numbers,  $Y_A$  and  $Y_B$  are a disjoint partition of  $Y$ ,  $\mathbf{1}_{Y_A}(y), \mathbf{1}_{Y_B}(y)$  are the corresponding characteristic functions. Denote by  $\Gamma = \partial Y_A \cap \partial Y_B$  the interface between the two phases  $A$  and  $B$ . We rely on the level set method [39] for the shape optimization of the interface  $\Gamma$ , as is now quite common in structural mechanics [8], [38], [52]. Under an 8-fold symmetry assumption for the unit cell  $Y$ , the dispersion tensor  $\mathbb{D}^*$  is characterized by two scalar parameters  $\alpha$  and  $\beta$ . We numerically compute Pareto fronts in the plane  $(\alpha, \beta)$  when minimizing or maximizing  $\mathbb{D}^*$  under two equality constraints for the phase proportions and for the homogenized (scalar) tensor  $a^*$ , as well as an inequality constraint for the perimeter of the interface  $\Gamma$ . Such an inequality constraint for the perimeter is not necessary (and indeed is not active at the optima) when maximizing dispersion. However, the perimeter constraint is required and active when minimizing dispersion since smaller details of the phase mixture yield smaller dispersion (this refinement process is stopped by the perimeter constraint). It turns out that computing these Pareto fronts is a delicate task since the optimization process is plagued by the existence of many local optima (in contrast, our algorithm easily finds global optima when optimizing the homogenized tensor  $a^*$ ). Therefore, we rely on a complicate strategy of continuation, re-initialization and non-convex approximation in order to obtain robust (hopefully global) optimal distributions of the two phases which minimize or maximize dispersion. Our main finding is that the upper Pareto front (which of course depends on the phase properties and proportions) seems to be a line segment. The corresponding optimal configurations are smooth and simple

geometric arrangements of the two phases. Note that the checkerboard pattern seems to be optimal for maximal  $\alpha$  and  $\beta$ . We conclude this brief description of our results by recognizing that other type of dispersive properties have already been optimized in a different context [32], [44], [50].

Let us now describe the contents of our paper. In Section 2 we recall the two-scale asymptotic expansion method for periodic homogenization, as introduced in [10], [12], [45]. We closely follow the presentation of [4]. The main result is Proposition 2.2 which gives (1.4) as a “high order homogenized equation”. In Proposition 2.5 we recall a result of [17] which states that the Burnett tensor  $\mathbb{D}^*$  is non-positive, making (1.4) an ill-posed equation. This inconvenient will be corrected later in Section 4.

In Section 3 we recall the classical theory of Bloch waves [12], [15], [17], [43], [54] which is an alternative method for deriving the homogenized problem (1.3), as well as the high order homogenized equation (1.4). The main result is Lemma 3.1, due to [16], [17], which proves that the Burnett tensor  $\mathbb{D}^*$  is the fourth-order derivative of the so-called first Bloch eigenvalue.

Section 4 explains how to correct equation (1.4) to make it well-posed (see Lemma 4.1). The main idea is a Boussinesq trick (i.e., replacing some space derivatives by time derivatives) which is possible because (1.4) is merely an approximation at order  $\varepsilon^4$ .

Section 5 presents some one-dimensional numerical simulations of wave propagation in a periodic medium. It compares the solutions of the original wave equation (1.1) with those of the homogenized equation (1.3) and the Boussinesq version of the high order homogenized equation (1.4). It demonstrates that, for long times of order  $\varepsilon^{-2}$ , the approximation is much better with (1.4) rather than with (1.3).

Section 6 discusses some properties of the Burnett tensor  $\mathbb{D}^*$ . First, we explain that, contrary to the homogenized tensor  $a^*$ , the fourth-order tensor  $\mathbb{D}^*$  depends on the scaling of the periodicity cell. More precisely, if the cell  $Y$  is scaled to be of size  $\kappa > 0$ , then  $\mathbb{D}^*$  is scaled as  $\kappa^2 \mathbb{D}^*$ . It implies that small heterogeneities yield small dispersion while large heterogeneities lead to large dispersion (see Lemma 6.3). Second, we prove that a standard 8-fold symmetry assumption of the coefficients  $a(y)$  in the unit cell  $Y$  (or of the two-phase geometry  $\mathbf{1}_{Y_A}(y), \mathbf{1}_{Y_B}(y)$ ) implies that the Burnett tensor  $\mathbb{D}^*$  is characterized by simply two scalar parameters.

Section 7 computes the shape derivative, i.e. the shape sensitivity, of the tensor  $\mathbb{D}^*$  with respect to the position of the interface  $\Gamma$ . Our main result is Theorem 7.3 which gives a rigorous shape derivative. From a numerical point of view, Theorem 7.3 is difficult to exploit because it involves jumps of discontinuous solution gradients through the interface  $\Gamma$ . Therefore, following [5], in Proposition 7.6 we compute a simpler shape derivative for a discretized version of  $\mathbb{D}^*$ .

Section 8 explains our numerical setting based on the level set algorithm of Osher and Sethian [39] and on a steepest descent optimization algorithm. Constraints on the volume, the perimeter and the the homogenized tensor  $A^*$  are enforced by means of Lagrange multipliers. We iteratively update the Lagrange multipliers so that the constraints are exactly satisfied at each iteration of the optimization algorithm.

Section 9 contains our numerical results on the optimization of the Burnett tensor  $\mathbb{D}^*$  with respect to the interface  $\Gamma$ . Since a first numerical test in Subsection 9.1 shows that dispersion can be minimized by a fine fragmentation of the two phase mixture (which is just stopped at a length-scale determined by the perimeter constraint), we later focus on determining the Pareto upper front for dispersion. It is not known if the set of dispersion tensor  $\mathbb{D}^*$  is convex or if its upper bound is a concave curve in the  $(\alpha, \beta)$  plane. Thus, we explain in Subsection 9.2 that a quadratic function of  $\alpha$  and  $\beta$  is optimized in order to be able to cope with a non-concave upper bound. In the same subsection we explain our intricate optimization strategy in order to avoid the many local optima that can be found. It is a combination of continuation, varying initializations and refinement process of the Pareto front. We are quite confident in our numerical approximation of the Pareto front since we checked it is insensitive to the choice of interface initializations, and of parameters for the minimized quadratic function. In Figure 16 we compare the upper Pareto front for various aspect ratios of the two phases, while in Figure 18 the comparison is made for various volume fractions. Eventually, Subsection 9.3 is devoted to the optimization of the other dispersion tensor  $d^*$  which is responsible for the dispersion of the source term in the right hand side of the high order homogenized equation (1.4).

Finally Section 10 is devoted to the (technical) proof of Theorem 7.3 which was stated in Section 7.

## Notations

In the sequel we shall use the following notations.

1.  $(e_1, \dots, e_d)$  denotes the canonical basis of  $\mathbb{R}^d$ .
2.  $Y = [0, 1]^d$  denotes the unit cube of  $\mathbb{R}^d$ , identified with the unit torus  $\mathbb{R}^d/\mathbb{Z}^d$ .

3.  $H_{\sharp}^1(Y)$  denotes the space of  $Y$ -periodic functions in  $H_{\text{loc}}^1(\mathbb{R}^d)$ .
4.  $H_{\sharp,0}^1(Y)$  denotes the subspace of  $H_{\sharp}^1(Y)$  composed of functions with zero  $Y$ -average.
5. The Einstein summation convention with respect to repeated indices is used.
6. All tensors are assumed to be symmetric, even if we do not write it explicitly. More precisely, if  $C$  is a  $n$ -order tensor  $C = (C_{i_1 \dots i_n})_{1 \leq i_1, \dots, i_n \leq d}$ , it is systematically identified with its symmetrized counterpart  $C^S$ , defined by

$$C^S = \left( \frac{1}{n!} \sum_{\sigma \in \mathfrak{S}_n} C_{\sigma(i_1) \dots \sigma(i_n)} \right)_{1 \leq i_1, \dots, i_n \leq d},$$

where  $\mathfrak{S}_n$  is the permutation group of order  $n$ .

7. If  $C$  is a  $n$ -order tensor, the notation  $C \nabla^n u$  means the full contraction

$$C \nabla^n u = \sum_{i_1, i_2, \dots, i_n=1}^d C_{i_1, i_2, \dots, i_n} \frac{\partial^n u}{\partial x_{i_1} \dots \partial x_{i_n}},$$

where  $C$  is indistinguishable from its symmetric counterpart  $C^S$ .

## 2 Two-scale asymptotic expansions

In this section we briefly recall the method of two-scale asymptotic expansion [10], [12], [45] and, in particular, explain how dispersion can be introduced in a so-called higher-order homogenized equation, as first proposed by [46], and studied by many others [48], [22], [30], [21], [9], [1], [2], [4].

The starting point of the method of two-scale asymptotic expansion is to assume that the solution of (1.1) is given by the following ansatz

$$u_\varepsilon(t, x) = \sum_{n=0}^{\infty} \varepsilon^n u_n \left( t, x, \frac{x}{\varepsilon} \right),$$

where  $y \rightarrow u_n(t, x, y)$  are  $Y$ -periodic. This ansatz is formal since, not only the series does not converge, but it lacks additional boundary layer terms in case of a bounded domain. Plugging this ansatz in (1.1) and using the chain rule lemma for each term

$$\nabla \left( u_n \left( t, x, \frac{x}{\varepsilon} \right) \right) = \left( \nabla_x u_n + \frac{1}{\varepsilon} \nabla_y u_n \right) \left( t, x, \frac{x}{\varepsilon} \right),$$

we deduce a cascade of equations which allow us to successively compute each term  $u_n(t, x, y)$ . To make this cascade of equations explicit, we introduce the following operators

$$\begin{cases} A_{yy} & := -\operatorname{div}_y (a(y) \nabla_y \cdot) \\ A_{xy} & := -\operatorname{div}_x (a(y) \nabla_y \cdot) - \operatorname{div}_y (a(y) \nabla_x \cdot) \\ A_{xx} & := -\operatorname{div}_x (a(y) \nabla_x \cdot), \end{cases} \quad (2.1)$$

which satisfy, for any function  $v(x, y)$ ,

$$-\operatorname{div} (a(\varepsilon^{-1}x) \nabla v(x, \varepsilon^{-1}x)) = (\varepsilon^{-2} A_{yy} v + \varepsilon^{-1} A_{xy} v + A_{xx} v) (x, \varepsilon^{-1}x). \quad (2.2)$$

Then, we deduce the cascade of equations

$$\left\{ \begin{array}{l} \text{order } \varepsilon^{-2}, \quad 0 = A_{yy}u_0 \\ \text{order } \varepsilon^{-1}, \quad 0 = A_{yy}u_1 + A_{xy}u_0 \\ \text{order } \varepsilon^0, \quad f = A_{yy}u_2 + A_{xy}u_1 + A_{xx}u_0 + \frac{\partial^2 u_0}{\partial t^2} \\ \text{order } \varepsilon^1, \quad 0 = A_{yy}u_3 + A_{xy}u_2 + A_{xx}u_1 + \frac{\partial^2 u_1}{\partial t^2} \\ \vdots \\ \text{order } \varepsilon^{n-2}, \quad 0 = A_{yy}u_n + A_{xy}u_{n-1} + A_{xx}u_{n-2} + \frac{\partial^2 u_{n-2}}{\partial t^2} \\ \vdots \end{array} \right. \quad (2.3)$$

These equations are solved successively by means of the following lemma, called Fredholm alternative (see [10], [12], [45] for a proof).

**Lemma 2.1.** *For  $g(y) \in L^2(Y)$ , consider the following problem*

$$\left\{ \begin{array}{l} -\operatorname{div}_y(a(y)\nabla_y w) = g \quad \text{in } Y, \\ y \mapsto w(y) \quad Y\text{-periodic.} \end{array} \right. \quad (2.4)$$

*It admits a solution  $w(y) \in H_{\#}^1(Y)$ , unique up to an additive constant, if and only if the right hand side satisfies the following compatibility condition*

$$\int_Y g(y) dy = 0.$$

Thanks to Lemma 2.1 we now deduce from (2.3) the formulas for successive terms  $u_n$  in the ansatz. These formulas will imply a separation of variables, namely each function  $u_n(t, x, y)$  is a sum of products of cell solutions depending only on  $y$  and on space derivatives of the homogenized solution  $u(t, x)$ . Before we start the study of the cascade of equations, we emphasize two important notations for the sequel. First, according to the Fredholm alternative of Lemma 2.1, all cell solutions, introduced below, have zero-average in the unit cell  $Y$ . Second, all tensors below are symmetric (i.e. invariant by a permutation of the indices) since they are contracted with the symmetric derivative tensors  $\nabla_x^k u(t, x)$ . Nevertheless, for the sake of simplicity in the notations, we do not explicitly symmetrize all tensors but the reader should keep in mind that they are indeed symmetric.

**Computation of  $u_0$ :** since the source term is zero, the solution is constant with respect to  $y$ ,

$$u_0(t, x, y) = u(t, x).$$

**Computation of  $u_1$ :** the source term satisfies the compatibility condition and by linearity we obtain

$$u_1(t, x, y) = -\chi_i(y) \frac{\partial u}{\partial x_i}(t, x) + \tilde{u}_1(t, x), \quad (2.5)$$

where  $\chi_i$  and  $\chi_\eta^{(1)} = \sum_{i=1}^d \eta_i \chi_i$  are solutions in  $H_{\#,0}^1(Y)$  of the equations

$$A_{yy} \chi_i = -\operatorname{div}_y(ae_i) \quad \text{and} \quad A_{yy} \chi_\eta^{(1)} = -\operatorname{div}_y(a\eta), \quad \text{for } \eta \in \mathbb{R}^d. \quad (2.6)$$

**Computation of  $u_2$ :** the third equation of (2.3) has a solution if and only if its source term has a zero  $Y$ -average, which leads to the homogenized equation

$$\frac{\partial^2 u}{\partial t^2} - \operatorname{div}(a^* \nabla u) = f, \quad (2.7)$$

where the homogenized symmetric matrix  $a^*$  is given by

$$a^* \eta := \int_Y (a\eta - a \nabla_y \chi_\eta^{(1)}) dy, \quad \text{for } \eta \in \mathbb{R}^d. \quad (2.8)$$

Inserting (2.5), the third equation of (2.3) becomes

$$\begin{aligned} A_{yy}u_2 &= -\frac{\partial^2 u}{\partial t^2} + f + \operatorname{div}_y (a e_j) \frac{\partial \tilde{u}_1}{\partial x_j} \\ &+ \left( a_{ij} - a \nabla_y \chi_i \cdot e_j - \operatorname{div}_y (\chi_i a e_j) \right) \frac{\partial^2 u}{\partial x_i \partial x_j}. \end{aligned} \quad (2.9)$$

Hence, defining for  $i, j \in \{1, \dots, d\}$

$$b_{ij} := a_{ij} - a \nabla_y \chi_i \cdot e_j - \operatorname{div}_y (\chi_i a e_j), \quad \text{with} \quad \int_Y b_{ij} = a_{ij}^*, \quad (2.10)$$

$u_2$  can be written as

$$u_2(t, x, y) = \chi_{ij}(y) \frac{\partial^2 u}{\partial x_i \partial x_j}(t, x) - \chi_i(y) \frac{\partial \tilde{u}_1}{\partial x_i}(t, x) + \tilde{u}_2(t, x), \quad (2.11)$$

where the functions  $\chi_{ij}$  and  $\chi_\eta^{(2)} := \chi_{ij} \eta_i \eta_j$  are the solutions in  $H_{\#,0}^1(Y)$  of the equations

$$A_{yy} \chi_{ij} = b_{ij} - \int_Y b_{ij} = b_{ij} - a_{ij}^* \quad \text{and} \quad A_{yy} \chi_\eta^{(2)} = b \eta \cdot \eta - a^* \eta \cdot \eta, \quad \text{for } \eta \in \mathbb{R}^d. \quad (2.12)$$

Note that only the symmetric part of  $b_{ij}$  plays a role in (2.9) and the same is true for  $\chi_{ij}$  in (2.11).

**Computation of  $u_3$ :** starting from here, namely for  $n \geq 3$ , the solvability condition of the Fredholm alternative for the existence of  $u_n$  is

$$\int_Y \left( \frac{\partial^2 u_{n-2}}{\partial t^2} - \operatorname{div}_x (a(y) (\nabla_x u_{n-2} + \nabla_y u_{n-1})) \right) dy = 0. \quad (2.13)$$

Thus, there exists a solution  $u_3$  in (2.3) if and only if (2.13) is satisfied for  $n = 3$  which, since the  $Y$ -averages of the cell solutions  $\chi_i$  are zero, leads to an equation for  $\tilde{u}_1$

$$\frac{\partial^2 \tilde{u}_1}{\partial t^2} - \operatorname{div} (a^* \nabla \tilde{u}_1) = C^* \nabla^3 u, \quad (2.14)$$

with a tensor  $C^*$  defined by

$$C_{ijk}^* := \int_Y (a \nabla_y \chi_{ij} \cdot e_k - a_{ij} \chi_k) dy. \quad (2.15)$$

It turns out, by symmetry in  $i, j, k$ , that this tensor vanishes,  $C^* = 0$  (see [35], [4]). Therefore, since its initial data vanishes, the function  $\tilde{u}_1$  vanishes too,

$$\tilde{u}_1(t, x) = 0. \quad (2.16)$$

Let us now compute  $u_3$  which, by (2.3), is a solution of

$$A_{yy}u_3 = -\frac{\partial^2 \tilde{u}_1}{\partial t^2}(t, x) + \chi_i(y) \frac{\partial^3 u}{\partial t^2 \partial x_i}(t, x) - A_{xy}u_2 - A_{xx}u_1. \quad (2.17)$$

Replacing  $u_2$  and  $u_1$  by their expressions (2.11) and (2.5), introducing the solutions  $w_k$  in  $H_{\#,0}^1(Y)$  of

$$A_{yy}w_k = \chi_k, \quad (2.18)$$

the solutions  $\chi_{ijk}$  in  $H_{\#,0}^1(Y)$  of

$$A_{yy} \chi_{ijk} = c_{ijk} - \int_Y c_{ijk}, \quad \text{for } i, j, k \in \{1, \dots, d\}, \quad (2.19)$$

where

$$c_{ijk} := a \nabla_y \chi_{jk} \cdot e_i + \operatorname{div}_y (\chi_{jk} a e_i) - a_{ij} \chi_k, \quad (2.20)$$

and using (2.6), (2.10),  $u_3$  can be written as

$$\begin{aligned} u_3(t, x, y) &= w_i(y) \frac{\partial^3 u}{\partial x_i \partial t^2}(t, x) + \chi_{ijk}(y) \frac{\partial^3 u}{\partial x_i \partial x_j \partial x_k}(t, x) \\ &+ \chi_{ij}(y) \frac{\partial^2 \tilde{u}_1}{\partial x_i \partial x_j}(t, x) - \chi_i(y) \frac{\partial \tilde{u}_2}{\partial x_i}(t, x) + \tilde{u}_3(t, x), \end{aligned} \quad (2.21)$$

**Equation of  $u_4$ :** there exists a solution  $u_4$  in (2.3) if and only if condition (2.13) for  $n = 4$  is satisfied. Replacing  $u_2$  and  $u_3$  by their formulas (2.11) and (2.21) leads to an equation for  $\tilde{u}_2$

$$\frac{\partial^2 \tilde{u}_2}{\partial t^2} - \operatorname{div}(a^* \nabla \tilde{u}_2) = C^* \nabla^3 \tilde{u}_1 + \mathbb{B}^* \nabla^4 u + \operatorname{div}\left(d^* \nabla \frac{\partial^2 u}{\partial t^2}\right), \quad (2.22)$$

where  $\mathbb{B}^*$  is defined by

$$\mathbb{B}_{ijkl}^* := \int_Y \left( a_{ij} \chi_{k\ell} + a_{im} \frac{\partial \chi_{k\ell j}}{\partial y_m} \right) dy, \quad \text{for } i, j, k, \ell \in \{1, \dots, d\}. \quad (2.23)$$

and

$$d_{ij}^* = \int_Y a \nabla_y w_i \cdot e_j dy = \int_Y \chi_i \chi_j dy, \quad \text{for } i, j \in \{1, \dots, d\}. \quad (2.24)$$

We simplify (2.22) by recalling that  $C^* = 0$  and using the homogenized equation (2.7) to replace  $\frac{\partial^2 u}{\partial t^2}$  by  $f + \operatorname{div}(a^* \nabla u)$ . Then, introducing the tensor

$$\mathbb{D}^* = -\mathbb{B}^* - a^* \otimes d^*, \quad (2.25)$$

we deduce that (2.22) is equivalent to

$$\frac{\partial^2 \tilde{u}_2}{\partial t^2} - \operatorname{div}(a^* \nabla \tilde{u}_2) = -\mathbb{D}^* \nabla^4 u + \operatorname{div}(d^* \nabla f). \quad (2.26)$$

We do not compute explicitly  $u_4$  (although it is possible) since our only interest in studying the equation for  $u_4$  is to find the homogenized equation (2.26) for  $\tilde{u}_2$ . We are now in a position to collect the above results and to give an approximate formula for the exact solution  $u_\varepsilon$  of (1.1)

$$u_\varepsilon(t, x) \approx u(t, x) + \varepsilon u_1\left(t, x, \frac{x}{\varepsilon}\right) + \varepsilon^2 u_2\left(t, x, \frac{x}{\varepsilon}\right), \quad (2.27)$$

where  $u$  is a solution of the homogenized equation (2.7),  $u_1$  is defined by (2.5) and  $u_2$  by (2.11). Each term,  $u_1$  and  $u_2$  is the sum of a zero  $Y$ -average contribution and of  $\tilde{u}_1$  and  $\tilde{u}_2$  defined by

$$\tilde{u}_1(t, x) = \int_Y u_1(t, x, y) dy \quad \text{and} \quad \tilde{u}_2(t, x) = \int_Y u_2(t, x, y) dy,$$

which are defined as the solutions of (2.14) and (2.26), respectively. Furthermore, we know from (2.16) that  $\tilde{u}_1(t, x) = 0$  is identically zero. Therefore, on average, (2.27) implies that

$$u_\varepsilon(t, x) \approx u(t, x) + \varepsilon^2 \tilde{u}_2(t, x) := v_\varepsilon(t, x). \quad (2.28)$$

It is possible to find a single approximate equation for  $v_\varepsilon$  by adding equation (2.7) with (2.26) multiplied by  $\varepsilon^2$ : it yields

$$\frac{\partial^2 v_\varepsilon}{\partial t^2} - \operatorname{div}(a^* \nabla v_\varepsilon) + \varepsilon^2 \mathbb{D}^* \nabla^4 v_\varepsilon = f + \varepsilon^2 \operatorname{div}(d^* \nabla f) + \mathcal{O}(\varepsilon^4). \quad (2.29)$$

Neglecting the term of order  $\varepsilon^4$  in (2.29) gives the ‘‘higher order’’ homogenized equation (1.4), as announced in the introduction.

We summarize our results in the following proposition.

**Proposition 2.2.** *The ‘‘high order’’ homogenized equation of the wave equation (1.1) is*

$$\frac{\partial^2 v_\varepsilon}{\partial t^2} - \operatorname{div}(a^* \nabla v_\varepsilon) + \varepsilon^2 \mathbb{D}^* \nabla^4 v_\varepsilon = f + \varepsilon^2 \operatorname{div}(d^* \nabla f). \quad (2.30)$$

**Remark 2.3.** *Writing an effective equation for a truncated version of the non oscillating ansatz has been studied in various settings (see [10], [46], [30], [21], [48]) under the name of ‘‘higher order homogenization’’. Proposition 2.2 gives the ‘‘second order’’ homogenized equation which is a proposed explanation of dispersive effects for wave propagation in periodic media [46], [30], [21], [1], [2] or of second gradient theory in mechanics [48].*

**Remark 2.4.** *Note that the initial data did not enter the entire asymptotic process which is purely formal at this stage.*



A fundamental property of the Burnett tensor  $\mathbb{D}^*$ , discovered by [17], is that it is non-positive.

**Proposition 2.5** ([17]). *The fourth-order tensor  $\mathbb{D}^*$ , defined by (3.3), satisfies for any  $\eta \in \mathbb{R}^d$*

$$\mathbb{D}^*(\eta \otimes \eta) : (\eta \otimes \eta) = - \int_Y a \nabla_y \left( \chi_\eta^{(2)} - \frac{1}{2} (\chi_\eta^{(1)})^2 \right) \cdot \nabla_y \left( \chi_\eta^{(2)} - \frac{1}{2} (\chi_\eta^{(1)})^2 \right) dy \leq 0, \quad (2.31)$$

where  $\chi_\eta^{(1)}$  is defined by (2.6) and  $\chi_\eta^{(2)}$  by (2.12).

**Remark 2.6.** *If the tensor  $\mathbb{D}^*$  were non-negative, equation (2.30) would be well-posed. Unfortunately,  $\mathbb{D}^*$  has the wrong sign, i.e. it is non-positive and (2.30) is thus not well-posed. We shall see in Section 4 how to modify it to make it well-posed by using a Boussinesq trick.*

**Remark 2.7.** *The tensor  $\mathbb{D}^*$  arises in the two-scale asymptotic expansion process as the coefficient fourth-order tensor of the fourth-order derivative  $\nabla^4 u$ . Recall from our notations that  $\mathbb{D}^* \nabla^4 u$  means the full contraction of both fourth-order tensors. Therefore, only the symmetric part of  $\mathbb{D}^*$  is accessible by this method. In other words,  $\mathbb{D}^*$  belongs to the class of fully symmetric fourth-order tensors which satisfy, for any permutation  $\sigma$  of  $\{i, j, k, l\}$ ,*

$$\mathbb{D}_{ijkl}^* = \mathbb{D}_{\sigma(i)\sigma(j)\sigma(k)\sigma(l)}^*.$$

*This class of fully symmetric fourth-order tensors is completely characterized by the knowledge of their quartic form*

$$\mathbb{D}^*(\eta \otimes \eta) : (\eta \otimes \eta) = \sum_{i,j,k,l=1}^d \mathbb{D}_{ijkl}^* \eta_i \eta_j \eta_k \eta_l.$$

*Indeed, differentiating the quartic form four times with respect to  $\eta_i, \eta_j, \eta_k, \eta_l$  allows us to recover the (symmetrized) coefficient  $\mathbb{D}_{ijkl}^*$ .*

In one space dimension, the formula for  $\mathbb{D}^*$  is simpler, as stated in the next lemma.

**Lemma 2.8** ([17]). *In one space dimension, we have  $\mathbb{D}^* = -a^* d^*$  where  $a^*$  is defined by (2.8) and  $d^*$  is defined by (2.24).*

### 3 Bloch wave method

Another method of homogenization is the so-called Bloch wave decomposition method [43], [54]. Its application to periodic homogenization is discussed in [12], [15]. It relies on a family of spectral problems for the operator  $A_{yy}$  in the unit cell  $Y$ . More precisely, for a given parameter  $\eta \in Y$ , we look for eigenvalues  $\lambda = \lambda(\eta)$  in  $\mathbb{R}$  and eigenvectors  $\phi = \phi(\eta)$  in  $H_{\#}^1(Y)$ , normalized by  $\|\phi\|_{L^2(Y)} = 1$ , satisfying

$$A(\eta)\phi = \lambda(\eta)\phi \quad \forall y \in Y,$$

where  $A(\eta)$  is the translated (or shifted) operator defined by

$$A(\eta) := - \left( \frac{\partial}{\partial y_k} + 2\pi i \eta_k \right) \left[ a_{k\ell} \left( \frac{\partial}{\partial y_\ell} + 2\pi i \eta_\ell \right) \right]. \quad (3.1)$$

with, of course,  $A(0) = A_{yy}$ . The above spectral problem for  $A(\eta)$  in the unit torus  $Y$ , the so-called Bloch problem, admits an infinite countable number of non-negative eigenvalues and corresponding normalized eigenfunctions [43], [54]. We are interested in the first eigenvalue  $\lambda_1(\eta)$  which is the relevant one in the homogenization process. When  $\eta = 0$ , one can check that  $\lambda_1(0) = 0$  is a simple eigenvalue of  $A(0) = A_{yy}$  with constants as eigenfunctions. Regular perturbation theory proves then that  $\lambda_1(\eta)$  is simple and analytic in a neighborhood of  $\eta = 0$ . We recall some results from [16], [17] about the fourth-order Taylor expansion of  $\lambda_1(\eta)$  at  $\eta = 0$ .

**Lemma 3.1.** *The first eigenvalue  $\lambda_1(\eta)$  admits the following fourth-order expansion:*

$$\lambda_1(\eta) = 4\pi^2 a^* \eta \cdot \eta + (2\pi)^4 \mathbb{D}^* \cdot (\eta \otimes \eta \otimes \eta \otimes \eta) + \mathcal{O}(|\eta|^6), \quad (3.2)$$

where  $\frac{1}{8\pi^2} \nabla_\eta^2 \lambda_1(0) = a^*$  is the homogenized matrix defined by (2.8) and  $\mathbb{D}^*$  is the symmetric fourth-order tensor  $\frac{1}{4!(2\pi)^4} \nabla_\eta^4 \lambda_1(0)$  (also called Burnett tensor) which is equivalently defined by

$$\mathbb{D}_{ijkl}^* := - \int_Y \left( a_{ij} \chi_{k\ell} + a_{im} \frac{\partial \hat{\chi}_{k\ell j}}{\partial y_m} \right) dy, \quad \text{for } i, j, k, \ell \in \{1, \dots, d\}, \quad (3.3)$$

where the functions  $\chi_{ij}$  are defined by (2.12) and  $\hat{\chi}_{ijk}$  are the solutions in  $H_{\sharp,0}^1(Y)$  of

$$A_{yy} \hat{\chi}_{ijk} = a_{ij}^* \chi_k + c_{ijk} - \int_Y c_{ijk} dy, \quad \text{for } i, j, k \in \{1, \dots, d\}, \quad (3.4)$$

where  $\chi_k$  are given by (2.6) and  $c_{ijk}$  are given by (2.20).

As usual, the tensor  $\mathbb{D}^*$  and the functions  $\hat{\chi}_{ijk}$  are understood as symmetrized (this is obvious for  $\mathbb{D}^*$  which arises as the fourth-order derivative of the eigenvalue  $\lambda_1$ ). Note that the functions  $\hat{\chi}_{ijk}$  are different from the previous ones  $\chi_{ijk}$  defined by (2.19) since  $\hat{\chi}_{ijk} = \chi_{ijk} + a_{ij}^* w_k$ .

**Remark 3.2.** As a by-product of Lemma 3.1 it was shown in [16] that the  $\eta$ -derivatives of the first eigenfunction  $\phi_1(y, \eta)$  coincide with the solutions of some cell problems.

A fundamental property of the Bloch waves is that they diagonalize the operator  $A_{yy}$  in  $L^2(\mathbb{R}^d)$ . More precisely, we have the following Bloch wave decomposition written in rescaled variables  $x = \varepsilon y$  and  $\xi = \eta/\varepsilon$ .

**Lemma 3.3.** Any function  $f \in L^2(\mathbb{R}^d)$  can be decomposed as

$$f(x) = \sum_{n \geq 1} \int_{\varepsilon^{-1}Y} \alpha_n^\varepsilon(\xi) \phi_n \left( \frac{x}{\varepsilon}, \varepsilon \xi \right) e^{2\pi i x \cdot \xi} d\xi \quad (3.5)$$

where

$$\alpha_n^\varepsilon(\xi) = \int_{\mathbb{R}^d} f(x) e^{-2\pi i x \cdot \xi} \overline{\phi_n \left( \frac{x}{\varepsilon}, \varepsilon \xi \right)} dx, \quad (3.6)$$

and  $\phi_n(y, \eta)$  is the  $n$ -th normalized eigenfunction of (3.1). Furthermore, it satisfies Parseval equality

$$\int_{\mathbb{R}^d} |f(x)|^2 dx = \sum_{n \geq 1} \int_{\varepsilon^{-1}Y} |\alpha_n^\varepsilon(\xi)|^2 d\xi. \quad (3.7)$$

We now explain how the Bloch wave method is used for the homogenization of the wave equation (1.1). First, we recall the definition of the Fourier transform  $\hat{f}(\xi)$  of a function  $f(x) \in L^2(\mathbb{R}^d)$

$$f(x) = \int_{\mathbb{R}^d} \hat{f}(\xi) e^{2\pi i x \cdot \xi} d\xi. \quad (3.8)$$

For simplicity, let us replace the fixed (with respect to  $\varepsilon$ ) initial data and source term in (1.1) by well-prepared initial data and source in terms of Bloch waves. Denoting by  $\hat{u}^{\text{init}}(\xi)$ ,  $\hat{v}^{\text{init}}(\xi)$  and  $\hat{f}(t, \xi)$  the Fourier transforms of  $u^{\text{init}}(x)$ ,  $v^{\text{init}}(x)$  and  $f(t, x)$  (in the sense of (3.8)), we introduce these new forcing term and initial data

$$f_\varepsilon(t, x) = \int_{\varepsilon^{-1}Y} \hat{f}(t, \xi) \phi_1 \left( \frac{x}{\varepsilon}, \varepsilon \xi \right) e^{2\pi i x \cdot \xi} d\xi, \quad (3.9)$$

$$u_\varepsilon^{\text{init}}(x) = \int_{\varepsilon^{-1}Y} \hat{u}^{\text{init}}(\xi) \phi_1 \left( \frac{x}{\varepsilon}, \varepsilon \xi \right) e^{2\pi i x \cdot \xi} d\xi, \quad v_\varepsilon^{\text{init}}(x) = \int_{\varepsilon^{-1}Y} \hat{v}^{\text{init}}(\xi) \phi_1 \left( \frac{x}{\varepsilon}, \varepsilon \xi \right) e^{2\pi i x \cdot \xi} d\xi,$$

and change (1.1) into

$$\begin{cases} \frac{\partial^2 u_\varepsilon}{\partial t^2} - \operatorname{div}(a_\varepsilon \nabla u_\varepsilon) = f_\varepsilon(t, x), \\ u_\varepsilon(0, x) = u_\varepsilon^{\text{init}}(x), \quad \frac{\partial u_\varepsilon}{\partial t}(0, x) = v_\varepsilon^{\text{init}}(x). \end{cases} \quad (3.10)$$

Similarly, using Lemma 3.3, we decompose the solution of (3.10) as

$$u_\varepsilon(t, x) = \int_{\varepsilon^{-1}Y} \hat{u}_1^\varepsilon(t, \xi) \phi_1 \left( \frac{x}{\varepsilon}, \varepsilon \xi \right) e^{2\pi i x \cdot \xi} d\xi, \quad (3.11)$$

Since the eigenbasis  $\{\phi_n\}$  diagonalizes the elliptic operator, equation (3.10) is reduced to a family of ordinary differential equations: for any  $\xi \in \varepsilon^{-1}Y$ ,  $\hat{u}_1^\varepsilon(t, \xi)$  is a solution of the following ordinary differential equation

$$\begin{cases} \frac{d^2 \hat{u}_1^\varepsilon}{dt^2} + \varepsilon^{-2} \lambda_1(\varepsilon \xi) \hat{u}_1^\varepsilon = \hat{f}(t, \xi), \\ \hat{u}_1^\varepsilon(0, \xi) = \hat{u}^{\text{init}}(\xi), \quad \frac{d\hat{u}_1^\varepsilon}{dt}(0, \xi) = \hat{v}^{\text{init}}(\xi). \end{cases} \quad (3.12)$$

Using the Taylor expansion (3.2) of  $\lambda_1$ , we deduce that

$$\begin{cases} \frac{d^2 \hat{u}_1^\varepsilon}{dt^2} + \left( 4\pi^2 a^* \xi \cdot \xi + \varepsilon^2 (2\pi)^4 \mathbb{D}^* \cdot (\xi \otimes \xi \otimes \xi \otimes \xi) \right) \hat{u}_1^\varepsilon = \hat{f}(t, \xi) + \mathcal{O}(\varepsilon^4), \\ \hat{u}_1^\varepsilon(0, \xi) = \hat{u}^{\text{init}}(\xi), \quad \frac{d\hat{u}_1^\varepsilon}{dt}(0, \xi) = \hat{v}^{\text{init}}(\xi). \end{cases} \quad (3.13)$$

At least formally, dropping the  $\mathcal{O}(\varepsilon^4)$  in the above equation,  $\hat{u}_1^\varepsilon(t, \xi)$  is well approximated by  $\hat{v}_\varepsilon(t, \xi)$  which is the Fourier transform of the solution  $v_\varepsilon(t, x)$  of the following high order homogenized equation

$$\begin{cases} \frac{\partial^2 v_\varepsilon}{\partial t^2} - \operatorname{div}(a^* \nabla v_\varepsilon) + \varepsilon^2 \mathbb{D}^* \cdot \nabla_x^4 v_\varepsilon = f(t, x), \\ v_\varepsilon(0, x) = u^{\text{init}}(x), \quad \frac{\partial v_\varepsilon}{\partial t}(0, x) = v^{\text{init}}(x), \end{cases} \quad (3.14)$$

This equations is identical to the “higher order” homogenized equation (1.4), or (2.29), except for the right hand side which does not feature the additional term  $\varepsilon^2 \operatorname{div}(d^* \nabla f)$ . This is due to our replacement of the original right hand side  $f$  by its well-prepared variant  $f_\varepsilon$ , defined by (3.9) (see [4] for a more complete explanation).

In any case, the differential operator of the “higher order” homogenized equation is the same whatever the method of derivation, be it two-scale asymptotic expansions or Bloch wave decomposition. Once again, the fourth-order tensor  $\mathbb{D}^*$  is a manifestation of dispersive effects in the wave propagation.

## 4 Boussinesq approximation

The high order homogenized equations (1.4), (2.30) and (3.14) are not well posed since, by virtue of Proposition 2.5, the tensor  $\mathbb{D}^*$  has the “wrong” sign (the bilinear form associated to the operator  $\mathbb{D}^* \nabla_x^4$  is non-positive). The goal of this section is to explain how to modify these equations in order to make them well-posed by using a classical Boussinesq trick (see *e.g.* [14] for historical references). This trick has been applied in recent works [22], [30], [21], [1], [2]. It is also well known in the study of continuum limits of discrete spring-mass lattices [31].

The key point is that both equations (3.13) and (2.29) are actually defined, up to the addition of a small remainder term of order  $\varepsilon^4$ . Therefore one can modify them adding any term of the same order  $\varepsilon^4$ , without altering their approximate validity. Let us explain the Boussinesq trick for (3.14) (the case of (2.30) is completely similar). We define the minimum value

$$m = \min_{|\xi|=1} \frac{\mathbb{D}^* \cdot (\xi \otimes \xi \otimes \xi \otimes \xi)}{a^* \xi \cdot \xi}, \quad (4.1)$$

which is a non-positive number  $m \leq 0$  because of Proposition 2.5 (if  $m > 0$  were positive, (3.14) would be well posed and there would be nothing to do). Introducing the non-negative second order tensor  $C = -m \operatorname{Id} \geq 0$ , we define a fourth order tensor  $\mathfrak{D}^*$  by

$$\mathfrak{D}^* \cdot (\xi \otimes \xi \otimes \xi \otimes \xi) = \mathbb{D}^* \cdot (\xi \otimes \xi \otimes \xi \otimes \xi) + (a^* \xi \cdot \xi) (C \xi \cdot \xi) \geq 0 \quad \forall \xi \in \mathbb{R}^d, \quad (4.2)$$

which is non-negative in view of (4.1). Then, the Fourier transform of (3.14)

$$\frac{d^2 \hat{v}_\varepsilon}{dt^2}(\xi) + 4\pi^2 (a^* \xi \cdot \xi) \hat{v}_\varepsilon(\xi) + \varepsilon^2 (2\pi)^4 \mathbb{D}^* \cdot (\xi \otimes \xi \otimes \xi \otimes \xi) \hat{v}_\varepsilon(\xi) = \hat{f}(\xi) \quad (4.3)$$

can be replaced by

$$\begin{aligned} (1 + \varepsilon^2 4\pi^2 C \xi \cdot \xi) \frac{d^2 \hat{v}_\varepsilon}{dt^2} + (4\pi^2 a^* \xi \cdot \xi + \varepsilon^2 (2\pi)^4 \mathfrak{D}^* \cdot (\xi \otimes \xi \otimes \xi \otimes \xi)) \hat{v}_\varepsilon \\ = \hat{f}(t, \xi) + \varepsilon^2 4\pi^2 (C \xi \cdot \xi) \hat{f}(t, \xi) + \mathcal{O}(\varepsilon^4), \end{aligned} \quad (4.4)$$

since truncating (4.3) implies

$$\frac{d^2 \hat{v}_\varepsilon}{dt^2}(\xi) + 4\pi^2 (a^* \xi \cdot \xi) \hat{v}_\varepsilon(\xi) = \hat{f}(\xi) + \mathcal{O}(\varepsilon^2).$$

By the inverse Fourier transform, applied to (4.4), we deduce the following equation

$$\frac{\partial^2 v_\varepsilon}{\partial t^2} - \varepsilon^2 \operatorname{div} \left( C \nabla \frac{\partial^2 v_\varepsilon}{\partial t^2} \right) - \operatorname{div}(a^* \nabla v_\varepsilon) + \varepsilon^2 \mathfrak{D}^* \cdot \nabla^4 v_\varepsilon = f - \varepsilon^2 \operatorname{div}(C \nabla f) + \mathcal{O}(\varepsilon^4),$$

which is well posed because  $C$  and  $\mathfrak{D}^*$  are non-negative.

We summarize our result in the following lemma.

**Lemma 4.1.** *Up to an error term of order  $\mathcal{O}(\varepsilon^4)$ , the high order homogenized equation (3.14) is equivalent to*

$$\frac{\partial^2 v_\varepsilon}{\partial t^2} - \varepsilon^2 \operatorname{div} \left( C \nabla \frac{\partial^2 v_\varepsilon}{\partial t^2} \right) - \operatorname{div} (a^* \nabla v_\varepsilon) + \varepsilon^2 \mathfrak{D}^* \cdot \nabla^4 v_\varepsilon = f - \varepsilon^2 \operatorname{div} (C \nabla f),$$

while the high order homogenized equation (1.4) is equivalent to

$$\frac{\partial^2 v_\varepsilon}{\partial t^2} - \varepsilon^2 \operatorname{div} \left( C \nabla \frac{\partial^2 v_\varepsilon}{\partial t^2} \right) - \operatorname{div} (a^* \nabla v_\varepsilon) + \varepsilon^2 \mathfrak{D}^* \cdot \nabla^4 v_\varepsilon = f - \varepsilon^2 \operatorname{div} ((C - d^*) \nabla f), \quad (4.5)$$

which are both well posed because  $C \geq 0$  and  $\mathfrak{D}^* \geq 0$ .

**Remark 4.2.** *In 1-d, by virtue of Lemma 2.8, we have  $\mathbb{D}^* = -a^* d^*$  with  $d^* > 0$ . Therefore, in 1-d it is possible to choose  $\mathfrak{D}^* = 0$  and  $C = d^*$ . Then, the right hand side of (4.5) is simply  $f$ , as in the usual homogenized equation (1.3).*

From a numerical point of view, (4.5) should be solved rather than the ill-posed original equation (1.4). Of course, any choice of matrix  $C$ , which makes (4.2) non-negative, is acceptable. Therefore, there is a whole family of higher order homogenized equation (4.5), all of them being equivalent up to order  $\varepsilon^4$ . In this context, the dispersive effect is measured by the fourth-order tensor  $\mathbb{D}^*$  and not by  $\mathfrak{D}^*$  alone.

It was proved in [30], [21] (see also [2]) that the solution  $v_\varepsilon$  of (4.5) provides an approximation of the exact solution  $u_\varepsilon$  of (1.1), up to an error term of order  $\varepsilon$  in the  $L_t^\infty(L_x^2)$ -norm for long times up to  $T\varepsilon^{-2}$ .

**Remark 4.3.** *The dispersive character of the high order homogenized equation (4.5) can easily be checked for plane-wave solutions, in the absence of any source term. Indeed, plugging in (4.5) a plane wave solution*

$$u(t, x) = \mathbf{u} e^{i(\omega t - \xi \cdot x)},$$

where  $\mathbf{u} \in \mathbb{R}$  is the amplitude,  $\omega \in \mathbb{R}^+$  the frequency and  $\xi \in \mathbb{R}^d$  the wave number, we obtain the relation dispersion

$$\omega(\xi) = \left( \frac{a^* \xi \cdot \xi + \varepsilon^2 \mathfrak{D}^* \cdot \xi \otimes \xi \otimes \xi \otimes \xi}{1 + \varepsilon^2 C \xi \cdot \xi} \right)^{1/2} \leq \sqrt{a^* \xi \cdot \xi}. \quad (4.6)$$

For  $\varepsilon |\xi| \ll 1$ , a Taylor expansion of (4.6) yields

$$\omega(\xi) = \sqrt{a^* \xi \cdot \xi} \left( 1 + \frac{\varepsilon^2}{2} \frac{\mathfrak{D}^* \cdot \xi \otimes \xi \otimes \xi \otimes \xi}{a^* \xi \cdot \xi} - \frac{\varepsilon^2}{2} C \xi \cdot \xi + \mathcal{O}(\varepsilon^4) \right).$$

Recall from (4.2) that  $\mathfrak{D}^* = \mathbb{D}^* + C \otimes a^* + R$ , with  $R \cdot \xi \otimes \xi \otimes \xi \otimes \xi \geq 0$  for any  $\xi$ . Denoting by  $\xi^0$  a minimizer in (4.1), we have  $R \cdot \xi^0 \otimes \xi^0 \otimes \xi^0 \otimes \xi^0 = 0$  and  $R \xi^0 \otimes \xi^0 \otimes \xi^0 = 0$  by minimality. Thus, in this optimal direction we deduce

$$\omega(\xi^0) = \sqrt{a^* \xi^0 \cdot \xi^0} \left( 1 + \frac{\varepsilon^2}{2} \frac{\mathbb{D}^* \cdot \xi^0 \otimes \xi^0 \otimes \xi^0 \otimes \xi^0}{a^* \xi^0 \cdot \xi^0} + \mathcal{O}(\varepsilon^4) \right).$$

From (4.6) we deduce the group velocity

$$V(\xi) = \frac{d\omega}{d\xi}(\xi) = \frac{a^* \xi}{\sqrt{a^* \xi \cdot \xi}} + \varepsilon^2 \delta(\xi) + \mathcal{O}(\varepsilon^4),$$

where the corrector term can be computed in the optimal direction

$$\delta(\xi^0) = -\frac{C \xi^0 \cdot \xi^0}{2} \frac{a^* \xi^0}{\sqrt{a^* \xi^0 \cdot \xi^0}} - \sqrt{a^* \xi^0 \cdot \xi^0} C \xi^0$$

Simplifying further to the isotropic case,  $a^* \xi^0 \cdot \xi^0 = a^* |\xi^0|^2$ , we obtain

$$\delta(\xi^0) = \frac{3m\sqrt{a^*}}{2} \xi^0 \quad \Rightarrow \quad V(\xi^0) = \sqrt{a^*} \frac{\xi^0}{|\xi^0|} \left( 1 + \frac{3m}{2} \varepsilon^2 + \mathcal{O}(\varepsilon^4) \right).$$

Since  $m \leq 0$  by virtue of (4.1), we deduce that, up to fourth order, the group velocity is smaller when taking into account dispersive effect. Furthermore, to reduce dispersion (namely to have  $V(\xi^0)$  as close as possible to  $\sqrt{a^* \xi^0 / |\xi^0|}$ ) is equivalent to minimize  $|m|$  or, in other words, to minimize the (absolute) value of  $\mathbb{D}^*$ .

## 5 Numerical simulation of the dispersive effect

To illustrate the dispersive effect and explain the role of the high order homogenized equation (4.5), we perform some numerical experiments in 1-d. Similar computations previously appeared in [46], [2], therefore our goal is purely pedagogical and illustrative. To simplify, the source term  $f$  is set to zero. Then, by virtue of Lemma 2.8 the one-dimensional high order homogenized equation (1.4) reads as follows:

$$\frac{\partial^2 v_\epsilon}{\partial t^2} - a^* \frac{\partial^2 v_\epsilon}{\partial x^2} - \epsilon^2 a^* d^* \frac{\partial^4 v_\epsilon}{\partial x^4} = 0.$$

By virtue of Lemma 4.1, this equation is equivalent, approximately up to an error of  $\mathcal{O}(\epsilon^4)$ , to

$$\frac{\partial^2 v_\epsilon}{\partial t^2} - \epsilon^2 C \frac{\partial^4 v_\epsilon}{\partial x^2 \partial t^2} - a^* \frac{\partial^2 v_\epsilon}{\partial x^2} + \epsilon^2 a^* (C - d^*) \frac{\partial^4 v_\epsilon}{\partial x^4} = 0, \quad (5.1)$$

where  $C \geq 0$  plays the role of a parameter (a scalar in 1-d). Following the test case of [2], in the numerical simulations, the periodic coefficient is

$$a\left(\frac{x}{\epsilon}\right) = \sqrt{2} + \sin\left(2\pi\left(\frac{x}{\epsilon} - \frac{1}{4}\right)\right),$$

with  $\epsilon = 0.05$ . Then, the homogenized tensor  $a^*$  and dispersive tensor  $d^*$  are  $a^* = 1$  and  $d^* = 0.00909633$ , respectively. The computational domain is  $\Omega = (-0.5, 0.5)$ , complemented with periodic boundary conditions, and we discretize it with a space step  $\Delta x = 1/2000$  and a time step  $\Delta t = 0.02 \times \Delta x$ . We use a leapfrog scheme in time and, in space, a fourth-order centred finite difference scheme for the diffusion term and a second order centred scheme for the dispersive term. We consider two different sets of initial condition which are non-oscillating. The first type of initial data features a zero initial velocity and triggers two waves (see Figure 1) propagating symmetrically in opposite directions:

$$v_\epsilon(t=0, x) = \exp\left(-\frac{x^2}{0.05}\right) - \frac{1}{|\Omega|} \int_\Omega \exp\left(-\frac{x^2}{0.05}\right) dx \quad \text{and} \quad \frac{\partial v_\epsilon}{\partial t}(t=0, x) = 0.$$

The second set of initial data yields a single wave (see Figure 2) for the standard homogenized equations, propagating with group velocity  $\sqrt{a^*}$ :

$$v_\epsilon(t=0, x) = \exp(-64x^2) - \frac{1}{|\Omega|} \int_\Omega \exp(-64x^2) dx \quad \text{and} \quad \frac{\partial v_\epsilon}{\partial t}(t=0, x) = -\frac{\sqrt{a^*}x}{32} \exp(-64x^2).$$

In Figure 1 and Figure 2, we compare the numerical solutions of the original wave equation (1.1), of the homogenized wave equation (1.3) and of the high order homogenized equation (5.1), for three different values of the  $C$  parameter, and at different times  $T$ . The five different curves are: case 1, the solution of (1.1); case 2, the solution of (1.3); case 3, the solution of (5.1) with  $C = d^*$ ; case 4, the solution of (5.1) with  $C = 2d^*$  and case 5, the solution of (5.1) with  $C = 4d^*$ .

We notice that all solutions are very close (in the supremum norm) for short times (say  $T = 1$ ) while for larger times (say  $T = 100$ ) only the solutions of the high order homogenized equation stay close to the "exact" solution (while the homogenized solution propagates at the wrong speed, a clear manifestation of dispersive effects not taken into account in (1.3)). At very long times (say  $T = 400$ ), the agreement between the exact and high order homogenized solutions is very good for the first example but less convincing for the second example: this may be due to the more complex profile of the solutions which may be more prone to numerical diffusion/dispersion that pollute their accuracy for such long times. In any case, the high order homogenized equation (5.1) is clearly a better approximation than the standard homogenized equation (1.3).

## 6 Some properties of the Burnett tensor $\mathbb{D}^*$

We first investigate the dependence of  $\mathbb{D}^*$  to the choice of the periodicity cell. For any integer  $k \geq 1$ , define the coefficients

$$a_k(y) = a(ky) \quad \text{in } Y,$$

which are just the periodic repetition of smaller cells of size  $1/k$  in the unit cell  $Y$ . The same microstructure or geometry can be modeled by  $a$  or  $a_k$  but with a different value of the small parameter. Define

$$\epsilon_k = k\epsilon.$$

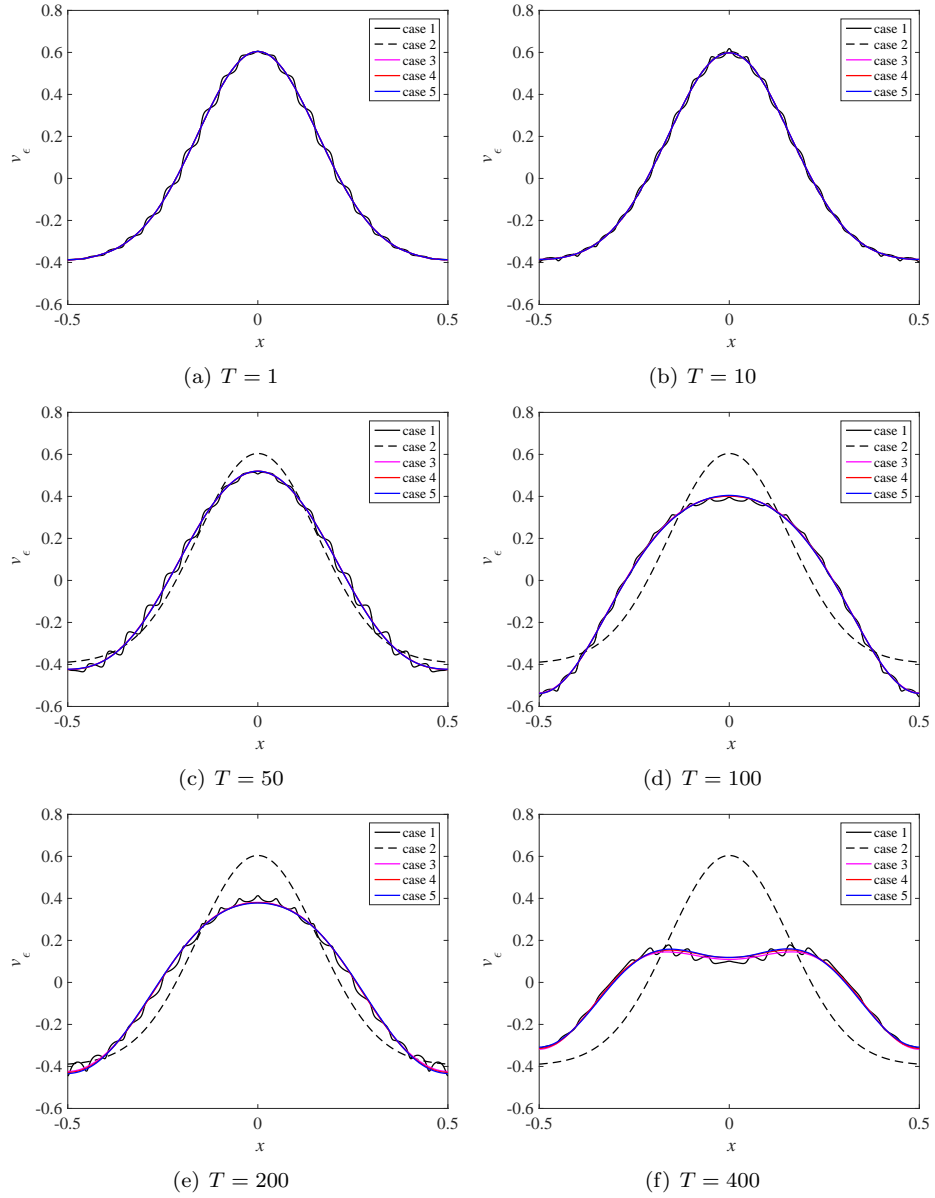


Figure 1: Plot of the solutions of equations (1.1), (1.3) and (5.1) with the first type of initial conditions.

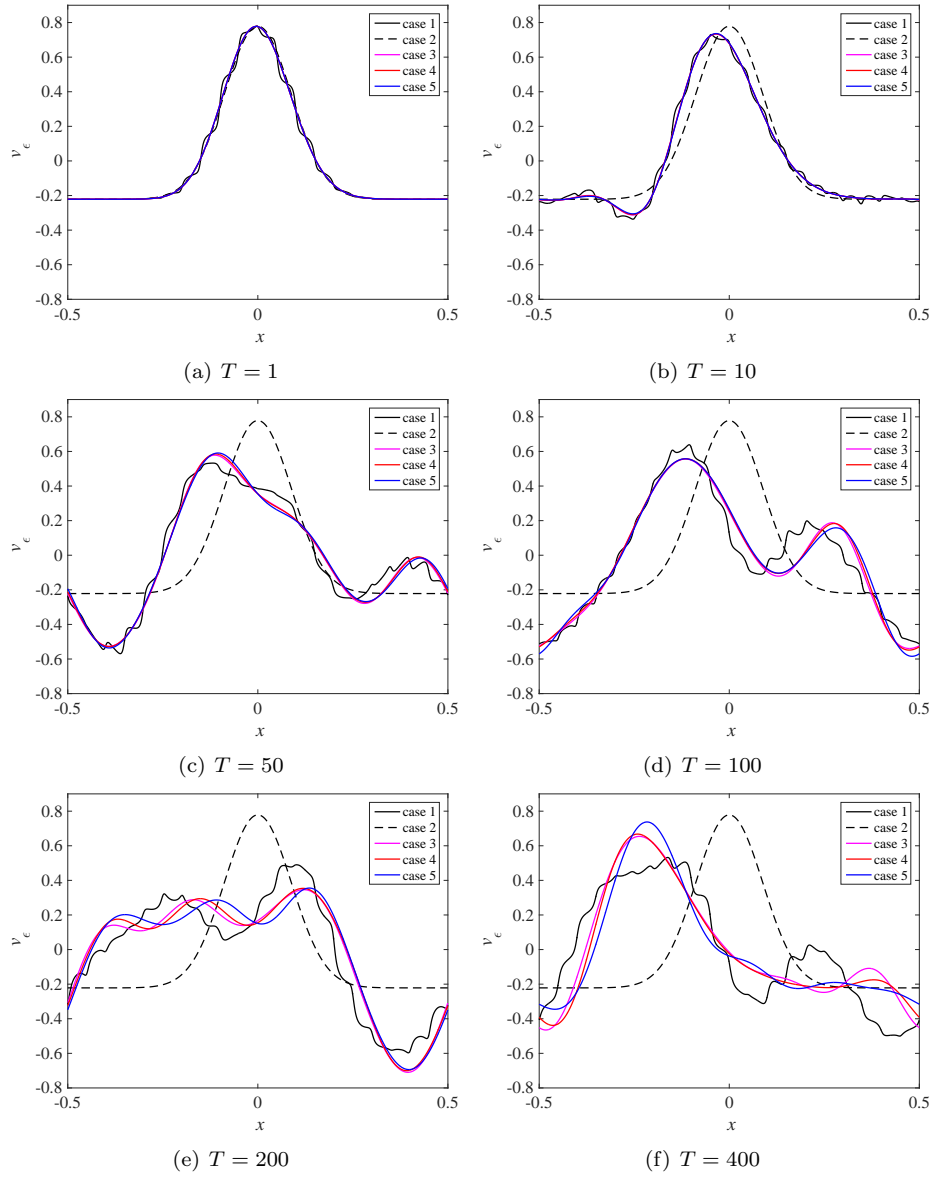


Figure 2: Plot of the solutions of equations (1.1), (1.3) and (5.1) with the second type of initial conditions.

Then we have

$$a\left(\frac{x}{\epsilon}\right) = a_k\left(\frac{x}{\epsilon_k}\right).$$

In other words, if more periodic patterns are present in the unit cell, there are less unit cells in the macroscopic domain and the ratio  $\epsilon_k$  is larger. One can reproduce the homogenization process of Section 2 with these new coefficients  $a_k$  and small parameters  $\epsilon_k$ . According to Proposition 2.2, the new high order homogenized equation of the wave equation (1.1) is

$$\frac{\partial^2 v_\epsilon}{\partial t^2} - \operatorname{div}(a_k^* \nabla v_\epsilon) + \epsilon_k^2 \mathbb{D}_k^* \nabla^4 v_\epsilon = f + \epsilon_k^2 \operatorname{div}(d_k^* \nabla f), \quad (6.1)$$

with new homogenized properties  $a_k^*, \mathbb{D}_k^*, d_k^*$  corresponding to the new coefficient  $a_k$ .

**Lemma 6.1.** *For any integer  $k \geq 1$ , one has*

$$a_k^* = a^*, \quad \mathbb{D}_k^* = k^{-2} \mathbb{D}^*, \quad d_k^* = k^{-2} d^*.$$

In other words,  $\epsilon_k^2 \mathbb{D}_k^* = \epsilon^2 \mathbb{D}^*$  and  $\epsilon_k^2 d_k^* = \epsilon^2 d^*$ .

*Proof.* Let us denote by  $\chi_i^k$  and  $\chi_{ij}^k$  the cell solutions for the coefficients  $a_k$ . It is easily seen that

$$\chi_i^k(y) = \frac{1}{k} \chi_i(ky), \quad \chi_{ij}^k(y) = \frac{1}{k^2} \chi_{ij}(ky),$$

from which we deduce the desired properties.  $\square$

**Remark 6.2.** *As a consequence of Lemma 6.1, the dispersion tensor can be made as small as desired (in norm) by considering smaller and smaller periodic patterns in the unit cell. However, there is no contradiction because the product  $\epsilon_k^2 \mathbb{D}_k^*$  is constant. In any case, there is no point in minimizing the norm or a negative linear combination of entries of  $\mathbb{D}^*$ , except if one adds a geometrical constraint (like an upper bound on the perimeter) which would prevent the unlimited fragmentation of the periodic microstructure.*

A similar result holds true if one considers a scaled version of the unit cell.

**Lemma 6.3.** *For any real number  $\kappa > 0$ , consider a scaled periodicity cell  $Z = (0, \kappa)^d$ . Introduce the scaled variable  $z := \kappa y$ , with  $y \in Y$ , and define the  $Z$ -periodic coefficients  $a_\kappa(z) := a(\frac{z}{\kappa})$ . Then, its homogenized coefficients satisfy*

$$a_\kappa^* = a^*, \quad \mathbb{D}_\kappa^* = \kappa^2 \mathbb{D}^*, \quad d_\kappa^* = \kappa^2 d^*.$$

*Proof.* Note first that, when computing homogenized formula on  $Z$ , one has to average on the cell  $Z$  which has volume  $\kappa^d$ . Let us denote by  $\chi_i^\kappa$  and  $\chi_{ij}^\kappa$  the cell solutions for the coefficients  $\tilde{a}$ . It is easily seen that

$$\chi_i^\kappa(z) = \kappa \chi_i\left(\frac{z}{\kappa}\right), \quad \chi_{ij}^\kappa(z) = \kappa^2 \chi_{ij}\left(\frac{z}{\kappa}\right),$$

from which we deduce the desired properties.  $\square$

**Remark 6.4.** *As a consequence of Lemma 6.3, if one can change the periodicity cell, then the dispersion tensor can be made as large (or small) as desired by considering larger (or smaller) periodicity cells. However, for a given physical configuration, there is no contradiction because the product  $\epsilon_\kappa^2 \mathbb{D}_\kappa^*$  is constant. In any case, if one fix the periodicity cell to be  $Y = (0, 1)^d$ , then one cannot use this scaling argument and the norm of  $\mathbb{D}^*$  can be bounded from above. Indeed, in 1-d, for a two-phase mixture of  $a_A, a_B$  in proportions  $\gamma, (1 - \gamma)$ , the following upper bound on  $-\mathbb{D}^*$  was proved in [19]*

$$-\mathbb{D}^* \leq \frac{1}{12} (a^*)^2 \gamma^2 (1 - \gamma)^2 (a_A^{-1} - a_B^{-1})^2,$$

and this upper bound is uniquely attained by a simple laminate of  $a_A, a_B$  with just one point-interface in the unit cell  $Y = (0, 1)$ .

We now consider the effect of rotations on the periodicity cell. The analytic formula of Lemma 6.5 will be useful to check some of our numerical results which feature equivalent shapes, up to rotations (see Remark 9.7).



**Lemma 6.5.** Let  $\mathcal{R}$  be a rotation matrix and consider the rotated variable  $z := \mathcal{R}y$ , as well as the rotated material properties  $\tilde{a}(z) := \mathcal{R}a(\mathcal{R}^T z)\mathcal{R}^T$ . Then, the homogenized properties of  $\tilde{a}(z)$  satisfy

$$\tilde{a}^* = \mathcal{R}a^*\mathcal{R}^T, \quad \tilde{d}^* = \mathcal{R}d^*\mathcal{R}^T, \quad \tilde{\mathbb{D}}^*(\xi \otimes \xi) \cdot (\xi \otimes \xi) = \mathbb{D}^*(\mathcal{R}\xi \otimes \mathcal{R}\xi) \cdot (\mathcal{R}\xi \otimes \mathcal{R}\xi).$$

*Proof.* Let us denote by  $\tilde{\chi}_i(z)$  and  $\tilde{\chi}_{ij}(z)$  the cell solutions for the coefficients  $\tilde{a}(z)$ . One can check that, for any vector  $\eta \in \mathbb{R}^d$ , we have  $\tilde{\chi}_\eta^{(1)}(z) = \chi_{\mathcal{R}\eta}^{(1)}(\mathcal{R}^T z)$  and  $\tilde{\chi}_\eta^{(2)}(z) = \chi_{\mathcal{R}\eta}^{(2)}(\mathcal{R}^T z)$ , from which we deduce the desired properties.  $\square$

In order to simplify the analysis of the fourth-order tensor  $\mathbb{D}^*$ , we choose to restrict the geometry of the periodic coefficients. From now on we make the following *8-fold symmetry* assumption on the periodic coefficient  $a(y)$ :

1.  $y \rightarrow a(y)$  is a scalar-valued function,
2.  $a$  is even in the sense that  $a = a \circ S_i$  for  $1 \leq i \leq d$ , where  $S_i$  is the symmetry operator defined by

$$S_i(y) = (y_1, \dots, y_{i-1}, -y_i, y_{i+1}, \dots, y_d),$$

3.  $a$  is 90°-rotationally invariant in the sense that  $a = a \circ P_{ij}$  for  $1 \leq i, j \leq d$ , where  $P_{ij}$  is the permutation operator defined by

$$P_{ij}(y) = P_{ij}(y_1, \dots, y_i, \dots, y_j, \dots, y_d) = (y_1, \dots, y_j, \dots, y_i, \dots, y_d).$$

Note that, by periodicity of the coefficients, one can consider that the unit cell is  $Y = (-1/2, +1/2)^d$  and the above assumption is equivalent to symmetries with respect to the principal hyperplanes (orthogonal to the main axis) and to the diagonal hyperplanes passing through the origin.

The following result is then easily proved (see e.g. section 3 in chapter 6 of [10]).

**Lemma 6.6.** Under the 8-fold symmetry assumption, if  $w$  is a  $Y$ -periodic solution of

$$-\operatorname{div}_y(a(y)\nabla_y w) = g \quad \text{in } Y,$$

then  $w \circ S_i$  is a  $Y$ -periodic solution of

$$-\operatorname{div}_y(a(y)\nabla_y(w \circ S_i)) = g \circ S_i \quad \text{in } Y,$$

and  $w \circ P_{ij}$  is a  $Y$ -periodic solution of

$$-\operatorname{div}_y(a(y)\nabla_y(w \circ P_{ij})) = g \circ P_{ij} \quad \text{in } Y.$$

**Proposition 6.7.** Under the 8-fold symmetry assumption, the dispersion tensor  $\mathbb{D}^*$  is characterized by two parameters  $\alpha, \beta \in \mathbb{R}$

$$\mathbb{D}^*(\eta \otimes \eta) : (\eta \otimes \eta) = -\alpha \sum_{i=1}^d \eta_i^4 - \beta \sum_{i,j=1, i < j}^d \eta_i^2 \eta_j^2, \quad (6.2)$$

with constant  $\alpha$  and  $\beta$ , independent of the indices  $i, j$  such that, for any  $i \neq j$ ,

$$\alpha := \int_Y a |\nabla \chi_{ii} - \chi_i \nabla \chi_i|^2 dy \quad (6.3)$$

$$\beta := \int_Y \left( 2a (\nabla \chi_{ii} - \chi_i \nabla \chi_i) \cdot (\nabla \chi_{jj} - \chi_j \nabla \chi_j) + a |\nabla \chi_{ij} + \nabla \chi_{ji} - \chi_i \nabla \chi_j - \chi_j \nabla \chi_i|^2 \right) dy. \quad (6.4)$$

*Proof.* The fact that, under the 8-fold symmetry assumption, the homogenized tensor  $a^*$  is scalar is classical. Using Lemma 6.6 one can check the following symmetry properties on the cell solutions

$$\chi_i \circ S_i = -\chi_i, \quad \chi_i \circ S_j = \chi_i \text{ for } i \neq j,$$

$$\chi_{ii} \circ S_k = \chi_{ii} \text{ for } 1 \leq k \leq d,$$

$$\text{for } i \neq j \quad \chi_{ij} \circ S_i = \chi_{ij} \circ S_j = -\chi_{ij}, \quad \chi_{ij} \circ S_k = \chi_{ij} \text{ for } k \neq i, j.$$

In particular,  $\chi_{ij}$  and  $(\chi_i\chi_j)$  have the same symmetry properties. On the other hand, we also have the following permutation properties

$$\chi_i \circ P_{ij} = \chi_j, \quad \chi_{ii} \circ P_{ij} = \chi_{jj}, \quad \chi_{ij} \circ P_{jk} = \chi_{ik} \quad \text{for } i \neq j.$$

From formula (2.31) in Proposition 2.5 we find that

$$\mathbb{D}_{ijkl}^* := - \left[ \int_Y a \nabla \left( \chi_{ij} - \frac{1}{2} \chi_i \chi_j \right) \cdot \nabla \left( \chi_{kl} - \frac{1}{2} \chi_k \chi_l \right) dy \right]^S, \quad (6.5)$$

where the index  $S$  means that  $\mathbb{D}_{ijkl}^*$  has to be symmetrized. Now, using the permutation properties of the cell functions, it is easily seen that

$$\mathbb{D}_{iii}^* = - \int_Y a |\nabla \chi_{ii} - \chi_i \nabla \chi_i|^2 dy \quad \text{does not depend on the direction } i.$$

Similarly,

$$2\mathbb{D}_{iij}^* = - \int_Y \left( 2a (\nabla \chi_{ii} - \chi_i \nabla \chi_i) \cdot (\nabla \chi_{jj} - \chi_j \nabla \chi_j) + a |\nabla \chi_{ij} + \nabla \chi_{ji} - \chi_i \nabla \chi_j - \chi_j \nabla \chi_i|^2 \right) dy$$

is independent of the couple of indices  $i \neq j$ . Let us show that all other entries of the tensor  $\mathbb{D}^*$  vanish. Consider, for example, the entry

$$\mathbb{D}_{iii}^* = - \int_Y a \nabla \left( \chi_{ii} - \frac{1}{2} \chi_i^2 \right) \cdot \nabla \left( \chi_{ij} - \frac{1}{2} \chi_i \chi_j \right) dy.$$

We decompose it as

$$\mathbb{D}_{iii}^* = \sum_{k=1}^d - \int_Y a d_{ii}^k d_{ij}^k dy \quad \text{with} \quad d_{ii}^k = \frac{\partial}{\partial y_k} \left( \chi_{ii} - \frac{1}{2} \chi_i^2 \right), \quad d_{ij}^k = \frac{\partial}{\partial y_k} \left( \chi_{ij} - \frac{1}{2} \chi_i \chi_j \right).$$

For  $k \neq i, j$ , the function  $d_{ii}^k = d_{ii}^k \circ S_j$  is even with respect to  $y_j$ , while  $d_{ij}^k = -d_{ij}^k \circ S_j$  is odd with respect to  $y_j$ . Therefore, the integrand  $a d_{ii}^k d_{ij}^k$  has zero average on  $Y$ . For  $k = i$ ,  $d_{ii}^i = d_{ii}^i \circ S_i$  is odd with respect to  $y_i$ , while  $d_{ij}^i = d_{ij}^i \circ S_i$  is even with respect to  $y_i$  (as the derivative of an odd function). Again, the integrand  $a d_{ii}^i d_{ij}^i$  has zero average on  $Y$ . Eventually, for  $k = j$ ,  $d_{ii}^j = d_{ii}^j \circ S_j$  is odd with respect to  $y_j$  (as the derivative of an even function), while  $d_{ij}^j = d_{ij}^j \circ S_j$  is even with respect to  $y_j$  (as the derivative of an odd function), and the integrand  $a d_{ii}^j d_{ij}^j$  has zero average on  $Y$ . This implies that  $\mathbb{D}_{iii}^* = 0$ .

A similar argument work for all other entries  $\mathbb{D}_{iijk}^*$ , with different  $i, j, k$ , and  $\mathbb{D}_{ijkl}^*$ , with different  $i, j, k, l$ . A key ingredient is always that  $\chi_{ij}$  and  $(\chi_i\chi_j)$  have the same symmetry properties. Therefore, we obtain the desired result.  $\square$

## 7 Shape optimization

### 7.1 Two phase periodic mixture

From now on we shall study dispersive effects for wave propagation in a two-phase periodic medium. More precisely, in the context of periodic homogenization we assume that the unit cell  $Y = (0, 1)^d$  is decomposed in two subdomains  $Y^A$  and  $Y^B$ , separated by a smooth interface  $\Gamma$  (see Figure 3). The subdomains  $Y^A$  and  $Y^B$  are filled with an isotropic liner material, which coefficients  $a^A$  and  $a^B$ , respectively. We consider only those mixtures which satisfy the 8-fold symmetry assumption of Section 6. Our ultimate goal is to find the set of all possible dispersion tensors  $\mathbb{D}^*$  with, possibly, a volume constraint for the two phases, a perimeter constraint (on the measure of  $\Gamma$ ) and a prescribed homogenized property  $a^*$ . In particular, we would like to know which microstructures in the unit cell yield minimal or maximal values of  $\mathbb{D}^*$ . To do so, we study the shape optimization problem which determines the optimal geometry in the unit cell  $Y$  in order to minimize some objective function depending on  $\mathbb{D}^*$  or more generally on the first and second-order cell problems (which give the value of  $\mathbb{D}^*$  by virtue of Proposition 2.5). More specifically, we consider coefficients defined by

$$a(y) = a^A \mathbf{1}_{Y^A}(y) + a^B \mathbf{1}_{Y^B}(y),$$

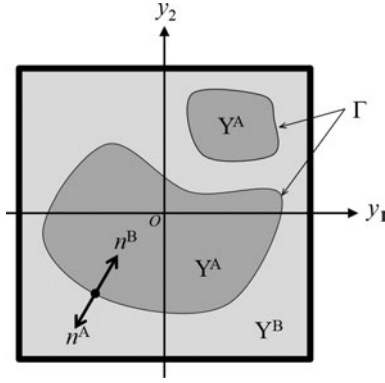


Figure 3: Periodicity cell of a two-phase composite

where  $a^A, a^B > 0$  are two constant positive real numbers,  $\mathbf{1}_{Y^A}(y), \mathbf{1}_{Y^B}(y)$  are the characteristic functions of  $Y^A$  and  $Y^B$ .

To optimize the dispersive properties of a periodic two-phase geometry, we consider the following objective function:

$$J(Y^A) = \int_Y \mathcal{J}(y, \{\chi_i\}_{1 \leq i \leq d}, \{\chi_{ij}\}_{1 \leq i, j \leq d}) dy, \quad (7.1)$$

where  $\mathcal{J}$  is a smooth function satisfying adequate growth conditions,  $\chi_i$  is the first order cell solution of (2.6),  $\chi_{ij}$  is the second order cell solution of (2.12). By virtue of Proposition 2.5, the entries of the dispersion tensor  $\mathbb{D}^*$  are of the type of (7.1).

We shall minimize the objective function  $J(Y^A)$  with constraints (all of them or just some of them) like volume fraction of  $Y^A$  and  $Y^B$ , perimeter or measure of  $\Gamma$ , homogenized tensor  $a^*$ . As is well known, optimal designs do not always exist in such problems [3], [37], [41], [51], unless some smoothness, geometrical or topological constraint is added. We shall not discuss this issue since our goal is rather numerical than theoretical.

## 7.2 Shape derivative in the multi-material problem

In order to minimize the objective function (7.1), a gradient based shape optimization algorithm [34, 41, 49] is applied. Most of the works on the Hadamard method for computing shape sensitivity are concerned with problems where the domain boundary is the optimization variable. However, here we rather optimize an interface between two materials and there are fewer works in this setting. Let us mention the cases of Darcy flows [13], conductivity problems [26, 40] and elasticity systems [5, 7, 29]. There are also some works concerned with the optimization of the homogenized tensor  $a^*$  in terms of the cell properties  $a(y)$  [11], [24], [47]. In this work, we follow the same approach but applied to the higher order cell problems in homogenization and to the dispersion tensor  $\mathbb{D}^*$ .

To begin with, we recall the approach of Murat and Simon [37] for shape differentiation. For a smooth reference open set  $\Omega$ , we consider domains of the type

$$\Omega_\theta = (\text{Id} + \theta)(\Omega),$$

with a vector field  $\theta \in W^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d)$  such that  $\theta$  is tangential on  $\partial\Omega$ .

**Definition 7.1.** *The shape derivative of  $J(\Omega)$  at  $\Omega$  is defined as the Fréchet derivative in  $W^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d)$  at 0 of the application  $\theta \rightarrow J((\text{Id} + \theta)(\Omega))$ . The following asymptotic expansion holds in the neighborhood of  $0 \in W^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d)$ :*

$$J((\text{Id} + \theta)(\Omega)) = J(\Omega) + J'(\Omega)(\theta) + o(\theta) \quad \text{with} \quad \lim_{\theta \rightarrow 0} \frac{|o(\theta)|}{\|\theta\|} = 0, \quad (7.2)$$

where the shape derivative  $J'(\Omega)$  is a continuous linear form on  $W^{1,\infty}(\mathbb{R}^d, \mathbb{R}^d)$ .

**Lemma 7.2.** *Let  $\Omega$  be a smooth bounded open set and  $\phi_1(x) \in W^{1,1}(\mathbb{R}^d)$  and  $\phi_2(x) \in W^{2,1}(\mathbb{R}^d)$  be two given functions. The shape derivatives of*

$$J_1(\Omega) = \int_\Omega \phi_1(x) dx \quad \text{and} \quad J_2(\Omega) = \int_{\partial\Omega} \phi_2(x) ds$$

are

$$J'_1(\Omega)(\theta) = \int_{\partial\Omega} \theta(x) \cdot n(x) \phi_1(x) ds$$

and

$$J'_2(\Omega)(\theta) = \int_{\partial\Omega} \theta(x) \cdot n(x) \left( \frac{\partial\phi_2(x)}{\partial n} + H(x)\phi_2(x) \right) ds,$$

for any  $\theta \in W^{1,\infty}(\mathbb{R}^d; \mathbb{R}^d)$ , respectively, where  $H$  is the mean curvature of  $\partial\Omega$ , defined by  $H = \operatorname{div} n$ ,  $n$  is the unit normal vector on  $\partial\Omega$  and  $ds$  is the  $(d-1)$ -dimensional measure along  $\partial\Omega$ .

**Theorem 7.3.** *The shape derivative of the objective function  $J$ , defined in (7.1) reads*

$$J'(Y^A)(\theta) = \int_{\Gamma} \mathcal{D} \left( \{\chi_i\}_{1 \leq i \leq d}, \{\chi_{ij}\}_{1 \leq i,j \leq d}, \{p_i\}_{1 \leq i \leq d}, \{p_{ij}\}_{1 \leq i,j \leq d} \right) \theta \cdot n ds, \quad (7.3)$$

with

$$\begin{aligned} \mathcal{D} &= [\mathcal{J}] + [a](\nabla\chi_i - e_i)_t \cdot \left( \nabla p_i + e_j \left( p_{ij} - \int_Y p_{ij} dy \right) \right)_t + [a](\nabla\chi_{ij} - \chi_j e_i)_t \cdot \nabla_t p_{ij} \\ &- [a^{-1}] \left( a(\nabla\chi_i - e_i) \cdot n \right) \left( a \left( \nabla p_i + e_j \left( p_{ij} - \int_Y p_{ij} dy \right) \right) \cdot n \right) - [a^{-1}] \left( a(\nabla\chi_{ij} - \chi_j e_i) \cdot n \right) \left( a \nabla p_{ij} \cdot n \right), \end{aligned}$$

where  $[*] = *^A - *^B$  denotes the jump through the interface  $\Gamma$  and  $n = n^A = -n^B$  is the unit normal vector of  $\Gamma$ . The suffix  $_t$  denotes the tangential component of a vector. The adjoint states  $p_i, p_{ij} \in H_{\#,0}^1(Y)$  are defined as the unique periodic solutions of the following adjoint problems:

$$-\operatorname{div}(a\nabla p_i) = -\frac{\partial \mathcal{J}}{\partial \chi_i} + \operatorname{div} \left( a e_j \left( p_{ij} - \int_Y p_{ij} \right) \right) + a e_j \cdot \nabla p_{ij} \quad \text{in } Y, \quad (7.4)$$

$$-\operatorname{div}(a\nabla p_{ij}) = -\frac{\partial \mathcal{J}}{\partial \chi_{ij}} \quad \text{in } Y. \quad (7.5)$$

**Remark 7.4.** *In the statement of Theorem 7.3 the Einstein summation convention with respect to repeated indices is used. The solutions of the adjoint equations (7.4) and (7.5) are defined up to an additive constant. They are unique in  $H_{\#,0}^1(Y)$ , namely when their average on  $Y$  is zero. Therefore, when this normalization condition is used, the integral term  $(\int_Y p_{ij})$  in (7.4) can safely be dropped.*

**Remark 7.5.** *The governing equation (2.12) of the second order corrector functions  $\chi_{ij}$  depends on the first order corrector functions  $\chi_i$ . Therefore, in numerical practice, the functions  $\chi_{ij}$  are computed after the functions  $\chi_i$ . On the other hand, the adjoint equation (7.4) for  $p_i$  depends on  $p_{ij}$ , while the other adjoint equation (7.5) depends merely on the corrector functions  $\chi_i$  and  $\chi_{ij}$ . Therefore, the second order adjoint functions  $p_{ij}$  are computed first, followed by the computation of the first order adjoint functions  $p_i$ . This peculiarity is similar to the backward character of the adjoint equation for a time dependent problem.*

The proof of Theorem 7.3 is obtained by a standard, albeit tedious, application of the Lagrangian method for shape derivation (see [7, 26, 40] for similar proofs). For the sake of completeness, it is given in Section 10.

### 7.3 Shape derivative of a discrete approximation

Although the formulation discussed in the previous subsection is satisfying from a mathematical point of view, its numerical implementation is quite tricky since it requires one of the two following delicate algorithms. A first possibility is to re-mesh at each iteration so that the mesh is fitted to the interface  $\Gamma$ : then, jumps, as well as continuous quantities, can be accurately computed (see section 6.4 in [6]). A second possibility is to fix the mesh and capture  $\Gamma$  by e.g. a level set function. In this latter case, only approximate jumps and continuous quantities at the interface can be computed (see [7]). Both approaches are not obvious to implement in practice. To avoid these difficulties, we use the approximated shape sensitivity in the multi-material setting proposed in [5], [33], [53]. In this formulation, the optimization problem is first discretized and second its shape sensitivity is derived. Let us introduce a finite-dimensional space of conforming finite elements  $V_h \subset H_{0,\#}^1(Y)$  in which are computed the approximate solutions  $\chi_i^h$  of (2.6),  $\chi_{ij}^h$  of (2.12),  $p_i^h$  of (7.4) and  $p_{ij}^h$  of (7.5). More precisely,  $\chi_i^h \in V_h$  and  $\chi_{ij}^h \in V_h$  are the unique solutions of, respectively,

$$\int_Y a(\nabla\chi_i^h - e_i) \cdot \nabla\phi^h dy = 0 \quad \forall \phi^h \in V_h, \quad (7.6)$$

$$\int_Y a(\nabla \chi_{ij}^h - \chi_i^h e_j) \cdot \nabla \phi^h dy = \int_Y (a_{ij} - a_{ij}^{*,h} - a \nabla \chi_i^h \cdot e_j) \phi^h dy \quad \forall \phi^h \in V_h. \quad (7.7)$$

The precise definitions of  $p_i^h$  and  $p_{ij}^h$  will be given in the proof below. Typically, these approximate solutions are of the type

$$\chi_i^h := \sum_{k=1}^{N_h} \chi_{i,k}(\Gamma) \varphi_k(x), \quad p_i^h := \sum_{k=1}^{N_h} p_{i,k}(\Gamma) \varphi_k(x)$$

where  $N_h$  is the dimension of  $V_h$ ,  $\varphi_k(x)$  are the shape functions and  $\chi_{i,k}(\Gamma)$ , resp.  $p_{i,k}(\Gamma)$ , are the nodal values of  $\chi_i^h$ , resp.  $p_i^h$ , which depend on the interface  $\Gamma$ . However, the basis functions  $\varphi_k$  are independent of  $\Gamma$  and, in particular, do not satisfy any special transmission conditions at the interface  $\Gamma$ . It implies that the state functions  $\chi_i^h$  are shape differentiable [5].

We introduce the discrete objective function

$$J_h(Y^A) = \int_Y \mathcal{J}(y, \{\chi_i^h\}_{1 \leq i \leq d}, \{\chi_{ij}^h\}_{1 \leq i, j \leq d}) dy. \quad (7.8)$$

**Proposition 7.6.** *The discrete objective function  $J_h$  is shape differentiable and its derivative reads*

$$J'_h(Y^A)(\theta) = \int_{\Gamma} \mathcal{D}_h(\{\chi_i^h\}_{1 \leq i \leq d}, \{p_i^h\}_{1 \leq i \leq d},) \theta \cdot n ds, \quad (7.9)$$

where

$$\begin{aligned} \mathcal{D}_h := [\mathcal{J}] + [a] & \left\{ (\nabla \chi_i^h - e_i) \cdot \nabla p_i^h + (\nabla \chi_{ij}^h - \chi_i^h e_j) \cdot \nabla p_{ij}^h + \nabla \chi_i^h \cdot e_j p_{ij}^h \right. \\ & \left. - \nabla \chi_j^h \cdot e_i \int_Y p_{ij}^h dy + e_i \cdot e_j \left( \int_Y p_{ij}^h dy - p_{ij}^h \right) \right\}. \end{aligned} \quad (7.10)$$

*Proof.* The proof follows that of Proposition 1.5 in [5]. We use the Lagrangian method, which introduces a Lagrangian  $\mathcal{L}_h$  as the sum of the objective function and of the constraints multiplied by Lagrange multipliers, namely the discrete variational formulations (7.6) and (7.7),

$$\begin{aligned} \mathcal{L}_h(\Gamma, \{\hat{\chi}_i^h\}_{1 \leq i \leq d}, \{\hat{\chi}_{ij}^h\}_{1 \leq i, j \leq d}, \{\hat{p}_i^h\}_{1 \leq i \leq d}, \{\hat{p}_{ij}^h\}_{1 \leq i, j \leq d}) := \\ \int_Y \mathcal{J}(y, \{\hat{\chi}_i^h\}_{1 \leq i \leq d}, \{\hat{\chi}_{ij}^h\}_{1 \leq i, j \leq d}) dy + \int_Y a(\nabla \hat{\chi}_i^h - e_i) \cdot \nabla \hat{p}_i^h dy \\ + \int_Y a(\nabla \hat{\chi}_{ij}^h - \hat{\chi}_i^h e_j) \cdot \nabla \hat{p}_{ij}^h dy - \int_Y (a_{ij} - a_{ij}^{*,h} - a \nabla \hat{\chi}_i^h \cdot e_j) \hat{p}_{ij}^h dy, \end{aligned} \quad (7.11)$$

where the functions  $\hat{\chi}_i^h, \hat{\chi}_{ij}^h, \hat{p}_i^h, \hat{p}_{ij}^h$  are any functions in  $V_h$  and with

$$a_{ij}^{*,h} = \int_Y (a_{ij} - a \nabla \hat{\chi}_j^h \cdot e_i) dy.$$

The space  $V_h$  is independent of the interface  $\Gamma$ . Therefore, the Lagrangian  $\mathcal{L}_h$  can be differentiated in the usual manner and its stationarity is going to give the optimality conditions of the optimization problem.

By definition the partial derivative of  $\mathcal{L}_h$  with respect to  $p_i^h$  and  $p_{ij}^h$  leads to the variational formulation (7.6) and (7.7). Next, the discrete adjoint equations are obtained by taking the partial derivative of  $\mathcal{L}_h$  with respect to the variables  $\chi_i^h$  and  $\chi_{ij}^h$ . It yields the following discrete variational formulations

$$\int_Y a(\nabla p_i^h + e_j p_{ij}^h) \cdot \nabla \phi^h dy + \int_Y \frac{\partial \mathcal{J}}{\partial \chi_i^h} \phi^h dy - \int_Y a \nabla p_{ij}^h \cdot e_j \phi^h dy = 0 \quad \forall \phi^h \in V_h, \quad (7.12)$$

$$\int_Y a \nabla p_{ij}^h \cdot \nabla \phi^h dy + \int_Y \frac{\partial \mathcal{J}}{\partial \chi_{ij}^h} \phi^h dy = 0 \quad \forall \phi^h \in V_h, \quad (7.13)$$

which are approximations of the exact variational formulations of (7.4) and (7.5).

Eventually, by a classical result (see e.g. Lemma 3.5 in [7]), the partial derivative of  $\mathcal{L}_h$  with respect to  $\Gamma$  is precisely the shape derivative of  $J_h$ . Applying Lemma 7.2 to the Lagrangian (7.11) leads to (7.10).  $\square$

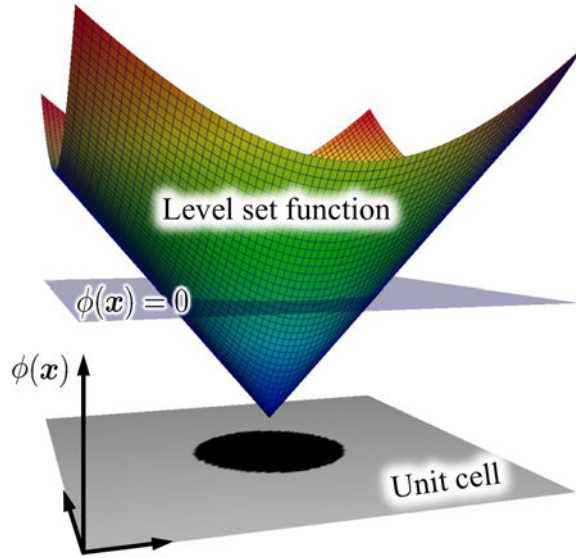


Figure 4: Level set function

## 8 Level set and optimization algorithms

In order to make it possible to change topology by merging boundaries during the shape optimization procedure, the level set method, introduced by Osher and Sethian [39], is used. As shown in Figure 4, the basic idea is that the boundary is represented as the zero iso-surface of a level set function  $\phi(y)$  and the subdomains are distinguished by the sign of the level set function  $\phi(x)$ . More precisely, the level set function  $\phi(y)$  is defined by

$$\forall y \in Y \quad \begin{cases} \phi(y) > 0 & \text{if } y \in Y^A \\ \phi(y) = 0 & \text{if } y \in \Gamma \\ \phi(y) < 0 & \text{if } y \in Y^B \end{cases}$$

Based on the level set representation, the approximate material property for the finite element analyses is defined as follows:

$$a_Y(y) := a^A + h_w(\phi(y))(a^B - a^A), \quad \forall y \in Y \quad (8.1)$$

where  $h_w : \mathbb{R} \rightarrow \mathbb{R}$  is a smooth monotone approximate Heaviside function:

$$\forall \tau \in \mathbb{R}, \quad h_w(\tau) := \begin{cases} 0 & \text{if } \tau < -w \\ \frac{1}{2} \left( 1 + \frac{\tau}{w} + \frac{1}{\pi} \sin\left(\frac{\pi\tau}{w}\right) \right) & \text{if } -w \leq \tau \leq w \\ 1 & \text{if } \tau > w \end{cases} \quad (8.2)$$

where the parameter  $w > 0$  is the width of the smoothed approximate interface. There is nothing critical in this interface smoothing process (for example, other functions  $h_w$  could be used), but it makes easier the finite element implementation. For instance, the boundary element method [27] is not required here.

In the level set method for shape optimization, the shape changes during the optimization is represented as an evolution of the level set function. That is, introducing fictitious time  $t \in [0, T]$  (that could be interpreted as a descent step), the shape evolution is obtained by solving the following Hamilton-Jacobi equation:

$$\frac{\partial \phi}{\partial t} + V |\nabla \phi| = 0, \quad t \in (0, T), x \in Y, \quad (8.3)$$

with periodic boundary conditions and where the normal velocity  $V$  is defined as a descent direction for the shape sensitivity

$$J'(Y^A)(\theta) = - \int_{\Gamma} v \theta \cdot n \, ds. \quad (8.4)$$

A typical simple choice is to define  $V$  as an extension of  $v$  (which is defined merely on the interface  $\Gamma$  by (8.4)) to the entire cell  $Y$ . However, it is well known that the shape sensitivity does not have sufficient smoothness [34]

and following a classical regularization process we replace  $V$  by a regularized variant  $V_{reg}$  which is defined as the unique solution in  $H_{\#}^1(Y)$  of

$$\int_Y \left( \epsilon_r^2 \nabla V_{reg} \cdot \nabla \tilde{V} + V_{reg} \tilde{V} \right) dy = \int_{\Gamma} v \tilde{V} dy \quad \text{for any } \tilde{V} \in H_{\#}^1(Y)$$

where  $\epsilon_r > 0$  is a regularization parameter, having the interpretation of a smoothing length (typically of the order of a few mesh cell size). In numerical practice, since the interface  $\Gamma$  is not exactly meshed, we replace the interface integral in the above variational formulation by a volume integral with a smoothed Dirac function  $\delta_w(d_{\Gamma}(y))$  where  $d_{\Gamma}$  is the signed distance function to the interface  $\Gamma$  and  $\delta_{\eta}$  is defined as follows:

$$\delta_{\eta}(\tau) := \begin{cases} 0 & \text{if } \tau < -\eta \\ \frac{1}{2\eta} \left( 1 + \cos\left(\frac{\pi\tau}{\eta}\right) \right) & \text{if } -\eta \leq \tau \leq \eta \\ 0 & \text{if } \tau > \eta \end{cases}$$

where  $\eta > 0$  is a small parameter.

In order to minimize numerical dissipation in solving the Hamilton-Jacobi equation (8.3), the level set function is reinitialized as the signed distance function  $d_{\Gamma}$  at each optimization iteration by solving

$$\frac{\partial \phi}{\partial t} + \text{sign}(\phi_0) (|\nabla \phi| - 1) = 0 \quad \text{in } Y$$

starting from the initial condition  $\phi_0(y)$ , the prior level set function. This equation, as well as the Hamilton-Jacobi equation (8.3), are solved by a standard second-order upwind explicit finite difference scheme.

In truth, we are performing constrained optimization so that the velocity  $V$  is not computed only in terms of the derivative of the objective function (8.4) but also in terms of the constraints derivatives. More precisely, we rely on a standard Lagrangian approach, i.e. we replace the objective function by the Lagrangian which is the sum of the objective function and of the constraints multiplied by Lagrange multipliers. These Lagrange multipliers are updated at each iteration in such a way that the constraints are exactly satisfied.

The optimization process goes on as follows. First, the level set function is initialized to represent an appropriate initial configuration. Second, iterations of a steepest descent method start. Each iteration is made of the following steps. In a first step, the governing equations are solved using the finite element method and the objective function is computed. If the objective function is converged, the optimization is stopped. If not, the adjoint equations are solved in a second step. In a third step, the Lagrange multipliers are estimated to satisfy the constraints and the resulting shape derivative of the Lagrangian is used to deduce the velocity  $V$  in (8.3) (this velocity is regularized as explained above). In a fourth step, the level set function is updated based on the Hamilton-Jacobi equation (8.3). Note that the Lagrange multiplier for the volume constraint is computed using Newton's method so that the volume constraint is exactly satisfied. Finally, if the objective function is improved and all constraints are satisfied, the time increment of the Hamilton-Jacobi equation is increased and we go back to the second step for the next optimization iteration. Otherwise, the time increment is decreased and we go back to the fourth step.

## 9 Numerical simulations

In our numerical simulations, we impose the 8-fold symmetry condition for the two-dimensional unit cell. As shown in Figure 5, the analysis domain is a quarter of the unit cell (for simplicity), while the design domain is one eighth of the unit cell, namely half of the analysis domain. The design is recovered on the other half of the analysis domain by symmetry with respect to the diagonal. The finite element analysis is performed with the **FreeFEM++** software [25]. The domain is meshed with triangular elements and we use  $\mathbb{P}1$  finite elements. The two phases are isotropic with material properties  $a_A = 10$  and  $a_B = 20$ , respectively. By our 8-fold symmetry assumption, the dispersive tensor  $\mathbb{D}^*$  is characterized by two scalar coefficients  $\alpha$  and  $\beta$  (see Proposition 6.7).

In all our numerical examples, we rely on a structured triangular mesh for the finite element analysis. This mesh is obtained from a regular squared mesh by dividing each square in four triangles along its diagonals. The squared mesh is used for the finite differences discretization of the Hamilton-Jacobi equation. The regularization parameter is set to  $\epsilon_r^2 = 0.05$ , the width of the approximated Heaviside function is set to  $w = 0.02$  and the width of the approximated Dirac function is  $\eta = 0.055$ .

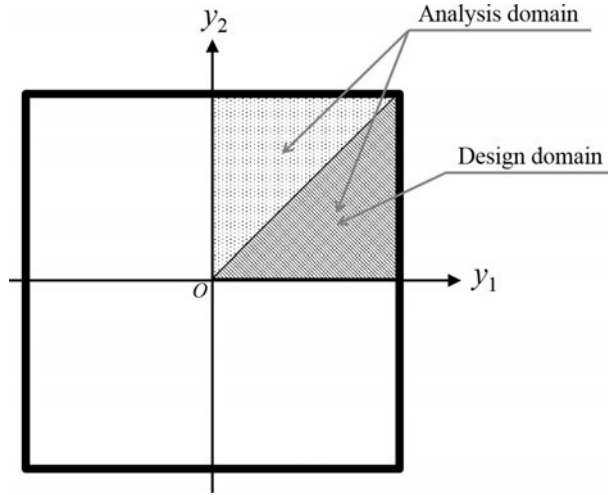


Figure 5: Analysis and design domain for the unit cell

### 9.1 Optimizing an energy associated to $\mathbb{D}^*$

In this subsection, as a first numerical test, we choose the specific direction  $\eta = (1, 1)$  and we minimize or maximize the energy  $\mathbb{D}^*(\eta \otimes \eta) : (\eta \otimes \eta)$  for the Burnett tensor  $\mathbb{D}^*$  with volume constraint, perimeter constraint and prescribing the homogenized tensor  $a^*$  as follows:

$$\begin{aligned}
 \min_{\Gamma} \quad \text{or} \quad \max_{\Gamma} \quad & J(\Gamma) := 2\alpha(\Gamma) + \beta(\Gamma) \\
 \text{subject to :} \quad & G_v(\Gamma) := \frac{\int_{Y^A} dy}{\int_{Y^A \cup Y^B} dy} - \overline{G}_v = 0 \\
 & G_p(\Gamma) := \int_{\Gamma} d\Gamma - \overline{G}_p \leq 0 \\
 & G_{a^*}(\Gamma) := \|a^*(\Gamma) - \overline{a^*}\| = 0
 \end{aligned}$$

where  $G_v$ ,  $G_p$  and  $G_{a^*}$  are constraint functions for the volume, the perimeter and the homogenized tensor  $a^*$ , respectively. The constants  $\overline{G}_v$ ,  $\overline{G}_p$  and  $\overline{a^*}$  are prescribed values for these constraints, respectively. We shall use the optimization algorithm of Section 8. However, it is not obvious to find an admissible initial configuration, satisfying all constraints. Therefore, we adopt the following four-step optimization procedure, starting from any initialization:

**Step 1:** optimize  $G_v$  alone to satisfy  $G_v = 0$ .

**Step 2:** minimize  $G_p$ , while keeping  $G_v = 0$ , to satisfy  $G_p \leq 0$ .

**Step 3:** minimize  $G_{a^*}$ , with the constraints  $G_v = 0$  and  $G_p \leq 0$ , to satisfy  $G_{a^*} = 0$ .

**Step 4:** optimize  $J(\Gamma)$ , with the constraints  $G_v = 0$ ,  $G_p \leq 0$  and  $G_{a^*} = 0$ .

In this subsection, we use a  $50 \times 50$  structured mesh for the analysis domain. The isotropic materials A and B have material properties  $a_A = 10$  and  $a_B = 20$ . The upper limit of the perimeter constraint is set to  $\overline{G}_p = 1.5$ . We consider two cases for the volume constraint: either  $\overline{G}_v = 0.9$  or  $\overline{G}_v = 0.1$ , which can be interpreted as material A being the inclusion in the first case, and material B being the inclusion in the second case. By symmetry, the homogenized tensor  $\overline{a^*}$  is isotropic and its prescribed scalar value is set to 10.705 in the first case and 18.72 in the second case. The relative error for judging whether the constraint function  $G_{a^*}$  is satisfied is set to  $5 \times 10^{-3}$ .

Figures 6 and 7 show initial and optimal configurations when material A (in black) is the inclusion and when material B (in white) is the inclusion, respectively. We test three different initial configurations, for which the optimal configurations may be quite different. Therefore, we guess there are many local optimal solutions for this type of optimization problems. In the minimization cases, the inclusions are changed to more complex and detailed shape. This is consistent with our Remark 6.2 which states that smaller inclusions yield smaller dispersion (in absolute value). On the other hand, in the maximization cases, the smaller inclusions may merge and give rise to simpler optimal shapes.



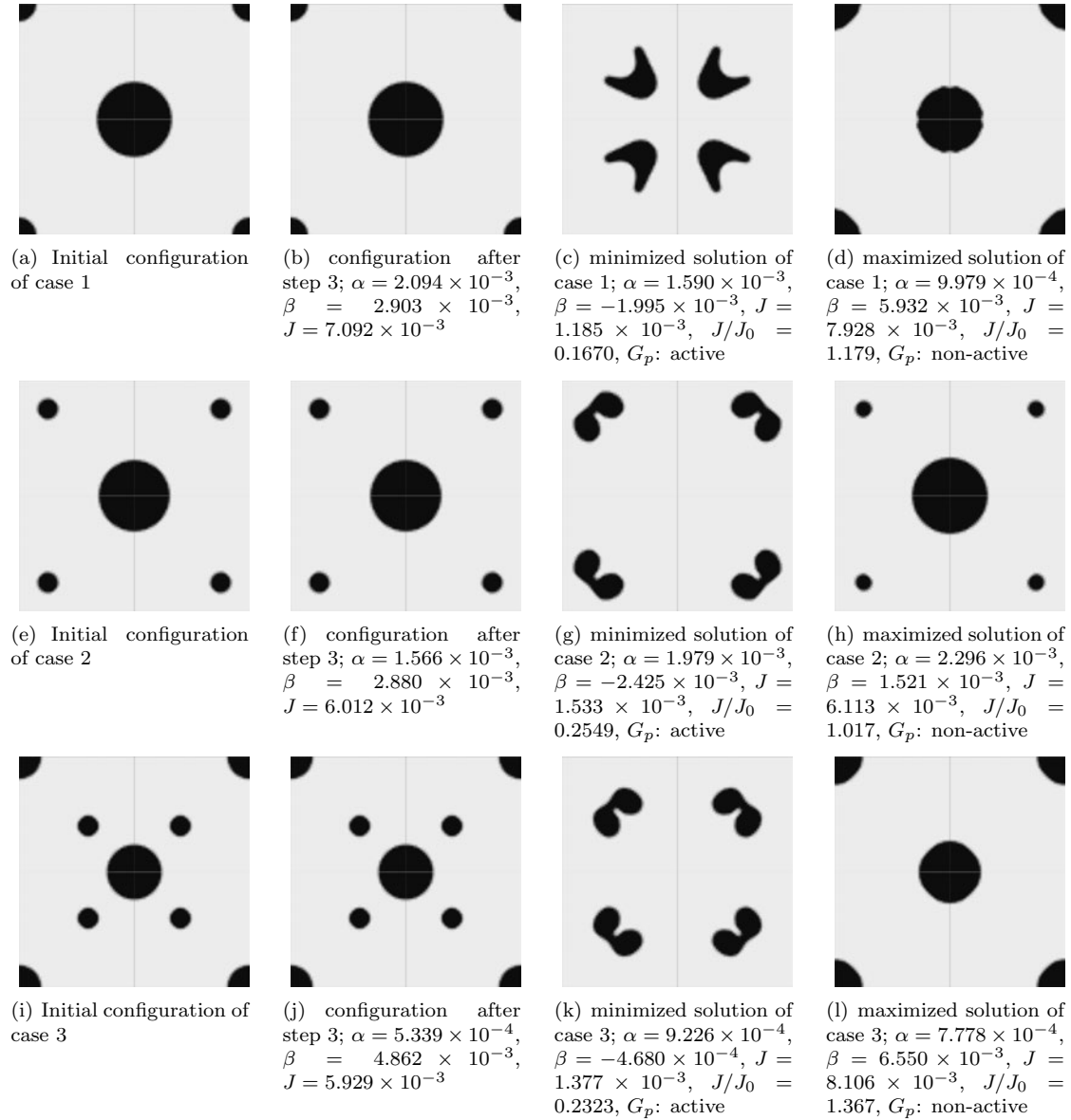


Figure 6: Volume fraction  $\overline{G}_v = 0.9$  (material A, in black, being the inclusion)

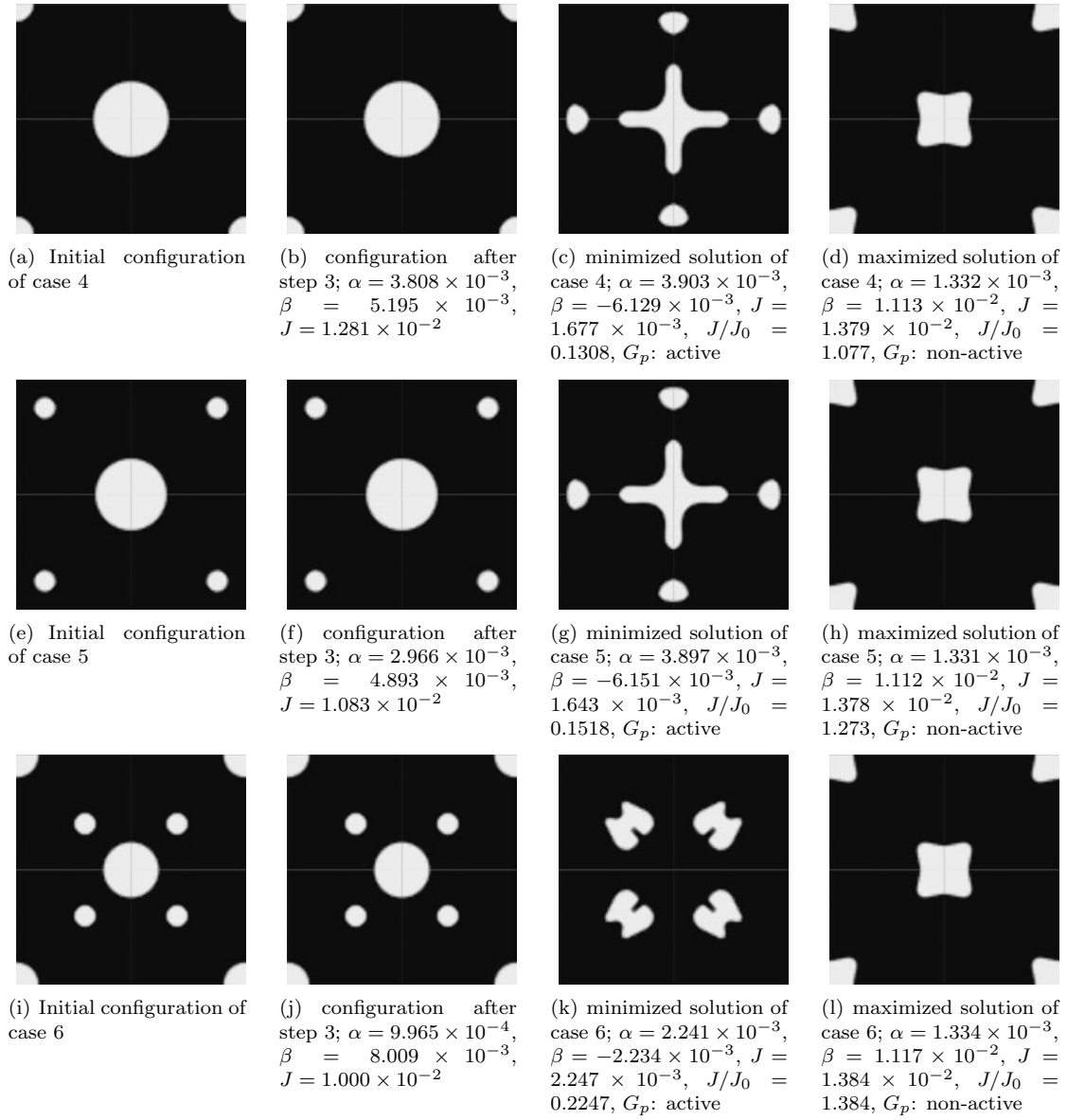


Figure 7: Volume fraction  $\overline{G}_v = 0.1$  (material B, in white, being the inclusion)

**Remark 9.1.** *The perimeter constraint is active in all cases of minimizing  $J(\Gamma)$  and non-active in all cases of maximizing  $J(\Gamma)$ . This is consistent with our Remark 6.2.*

## 9.2 Upper bounds on the dispersive effect

The goal of this subsection is to numerically find upper bounds on the coefficient  $\alpha$  and  $\beta$  of the isotropic dispersive tensor  $\mathbb{D}^*$  defined in Proposition 6.7. Recall that this isotropy condition is the result of our 8-fold symmetry assumption in the unit cell. Without this assumption, the tensor  $\mathbb{D}^*$  is characterized by 5 independent coefficients in  $2-d$  (and 15 in  $3-d$ ), which seriously complexifies the task of finding upper bounds. We restrict ourselves to upper bounds since, by virtue of Remark 6.2, an optimal lower bound on  $-\mathbb{D}^*$  is zero (which is achieved by taking smaller and smaller repetition of the same microstructure in the unit cell). Of course, non trivial lower bounds could be found if one adds a perimeter constraint but we did not explore this issue and instead focus only on upper bounds.

We use numerical (gradient-based) optimization to find such upper bounds and, more precisely, the Pareto front in the  $(\alpha, \beta)$  plane for given phase properties and proportions. Our goal is thus to obtain the curve of upper bounds for all possible  $(\alpha, \beta)$ , which is alike the celebrated Hashin-Shtrikman bounds [23] but for dispersive effects. If the set of all possible  $(\alpha, \beta)$  were convex, then the upper Pareto front could be obtained by maximizing all possible linear convex combination of  $\alpha$  and  $\beta$ :

$$\begin{aligned} \max_{\Gamma} \quad & \theta\alpha(\Gamma) + (1-\theta)\beta(\Gamma) \\ \text{subject to:} \quad & G_v(\Gamma) = 0, G_p(\Gamma) \leq 0, G_{a^*}(\Gamma) = 0 \end{aligned} \quad (9.1)$$

where  $\theta \in [0, 1]$  is a parameter, the phase proportion and the homogenized tensor  $a^*$  are constrained, and a perimeter constraint is added on the interface  $\Gamma$  to exclude too fragmented configurations. Unfortunately, it is not known whether the set of all possible  $(\alpha, \beta)$  is convex or not and solving (9.1) for different values of  $\theta \in [0, 1]$  would yield an upper bound merely on the convex envelope of this unknown set.

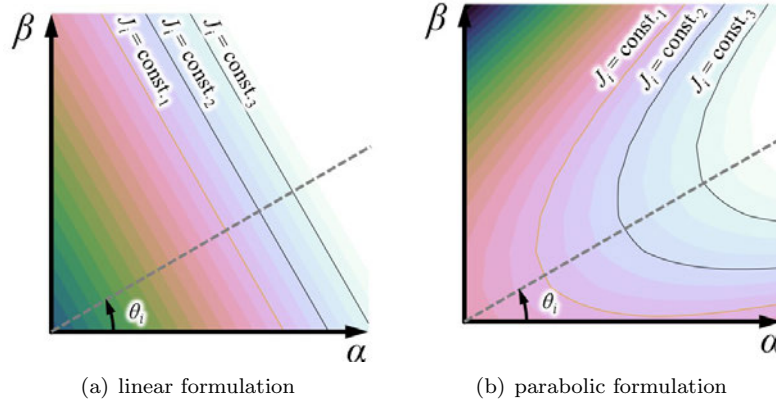


Figure 8: Comparison between the linear and parabolic formulations; the contour colors represent values of the objective function with  $\text{const.}_1 < \text{const.}_2 < \text{const.}_3$ .

In order to capture a possibly non-convex upper bound, we modify (9.1) by replacing the linear objective function by a rotated quadratic one. The main idea (see Figure 8) is to locally approximate the Pareto front by parabolas the main axis of which is oriented by an angle  $\theta$  with respect to the horizontal axis. Discretizing uniformly the angle as  $0 \leq \theta_i \leq \frac{\pi}{2}$ , for  $i = 1, 2, \dots, n$ , we introduce rotated coordinates

$$\begin{pmatrix} \alpha_{\theta_i} \\ \beta_{\theta_i} \end{pmatrix} := \begin{pmatrix} \cos \theta_i & \sin \theta_i \\ -\sin \theta_i & \cos \theta_i \end{pmatrix} \begin{pmatrix} \alpha_N \\ \beta_N \end{pmatrix}, \quad \begin{pmatrix} \alpha_N \\ \beta_N \end{pmatrix} := \begin{pmatrix} \frac{\alpha(\Gamma) - \alpha_{\min}}{\alpha_{\max} - \alpha_{\min}} \\ \frac{\beta(\Gamma) - \beta_{\min}}{\beta_{\max} - \beta_{\min}} \end{pmatrix},$$

where  $\alpha_{\max}, \alpha_{\min}, \beta_{\max}, \beta_{\min}$  are maximum and minimum values for  $\alpha$  and  $\beta$  (see Remark 9.2 for their evaluation). Therefore,  $\alpha_N$  and  $\beta_N$  represent normalized  $\alpha$  and  $\beta$ . Then, we replace the liner formulation (9.1) by the following

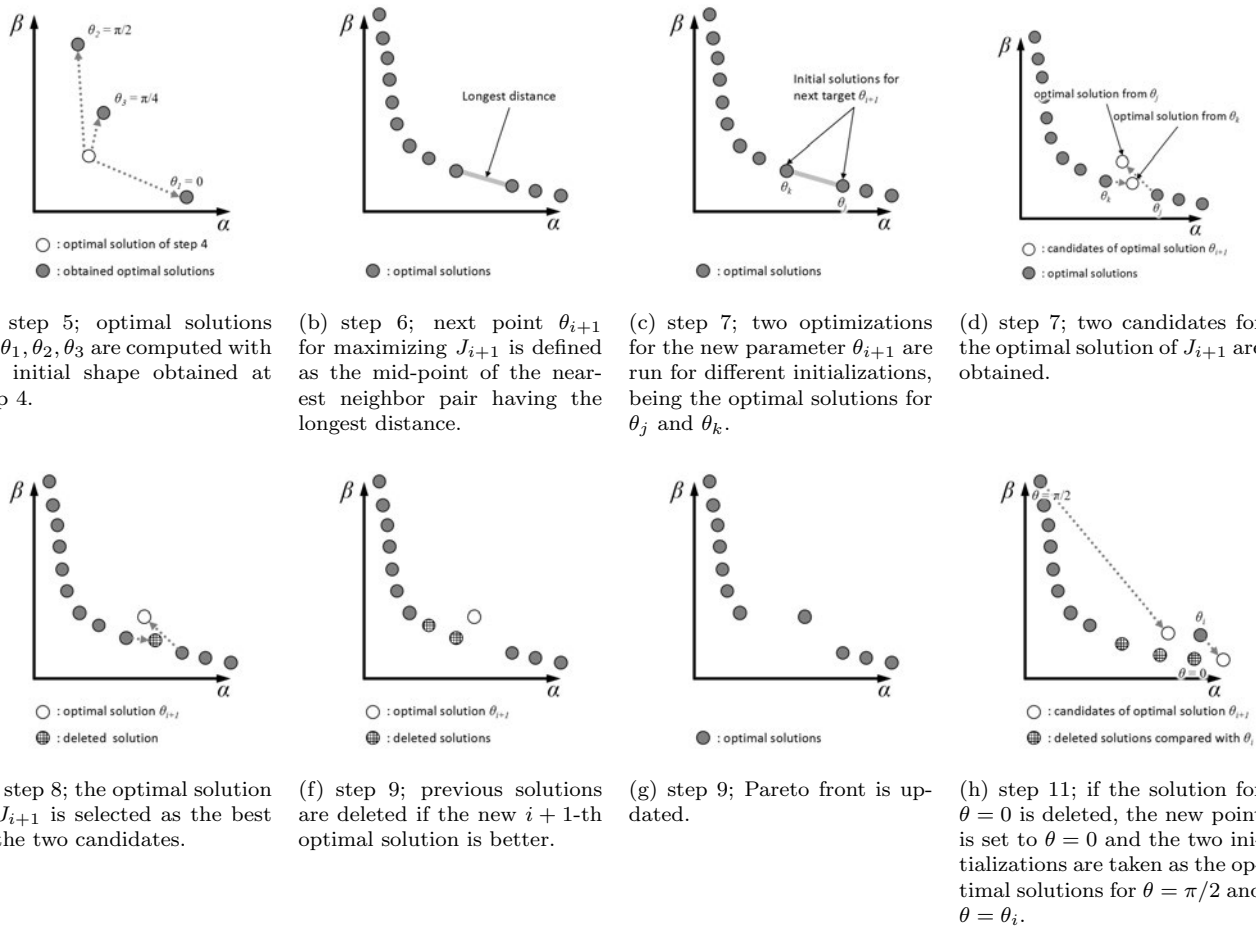


Figure 9: Optimization strategy for obtaining a Pareto front of optimal solutions

parabolic formulation

$$\begin{aligned} \max_{\Gamma} \quad & J_i(\Gamma) := \alpha\theta_i - c_p\beta_{\theta_i}^2 \\ \text{subject to:} \quad & G_v(\Gamma) = 0, G_p(\Gamma) \leq 0, G_{a^*}(\Gamma) = 0 \end{aligned} \quad (9.2)$$

where  $c_p > 0$  is a parameter for the parabola (we shall discuss its choice in a next subsection). For sufficiently large values of  $c_p$  we expect that such parabolas can better fit the possibly non-convex shape of the  $(\alpha, \beta)$  set, although one can easily imagine non-convex (but highly unlikely) shapes that cannot be approached from the outside by parabolas. All the numerical results in this section have been obtained by using the parabolic formulation (9.2) with a complicated discretization and initialization strategy for the angle  $\theta$  that we now describe (see Figure 9).

Since the results of the previous subsection have shown evidence of possible local maxima for (9.2), we devise a strategy to avoid as much as possible the effect of local optima and blind initializations in the optimization process. The details of our optimization strategy are as follows:

**step 1:** The level set function is initialized and the parabola parameter  $c_p$  is defined.

**step 2:** The function  $G_v(\Gamma)$  is optimized until satisfying  $G_v(\Gamma) = 0$ .

**step 3:** The function  $G_p(\Gamma)$  is minimized with  $G_v(\Gamma) = 0$  until satisfying  $G_p \leq 0$ .

**step 4:** The function  $G_{a^*}(\Gamma)$  is minimized with  $G_v = 0$  and  $G_p \leq 0$  until satisfying  $G_{a^*} = 0$ .

**step 5:** As shown in Figure 9 (a), starting from the optimal shape of step 4, three optimal solutions are computed by maximizing  $J_i|_{i=1,2,3}$  with constraints, for the angular parameters  $\theta_1 = 0$ ,  $\theta_2 = \pi/2$  and  $\theta_3 = \pi/4$ . This is the end of the initialization and we now start iterating by adding more and more angles  $\theta_i$  and removing those which are not optimal (step 9).

**step 6:** A new discrete angle  $\theta_{i+1}$  is defined which yields a new objective function  $J_{i+1}$  to be maximized. To begin with, find a "nearest neighbor" pair  $(\theta_j, \theta_k)$  such that their corresponding optimal values  $(\alpha_j, \beta_j)$  and  $(\alpha_k, \beta_k)$  are farthest apart, where the distance is measured by the Euclidean distance in the normalized  $\alpha_N$ - $\beta_N$  coordinate system, as shown in Figure 9(b). The pair  $(\theta_j, \theta_k)$  is said to be "nearest neighbor" if no other angle  $\theta_i$  lies between them (the angles are not labeled in a monotone order). The next angular parameter  $\theta_{i+1}$  is then defined as the mid-point of this pair:

$$\theta_{i+1} := \frac{\theta_j + \theta_k}{2}.$$

To avoid a too fine local search, we do not consider too close pairs  $(\theta_j, \theta_k)$  such that  $|\theta_j - \theta_k| < \epsilon_\theta$ , where  $\epsilon_\theta > 0$  is set to  $1 \times 10^{-4}$ .

**step 7:** To find an optimal solution for the new discrete angle  $\theta_{i+1}$ , we run two different optimization calculations. As shown in Figure 9(c), the initialization of each run is the optimal solution for  $\theta_j$  or  $\theta_k$ , respectively. We thus obtain two candidates for being the optimal solution associated to  $\theta_{i+1}$ , as shown in Figure 9(d).

**step 8:** The values of the objective function  $J_{i+1}$  for the two candidates computed in step 7 are compared. The candidate with the largest value is kept as the optimal solution for  $\theta_{i+1}$ , while the other candidate is deleted, as shown in Figure 9(e).

**step 9:** To update the Pareto front, the newly obtained optimal solution  $\Gamma_{i+1}$  is compared with previous optimal solutions  $\Gamma_l$ ,  $l = 1, 2, \dots, i$ . More precisely (see Figure 9(f) and (g)), a previous solution  $\Gamma_l$  is deleted from the Pareto front if it is inferior to  $\Gamma_{i+1}$  in the sense that:

$$J_l(\Gamma_l) \leq J_l(\Gamma_{i+1}).$$

**step 10:** The iterative optimization process is stopped if a maximal number  $n_\theta$  of discrete angles have been created. Go to step 11, if one of the end points ( $\theta = 0$  or  $\theta = \pi/2$ ) of the Pareto front is deleted in step 9. Otherwise, iterate by going back to step 6.

**step 11:** If the end point  $\theta = 0$  was deleted, then the next new discrete angle is  $\theta_{i+1} = 0$ . We perform step 7 with two initializations, corresponding to the optimal solutions for  $\theta = \pi/2$  and  $\theta_i$ , as shown in Figure 9(h). After that go back to step 8. A symmetric argument is used in the case of the other end point  $\theta = \pi/2$ .

In this subsection, we use a  $90 \times 90$  structured mesh for the analysis domain. The prescribed homogenized coefficient is set to  $\bar{a}^* = 14.141667$ . The relative error for judging whether the constraint function  $G_{a^*}$  is satisfied is set to  $5 \times 10^{-3}$ . The upper limit of the perimeter constraint is set to  $G_p = 5.0$ . The perimeter constraint is never active for the final results of the Pareto front. However, it is active and useful to improve convergence for the intermediate results.

**Remark 9.2.** *The maximum and minimum values  $\alpha_{\max}$ ,  $\alpha_{\min}$ ,  $\beta_{\max}$ ,  $\beta_{\min}$  are a priori unknown values. We initialize them as:*

$$\alpha_{\max} := 1.2\alpha_0, \quad \alpha_{\min} := 0.9\alpha_0, \quad \beta_{\max} := 1.2|\beta_0|, \quad \beta_{\min} := \beta_0 \quad (9.3)$$

*where  $\alpha_0$ ,  $\beta_0$  are the values of  $\alpha$  and  $\beta$  at the optimal solution of step 4. Then, in the successive iterations, these values are constantly updated as the maximal or minimal values of the previous computations.*

**Remark 9.3.** *The strategy is quite complex to implement and required a high computational cost at first glance. However, the set of optimal solutions is easily updated by storing the connectivity of the Pareto optimal solutions. In other words, the Pareto front is seen as a curve discretized by the boundary element method. Therefore, the strategy could be extended to the three dimensional case where the length of a segment is replaced by the area of a fictitious triangular element.*

**Remark 9.4.** *We tried the same strategy applied to the linear formulation (9.1) instead of the parabolic formulation. However, we could never obtain the same results and always finished with suboptimal solutions. A possible explanation is the presence of local minima together with non-convex level lines of the objective function.*

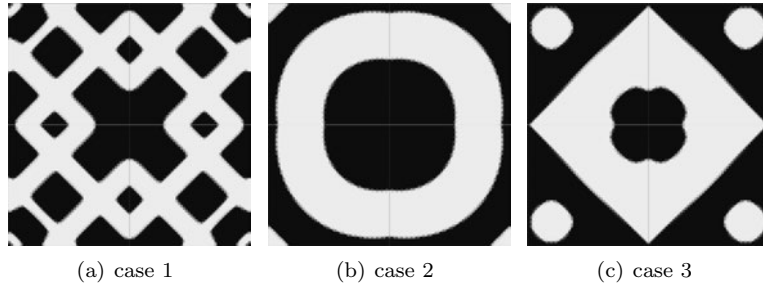


Figure 10: Initial configurations

### 9.2.1 Comparison of initializations

Since we expect possible local minima at each individual optimization step, it is important to check that our strategy avoids, as much as possible, local minima for the final result, i.e. for the Pareto front. Therefore, we examine the effect of different initial configurations in step 1 upon the estimated upper bound on the set of dispersive coefficients  $(\alpha, \beta)$ . Here, the parabolic parameter  $c_p$  is set to 50. Figure 10 displays three different initial configurations. The volume constraint is set to  $\overline{G}_v = 0.5$ , meaning that both phases have the same proportion. The following numerical tests are also an opportunity to check the convergence of the Pareto front as the number of discrete angles  $\theta_i$ , and therefore optimal shapes  $\Gamma_i$ , are added to the Pareto front.

Figure 11 shows the obtained Pareto fronts for each case of initial configuration in Figure 10. As can be clearly seen, the estimated Pareto fronts expand as the number of points  $n_\theta$  increase during the optimization process and, though very different initially, they almost overlap each other in the end, say for  $n_\theta = 500$ , see Figure 11 (f). As a conclusion, we claim that our final upper Pareto front is numerically insensitive to the the initial configuration.

A striking feature of the obtained upper Pareto front in Figure 12 is that it is almost linear. We estimate that the Pareto front is thus the line  $2\alpha + \frac{1}{2}\beta \approx 0.06$ . Although the Pareto front is almost linear, we confirm that, in practice, the non-convexity of the curve must be taken into account during the optimization process, since intermediate results are not convex curves.

Figure 12 features some of the obtained optimal shapes at  $n_\theta = 500$  (each line of subfigures corresponds to a different initialization of Figure 10). Clearly, the optimal configurations for each angular parameter  $\theta_i$  are very similar, regardless of the initial configuration (note that the shapes of the second line are identical to the ones of the first and third lines by a simple translation in the periodicity cell).

**Remark 9.5.** *Although the perimeter constraint is active at several intermediate optimal solutions, the constraint is non-active on the final upper bound in all examples.*

**Remark 9.6.** *The optimal shape for the end point of the Pareto front, corresponding to maximal  $\alpha$ , looks like a checkerboard pattern (see the left column of Figure 12). Therefore, we computed the values of  $\alpha$  and  $\beta$  for the checkerboard pattern of Figure 13(a) and found that indeed it yields a value  $\alpha = 2.86 \times 10^{-2}$  which is maximal. Therefore, we conjecture that this checkerboard pattern is an optimal configuration for maximal  $\alpha$ .*

*Although it is less obvious, the optimal shape for maximal  $\beta$  in Figure 12(l) looks like another checkerboard pattern. We again computed the values of  $\alpha$  and  $\beta$  for the checkerboard pattern of Figure 13(b) and obtained the largest value  $\beta = 5.52 \times 10^{-2}$ . Again we conjecture that this checkerboard pattern is an optimal configuration for maximal  $\beta$ .*

*We must admit that those checkerboard patterns are not clearly attained by our numerical optimization, which is a clear limitation of our approach. One reason is that it may be difficult to reproduce sharp corners with an algorithm based on (smooth) shape differentiation. Another reason is the presence of many local optima.*

**Remark 9.7.** *The two checkerboard configurations in Figure 13 are equivalent after rotation by  $\pi/4$  and scaling by a factor of  $\sqrt{2}$ . Since in Section 6 we obtained formulas for the dispersion tensor  $\mathbb{D}^*$  of a scaled or rotated microstructure, it is tempting to validate our numerical computations by checking the validity of these formulas. First, by virtue of Lemma 6.5, under the 8-fold symmetry assumption and for a rotation  $\mathcal{R}$  of angle  $\frac{\pi}{4}$ , the dispersive tensor after rotation is given by:*

$$\tilde{\mathbb{D}}^*(\eta \otimes \eta \otimes \eta \otimes \eta) = \mathbb{D}^*(\mathcal{R}\eta \otimes \mathcal{R}\eta \otimes \mathcal{R}\eta \otimes \mathcal{R}\eta) = -\frac{1}{4}(2\alpha + \beta)(\eta_1^4 + \eta_2^4) - \frac{1}{2}(6\alpha - \beta)\eta_1^2\eta_2^2.$$

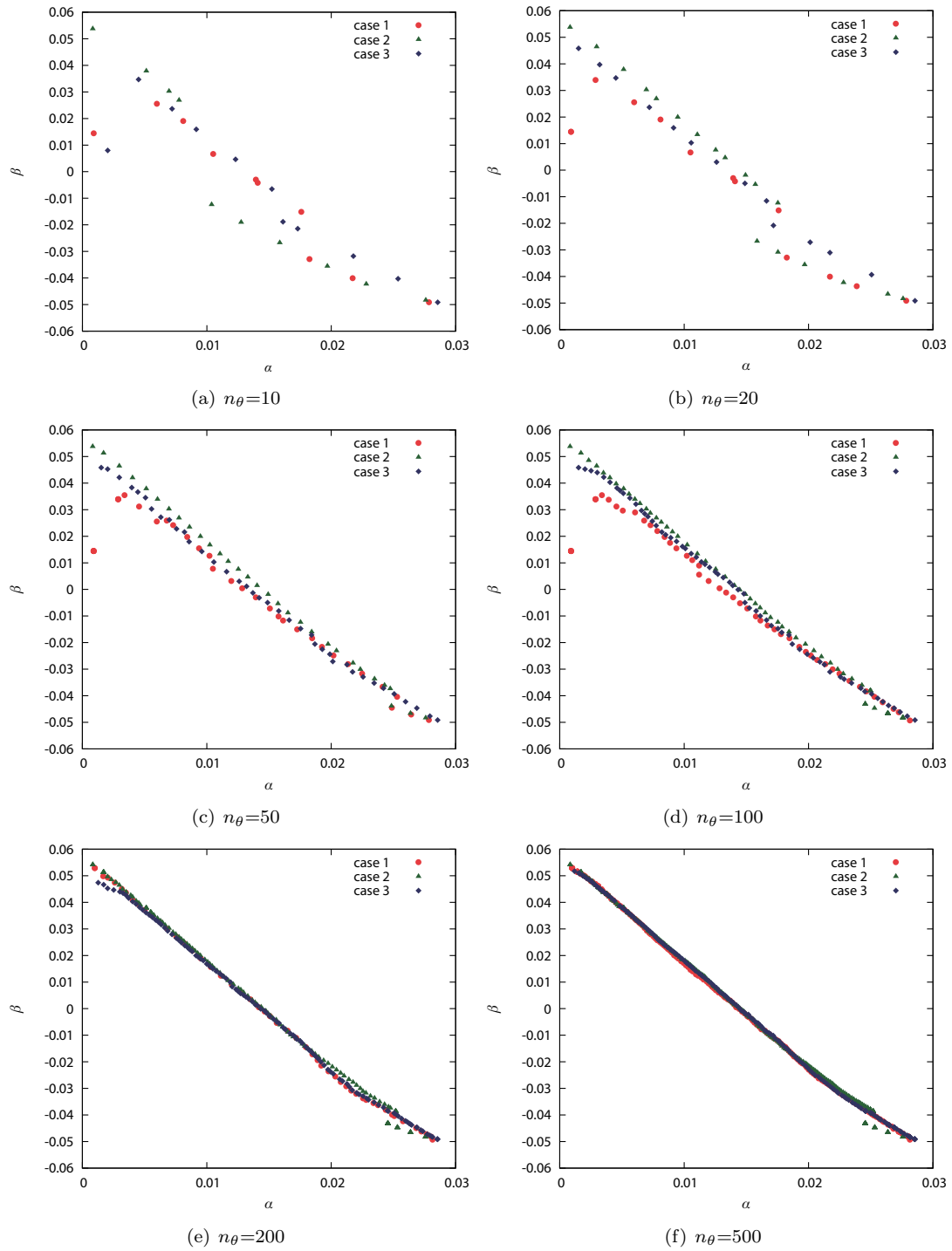


Figure 11: Upper bound or Pareto front in the  $(\alpha, \beta)$  plane. Dependency on the initial configurations (cases 1, 2 and 3)

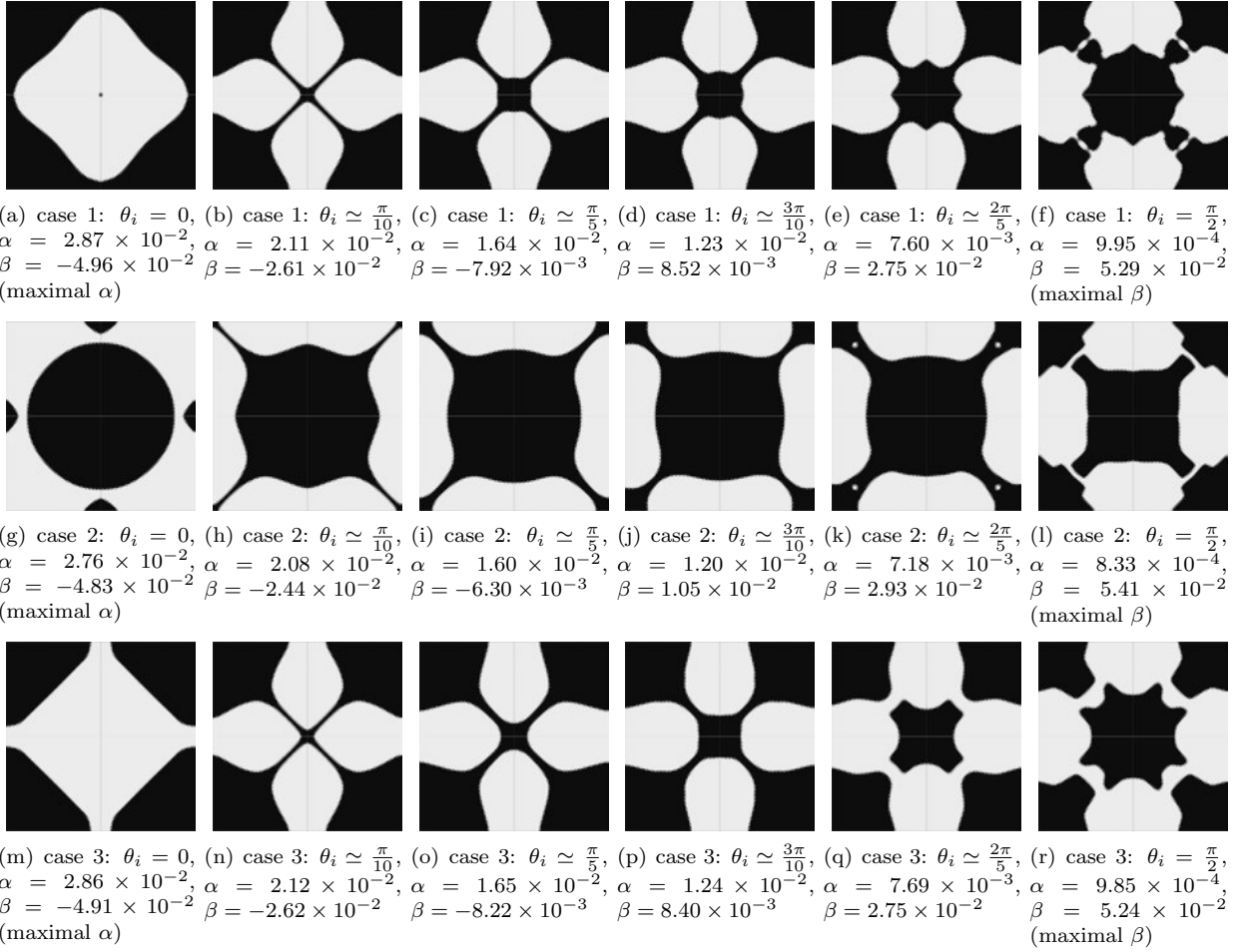
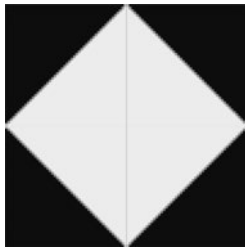
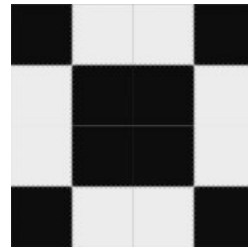


Figure 12: Optimal shapes of the upper Pareto front at  $n_\theta=500$



(a) maximal  $\alpha$ :  $\alpha = 2.86 \times 10^{-2}$ ,  $\beta = -4.97 \times 10^{-2}$



(b) maximal  $\beta$ :  $\alpha = 9.26 \times 10^{-4}$ ,  $\beta = 5.52 \times 10^{-2}$

Figure 13: Checkerboard patterns, conjectured to yield maximal  $\alpha$  (left) and maximal  $\beta$  (right)



	Figure 13 (a)		Figure 13 (b)	
	$\alpha$	$\beta$	$\alpha$	$\beta$
original	$2.86 \times 10^{-2}$	$-4.97 \times 10^{-2}$	$9.26 \times 10^{-4}$	$5.52 \times 10^{-2}$
rotation	$1.87 \times 10^{-3}$	$1.11 \times 10^{-1}$	$1.43 \times 10^{-2}$	$-2.48 \times 10^{-2}$
scaling	$9.32 \times 10^{-4}$	$5.53 \times 10^{-2}$	$2.85 \times 10^{-2}$	$-4.96 \times 10^{-2}$

Table 1: Comparison of the checkerboards of Figure 13 and their rotated and scaled versions

Second, since the scaling factor between the two configurations of Figure 13 is  $\sqrt{2}$ , by virtue of Lemma 6.3, the dispersive tensors have a ratio of 2. We confirm that our numerical results of Figure 13 satisfy those formulas, within an error of less than 1%, as can be checked in Table 1.

### 9.2.2 Sensitivity to the parabolic parameter $c_p$

We investigate the dependence of the Pareto front to the parameter  $c_p$  which enters the definition of the objective function in (9.2). The previous computations were performed with  $c_p = 50$ . We now consider various values  $c_p = 1, 30, 60, 80$  for the sole initialization of case 1 in Figure 10(a). As in Figure 11 we compute the Pareto front in the  $(\alpha, \beta)$  plane for these different values of  $c_p$ .

Figure 14 shows the obtained Pareto fronts for each cases and for increasing numbers  $n_\theta$  of points. Except for the case  $c_p = 1$ , all fronts are almost identical for a large number of points  $n_\theta = 500$ . Therefore, our Pareto front does not depend on this parameter, provided it is not too small. The bad behavior in the case  $c_p = 1$  is similar to the observed bad behavior for the linear formulation of the optimization problem (9.1).

Figure 15 displays the obtained optimal configurations at  $n_\theta = 500$ . Except in the caser of Figure 15(n), all shapes are very similar, regardless of the parameter  $c_p$ .

### 9.2.3 Dependence on the aspect ratio of the material properties $a_B/a_A$

We examine how the upper bound varies with the aspect ratio of the material properties  $a_B/a_A$ . We compare the previous case,  $a_A = 10$  and  $a_B = 20$ , called case 1 in the sequel, to the new case 2 for which  $a_A = 10$  and  $a_B = 25$  and case 3 for which  $a_A = 10$  and  $a_B = 30$ . The prescribed value for the volume constraint is  $\overline{G}_v = 0.5$  and the parabolic parameter is  $c_p = 50$ . The prescribed homogenized tensors are set to  $a^* = 14.141667$  in case 1,  $a^* = 15.812500$  in case 2 and  $a^* = 17.32500$  in case 3, where each value is the middle point between its upper and lower bounds (given by the Hashin-Shtrikman bounds). Figure 16 shows the two obtained Pareto fronts which are quite different. Both seem to be linear curves and the range of dispersion is larger for a larger aspect ratio. Remark that, as the ratio  $a_B/a_A$  converges to 1, one can easily show that the first and second order cell functions  $\chi_i, \chi_{ij}$  converge to 0 and, therefore, the dispersion tensor  $\mathbb{D}^*$  converges to 0 too. This is consistent with our numerical observation that the range of  $\mathbb{D}^*$  is smaller for smaller aspect ratio.

Figure 17 displays the obtained optimal configurations for case 2 and case 3. By comparison with Figure 12, the optimal shapes of case 1, case 2 and case 3 are very similar.

### 9.2.4 Dependence on the volume constraint $\overline{G}_v$

Next, we examine the effect of the phases proportion (or volume constraint  $\overline{G}_v$ ) upon the estimated upper bound on the parameters  $\alpha$  and  $\beta$  of the dispersive tensor. The values are set to  $\overline{G}_v = 0.5$  in case 1,  $\overline{G}_v = 0.6$  in case 2 and  $\overline{G}_v = 0.7$  in case 3, respectively. The material properties are set to  $a_A = 10$  and  $a_B = 20$ . The prescribed tensors are set to  $a^* = 14.141667$  in case 1,  $a^* = 13.197500$  in case 2 and  $a^* = 12.324720$  in case 3, where each value is the middle point between its upper and lower Hashin-Shtrikman bounds. Figure 18 shows the obtained Pareto optimal solutions for each cases.

As can be seen, the obtained curves are almost linear. The range of the Pareto front is maximal for  $\overline{G}_v = 0.5$  (its range should converge to the single point 0 when  $\overline{G}_v$  converges to 0 or 1, i.e. in the limit of pure phases). Figure 19 shows the obtained optimal configurations for the upper bound in cases 2 and 3. By comparison with Figure 12, the optimal configurations of cases 1, 2 and 3 are similar.

**Remark 9.8.** *It is easy to prove that, when the volume fraction  $\overline{G}_v$  tends to 0 or to 1, then the periodic coefficients  $a(y)$  converge to a constant, while the first and second order cell functions  $\chi_i, \chi_{ij}$  converge to 0. Therefore, the dispersion tensor  $\mathbb{D}^*$  converges to 0 too. This is consistent with our numerical observation that the range of  $\mathbb{D}^*$  is smaller for  $\overline{G}_v$  closer to 0 or 1.*

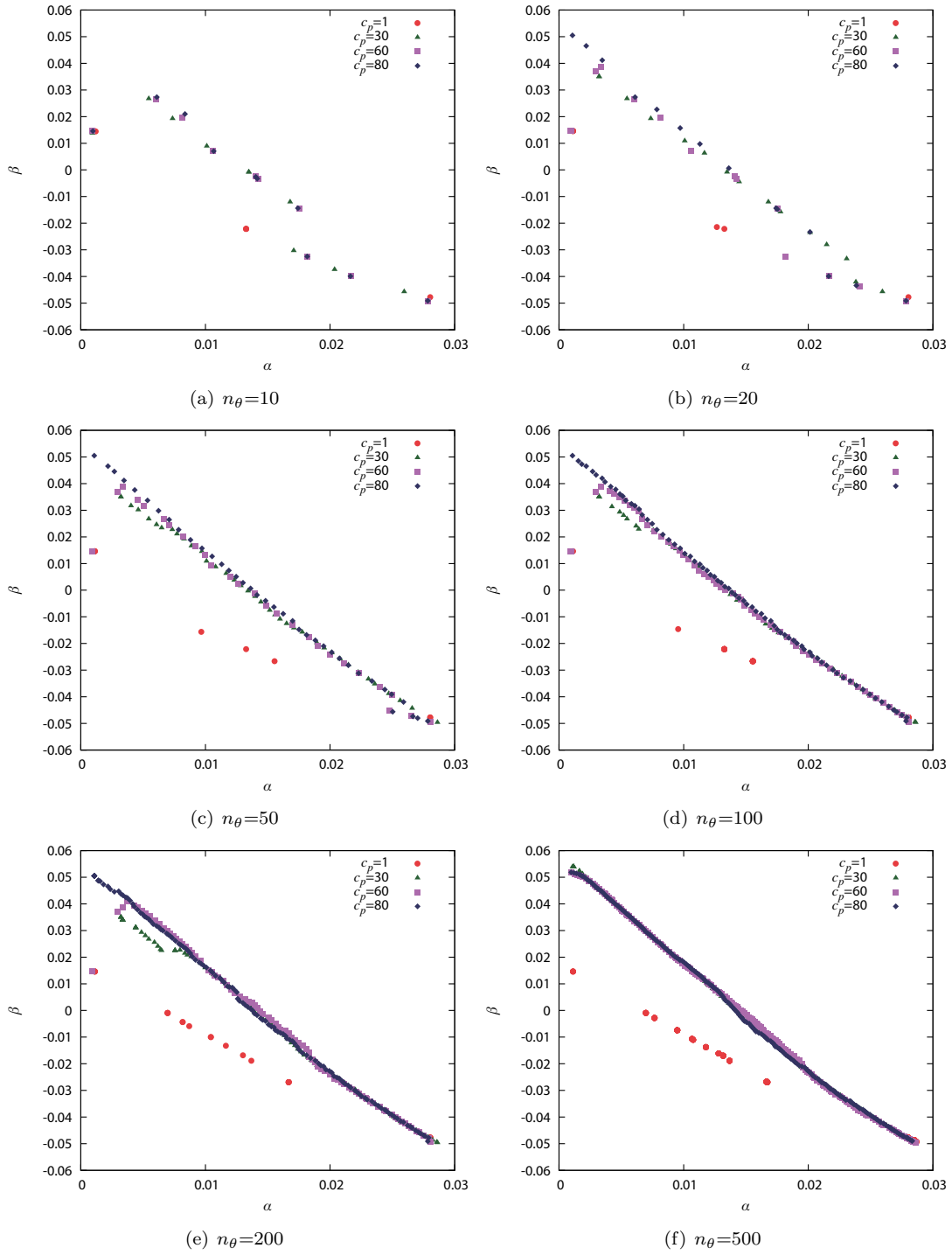


Figure 14: Sensitivity of the Pareto front to the parabolic parameter  $c_p$ , for increasing numbers  $n_\theta$  of points

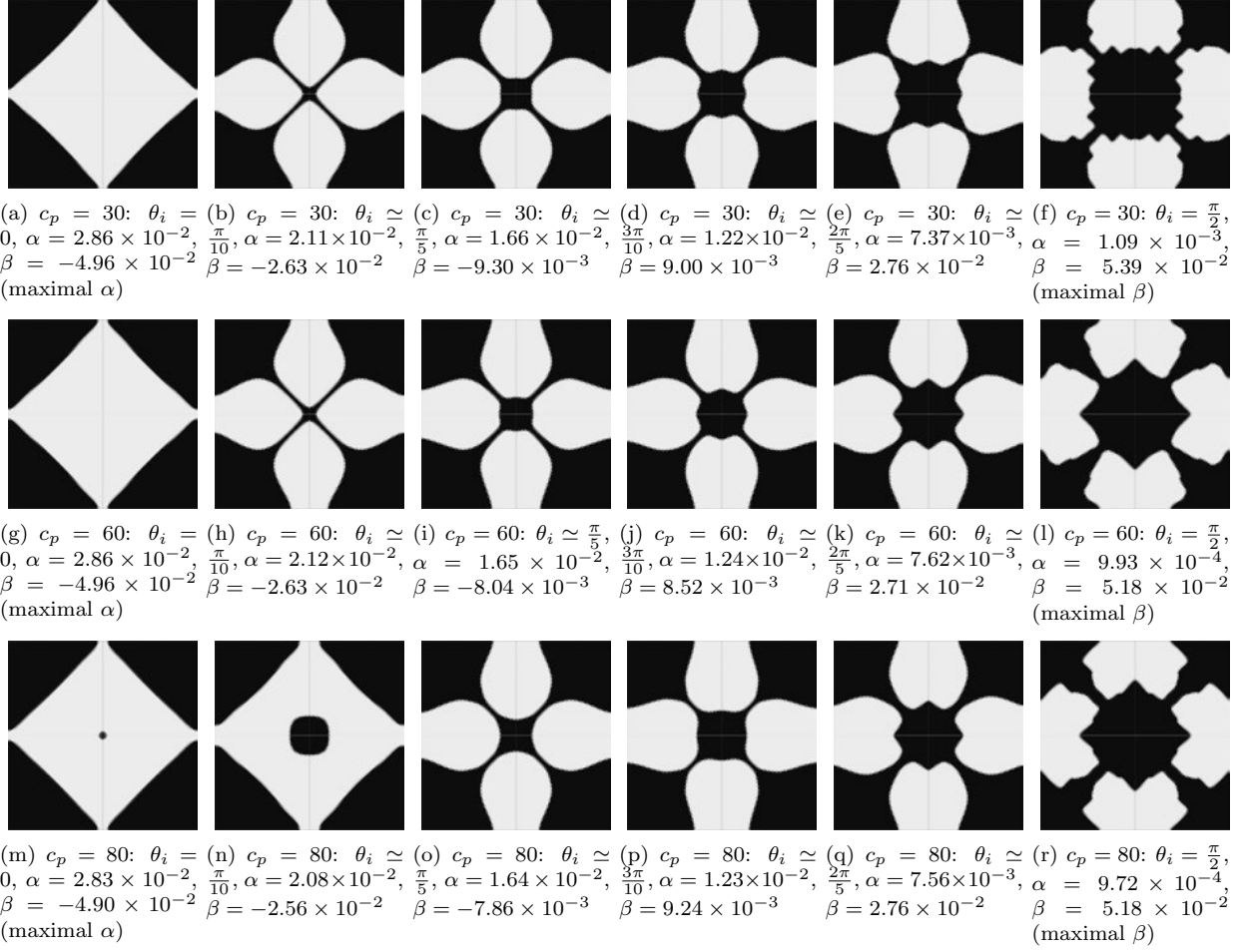


Figure 15: Optimal configurations at  $n_\theta = 500$  for  $c_p = 30, 60, 80$

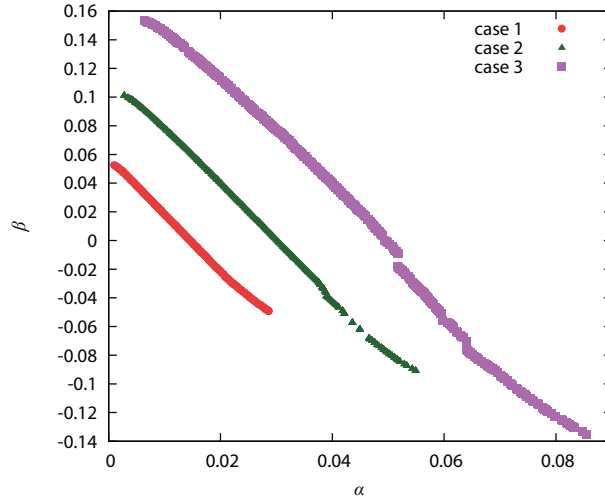
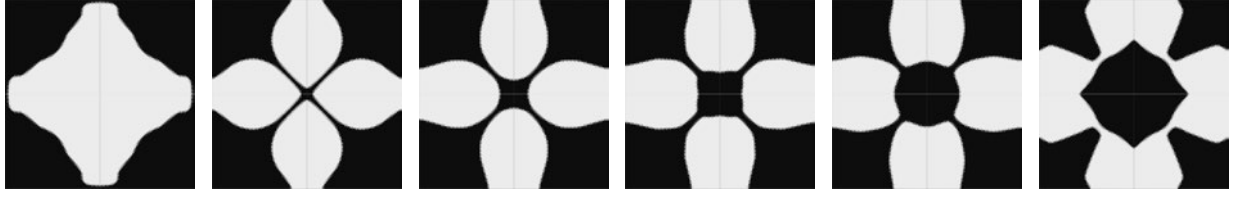
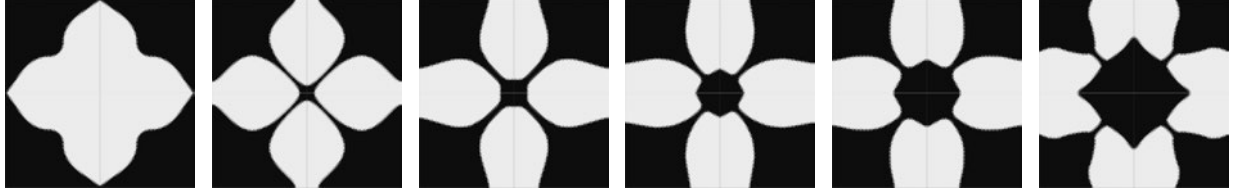


Figure 16: Pareto front of the upper bound on dispersion: case 1 ( $a_A = 10$ ,  $a_B = 20$ ), case 2 ( $a_A = 10$ ,  $a_B = 25$ ) and case 3 ( $a_A = 10$ ,  $a_B = 30$ )



(a) case 2,  $\theta_i = 0$ , (b) case 2,  $\theta_i \simeq \frac{\pi}{10}$ , (c) case 2,  $\theta_i \simeq \frac{\pi}{5}$ , (d) case 2,  $\theta_i \simeq \frac{3\pi}{10}$ , (e) case 2,  $\theta_i \simeq \frac{2\pi}{5}$ , (f) case 2,  $\theta_i = \frac{\pi}{2}$ ,  
 $\alpha = 5.50 \times 10^{-2}$ ,  $\alpha = 4.09 \times 10^{-2}$ ,  $\alpha = 3.26 \times 10^{-2}$ ,  $\alpha = 2.44 \times 10^{-2}$ ,  $\alpha = 1.47 \times 10^{-2}$ ,  $\alpha = 2.79 \times 10^{-3}$ ,  
 $\beta = -9.10 \times 10^{-2}$   $\beta = -4.56 \times 10^{-2}$   $\beta = -9.86 \times 10^{-3}$   $\beta = 2.26 \times 10^{-2}$   $\beta = 6.06 \times 10^{-2}$   $\beta = 1.01 \times 10^{-1}$   
(maximal  $\alpha$ ) (maximal  $\beta$ )



(g) case 3,  $\theta_i = 0$ , (h) case 3,  $\theta_i \simeq \frac{\pi}{10}$ , (i) case 3,  $\theta_i \simeq \frac{\pi}{5}$ , (j) case 3,  $\theta_i \simeq \frac{3\pi}{10}$ , (k) case 3,  $\theta_i \simeq \frac{2\pi}{5}$ , (l) case 3,  $\theta_i = \frac{\pi}{2}$ ,  
 $\alpha = 8.55 \times 10^{-2}$ ,  $\alpha = 6.39 \times 10^{-2}$ ,  $\alpha = 5.19 \times 10^{-2}$ ,  $\alpha = 4.06 \times 10^{-2}$ ,  $\alpha = 2.59 \times 10^{-2}$ ,  $\alpha = 6.33 \times 10^{-3}$ ,  
 $\beta = -1.36 \times 10^{-1}$   $\beta = -7.01 \times 10^{-2}$   $\beta = -9.37 \times 10^{-3}$   $\beta = 3.71 \times 10^{-2}$   $\beta = 9.16 \times 10^{-2}$   $\beta = 1.54 \times 10^{-1}$   
(maximal  $\alpha$ ) (maximal  $\beta$ )

Figure 17: Configurations of the Pareto front in case 2 ( $a_A = 10$ ,  $a_B = 25$ ) and case 3 ( $a_A = 10$ ,  $a_B = 30$ )

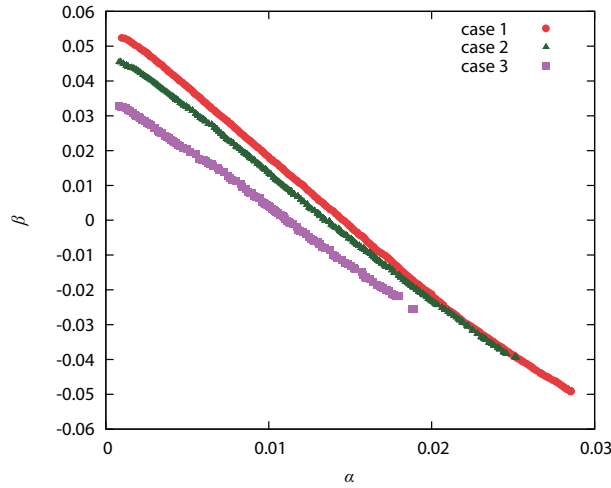
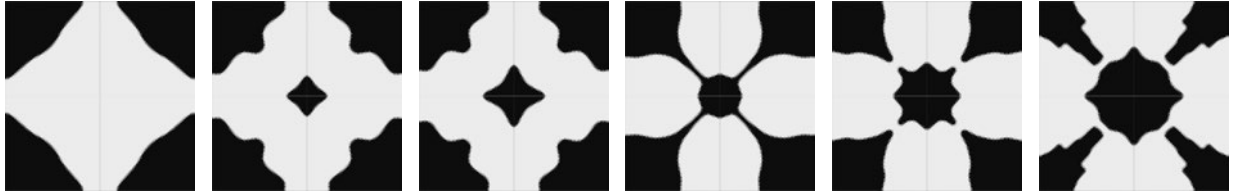
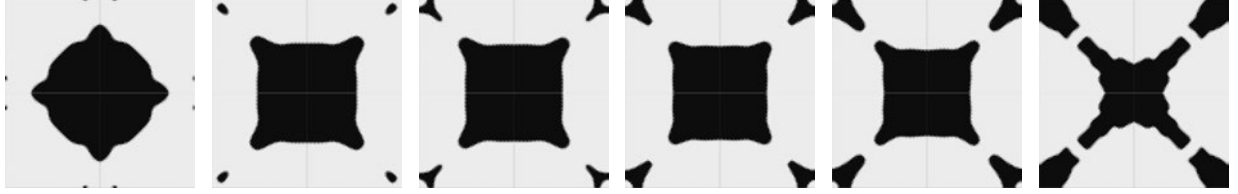


Figure 18: Upper bounds in different prescribed value of volume constraint



(a) case 2,  $\theta_i = 0$ ,  $\alpha = 2.52 \times 10^{-2}$ ,  $\beta = -3.95 \times 10^{-2}$  (maximal  $\alpha$ )  
 (b) case 2,  $\theta_i \simeq \frac{\pi}{10}$ ,  $\alpha = 1.95 \times 10^{-2}$ ,  $\beta = -2.12 \times 10^{-2}$   
 (c) case 2,  $\theta_i \simeq \frac{\pi}{5}$ ,  $\alpha = 1.51 \times 10^{-2}$ ,  $\beta = -5.42 \times 10^{-3}$   
 (d) case 2,  $\theta_i \simeq \frac{3\pi}{10}$ ,  $\alpha = 1.11 \times 10^{-2}$ ,  $\beta = 9.30 \times 10^{-3}$   
 (e) case 2,  $\theta_i \simeq \frac{2\pi}{5}$ ,  $\alpha = 6.81 \times 10^{-3}$ ,  $\beta = 2.61 \times 10^{-2}$   
 (f) case 2,  $\theta_i = \frac{\pi}{2}$ ,  $\alpha = 8.05 \times 10^{-4}$ ,  $\beta = 4.55 \times 10^{-2}$  (maximal  $\beta$ )



(g) case 3,  $\theta_i = 0$ ,  $\alpha = 1.88 \times 10^{-2}$ ,  $\beta = -2.55 \times 10^{-2}$  (maximal  $\alpha$ )  
 (h) case 3,  $\theta_i \simeq \frac{\pi}{10}$ ,  $\alpha = 1.48 \times 10^{-2}$ ,  $\beta = -1.26 \times 10^{-2}$   
 (i) case 3,  $\theta_i \simeq \frac{\pi}{5}$ ,  $\alpha = 1.15 \times 10^{-2}$ ,  $\beta = -1.08 \times 10^{-3}$   
 (j) case 3,  $\theta_i \simeq \frac{3\pi}{10}$ ,  $\alpha = 8.61 \times 10^{-3}$ ,  $\beta = 8.87 \times 10^{-3}$   
 (k) case 3,  $\theta_i \simeq \frac{2\pi}{5}$ ,  $\alpha = 5.19 \times 10^{-3}$ ,  $\beta = 1.97 \times 10^{-2}$   
 (l) case 3,  $\theta_i = \frac{\pi}{2}$ ,  $\alpha = 7.95 \times 10^{-4}$ ,  $\beta = 3.28 \times 10^{-2}$  (maximal  $\beta$ )

Figure 19: Optimal configurations on the upper bound in case 2 and case 3 (comparison of  $\overline{G_v}$ )

### 9.3 Optimizing the dispersion coefficient $d^*$ for the source term

In the high order homogenized equation (1.4), another source of dispersion comes from the source term which is perturbed by a second order derivative, the coefficient of which is the matrix  $d^*$ , defined by (2.24). The goal of this subsection is to investigate the range of this matrix  $d^*$ . Under our 8-fold symmetry assumption, the tensor  $d^*$  is a scalar matrix, i.e., can be rewritten as follows:

$$d_{ij}^* = \gamma \text{Id} \quad \text{with} \quad \gamma := \int_Y \chi_1^2 dy.$$

Therefore, we minimize or maximize this dispersion coefficient  $\gamma$  with volume constraint, perimeter constraint and prescribing the homogenized tensor  $a^*$ , as follows:

$$\begin{aligned} \min_{\Gamma} \quad \text{or} \quad \max_{\Gamma} \quad & J(\Gamma) := \gamma(\Gamma) \\ \text{subject to:} \quad & G_v(\Gamma) = 0, \quad G_p(\Gamma) \leq 0, \quad G_{a^*}(\Gamma) = 0 \end{aligned}$$

We rely on the optimization algorithm of Section 8 and adopt the four-step optimization procedure of subsection 9.1, starting from the same initializations. We use a  $50 \times 50$  structured mesh for the analysis domain. The isotropic materials A and B have material properties  $a_A = 10$  and  $a_B = 20$ . The upper limit of the perimeter constraint is set to  $\overline{G_p} = 1.5$ . We consider two cases for the volume constraint: either  $\overline{G_v} = 0.9$  or  $\overline{G_v} = 0.1$ , which can be interpreted as material A being the inclusion in the first case, and material B being the inclusion in the second case. By symmetry, the homogenized tensor  $\overline{a^*}$  is isotropic and its prescribed scalar value is set to 10.705 in the first case and 18.72 in the second case. The relative error for judging whether the constraint function  $G_{a^*}$  is satisfied is set to  $5 \times 10^{-3}$ .

Figures 20 and 21 show initial and optimal configurations when material A (in black) is the inclusion and when material B (in white) is the inclusion, respectively. When minimizing  $\gamma$ , the inclusions are fragmented with smaller and more complex details (which depend on the value of the perimeter constraint which is always active). This is consistent with our Remark 6.2 which states that smaller inclusions yield smaller dispersion. On the other hand, when maximizing  $\gamma$ , we obtain a single smooth inclusion in the unit cell and the perimeter constraint is not active.

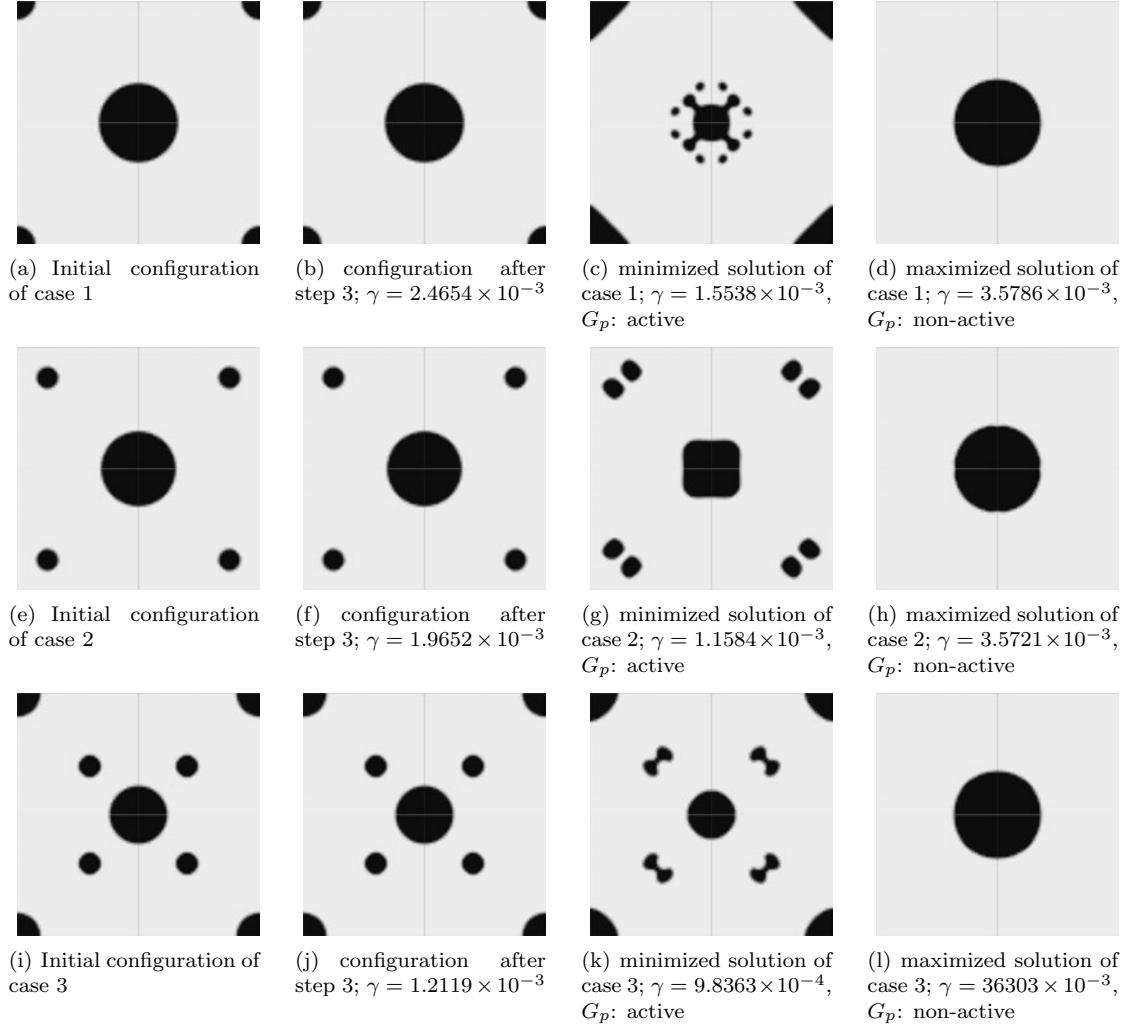


Figure 20: Optimal shapes for  $d^*$  with volume fraction  $\overline{G_v} = 0.9$  (material A, in black, being the inclusion)

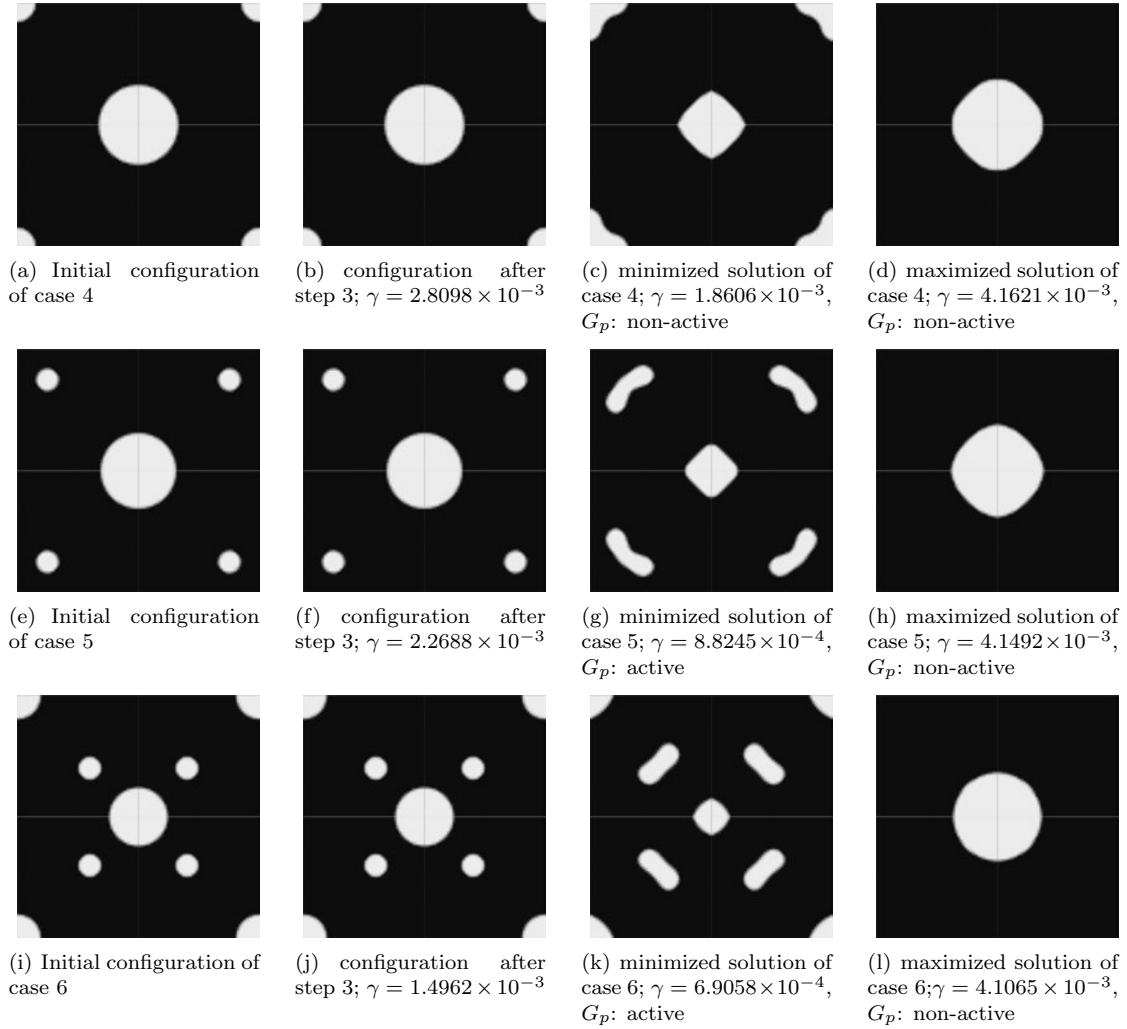


Figure 21: Optimal shapes for  $d^*$  with volume fraction  $\overline{G_v} = 0.1$  (material B, in white, being the inclusion)

## 10 Proof of Theorem 7.3

This section is devoted to the proof of Theorem 7.3, which gives the shape derivative of an objective function depending on the first-order and second-order cell solutions. As usual, we rely on the Lagrangian method, which introduces a Lagrangian as the sum of the objective function and of the constraints multiplied by Lagrange multipliers. The governing equations are treated as constraints and the corresponding Lagrange multipliers are precisely the adjoint states at optimality. Eventually, the shape derivative is obtained as a simple partial derivative of the Lagrangian using Lemma 7.2. There is a slight difficulty since the solutions  $\chi_i$  and  $\chi_{ij}$  of the cell problems are not shape differentiable (in the sense of Definition 7.1) because their normal derivatives are discontinuous across the interface  $\Gamma$  (see [7, 40]). However, when restricted to each subdomain  $Y^A$  and  $Y^B$ , the functions  $\chi_i$  and  $\chi_{ij}$  are shape differentiable. Therefore, we shall rewrite the cell equations as transmission problems and define an adequate Lagrangian which will involve additional Lagrange multipliers for the interface transmission conditions (see [7, 40]).

Thus, we introduce the restrictions  $\chi_i^A$  and  $\chi_{ij}^A$  to  $Y^A$ , and  $\chi_i^B$  and  $\chi_{ij}^B$  to  $Y^B$ , of the solutions  $\chi_i$  of (2.6) and  $\chi_{ij}$  of (2.12),  $1 \leq i, j \leq d$ . They satisfy the following transmission problems:

$$\begin{cases} -\operatorname{div}(a^A \nabla \chi_i^A) = -\operatorname{div}(a^A e_i) & \text{in } Y^A \\ -\operatorname{div}(a^B \nabla \chi_i^B) = -\operatorname{div}(a^B e_i) & \text{in } Y^B \\ \chi_i^A = \chi_i^B & \text{on } \Gamma \\ a^A (\nabla \chi_i^A - e_i) \cdot n^A = -a^B (\nabla \chi_i^B - e_i) \cdot n^B & \text{on } \Gamma, \end{cases} \quad (10.1)$$

and

$$\begin{cases} -\operatorname{div}(a^A \nabla \chi_{ij}^A) = a_{ij}^A - a^A \nabla \chi_j \cdot e_i - \operatorname{div}(\chi_j^A a^A e_i) - a_{ij}^* & \text{in } Y^A \\ -\operatorname{div}(a^B \nabla \chi_{ij}^B) = a_{ij}^B - a^B \nabla \chi_j \cdot e_i - \operatorname{div}(\chi_j^B a^B e_i) - a_{ij}^* & \text{in } Y^B \\ \chi_{ij}^A = \chi_{ij}^B & \text{on } \Gamma \\ a^A (\nabla \chi_{ij}^A - \chi_j^A e_i) \cdot n^A = -a^B (\nabla \chi_{ij}^B - \chi_j^B e_i) \cdot n^B & \text{on } \Gamma, \end{cases} \quad (10.2)$$

which, taking into account periodicity condition, are equivalent to (2.6) and (2.12), respectively. Of course, the functions  $\chi_i^A$ ,  $\chi_{ij}^A$  and  $\chi_i^B$ ,  $\chi_{ij}^B$  are smooth on their subdomains, namely  $\chi_i^A, \chi_{ij}^A \in H^2(Y^A)$  and  $\chi_i^B, \chi_{ij}^B \in H^2(Y^B)$ . If  $\Gamma$  is smooth, they are restrictions to their subdomains of smooth functions in the entire unit cell  $Y$ . Then, for  $1 \leq i, j \leq d$ , for any functions  $\hat{\chi}_i^A, \hat{\chi}_{ij}^A, \hat{\chi}_i^B, \hat{\chi}_{ij}^B$  in  $H_{\#}^1(Y)$ , for any bulk "Lagrange multiplier" functions  $\hat{p}_i^A, \hat{p}_{ij}^A, \hat{p}_i^B, \hat{p}_{ij}^B$  in  $H_{\#}^1(Y)$  and interface "Lagrange multiplier" functions  $\hat{\lambda}_i, \hat{\lambda}_{ij}, \hat{\gamma}_i, \hat{\gamma}_{ij}$  in  $H_{\#}^1(Y)$ , we define a Lagrangian  $\mathcal{L}$  as follows:

$$\begin{aligned} \mathcal{L}(\Gamma, \{\hat{\chi}_i^A\}, \{\hat{\chi}_{ij}^A\}, \{\hat{\chi}_i^B\}, \{\hat{\chi}_{ij}^B\}, \{\hat{p}_i^A\}, \{\hat{p}_{ij}^A\}, \{\hat{p}_i^B\}, \{\hat{p}_{ij}^B\}, \{\hat{\lambda}_i\}, \{\hat{\lambda}_{ij}\}, \{\hat{\gamma}_i\}, \{\hat{\gamma}_{ij}\}) := \\ \int_{Y^A} \mathcal{J}(y, \{\hat{\chi}_i^A\}, \{\hat{\chi}_{ij}^A\}) dy + \int_{Y^B} \mathcal{J}(y, \{\hat{\chi}_i^B\}, \{\hat{\chi}_{ij}^B\}) dy \\ - \int_{Y^A} \operatorname{div}(a^A (\nabla \hat{\chi}_i^A - e_i)) \hat{p}_i^A dy - \int_{Y^B} \operatorname{div}(a^B (\nabla \hat{\chi}_i^B - e_i)) \hat{p}_i^B dy \\ + \int_{\Gamma} \hat{\lambda}_i (\hat{\chi}_i^A - \hat{\chi}_i^B) ds + \int_{\Gamma} \hat{\gamma}_i (a^A (\nabla \hat{\chi}_i^A - e_i) \cdot n^A + a^B (\nabla \hat{\chi}_i^B - e_i) \cdot n^B) ds \\ - \int_{Y^A} \left( \operatorname{div}(a^A (\nabla \hat{\chi}_{ij}^A - \hat{\chi}_j^A e_i)) + a_{ij}^A - \hat{a}_{ij}^* - a^A \nabla \hat{\chi}_j^A \cdot e_i \right) \hat{p}_{ij}^A dy \\ - \int_{Y^B} \left( \operatorname{div}(a^B (\nabla \hat{\chi}_{ij}^B - \hat{\chi}_j^B e_i)) + a_{ij}^B - \hat{a}_{ij}^* - a^B \nabla \hat{\chi}_j^B \cdot e_i \right) \hat{p}_{ij}^B dy \\ + \int_{\Gamma} \hat{\lambda}_{ij} (\hat{\chi}_{ij}^A - \hat{\chi}_{ij}^B) ds + \int_{\Gamma} \hat{\gamma}_{ij} (a^A (\nabla \hat{\chi}_{ij}^A - \hat{\chi}_j^A e_i) \cdot n^A + a^B (\nabla \hat{\chi}_{ij}^B - \hat{\chi}_j^B e_i) \cdot n^B) ds, \end{aligned} \quad (10.3)$$

where the integrals on the interface  $\Gamma$ , involving the Lagrange multipliers  $\hat{\lambda}_i, \hat{\lambda}_{ij}, \hat{\gamma}_i, \hat{\gamma}_{ij}$ , are meant to enforce the transmission conditions on  $\Gamma$ , as they appear in (10.1), (10.2). The coefficient  $\hat{a}_{ij}^*$  is defined in terms of  $\hat{\chi}_i$  by a formula similar to (2.8) for the homogenized tensor, namely

$$\hat{a}_{ij}^* = \int_Y (a e_i - a \nabla \hat{\chi}_i) \cdot e_j dy.$$



Of course, in (10.3) the summation convention is used for  $1 \leq i, j \leq d$ . The notation  $\hat{\cdot}$  on top of each function means that it is not the optimal function, but any function in  $H_{\#}^1(Y)$ . At optimality, we shall remove the  $\hat{\cdot}$  and we shall recover  $\chi_i, \chi_{ij}$ , solutions of (10.1), (10.2), and  $p_i, p_{ij}$ , solutions of (7.4), (7.5), respectively. The important fact in the definition of  $\mathcal{L}$  is that all variables  $\hat{\chi}$  and  $\hat{p}$  are independent of  $\Gamma$  and defined in the fixed space  $H_{\#}^1(Y)$ . Therefore, the stationarity of the Lagrangian is going to give the optimality conditions of the optimization problem.

In a first (and easy) step, taking the partial derivatives of the Lagrangian with respect to the Lagrange multipliers  $p, \lambda$  and  $\gamma$  yield the state equations (10.1), (10.2). This is an obvious computation since the Lagrangian is linear with respect to  $p, \lambda$  and  $\gamma$ , and we skip it.

In a second step, taking the partial derivatives of the Lagrangian with respect to the variables  $\chi$  leads to the adjoint equations (7.4), (7.5). In other words, for a direction of derivation  $\psi = (\psi^A, \psi^B)$  with  $\psi^{A,B} \in H_{\#}^1(Y)$ , we ask that  $p_i$  be determined by

$$\left\langle \frac{\partial \mathcal{L}}{\partial \chi_i}, \psi \right\rangle = 0, \quad (10.4)$$

and  $p_{ij}$  by

$$\left\langle \frac{\partial \mathcal{L}}{\partial \chi_{ij}}, \psi \right\rangle = 0. \quad (10.5)$$

Let us explain the details for (10.5). We choose a test function  $\psi$ , with values  $\psi^A$  in  $Y^A$  and  $\psi^B$  in  $Y^B$ , which a priori is not continuous through the interface  $\Gamma$ . We find

$$\begin{aligned} \left\langle \frac{\partial \mathcal{L}}{\partial \chi_{ij}}, \psi \right\rangle &= \int_{Y^A} \frac{\partial \mathcal{J}}{\partial \chi_{ij}^A} \psi^A dy - \int_{Y^A} \operatorname{div}(a^A \nabla \psi^A) p_{ij}^A dy + \int_{Y^B} \frac{\partial \mathcal{J}}{\partial \chi_{ij}^B} \psi^B dy - \int_{Y^B} \operatorname{div}(a^B \nabla \psi^B) p_{ij}^B dy \\ &+ \int_{\Gamma} \lambda_{ij} (\psi^A - \psi^B) ds + \int_{\Gamma} \gamma_{ij} (a^A \nabla \psi^A \cdot n^A + a^B \nabla \psi^B \cdot n^B) ds = 0. \end{aligned}$$

After two integration by parts, recalling the notation  $[\cdot]$  for the jump through  $\Gamma$ , we deduce

$$\begin{aligned} \left\langle \frac{\partial \mathcal{L}}{\partial \chi_{ij}}, \psi \right\rangle &= \int_{Y^A} \left( \frac{\partial \mathcal{J}}{\partial \chi_{ij}^A} - \operatorname{div}(a^A \nabla p_{ij}^A) \right) \psi^A dy + \int_{Y^B} \left( \frac{\partial \mathcal{J}}{\partial \chi_{ij}^B} - \operatorname{div}(a^B \nabla p_{ij}^B) \right) \psi^B dy \\ &+ \int_{\Gamma} \lambda_{ij} [\psi] ds + \int_{\Gamma} \gamma_{ij} [a \nabla \psi \cdot n] ds + \int_{\Gamma} ([\psi a \nabla p_{ij} \cdot n] - [p_{ij} a \nabla \psi \cdot n]) ds = 0. \end{aligned} \quad (10.6)$$

Now we choose the test function  $\psi$  which satisfies successively five different types of conditions at the interface  $\Gamma$ .

1. Assume  $\psi$  to have compact support in  $Y^A$ . It gives

$$\left\langle \frac{\partial \mathcal{L}}{\partial \chi_{ij}}, \psi \right\rangle = \int_{Y^A} \frac{\partial \mathcal{J}}{\partial \chi_{ij}^A} \psi dy - \int_{Y^A} \operatorname{div}(a^A \nabla \psi) p_{ij}^A dy = 0,$$

which, by integration by parts, since  $\psi$  and its gradient vanish on  $\Gamma$ , leads to

$$-\operatorname{div}(a^A \nabla p_{ij}^A) = -\frac{\partial \mathcal{J}}{\partial \chi_{ij}^A} \quad \text{in } Y^A. \quad (10.7)$$

A symmetric computation works for  $Y^B$ . Because of (10.7) the two integrals on  $Y^A$  and  $Y^B$  cancel in (10.6).

2. Assume  $a^A \nabla \psi^A \cdot n^A = a^B \nabla \psi^B \cdot n^B = 0$  and  $[\psi] = 0$  on  $\Gamma$ . Thus (10.6) reduces to

$$\int_{\Gamma} \psi [a \nabla p_{ij} \cdot n] ds = 0,$$

for any value of the trace  $\psi = \psi^A = \psi^B$  on  $\Gamma$ . This implies  $[a \nabla p_{ij} \cdot n] = 0$ .

3. Assume just  $a^A \nabla \psi^A \cdot n^A = a^B \nabla \psi^B \cdot n^B = 0$ . Then (10.6) reduces to

$$\int_{\Gamma} [\psi] (\lambda_{ij} + a \nabla p_{ij} \cdot n) ds = 0,$$

for any value of the jump  $[\psi]$  on  $\Gamma$ . This gives the optimal value of the Lagrange multiplier  $\lambda_{ij} = -a^A \nabla p_{ij}^A \cdot n^A = a^B \nabla p_{ij}^B \cdot n^B$ .

4. Assume now  $[a\nabla\psi \cdot n] = 0$ . This time, (10.6) reduces to

$$\int_{\Gamma} a\nabla\psi \cdot n [p_{ij}] ds = 0,$$

for any value of the trace  $a\nabla\psi \cdot n = a^A\nabla\psi^A \cdot n^A = a^B\nabla\psi^B \cdot n^B$ . Thus we find  $[p_{ij}] = 0$  on  $\Gamma$ .

5. Finally if  $\psi$  does not satisfy any condition at the interface, (10.6) reduces to

$$\int_{\Gamma} [a\nabla\psi \cdot n] (\gamma_{ij} - p_{ij}) ds = 0,$$

which gives the optimal value of the Lagrange multiplier  $\gamma_{ij} = p_{ij}$  on  $\Gamma$ .

The five above items imply that  $p_{ij}$  is indeed a solution of (7.5) and furthermore the optimal Lagrange multipliers are determined.

Similarly for (10.4), taking a test function  $\psi$ , with discontinuous values  $\psi^A$  in  $Y^A$  and  $\psi^B$  in  $Y^B$ , after integrating by parts, we find

$$\begin{aligned} \left\langle \frac{\partial \mathcal{L}}{\partial \chi_i}, \psi \right\rangle &= \int_{Y^A} \left( \frac{\partial \mathcal{J}}{\partial \chi_i^A} - \operatorname{div}(a^A \nabla p_i^A) - \operatorname{div}(a^A p_{ij}^A e_j) - a^A \nabla p_{ij}^A \cdot e_j + \left( \int_Y p_{ij} \right) \operatorname{div}(a^A e_j) \right) \psi^A dy \\ &+ \int_{Y^B} \left( \frac{\partial \mathcal{J}}{\partial \chi_i^B} - \operatorname{div}(a^B \nabla p_i^B) - \operatorname{div}(a^B p_{ij}^B e_j) - a^B \nabla p_{ij}^B \cdot e_j + \left( \int_Y p_{ij} \right) \operatorname{div}(a^B e_j) \right) \psi^B dy \\ &+ 2 \int_{\Gamma} [a\psi p_{ij} e_j \cdot n] ds + \int_{\Gamma} ([\psi a \nabla p_i \cdot n] - [p_i a \nabla \psi \cdot n]) ds - \left( \int_Y p_{ij} \right) \int_{\Gamma} [\psi a e_j \cdot n] ds \\ &+ \int_{\Gamma} \lambda_i [\psi] ds + \int_{\Gamma} \gamma_i [a \nabla \psi \cdot n] ds - \int_{\Gamma} \gamma_{ij} [a \psi e_j \cdot n] ds = 0, \end{aligned} \quad (10.8)$$

where the terms in factor of  $(\int_Y p_{ij})$  come from the differentiation of  $\hat{a}_{ij}^*$ .

1. Taking  $\psi$  to have compact support in  $Y^A$ , or in  $Y^B$ , yields the bulk equation (7.4) away from  $\Gamma$ .
2. Assume  $a^A \nabla \psi^A \cdot n^A = a^B \nabla \psi^B \cdot n^B = 0$  and  $[\psi] = 0$  on  $\Gamma$ . Thus (10.8) reduces to

$$2 \int_{\Gamma} \psi [a p_{ij} e_j \cdot n] ds + \int_{\Gamma} \psi [a \nabla p_i \cdot n] ds - \left( \int_Y p_{ij} \right) \int_{\Gamma} \psi [a e_j \cdot n] ds - \int_{\Gamma} \gamma_{ij} \psi [a e_j \cdot n] ds = 0$$

for any value of the trace  $\psi = \psi^A = \psi^B$  on  $\Gamma$ . Since we already know that  $\gamma_{ij} = p_{ij}$ , this implies  $[a(\nabla p_i + p_{ij} e_j - (\int_Y p_{ij}) e_j) \cdot n] = 0$ .

3. Assume just  $a^A \nabla \psi^A \cdot n^A = a^B \nabla \psi^B \cdot n^B = 0$ . Then (10.8) reduces to

$$\int_{\Gamma} [\psi] \left( \lambda_i + a(\nabla p_i + p_{ij} e_j - (\int_Y p_{ij}) e_j) \cdot n \right) ds = 0,$$

for any value of the jump  $[\psi]$  on  $\Gamma$ . This gives the optimal value of the Lagrange multiplier  $\lambda_i = -a(\nabla p_i + p_{ij} e_j - (\int_Y p_{ij}) e_j) \cdot n$ .

4. Assume now  $[a\nabla\psi \cdot n] = 0$ . This time, (10.8) reduces to

$$\int_{\Gamma} a\nabla\psi \cdot n [p_i] ds = 0,$$

for any value of the trace  $a\nabla\psi \cdot n$ . Thus we find  $[p_i] = 0$  on  $\Gamma$ .

5. Finally if  $\psi$  does not satisfy any condition at the interface, (10.8) reduces to

$$\int_{\Gamma} [a\nabla\psi \cdot n] (\gamma_i - p_i) ds = 0,$$

which gives the optimal value of the Lagrange multiplier  $\gamma_i = p_i$  on  $\Gamma$ .

The five above items imply that  $p_i$  is indeed a solution of (7.4) and furthermore the optimal Lagrange multipliers are determined.

Finally, in a third and final step, we apply Lemma 7.2 to the computation of the partial derivative of the Lagrangian  $\mathcal{L}$  with respect to  $Y^A$  in the direction  $\theta$ . By a classical result (see e.g. Lemma 3.5 in [7]), we know that this partial derivative of  $\mathcal{L}$  is precisely the shape derivative of  $J$ ,

$$J'(Y^A)(\theta) = \left\langle \frac{\partial \mathcal{L}}{\partial \Gamma}, \theta \right\rangle \left( \Gamma, \{\chi_i^{A,B}\}, \{\chi_{ij}^{A,B}\}, \{p_i^{A,B}\}, \{p_{ij}^{A,B}\}, \{\lambda_i\}, \{\lambda_{ij}\}, \{\gamma_i\}, \{\gamma_{ij}\} \right)$$

when the right hand side is evaluated at the optimal state and adjoint solutions (i.e. without  $\hat{\cdot}$ ). It remains to compute the partial shape derivative of  $\mathcal{L}$  and to show that it is equal to the right hand side of (7.3).

As a preliminary step, we perform an integration by parts in the definition (10.3) of the Lagrangian to make it more symmetric. It yields

$$\begin{aligned} \mathcal{L} \left( \Gamma, \{\hat{\chi}_i^{A,B}\}, \{\hat{\chi}_{ij}^{A,B}\}, \{\hat{p}_i^{A,B}\}, \{\hat{p}_{ij}^{A,B}\}, \{\hat{\lambda}_i\}, \{\hat{\lambda}_{ij}\}, \{\hat{\gamma}_i\}, \{\hat{\gamma}_{ij}\} \right) &:= \int_{Y^A} \mathcal{J}^A dy + \int_{Y^B} \mathcal{J}^B dy \quad (10.9) \\ &+ \int_{Y^A} a^A (\nabla \hat{\chi}_i^A - e_i) \cdot \nabla \hat{p}_i^A dy + \int_{Y^B} a^B (\nabla \hat{\chi}_i^B - e_i) \cdot \nabla \hat{p}_i^B dy \\ &+ \int_{\Gamma} \hat{\lambda}_i (\hat{\chi}_i^A - \hat{\chi}_i^B) ds + \int_{\Gamma} \left( (\hat{\gamma}_i - \hat{p}_i^A) a^A (\nabla \hat{\chi}_i^A - e_i) \cdot n^A + (\hat{\gamma}_i - \hat{p}_i^B) a^B (\nabla \hat{\chi}_i^B - e_i) \cdot n^B \right) ds \\ &+ \int_{Y^A} \left( a^A (\nabla \hat{\chi}_{ij}^A - \hat{\chi}_j^A e_i) \cdot \nabla \hat{p}_{ij}^A + (-a_{ij}^A + \hat{a}_{ij}^* + a^A \nabla \hat{\chi}_j^A \cdot e_i) \hat{p}_{ij}^A \right) dy \\ &+ \int_{Y^B} \left( a^B (\nabla \hat{\chi}_{ij}^B - \hat{\chi}_j^B e_i) \cdot \nabla \hat{p}_{ij}^B + (-a_{ij}^B + \hat{a}_{ij}^* + a^B \nabla \hat{\chi}_j^B \cdot e_i) \hat{p}_{ij}^B \right) dy \\ &+ \int_{\Gamma} \hat{\lambda}_{ij} (\hat{\chi}_{ij}^A - \hat{\chi}_{ij}^B) ds + \int_{\Gamma} \left( (\hat{\gamma}_{ij} - \hat{p}_{ij}^A) a^A (\nabla \hat{\chi}_{ij}^A - \hat{\chi}_j^A e_i) \cdot n^A + (\hat{\gamma}_{ij} - \hat{p}_{ij}^B) a^B (\nabla \hat{\chi}_{ij}^B - \hat{\chi}_j^B e_i) \cdot n^B \right) ds. \end{aligned}$$

To obtain the shape derivative, Lemma 7.2 is applied to the Lagrangian (10.9) and the resulting expression is evaluated at the optimal states, adjoints and Lagrange parameters (i.e. without  $\hat{\cdot}$ ). This leads to

$$\begin{aligned} \left\langle \frac{\partial \mathcal{L}}{\partial \Gamma}, \theta \right\rangle &= \int_{\Gamma} (\mathcal{J}^A - \mathcal{J}^B) \theta \cdot n ds \quad (10.10) \\ &+ \int_{\Gamma} (a^A (\nabla \chi_i^A - e_i) \cdot \nabla p_i^A - a^B (\nabla \chi_i^B - e_i) \cdot \nabla p_i^B) \theta \cdot n ds \\ &+ \int_{\Gamma} \left( a^A (\nabla \chi_{ij}^A - \chi_j^A e_i) \cdot \nabla p_{ij}^A + (-a_{ij}^A + a_{ij}^* + a^A \nabla \chi_j^A \cdot e_i) p_{ij}^A \right) \theta \cdot n ds \\ &- \int_{\Gamma} \left( a^B (\nabla \chi_{ij}^B - \chi_j^B e_i) \cdot \nabla p_{ij}^B + (-a_{ij}^B + a_{ij}^* + a^B \nabla \chi_j^B \cdot e_i) p_{ij}^B \right) \theta \cdot n ds \\ &+ \int_{\Gamma} \lambda_i \frac{\partial (\chi_i^A - \chi_i^B)}{\partial n} \theta \cdot n ds \\ &+ \int_{\Gamma} \left( \frac{\partial (\gamma_i - p_i^A)}{\partial n} a^A (\nabla \chi_i^A - e_i) \cdot n^A + \frac{\partial (\gamma_i - p_i^B)}{\partial n} a^B (\nabla \chi_i^B - e_i) \cdot n^B \right) \theta \cdot n ds \\ &+ \int_{\Gamma} \lambda_{ij} \frac{\partial (\chi_{ij}^A - \chi_{ij}^B)}{\partial n} \theta \cdot n ds \\ &+ \int_{\Gamma} \left( \frac{\partial (\gamma_{ij} - p_{ij}^A)}{\partial n} a^A (\nabla \chi_{ij}^A - \chi_j^A e_i) \cdot n^A + \frac{\partial (\gamma_{ij} - p_{ij}^B)}{\partial n} a^B (\nabla \chi_{ij}^B - \chi_j^B e_i) \cdot n^B \right) \theta \cdot n ds, \end{aligned}$$

where we have already taken into account the continuity on  $\Gamma$  of the functions  $\chi$  and  $p$ , as well as the optimal values of the Lagrange parameters  $\gamma = p$  (in particular, it cancels all terms in factor of the mean curvature  $H$  in the shape derivatives of surface integrals). We simplify (10.10) by recalling the normal flux interface conditions

from (10.1) and (10.2), and the optimal value of  $\lambda$ . It yields

$$\begin{aligned}
\left\langle \frac{\partial \mathcal{L}}{\partial \Gamma}, \theta \right\rangle &= \int_{\Gamma} [\mathcal{J}] \theta \cdot n \, ds + \int_{\Gamma} [a(\nabla \chi_i - e_i) \cdot \nabla p_i] \theta \cdot n \, ds \\
&+ \int_{\Gamma} [a(\nabla \chi_{ij} - \chi_j e_i) \cdot \nabla p_{ij}] \theta \cdot n \, ds + \int_{\Gamma} [-a_{ij} + a \nabla \chi_j \cdot e_i] p_{ij} \theta \cdot n \, ds \\
&- \int_{\Gamma} a(\nabla p_i + p_{ij} e_j - (\int_Y p_{ij}) e_j) \cdot n \left[ \frac{\partial \chi_i}{\partial n} \right] \theta \cdot n \, ds \\
&- \int_{\Gamma} \left[ \frac{\partial p_i}{\partial n} \right] a(\nabla \chi_i - e_i) \cdot n \theta \cdot n \, ds \\
&- \int_{\Gamma} a \nabla p_{ij} \cdot n \left[ \frac{\partial \chi_{ij}}{\partial n} \right] \theta \cdot n \, ds \\
&- \int_{\Gamma} \left[ \frac{\partial p_{ij}}{\partial n} \right] a(\nabla \chi_{ij} - \chi_j e_i) \cdot n \theta \cdot n \, ds,
\end{aligned} \tag{10.11}$$

Several integrands in (10.11) are discontinuous across the boundary  $\Gamma$ . To make their values more precise, we rewrite them into continuous normal and tangential components, so the jumps appear only on the coefficient  $a$  and its inverse  $a^{-1}$ . For example

$$\begin{aligned}
[a(\nabla \chi_i - e_i) \cdot \nabla p_i] &= a(\nabla \chi_i - e_i) \cdot n \left[ \frac{\partial p_i}{\partial n} \right] + [a(\nabla \chi_i - e_i)_t] \cdot \nabla_t p_i \\
&= [a^{-1}] \left( a(\nabla \chi_i - e_i) \cdot n \right) \left( a \nabla p_i \cdot n \right) + [a] (\nabla_t \chi_i - e_i) \cdot \nabla_t p_i
\end{aligned}$$

and

$$[a(\nabla \chi_{ij} - \chi_j e_i) \cdot \nabla p_{ij}] = [a^{-1}] \left( a(\nabla \chi_{ij} - \chi_j e_i) \cdot n \right) \left( a \nabla p_{ij} \cdot n \right) + [a] (\nabla_t \chi_{ij} - \chi_j e_i) \cdot \nabla_t p_{ij}.$$

Some calculations leads to

$$\begin{aligned}
\left\langle \frac{\partial \mathcal{L}}{\partial \Gamma}, \theta \right\rangle &= \int_{\Gamma} [\mathcal{J}] \theta \cdot n \, ds + \int_{\Gamma} [a] (\nabla \chi_i - e_i)_t \cdot (\nabla p_i + e_i \tilde{p}_{ij})_t \theta \cdot n \, ds \\
&- \int_{\Gamma} [a^{-1}] \left( a(\nabla \chi_i - e_i) \cdot n \right) \left( a(\nabla p_i + e_j \tilde{p}_{ij}) \cdot n \right) \theta \cdot n \, ds \\
&+ \int_{\Gamma} [a] (\nabla \chi_{ij} - \chi_j e_i)_t \cdot \nabla_t p_{ij} \theta \cdot n \, ds \\
&- \int_{\Gamma} [a^{-1}] \left( a(\nabla \chi_{ij} - \chi_j e_i) \cdot n \right) \left( a \nabla p_{ij} \cdot n \right) \theta \cdot n \, ds,
\end{aligned} \tag{10.12}$$

with  $\tilde{p}_{ij} = p_{ij} - \int_Y p_{ij}$ . This finishes the proof of Theorem 7.3.

## Acknowledgement

The authors acknowledge the help of Professor Olivier Pantz (Université de Nice) in the implementation of their program in FreeFem++. T. Y. was able to conduct overseas research in France supported by JSPS postdoctoral fellowship for research abroad. The authors sincerely appreciate this support. G. A. is a member of the DEFI project at INRIA Saclay Ile-de-France.

## References

- [1] A. ABDULLE, M. GROTE, CH. STOHRER: *Finite element heterogeneous multiscale method for the wave equation: long-time effects*, Multiscale Model. Simul. **12** (2014), pp.1230-1257.
- [2] A. ABDULLE & T. N. POUCHON: *A priori error analysis of the finite element heterogeneous multiscale method for the wave equation in heterogeneous media over long time*, preprint 2015.

- [3] G. ALLAIRE: *Shape Optimization by the Homogenization Method*, Applied Mathematical Sciences **146**, Springer-Verlag, New-York, 2002, 456 pp.
- [4] G. ALLAIRE, M. BRIANE & M. VANNINATHAN: *A comparison between two-scale asymptotic expansions and Bloch wave expansions for the homogenization of periodic structures*, to appear in SEMA Journal. HAL preprint: hal-01215580 (October 2015).
- [5] G. ALLAIRE, C. DAPOGNY, G. DELGADO, G. MICHAELIDIS: *Multi-phase structural optimization via a level set method*, ESAIM: Control, Optimisation and Calculus of Variations, **20** (2014), 576–611.
- [6] G. ALLAIRE, CH. DAPOGNY, P. FREY: *Shape optimization with a level set based mesh evolution method*, CMAME **282**, 22-53 (2014).
- [7] G. ALLAIRE, F. JOUVE, N.V. GOETHEM: *Damage and fracture evolution in brittle materials by shape optimization methods*, J. of Comput. Phys., **230** (2011), 5010–5044.
- [8] G. ALLAIRE, F. JOUVE, A.-M. TOADER: *Structural optimization using sensitivity analysis and a level set method*, J. Comp. Phys., Vol 194/1, 363–393 (2004).
- [9] G. ALLAIRE, M. PALOMBARO, J. RAUCH: *Diffractive Geometric Optics for Bloch Wave Packets*, Archive Rat. Mech. Anal., **202**, pp.373-426 (2011).
- [10] N. BAKHVALOV, G. PANASENKO: *Homogenization: Averaging Processes in Periodic Media*, Kluwer, Dordrecht (1989).
- [11] C. BARBAROSIE: *Shape optimization of periodic structures*, Computational Mechanics, **30**, pp.235-246 (2003).
- [12] A. BENSOUSSAN, J.-L. LIONS & G. PAPANICOLAOU: *Asymptotic analysis for periodic structures*, corrected reprint of the 1978 original, AMS Chelsea Publishing, Providence, RI, 2011, 398 pp.
- [13] C. BERNARDI, O. PIRONNEAU: *Sensitivity of Darcy's law to discontinuities*, Chinese Annals of Mathematics. **24**, pp. 205–214 (2003).
- [14] C. I. CHRISTOV, G. A. MAUGIN & M. G. VELARDE: *Well-posed Boussinesq paradigm with purely spatial higher-order derivatives*, Phys. Rev. E, **54**, pp.3621-3638 (1996).
- [15] C. CONCA AND M. VANNINATHAN: *Homogenization of periodic structures via Bloch decomposition*, SIAM J. Appl. Math., **57** (1997), 1639-1659.
- [16] C. CONCA, R. ORIVE & M. VANNINATHAN: *Bloch approximation in homogenization and applications*, SIAM J. Math. Anal., **33** (5) (2002), 1166-1198.
- [17] C. CONCA, R. ORIVE & M. VANNINATHAN: *On Burnett coefficients in periodic media*, J. Math. Phys., **47** (3) (2006), 11 pp.
- [18] C. CONCA, J. SAN MARTÍN, L. SMARANDA & M. VANNINATHAN: *On Burnett coefficients in periodic media in low contrast regime*, J. Math. Phys., **49** (5) (2008), 23 pp.
- [19] C. CONCA, J. SAN MARTÍN, L. SMARANDA & M. VANNINATHAN: *Optimal bounds on dispersion coefficient in one-dimensional periodic media*, Math. Models Methods Appl. Sci., **19** (9) (2009), 1743-1764.
- [20] C. CONCA, J. SAN MARTÍN, L. SMARANDA & M. VANNINATHAN: *Burnett coefficients and laminates*, Appl. Anal., **91** (6) (2012), 1155-1176.
- [21] T. DOHNAL, A. LAMACZ & B. SCHWEIZER: *Bloch-wave homogenization on large time scales and dispersive effective wave equations*, Multiscale Model. Simul., **12** (2) (2014), 488-513.
- [22] J. FISH, W. CHEN, & G. NAGAI: *Non-local dispersive model for wave propagation in heterogeneous media: multi-dimensional case*, Internat. J. Numer. Methods Engrg., **54**(3):347-363, 2002.
- [23] Z. HASHIN, S. SHTRIKMAN: *A variational approach to the theory of the elastic behaviour of multiphase materials* Journal of the Mechanics and Physics of Solids. **11**, pp.127–140 (1963).
- [24] J. HASLINGER, J. DVORAK: *Optimum composite material design*, RAIRO M2AN, **29**, pp.657-686 (1995).

- [25] F. HECHT: *New development in FreeFem++*, J. Numer. Math. 20 (2012), no. 3-4, 251-265.
- [26] F. HETTLICH, W. RUNDELL: *The determination of a discontinuity in a conductivity from a single boundary measurement*, Inverse Problems. **14**, p. 67 (1998).
- [27] H. ISAKARI, K. KURIYAMA, S. HARADA, T. YAMADA, T. TAKAHASHI, T. MATSUMOTO: *A topology optimisation for three-dimensional acoustics with the level set method and the fast multipole boundary element method*, Mechanical Engineering Journal, **1**, p. CM0039 (2014).
- [28] V. JIKOV, S. KOZLOV, O. OLEINIK: *Homogenization of Differential Operators and Integral Functionals*; Springer Verlag: New York, Berlin, 1994.
- [29] A.L. KARCHEVSKY, *Reconstruction of pressure velocities and boundaries of thin layers in thinly-stratified layers*, J. Inverse Ill-Posed Probl. 18 (2010), no. 4, 371-388.
- [30] A. LAMACZ: *Dispersive effective models for waves in heterogeneous media*, Math. Models Methods Appl. Sci., **21** (9) (2011), 1871-1899.
- [31] M. LOMBARDO, H. ASKES: *Elastic wave dispersion in microstructured membranes*, Proc. R. Soc. A (2010) **466**, pp.1789-1807.
- [32] K. MCKINNON, N. SMYTH, A. WORTHY: *Optimization of soliton amplitude in dispersion-decreasing nonlinear optical fibers*, J. Opt. Soc. Amer. B Opt. Phys. 16 (1999), no. 3, 441-447.
- [33] Y. MEI AND X. WANG, *A level set method for structural topology optimization with multi-constraints and multi-materials*, Acta Mechanica Sinica, Vol.20, No.5, (2004).
- [34] B. MOHAMMADI, O. PIRONNEAU: *Applied Shape Optimization for Fluids*, Oxford University Press, Oxford **28** (2001).
- [35] S. MOSKOW & M. VOGELIUS: *First order corrections to the homogenized eigenvalues of a periodic composite medium. A convergence proof*, Proc. Roy. Soc. Edinburg, 127, pp.1263-1295 (1997).
- [36] F. MURAT: *H-convergence*, Séminaire d'Analyse Fonctionnelle et Numérique, 1977-78, Université d'Alger, multicopied, 34 pp. English translation : F. MURAT & L. TARTAR, *H-convergence*, Topics in the Mathematical Modelling of Composite Materials, ed. by L. Cherkaev & R.V. Kohn, Progress in Nonlinear Differential Equations and their Applications, **31**, Birkhäuser, Boston (1998), 21-43.
- [37] F. MURAT J. SIMON: *Sur le contrôle par un domaine géométrique*, Technical Report RR-76015, Laboratoire d'Analyse Numérique, Université Pierre et Marie Curie (1976).
- [38] S. OSHER, F. SANTOSA: *Level set methods for optimization problems involving geometry and constraints: frequencies of a two-density inhomogeneous drum*. J. Comp. Phys., **171**, 272-288 (2001).
- [39] S. OSHER, J.A. SETHIAN: *Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations*, J. of Comput. Phys., **79** (1) (1988), 12-49.
- [40] O. PANTZ: *Sensibilité de l'équation de la chaleur aux sauts de conductivité*, C.R. Acad. Dci. Paris, Ser. 1, **341** (2005), 333-337.
- [41] O. PIRONNEAU: *Optimal Shape Design for Elliptic Systems*, Springer-Verlag, New York (1984).
- [42] J. RAUCH: *Hyperbolic partial differential equations and geometrics optics*, volume 133 of *Graduate Studies in Mathematics*, American Mathematical Society (2012).
- [43] M. REED, B. SIMON: *Methods of modern mathematical physics*, Academic Press, New York (1978).
- [44] J. RIISHEDE, O. SIGMUND: *Inverse design of dispersion compensating optical fiber using topology optimization*, J. Opt. Soc. Amer. B Opt. Phys. 25 (2008), no. 1, 88-97.
- [45] E. SANCHEZ-PALENCIA: *Non-Homogeneous Media and Vibration Theory* Springer Lecture Notes in Physics **129** (1980).
- [46] F. SANTOSA, W. SYMES: *A dispersive effective medium for wave propagation in periodic composites*, SIAM J. Appl. Math., **51** (1991), 984-1005.

- [47] O. SIGMUND: *Materials with prescribed constitutive parameters: an inverse homogenization problem*, Int. J. Solids Struct. 31, pp.2313-2329 (1994).
- [48] V. P. SMYSHLYAEV, K. D. CHEREDNICHENKO: *On rigorous derivation of strain gradient effects in the overall behaviour of periodic heterogeneous media*, The J. R. Willis 60th anniversary volume. *J. Mech. Phys. Solids*, **48** (6-7) (2000), 1325-1357.
- [49] J. SOKOLOWSKI, J.P. ZOLESIO: *Introduction to Shape Optimization: Shape Sensitivity Analysis*, Springer Berlin Heidelberg, Berlin (1992).
- [50] R. STAINKO, O. SIGMUND: *Tailoring dispersion properties of photonic crystal waveguides by topology optimization*, Waves Random Complex Media 17 (2007), no. 4, 477–489.
- [51] L. TARTAR: *The General Theory of Homogenization: A Personalized Introduction*, Lecture Notes of the Unione Matematica Italiana, Springer-Verlag, Berlin Heidelberg 2009, pp. 471.
- [52] M.Y. WANG, X. WANG, D. GUO: *A level set method for structural topology optimization*, Comput. Methods Appl. Mech. Engrg., **192**, 227-246 (2003).
- [53] M. WANG AND X. WANG, *Color level sets: a multi-phase method for structural topology optimization with multiple materials*, Comput. Methods Appl. Mech. Engrg. 193, 469-496 (2004).
- [54] C. WILCOX: *Theory of Bloch waves*, J. Anal. Math. **33**, pp. 146–167 (1978).